First published: 10 August 2016
Search executed: Dec. 2013 – Jan. 2015

Interventions to Improve the Labour Market Outcomes of Youth: A Systematic Review of Training, Entrepreneurship Promotion, Employment Services and Subsidized Employment Interventions[1]

Jochen Kluve, Susana Puerto, David Robalino, Jose Manuel Romero, Friederike Rother, Jonathan Stöterau, Felix Weidenkaff, Marc Witte

[1] Note: This is an unedited version of a forthcoming publication by the International Labour Organization (ILO).

Title	Interventions to improve the labour market outcomes of youth: A systematic review of training, entrepreneurship promotion, employment services and subsidized employment interventions
Authors	Kluve, Jochen[1] Puerto, Susana[2] Robalino, David[3] Romero, Jose Manuel[3] Rother, Friederike[3] Stöterau, Jonathan[1] Weidenkaff, Felix[2] Witte, Marc[4] [1] Humboldt University Berlin and RWI [2] International Labour Organization [3] The World Bank [4] University of Oxford
No. of pages	293
Last updated	25 April 2016
Citation	Kluve J, Puerto S, Robalino D, Romero J M, Rother F, Stöterau J, Weidenkaff F, Witte M. (2016). Interventions to improve the labour market outcomes of youth: a systematic review of training, entrepreneurship promotion, employment services, and subsidized employment interventions.
Copyright	
Roles and responsibilities	Susana Puerto will be responsible for updating this review as additional evidence and resources become available.
Sources of support	• International Initiative for Impact Evaluation (3ie): Systematic Review Call 5 through funding from the Canadian International Development Agency (CIDA) • The Youth Employment Network • The International Labour Organization • The World Bank • Leibniz Association
Declarations of interest	The authors have no vested interest in the outcomes of this review, nor any incentive to represent findings in a biased manner.
Corresponding authors	Susana Puerto International Labour Organization Switzerland E-mail: puerto-gonzalez@ilo.org Jochen Kluve Humboldt University of Berlin and RWI Germany E-mail: jochen.kluve@hu-berlin.de Full list of author information is available at the end of the article

Table of contents

TABLES

BOXES

FIGURES

Abstract

Background – Today's labour market is a challenging arena for young people. Over 73 million youth are currently unemployed and many more are affected by vulnerable employment and working poverty. Youth remain highly susceptible to changing patterns in the world of work and experience slow and difficult transitions to stable jobs. What works to support them in the labour market? This is one of the most common and pressing questions posed by policymakers and practitioners today.

Methods – This systematic review addresses this question by synthesizing empirical evidence on the labour market outcomes of youth employment interventions worldwide. Interventions comprised skills training, entrepreneurship promotion, employment services, and subsized employment. Outcomes of interest included employment, earnings and business performance. A comprehensive systematic search for relevant evidence across more than 70 sources, using search terms in English, French, German, Portuguese and Spanish, identified over 30,000 records that were screened according to the review's inclusion criteria. For the selected studies that met the inclusion criteria, treatment effect estimates were coded and standardized mean differences (SMDs) were computed. The search process was completed in January 2015. The analysis explores the interventions' overall effectiveness and the roles that context, evaluation and programme design and implementation play in determining impact.

Results – A total of 113 counterfactual-based impact evaluations were identified, encompassing a unique set of evaluation methods, interventions and geographical coverage. Meta-analysis methods were employed to synthesize the evidence, based on 2,259 imputed effect sizes (SMDs). Overall, empirical results indicate positive treatment effects that are statistically different from zero on labour market outcomes. In other words, investing in young people through active labour market programmes (ALMPs) pays off with positive impacts particularly on employment and earnings outcomes. This impact does not take effect immediately and is more pronounced among low- and middle-income countries than among high-income countries.

Implications – Active measures to support the (re)intregration of young women and men into the labour market succed in enhancing employment and earnings

outcomes and have potential to increase human capital and employment prospects in the long-term. The evidence suggests that the type and design of youth employment intervention is important and is strongly influenced by the income level of the country. Programme evaluations generally show larger effect sizes in low- and middle-income countries than in high-income countries. Multi-pronged measures are shown to be effective in tackling the many barriers to success facing youth in the labour market, particularly in low- and middle-income countries, where skills training and entrepreneurship interventions have prompted significant improvements in youth employment and earnings outcomes. Targeting towards disadvantaged youth and features such as participant profiling, participant engagement mechanisms and incentives for service providers are positively correlated with a larger magnitude of impact. The results appear robust in terms of the quality of the underlying evidence. The review did not find differential treatment effects by gender or age.

Executive summary

BACKGROUND

The youth of today represent a vast potential for inclusive growth and development. If youth are given the opportunity to build appropriate skills and access decent employment, they can help to accelerate progress on the 2030 Agenda for Sustainable Development and engage in meaningful work that benefits them, their families and society as a whole.

Unfortunately, decent jobs are not a feasible prospect for all young women and men. Today, over 73 million young people are unemployed worldwide. Youth unemployment stands at a much higher level than the average unemployment rate for adults, in some cases over three times as high. Moreover, two out of five young people in the labour force are either working but poor or unemployed. The youth employment challenge is therefore not only about job creation, but also – and especially – about enhancing the quality of jobs for youth.

Youth's gloomy prospects in the labour market embody a massive waste of potential and a threat to social cohesion. Understanding what works to improve their labour market outcomes is therefore of paramount importance and a development priority for all countries and regions.

OBJECTIVES

The aim of this systematic review was to investigate the impact of youth employment interventions on the labour market outcomes of young people. The review looked at the available evaluation evidence in a systematic and rigorous manner in order to fill the knowledge gaps relating to the effectiveness of the various types of youth-targeted interventions in different contexts. The interventions under review comprised training and skills development, entrepreneurship promotion, employment services and subsidized employment. Outcomes of interest included an array of measures across employment, earnings and business performance outcomes.

The following research questions provided the framework for the analysis to establish which measures were effective – a process that will ultimately help decision-makers in determining their youth employment portfolio:

1. What are the impacts of youth employment interventions on the labour market outcomes of youth?

2. Which of these interventions are most effective?

SEARCH METHODS

The search for relevant literature was based on a variety of sources in order to ensure that both published and unpublished studies ("grey literature") relevant to the research question were included in the search process. The search process included both a primary search (i.e., searching of a wide range of general and specialized databases) and a complementary search (i.e., hand-searching of relevant websites, searching of dissertations, theses and grey literature databases, citation tracking, screening of reference lists and contacting authors and experts). The in-depth complementary search was an asset to the systematic review and allowed the identification of several non-published studies. The process included search terms in English, French, German, Portuguese and Spanish. For each source, the review team tested and documented several search strategies and identified one or more preferred search strategies, which yielded a comprehensive and precise set of potentially relevant results. The search process was completed in January 2015.

SELECTION CRITERIA

To be eligible for inclusion in the review, studies must have:

1. *evaluated an active labour market programme* (ALMP) that included at least one of the following categories of interventions: training and skills development, entrepreneurship promotion, employment services and/or subsidized employment;
2. *investigated programmes that were designed for – or targeted primarily – young women and men* aged between 15 and 35;
3. *reflected completed experimental and quasi-experimental evaluations* measuring impacts on eligible labour market outcomes; and
4. *reported at least one eligible outcome variable* measuring employment (e.g., probability of employment, hours worked, duration in unemployment), earnings (e.g., reported earnings, wages, consumption) or business performance (e.g., profits, sales).

In addition to the above inclusion criteria, the review focused on studies with a publication date between 1990 and 2014. No language restrictions were applied.

DATA COLLECTION AND ANALYSIS

The review relied on a comprehensive systematic search across more than 70 sources, including literature databases and a large number of websites, which allowed the identification of grey literature. A coding tool and manual were developed in order to guide a harmonised data extraction process. Treatment effect estimates were coded across all studies that met the inclusion criteria, along with other parameters and intervention characteristics deemed relevant for the analysis. Additional, non-reported information was retrieved from authors of the primary studies, supporting the computation of standardized mean differences (SMDs) effect sizes. The SMDs captured the relative magnitude of the treatment effect in a dimensionless way, which was therefore comparable across outcomes and studies. Effect sizes were summarized within and across reports to one effect size per outcome for each study.

Random-effects meta-analysis methods were employed to synthesize and compare effect sizes reported in the primary studies. Subsequently, multivariate meta-regression models were estimated and information about intervention-level, study-level and country-level characteristics were included to assess factors associated with the magnitude of reported effect size estimates.

RESULTS

The primary and complementary searches identified 32,117 records, of which a total of 1,141 records were selected for full text screening. The subsequent selection process led to a sample of 113 reports which were considered to be of adequate content and methodological rigour for inclusion in the meta-analysis.

The 113 reports represented 107 interventions. The evidence base spanned 31 countries and covered 55 skills training interventions, 15 entrepreneurship promotion interventions, ten employment services interventions and 21 subsidized employment interventions. There were six interventions for which no clear main category of intervention could be established. A large share of the evidence derived from recent publications, with nearly half of the sample produced after 2010. Evaluation designs varied, with 47 per cent of reports relying on experimental designs, 10 per cent on natural experiments and 44 per cent on quasi-experimental evaluations. Many of the most recent studies were experimental evaluations of interventions implemented in low- and middle-income countries, notably from Africa and Latin America and the Caribbean.

The comprehensive systematic search led to the identification and coding of a total of 3,629 treatment effect estimates. These estimates, along with further information reported and/or retrieved from authors of the primary studies and imputation of missing information, allowed the computation of 2,259 SMDs. Table 1 reports SMDs and other key parameters across outcomes and interventions.

Table 1: Summary of results by main category of intervention

Interventions	Standardized mean difference	Standard effect size errors	95% confidence interval		I^2	Number of SMDs	Number of interventions	Sample size	Mean difference	Control outcome	Treatment outcome	Percentage change
Employment outcomes												
Skills training	0.05	0.01	0.02	0.07	64.77	904	67	3 439 703	7.78	38.74	38.1	0.08
Entrepreneurship promotion	0.16	0.05	0.06	0.26	71.41	43	7	61 502	3.11	6.14	7.61	0.24
Employment services	0.01	0.01	-0.02	0.04	0	104	10	2 340 789	-1.09	19.16	14.41	0.02
Subsidized employment	0.02	0.02	-0.01	0.06	50	193	16	32 198 189	0.51	9.76	6.4	0
Unspecified	0.03	0.04	-0.04	0.1	0	86	5	178 863	-2.65	53.56	48.89	0.1
Overall	**0.04**	**0.01**	**0.03**	**0.06**	**63.66**	**1 330**	**105**	**38 219 046**	**4.67**	**30.65**	**29.14**	**0.07**
Earnings outcomes												
Skills training	0.07	0.01	0.05	0.08	85.7	495	60	2 045 960	3,781	15 488.55	17 274.87	0.11
Entrepreneurship promotion	0.09	0.04	0.01	0.18	63.81	50	12	42 530	2,347	2 744	3 640	0.22
Employment services	0.01	0.01	0	0.02	0	36	8	194 713	-17	226	186	0
Subsidized employment	-0.01	0.02	-0.05	0.03	61.24	57	9	10 358 155	-1524.22	7 057	5 536	-0.01
Unspecified	*Dropped from analysis*											
Overall	**0.05**	**0.01**	**0.03**	**0.06**	**81.64**	**670**	**92**	**12 696 812**	**2 084.35**	**10 355.27**	**11 210.02**	**0.09**
Business performance outcomes												
Skills training	*Dropped from analysis*											
Entrepreneurship promotion	0.1	0.05	0	0.19	39.09	162	10	58 519	47.72	371.77	395.75	0.15
Overall	**0.03**	**0.04**	**-0.05**	**0.12**	**48.83**	**169**	**14**	**62 905**	**-109.75**	**499.91**	**349.6**	**-0.06**

These are some of the key results from the systematic review and meta-analysis:

1. **Investing in youth through active labour market measures pays off.** Youth employment interventions lead to positive outcomes, increasing employment and earnings of participating youth. The positive effect on employment was captured by an overall SMD effect size of 0.04 – with a confidence interval (CI) of (0.03, 0.06) and a percentage change of the intervention over the control group mean (labelled pc for brevity) of 0.07, demonstrating that young people who were exposed to a youth employment intervention had better employment outcomes than those who were not.

   Impacts on earnings were also positive and statistically significant, with an effect size of 0.05 SMD (CI = 0.03, 0.06 and pc = 0.09), proving the responsiveness of earnings outcomes to measures that activate youth in the labour market. Evidence of youth employment programme effects on business performance outcomes was limited and the effect size was not statistically significant (SMD = 0.03; CI = -0.05, 0.12; pc = -0.06); however, when entrepreneurship promotion interventions were considered in isolation, the impact was larger and significant, at 0.10 SMD (CI = 0, 0.19 and pc = 0.15).

2. **It is, however, an investment that needs time to grow.** Youth employment interventions are an investment of long time horizon. Positive changes in labour market outcomes are more often observed a year after exposure to the intervention. This finding highlights the importance of regarding youth employment interventions as a long-term investment and calls for programme monitoring and evaluation that tracks outcomes into the future.

3. **Programme impacts conceal major contextual differences.** The meta-analysis showed important differences in the magnitudes of impact across outcomes and interventions. Despite the strong similarities across included studies, the differences in impact were not always driven by chance. Tests for heterogeneity demonstrated substantial variation in the effect size magnitude due to:

   - Country context,
   - Intervention design, and
   - Profile and characteristics of programme beneficiaries.

4. **The underlying evidence base varies by country income level.** Intervention characteristics and research designs differ significantly between high-income and low- or middle-income countries. A large proportion of the evidence from high-income countries derived from quasi-experimental studies of national programmes, implemented in collaboration with government

organizations. In contrast, the evidence from low- and middle-income countries was predominantly based on experimental impact evaluations of rather small-scale, targeted interventions, which were often implemented by NGOs or international organizations.

5. **Impact is higher in low- or middle-income countries than in high-income countries.** This result holds for employment and earnings outcomes and after controlling for study design. It points to a factual difference across country contexts: Being unemployed or unskilled in a high-income country – where labour demand is skill intensive – puts youth at a distinct disadvantage in comparison to a cohort that is, on average, well educated. While ALMPs help these youth to (re)connect to the labour market, they do not fully compensate for any failure to acquire knowledge or skills earlier in the education system. In lower income countries, with large cohorts of disadvantaged youth, marginal investments in skills and employment opportunities lead to larger changes in outcomes.

6. **In *low- and middle-income countries*, skills training and entrepreneurship interventions produce the greatest impacts.** Skills training and entrepreneurship promotion interventions yield positive results, on average, especially in terms of income gains. This is an important finding, which highlights the merits of combining both supply- and demand-side interventions to support youth. It also provides tangible evidence about the effect of interventions that aim to build human capital.

   Despite their large effect, entrepreneurship promotion interventions also exhibited great variability (see wide confidence intervals in Table 1), which calls for further research to enhance the consistency of the results from this intervention type.

7. **In *low-and middle-income countries*, measures that provide multiple services and programme components to youth lead to better outcomes.** Interventions that combine different components report higher outcomes than otherwise. The same relationship is not seen among interventions implemented in high-income countries.

8. **In *high-income countries*, the role of intervention types is less decisive.** Based on the random-effects meta-regression, no single type of intervention provided clear evidence of a significant effect on the employment or earnings outcomes of youth in high-income countries. Skills training appeared slightly more likely to effectuate some (albeit small) impact on employment or earnings, but the difference in comparison to other intervention types was generally not significant.

9. **Intervention design matters and tends to drive results much more than the type of intervention. Participant profiling, monitored programme participation and incentives offered to programme participants and service providers are key determinants of success**. In other words, the "how" seemed to be more important than the "what". This was particularly pronounced in high-income countries where the presence of a system to profile participants, mechanisms to engage participants or incentives for service providers to perform efficiently and achieve targets yielded larger impacts.

10. **Targeting the most disadvantaged youth increases programme effect.** Across measures of targeting, a focus on low-income youth, those with low levels of education or exhibiting strong disadvantages in the labour market triggers higher employment and earnings gains for youth across all country income levels. **There are however no differential effects by gender.** The analysis by gender generates less conclusive findings. While the overall effect size for employment and earnings appears to be larger for young women than for young men, the review discerned no strong patterns in the multivariate regression analysis to suggest that targeting women only will lead to better outcomes.

11. **There is no clear indication about the role of soft skills or that of different implementing actors.** The systematic review captured information about the type of skills delivered to young people and found no particular connection between soft skills and better labour market outcomes. Similarly, there was no systematic evidence about the role of public, private or civil entities in the implementation of a youth employment programme. More impact research is needed to account for these design and implementation features.

12. **The above results appear robust in terms of the quality of the underlying evidence**, as well as across different assumptions and model specifications. Most importantly, they held up under a restricted sample of experimental impact evaluations. While there was some evidence of publication (or small-sample) bias, it was controlled for in the meta-regression analysis, demonstrating that the results were not driven by sample composition bias.

CONCLUSIONS

The extent and urgency of the youth employment challenge and the level of global attention currently being given to this topic calls for more and better evidence-based action. Accordingly, this systematic review sought to examine the empirical evidence in order to understand what drives the success (or failure) of youth employment

interventions. Investments in youth employment will continue, and even increase, as countries embark on the implementation of the 2030 Agenda for Sustainable Development; therefore, this review focused on identifying "what works" and, as far as possible, "how".

Achieving an understanding of the "how" element is not an easy task. Frequently, impact evaluations do not assess relative effectiveness and, even more often, reports and papers fail to describe the underlying theory of change and observed transmission mechanisms behind an intervention. In some other cases, there is limited information about the characteristics of programme participants in the evaluation sample and their comparison group as well as fractional discussions about the occurrence and causes of attrition. Much remains to be done to improve reporting standards and advocate for more and better evidence examining the impact of youth employment interventions. The quality of the primary studies determines the quality of the systematic review and any subsequent synthesis of the evidence.

This systematic review builds on a growing base of studies measuring the impact of youth employment interventions and offers a rigorous synthesis and overall balance of empirical evidence taking into account the quality of the underlying research. The review is systematic through a clearly defined and transparent inclusion and exclusion criteria, an objective and extensive search, a punctual data extraction process, a standardized statistical testing and analysis, and a thorough reporting of findings. These elements and underlying methods and tools were laid out and reviewed in the protocol (Kluve et al. 2014).

The review shows that youth employment interventions do improve labour market outcomes of youth. It is an investment that pays off, but not immediately. A comparison of short- vs. long-term estimates indicates that outcomes are greater when measured at least a year after exposure to the intervention. While this result applies globally, it also conceals important contextual differences between developed and developing countries.

The evidence shows a significant impact gap across country income levels. While ALMPs in high-income countries can integrate disadvantaged young people into the labour market, they are not able to fully compensate for a lack of skills or other areas where youth failed to gain sufficient benefit from the education system. On the other hand, youth-targeted ALMPs in low- and middle-income countries do lead to meaningful impacts on both employment and earnings outcomes. Specifically, skills training and entrepreneurship promotion interventions appear to yield positive results on average. This is an important finding, which points to the potential benefits of combining supply- and demand-side interventions to support youth in the labour market.

The evidence also calls for careful design of youth employment interventions. The "how" seems to be more important than the "what" and, in this regard, targeting

disadvantaged youth as well as providing incentives for participation of youth, appropriate profiling mechanisms and schemes to motivate service providers to perform effectively appear to act as key factors of success.

The later emphasises the ability of specific design features within employment interventions to affect individual behaviours – in this case among both young people and service providers. It also implies – and calls for – sensible interpretation of the results. The findings from this review need to be discussed vis-à-vis the local and national context and should be complemented by a long-term and holistic commitment towards youth development.

The review also supported the identification of important evidence gaps:

- More and better evidence is needed on employment services, wage subsidies and public employment programmes for youth, particularly in low- and middle-income countries. It is important to note as well that despite the large and significant magnitude of effect of entrepreneurship promotion interventions, the evidence base is still limited and exhibits high variance. Therefore, more evidence will enhance the accuracy of the synthetized results.

- While the review highlighted a growing evaluation evidence from youth employment programmes implemented in Sub-Saharan Africa, it also reported very limited information from the Middle East and North Africa, South Asia and East Asia and the Pacific. These are regions were more targeted action to expand the evidence base should be considered.

- Similarly, more research is needed on intermediate outcomes. This is linked to the importance of improving research reporting standards and expanding the scope of outcomes of interest in order to better synthesize evidence about how interventions impact knowledge, skills, attitudes, and behaviours. More and better information on these intermediate outcomes will improve overall understanding about the causality and pathways of change between the intervention and the final outcomes.

- Soft skills are highly demanded by employers today. Their role in generating better outcomes is yet to be corroborated and more inquiry is required to understand their role in the causal chain as well as their interaction with more technical skills sets.

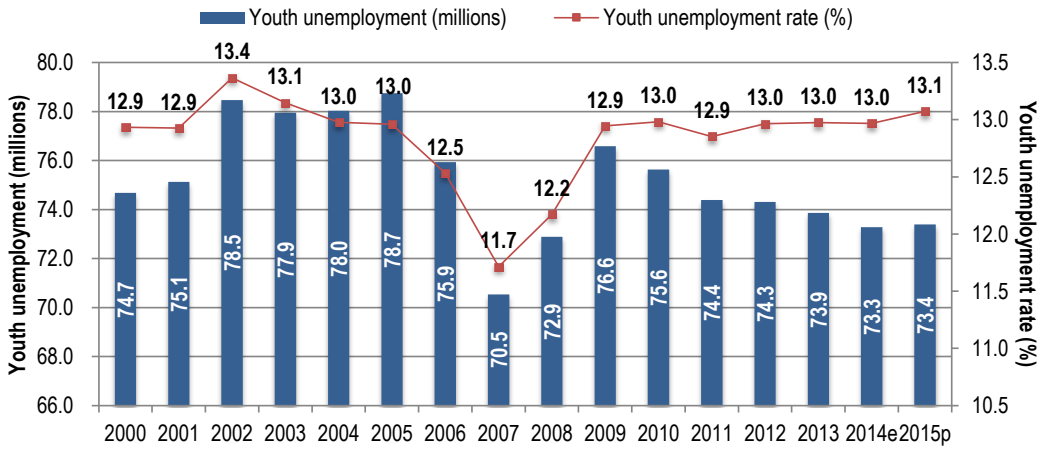
- Last, despite the rigorousness of the systematic review, findings will still be incomplete without the availability of cost information. The applicability of the evidence hinges not only on its internal and external validity but also on its feasibility. More information is needed on programme costs as well as systematic comparisons against programme effects. What may look highly effective may in fact be too expensive to replicate or scale up.

1 Background

1.1 THE RESEARCH PROBLEM: WHY YOUTH EMPLOYMENT?

The economic crisis brought about a swift reversal of the gradual declining trend in global youth unemployment rates observed between 2002 and 2007. The rapid increase in youth unemployment between 2007 and 2010 led to youth's discouragement and withdrawal in significant numbers from the labour force. It is estimated that nearly 6.4 million youth worldwide moved into inactivity in response to the crisis while many others continue to work yet live in poverty (ILO, 2012).

Figure 1: Global youth unemployment and unemployment rate, 2000–2015p



Source: Authors, based on ILO, *Trends Econometric Models*, April 2015, e = estimate; p = projection.

The youth employment crisis has become a stubbornly persistent reality in all regions and in nearly every country. Of the estimated 200 million unemployed people today, about 37 per cent – more than 73 million – are between the ages of 15 and 24. This translates into a global youth unemployment rate that has settled at 13.0 per cent during the period 2012 to 2014. While it is expected to remain relatively constant in the near future, this rate is still well above its pre-crisis level of 11.7 per cent (see Figure 1).

According to the 2015 *Global Employment Trends for Youth* report of the International Labour Organization, youth remain overrepresented among the unemployed and shaken by the changing patterns in the labour market. Two-fifths

(42.6 per cent) of the global youth labour force were reported as being unemployed or in working poverty in 2013. Regional youth unemployment trends remain fairly mixed. Most notably, the youth unemployment rates in the Middle East and North Africa (MENA) continue to be the highest worldwide, at 28.2 and 30.5 per cent for 2014, respectively. These figures stand out in comparison to other regions where the unemployment rate ranges from 10 to 20 per cent. In spite of the important achievements in boosting access to education and levels of educational attainment in the MENA region, today more than one in four active youth do not have a job (ILO, 2015a).

After being hit hard by the economic crisis, youth unemployment levels in Developed Economies and the European Union have seen some recent regional improvements, with the youth unemployment rate decreasing from 18.0 to 16.6 per cent, between 2012 and 2014. However, these improvements mask some difficult macroeconomic dynamics in certain countries, which are currently being further aggravated by conflict-driven migration. Six countries stand out in this respect, with unemployment rates of over 30 per cent, namely Croatia, Cyprus, Greece, Italy, Portugal and Spain.

Asian regions and sub-Saharan Africa continue to present relatively low unemployment rates among youth, although these statistics are all too often a reflection of the fact that youth cannot afford not to work and, as a matter of necessity, engage in poor quality and insecure jobs.

The challenge is not trivial since the "demographic dividend" can become a source of instability if young people around the world continue to face disappointing prospects in their job search. Unemployment depreciates human capital and has a significant negative influence on health, happiness, crime levels and socio-political stability (Bell and Blanchflower, 2009). Failing to address unemployment and underemployment among youth may contribute to the loss of human capital and an increase in social discontent.

Addressing the youth employment challenge continues to rank high in both international and local development priorities. The 2030 Agenda for Sustainable Development has placed the importance and urgency of achieving full and productive employment and decent work for all squarely at the centre of the new development vision, with youth explicitly identified as a key target group (Box 1).

It is therefore crucial to gather evidence to support the implementation of the 2030 Agenda. Yet very few rigorous overview and cross-country studies review and analyse the impact of youth employment programmes and what determines their success in different contexts. Even though the number of single-programme evaluations providing rigorous evidence on the effectiveness of active labour market programmes (ALMPs) has increased over the past decade, many fundamental questions remain unaddressed – particularly regarding the key issues: Which programmes work for a given target group, and under what circumstances? What

are the crucial design features necessary for youth employment programmes to be effective?

Box 1: Youth employment and the 2030 Agenda for Sustainable Development

Sustainable Development Goal 8 and a series of aspirational targets recognize today's employment challenge and open pathways for specific action on youth employment.

Key youth employment related targets in the 2030 Agenda:

- 4.4: "By 2030, substantially increase the number of youth and adults who have relevant skills, including technical and vocational skills, for employment, decent jobs and entrepreneurship";
- 8.3: "Promote development-oriented policies that support productive activities, decent job creation, entrepreneurship, creativity and innovation, and encourage the formalization and growth of micro-, small- and medium-sized enterprises, including through access to financial services";
- 8.5: "By 2030, achieve full and productive employment and decent work for all women and men, including for young people and persons with disabilities, and equal pay for work of equal value";
- 8.6: "By 2020, substantially reduce the proportion of youth not in employment, education or training";
- 8.b: "By 2020, develop and operationalize a global strategy for youth employment and implement the Global Jobs Pact of the International Labour Organization"; and
- 9.3: "Increase the access of small-scale industrial and other enterprises, in particular in developing countries, to financial services, including affordable credit, and their integration into value chains and markets".

Source: https://sustainabledevelopment.un.org/sdgs [18 Feb. 2016].

1.2 THE INTERVENTION: ALMPS FOR YOUTH

In support of more and better programmes and policies for the promotion of youth employment, this systematic review examines labour market interventions that fall into the category of ALMPs, which are further defined as

> all social expenditure (other than education) which is aimed at the improvement of the beneficiaries' prospect of finding gainful employment or to otherwise increase their earnings capacity. This category includes spending on public employment services and administration, labour market training, special programmes for youth when in transition from school to work, labour market programmes to provide or promote employment for unemployed and other persons (excluding young and disabled persons) and special programmes for the disabled (OECD, 2013).

ALMPs require active participation in programmes that enhance labour market integration, a requirement which differentiates them from other labour market – and social protection – policies, such as unemployment insurance schemes and non-conditional transfers. In the case of ALMPs, the economic rationale relies on market

clearing (i.e., achieving a match between labour demand and supply) and market efficiency (for instance, through job-search assistance, provision of labour market information and pre-screening of programme applicants). ALMPs can also enhance labour supply by providing training, foster labour demand through labour-intensive public employment programmes, entrepreneurship and self-employment measures, or alter the structure of demand by offering employment subsidies (Auer et al., 2008).

ALMPs considered in the systematic review are clustered in the following typology of interventions:

1. Training and skills development
2. Entrepreneurship promotion
3. Employment services
4. Subsidized employment.

Although the focus of ALMPs tends to be on economic relevance, they can have important social and political dimensions (Betcherman, Dar & Olivas, 2004). ALMPs can foster the social inclusion of disadvantaged groups while signalling a willingness on the part of politicians to engage with their specific problems.

1.3 HOW THE ALMPS ARE SUPPOSED TO WORK

This section offers some theoretical underpinning to the ways in which the interventions included in this systematic review may improve the labour market outcomes of youth. The underlying assumption of programmes is that participation in ALMPs will ultimately improve the employment and earnings outcomes of participants, as well as the performance of those businesses that programme participants start or already own.

Exposure to ALMPs is expected to create a spillover effect among non-participants, as well as general equilibrium effects throughout the economy. While some of these spillovers may positively affect overall employment outcomes, in certain cases ALMPs can have a negative impact on the performance of non-participants. For example, there is evidence that wage subsidy programmes can lead to substitution effects (with subsidized workers replacing non-subsidized workers) and windfall effects (when part of the subsidies go to workers who would have been hired in any case), thereby decreasing the overall employment impact of the programme. To address this issue, increased attention must be given to programme design features such as the establishment of conditionalities for employers (Almeida, Orr & Robalino, 2014).

This section summarizes the theories of change behind ALMPs for youth, aiming to map out the relationship between: (i) the resources that are invested ("Inputs"); (ii)

the intervention that takes place, including the different activities that may be part of the intervention ("Activities"); (iii) the individual-level competencies and constraints (such as knowledge, attitudes and behaviours) which are directly affected by the intervention ("Outputs"); and, finally, (iv) the individual labour market outcomes that can be measured as part of an impact evaluation study ("Outcomes"). Key assumptions are also made to determine whether any given event in the sequence actually yields the expected changes in labour market outcomes. Once the theories of change are clear, the systematic review examines whether the evidence supports the expected causality and impact across the selected intervention types, namely: training and skills development, entrepreneurship promotion, employment services and subsidized employment.

Building on existing literature, operational manuals and programme information, this section describes each intervention and its underlying theory of change. Even though labour market programmes often combine interventions from different categories, the results chains for each category have been separated to provide further transparency in the assumptions and support the interpretation of results to reveal potential causal mechanisms.

In the interests of a well-defined intervention description, those activities and outputs that are not strictly linked to labour market effects have been omitted. Similarly, a narrow focus has been adopted on individual-level labour market outcomes, leaving aside other potential side-effects, such as increased psychosocial well-being. For simplicity, higher level or "longer term" outcomes – such as poverty reduction, economic growth or democratization – are not explicitly shown in the chain of effects, nor are potential general equilibrium effects that may reduce the macroeconomic effectiveness of an intervention. Nonetheless, most of the programmes under scrutiny have broader macroeconomic effects, which will play an important role when scaling up or replicating the programme. In fact, some of the interventions may explicitly target higher level (economy-wide) macroeconomic outcomes, such as social protection aspects (e.g., public employment programmes may be designed to smooth consumption during recessions or crises).

1.3.1 Training and skills development

Education and skills are considered a core factor in determining young people's opportunities in the labour market (Biavaschi et al., 2012). Skills training programmes are therefore the most widely used labour market intervention for young people worldwide and are increasingly delivered as a complement to other labour market measures (Betcherman, Godfrey, Puerto, Rother & Stavreska, 2007; Fares & Puerto, 2009). Training and skills development comprises programmes outside the formal education system that offer skills training to young people in order to improve their employability and facilitate their transition into the labour

market.[2] The objective of skills training programmes is to develop the employment-relevant skills of jobseekers. Broadly speaking, these skills refer to a set of job-specific technical skills, but also include non-technical soft skills, such as self-management, teamwork and communication. Increasingly, employers across the world are placing higher value on these non-technical skills than on technical competencies (Manpower Group, 2013; Cunningham, Sanchez-Puerta & Wuermli, 2010; Youth Employment Network & International Youth Foundation, 2009).

This analysis classifies training programmes according to the skill set which they target (Table 2):

1. First, training programmes that address a lack of trade- or job-specific technical skills demanded by employers. Such skills range from manual skills to computer literacy. Technical skills training programmes often include an on-the-job training component in order to increase practical work experience (i.e., by placing participants in internships, workplace training or apprenticeship schemes).

2. Second, business skills training, which is often provided as an element of programmes that aim to increase entrepreneurial activities among youth. Such entrepreneurial training programmes cover a wide variety of factors that are believed to determine business success (ranging from financial skills to problem-solving skills).

3. Third, literacy and numeracy programmes, which are designed to teach basic skills or cognitive abilities to youth who had not acquired them by the time they left school (sometimes called "second-chance programmes").

4. Finally, programmes that improve non-technical skills, such as behavioural skills, life skills or soft skills of jobseekers.

Technical training programmes are popular in development cooperation because many developing countries experience a skills mismatch between their labour force and emerging segments of their economies. However, pure training programmes have not proven to be particularly successful in many contexts (Betcherman et al., 2004). Therefore, most recent programmes tend to combine skills training with other types of interventions; for example, on-the-job training or employment services (Cunningham et al., 2010; Fares & Puerto, 2009). An example of a skills training programme is provided in Box 2.

[2] The review excluded studies of formal training programmes, such as evaluations of dual systems in Austria, Germany and Switzerland.

Table 2: Training and skills development interventions: Results chain

Inputs	Activities		Outputs	Outcomes
1. Budget 2. Staff 3. Local counterparts 4. Trainers 5. Partnerships 6. Facilities 7. Equipment 8. Supplies 9. Technical expertise 10. Curricula	Technical skills training	1. Provision of skills training (e.g., distance or classroom training) 2. Placement in workplace training (e.g., internships, on-the-job training schemes) 3. Placement in apprenticeship schemes 4. Provision of financial incentives to young apprentices and employers providing apprenticeship training	Improved technical competencies in a specific trade	1. Increased probability of employment 2. Reduced time to find job/ shorter unemployment duration/ greater efficiency in the job search 3. Increased ability to retain job/longer job duration (hours worked) 4. Better quality of employment (contract type, job type) 5. Increased earnings or consumption
	Business skills training	As above	1. Improved management skills 2. Improved understanding of business mechanisms 3. Improved financial literacy	1. As above 2. Increased business performance (efficiency, profits, investments, output of entrepreneurs)[3]
	Literacy or numeracy skills training	As above	1. Improved reading and writing skills 2. Improved mathematical skills	
	Behavioural, life skills or soft skills training	As above	1. Improved psychosocial characteristics 2. Improved decision-making skills 3. Improved communication and teamwork skills 4. Increased self-management and self-esteem 5. Improved physical and mental health	As above

[3] Additional entrepreneurship-related outcomes are listed in Section 1.3.2 below.

Inputs	Activities		Outputs	Outcomes
Assumptions				
		1. Target group participates in training (there is awareness about the programme's existence) 2. Contracted training institutions conduct training and link participants to employers (if conditionality exists) 3. Employers train young people on the job and offer placement (if conditionality exists)	1. Participants attend and complete the training 2. Training addresses participants' constraints (e.g., existing skill shortages) as well as the constraints of the labour market 3. Participants learn in training/training increases skill level/training is well matched to interests and abilities of participants 4. Training induces expected behavioural and attitudinal change	1. Existing labour demand for skilled labour 2. Learned skills match labour market needs/demand 3. No stigmatizing effects 4. Training completion and related certificate signals acquisition of increased level of skill and higher (expected/observed) productivity 5. Employers value certified training 6. Participants gain recognized and valued qualifications 7. Adequate economic, social, institutional and administrative conditions are in place

A number of conditions determine whether skills training programmes are successful in bringing *additional* youth into work – most notably, correlation between the skills offered by a training programme and those demanded by the market. To this end, some programmes introduce a market-based (or bottom-up) approach in programme design. The application of this approach enables training curricula and programme components to respond much more effectively to the needs of employers (in both private and public sectors) and communities in a demand-driven fashion.

Furthermore, the success of all these interventions relies on the assumption that the (correct) target group participates in the training and that the training is appropriate and conducted in a way which actually augments the skill sets that are relevant to the labour market. Finally, a crucial element may be the award of a legitimate certificate on successful completion of a programme to prove the acquisition of increased knowledge and skills to potential employers in the job market.

Box 2: Training and skills development: Juventud y Empleo in the Dominican Republic

The Youth and Employment Programme, Juventud y Empleo (JE), in the Dominican Republic represents an innovative model of an ALMP to improve employability and human capital of young people between the ages of 16 and 29 who did not complete high school. The programme provided young people with vocational training (150 hours) and basic or life skills training (75 hours) combined with internships in private sector firms (240 hours). The programme was managed by the Ministry of Labour in cooperation with the National Institute of Technical and Vocational Training (Instituto Nacional de Formación Técnico Profesional) and with financial support from the Inter-American Development Bank. Training services were provided by private training institutions.

The programme came into operation in 2001 and was the first job-training programme in Latin America and the Caribbean to incorporate a randomized evaluation component in the project design. The first impact evaluation showed limited impacts on employment and wages, which led to changes in the programme to focus on working more closely with the private sector and providing a stronger life skills component. Further evaluation results showed that the programme had a positive impact on job formality for men and a positive effect on monthly earnings among those who were employed. In addition, the programme was effective in reducing teenage pregnancy and showed a positive impact in various measures of non-cognitive skills.

Sources: based on information available at: www.youth-employment-inventory.org [12 Oct. 2014]; Card, Ibarrarán, Regalia, Rosas-Shady and Soares, 2011; Ibarrarán, Ripani, Taboada, Villa and García, 2014.

1.3.2 Entrepreneurship promotion

Innovative entrepreneurial activities can promote job-rich growth and accelerate economic diversification paths through productivity and competitiveness. Entrepreneurship returns to economic development are maximized within business environments that are amenable to innovation and creativity and provide appropriate regulations, access to infrastructure services and finance (ILO, 2015b). However, entrepreneurship also carries substantial risks of failure and has the potential to contribute to job losses if increased productivity and competition leads to layoffs in existing enterprises (Kritikos, 2014).

Entrepreneurs are important income providers and job creators. They benefit booming economies by challenging existing enterprises to innovate and compete in order to keep up with rapidly changing technologies and global markets. They also benefit economies that are suffering from slow job growth or stagnation by boosting labour demand, developing innovative goods and services and stimulating competition.

Depending on the context, entrepreneurs can be driven by choice or by necessity. Entrepreneurs by choice select entrepreneurship over other employment options in order to increase their income or become more independent. Entrepreneurs by necessity, also known as subsistence entrepreneurs, face a market situation with insufficient labour demand and therefore lack formal employment opportunities, exposing their entrepreneurial ventures to the low productivity and precarious working conditions that prevail in the informal economy.

The enterprise size and its corresponding ability to grow and to create jobs also help to identify the rare "transformation entrepreneurs" or "gazelles". These are the few entrepreneurs whose enterprises grow to become larger enterprises and generate most of the new jobs. Their high-growth enterprises create jobs and income for others, beyond the scope of an individual's subsistence needs (Cho, Robalino & Watson, 2014). In contrast, the enterprises of subsistence entrepreneurs usually do not grow, but provide income and employment for the owner of the micro-enterprise and their immediate family.

Entrepreneurship promotion programmes considered for this systematic review aim to lower the barriers and costs associated with young unemployed and underemployed people planning to establish or maintain a business. Since the scope of formal wage employment is often limited in developing countries, increasing (formal) self-employment among the labour force is considered an important anti-poverty strategy (Gindling & Newhouse, 2012). Because self-employed and small-scale entrepreneurs often face numerous internal and external constraints, a multitude of measures exist to support the process.

Access to capital is often a primary constraint for young entrepreneurs. Schoof (2006) identifies a number of constraints to accessing start-up finance. These range from inadequate personal savings and resources to a lack of securities and credibility, insufficient business experience and skills, strict credit-scoring methodologies and regulations, among others. Accordingly, many entrepreneurship programmes address the lack of access to (affordable) finance faced by young entrepreneurs. The review team disaggregated such programmes into three types:

1. Those providing or facilitating access to credit (including microfinance programmes)

2. Those providing start-up grants

3. Those fostering microfranchising mechanisms.

ALMPs that facilitate access to finance often provide technical training and advice and support setting up partnerships and capacity-building schemes with (and for) microfinance institutions (MFIs) and banks.

In addition to access to finance, some programmes offer training on business and management skills as well as business advisory services and mentoring for soon-to-be or already self-employed youth. Finally, some interventions aim to reduce the barriers to business creation by assisting prospective entrepreneurs to enter established markets or existing value chains. The abovementioned interventions and their results chain are shown in Table 3. Some skills training programmes (as described in Section 1.3.1 above) incorporate features of entrepreneurship training and specific skills relevant for starting or maintaining a business.

Many entrepreneurship programmes take a multi-component approach; for example, combining access to credit with business skills training or the provision of post-programme consultation (i.e., mentoring and coaching).

Primarily, entrepreneurship programmes increase employment through their direct effect on the soon-to-be self-employed participant. The assumption is that beneficiaries actually plan to set up a new business after receiving credit and/or training (i.e., that targeted and trained individuals have been appropriately selected for the programme) and that they would not have done so without the intervention.

In order to generate additional jobs, entrepreneurship programmes have to assume that the intervention leads to either (i) increased marginal productivity of the input labour or (ii) increased output and profits resulting in additional investments and labour demand. To achieve this end, the training must suit the context and knowledge of the participants. Beneficiaries then have to apply the training or credit to their business and thereby increase performance and competitiveness.[4] Whether or not an entrepreneur will finally hire additional workers may also depend on the macroeconomic and labour market environment. Box 3 describes the programme Start and Improve Your Business (SIYB), a widely used and adapted entrepreneurship training package designed by the International Labour Organization (ILO) and tailored for youth.

[4] It is important to note that the theory of change analysis does not provide details on potential general equilibrium effects, such as substitution.

Table 3: Entrepreneurship interventions: Results chain

Inputs	Activities		Outputs	Outcomes
1. Budget 2. Staff 3. Local counterparts 4. Trainers 5. Partnerships 6. Facilities 7. Equipment 8. Supplies 9. Technical expertise 10. Curricula	Business and management training / business advisory services / mentoring and coaching	1. Training delivered 2. Advice delivered	1. Increase entrepreneurial impetus 2. Improved business and management skills (e.g., accounting practice, stocks management, investments) 3. Improved understanding of business mechanisms 4. Improved financial literacy/behaviour 5. Improved understanding of business practices, laws and regulations 6. Improved knowledge of business possibilities 7. Reduced risk/uncertainty in starting a business	1. Increased employment probability or number of hours worked 2. Increased earnings or consumption among young entrepreneurs 3. Business started 4. Increased business investment, performance and competitiveness (e.g., profits, sales, capital and investment, business survival)
	Access to markets and value chains	1. Support business networks 2. Provide technology necessary for value chain inclusion	1. Increased knowledge of markets and networks 2. Increased access to business networks and supply chains	As above
	Credit or access to credit	1. Provide credit to young entrepreneurs 2. Match entrepreneurs with credit agencies	1. Increased access to adequate financial services 2. Lower costs for finance 3. Higher probability of obtaining a loan, insurance or savings	As above
	Grants (monetary or in-kind)	Provide grants to young entrepreneurs	Beneficiaries possess sufficient capital to start a business	As above
	Microfranchising	1. Match participants with	Increased incentives (lower barriers) to start	As above

Inputs	Activities		Outputs	Outcomes
		franchisors 2. Intermediate between franchisors and potential franchisees 3. Distribute information about franchising 4. Assist in setting up franchise business 5. Support existing franchisees	own business/franchise	
Assumptions				
		1. Content, intensity and delivery of services is tailored to the needs of the target group and to the programme objective 2. Correct group is interested in the intervention and is targeted (e.g., participants are credit constrained) 3. Target group participates in programme and completes entire programme cycle	1. Participants learn from training and advisory service (sufficient skill level) 2. Training and advice prompted expected behavioural change 3. Credit or grant is used for enterprise 4. Credit agency/franchisor does not exploit entrepreneur	1. Created and supported businesses meet existing consumer demand 2. Adequate regulatory and business environment 3. Fertile macroeconomic environment 4. Adequate economic, social, institutional and administrative conditions 5. Start-ups benefit from additional investment/credit/networks 6. Credit or grant is used for productive investments

Box 3: Entrepreneurship promotion: Start and Improve Your Business

The Start and Improve Your Business (SIYB) programme is a management-training programme with a focus on starting and improving small businesses as a strategy for creating more and better employment in developing and transitional economies. The SIYB programme is a system of interrelated training packages and supporting materials for small-scale entrepreneurs. The programme is designed by the ILO and implemented with support from certified trainers in partner institutions in more than 100 countries with an estimated outreach of 6 million trainees. Initially developed in the 1980s, it has now been translated into more than 40 languages. The Start Your Business (SYB) package provides a five-day training course for potential entrepreneurs with concrete and feasible business ideas and proposes a follow-up programme including counselling sessions. SYB assists participants to develop a business plan with a marketing strategy, a staffing plan and a cost plan.

The 2011 SIYB Global Tracer Study found that in new businesses started after the training, on average, three jobs were generated. In Uganda, a randomized control trial (Fiala, 2014) providing mainly young business owners with loans, cash grants and the SYB training module or a combination of these components showed that, six and nine months after the interventions, men with access to loans with business skills training reported 54 per cent greater profits.

Sources: based on information available at: www.ilo.org/siyb [19 Feb. 2016]; van Lieshout, Sievers & Aliyev, 2012; Fiala, 2014; Majurin, 2014.

1.3.3 Employment services

Employment services programmes are generally based on the (matching and) intermediation approach to active labour market policy. Interventions within employment services are shown in Table 4. *Job-placement programmes* acknowledge the existence of information asymmetries and, particularly, incompleteness of information in the labour market. Hence, these programmes aim to improve the job-matching process by providing information and support to both sides of the labour market. On the one hand, they inform young jobseekers about suitable job opportunities (a service which is of particular relevance to youth who have only recently entered the labour market and are experiencing difficulties in marketing themselves or lack the knowledge, information and networks to find job openings) and, on the other hand, they provide information to potential employers about unemployed youth. The underlying idea is to facilitate the matching of employment opportunities with jobseekers while reducing the costs and risks to employers connected with recruiting young people.

The second type of intervention, *job-search assistance services*, includes job-search training, educational or career guidance, counselling and monitoring programmes. Such programmes primarily target disadvantaged or demotivated youth who are disconnected from the labour market. Their primary aim is to improve the intensity, motivation and effectiveness of participants' job-searches.

Table 4: Employment services interventions: Results chain

Inputs	Activities		Outputs	Outcomes
1. Budget 2. Staff 3. Local counterparts 4. Trainers 5. Partnerships 6. Facilities 7. Equipment 8. Supplies 9. Technical expertise 10. Curricula	Job placement/ intermediation services	1. Provide job placement services used by unemployed 2. Assess and match jobseekers and potential employers (broker information) 3. Market disadvantaged jobseekers to employers 4. Match unemployed with job vacancies	1. Improved matching of jobseekers and employers 2. Increased intensity (motivation) and efficiency of job-search	1. Increased labour-market participation 2. Increased probability of employment 3. Reduced time to find job/shorter unemployment duration 4. Increased ability to keep a job/longer job duration/increased number of hours worked 5. Better quality of employment (contract type) 6. Increased earnings or consumption
	Job counselling/ job-search assistance/ mentoring	1. Provide career and personal development advice 2. Provide job-search advice or training 3. Ensure active and efficient job-search	1. Participants better informed about labour market (i.e., qualifications in demand and where jobs are to be found) 2. Improved job-search skills 3. Increased intensity, motivation and efficiency of job-search 4. More informed decisions about investment in education	As above
	Financial assistance for job search	Provide credit or grants/stipends connected to job-search and job-acceptance (e.g., transport, childcare)	Greater ability to find and accept jobs (e.g., enhanced mobility)	As above

Inputs	Activities		Outputs	Outcomes
Assumptions				
		1. Target group (unemployed and employers) takes up the service offer (there is awareness about the programme's existence) 2. Participants complete/attend the programme 3. Participants comply with conditionalities and service requirements 4. Service matches the needs and abilities of participants	1. Correct target group identified (participants are constrained by lack of job-search skills) 2. Participants are motivated to search and take up work 3. Matched workers are able to do the job 4. Participants learn on the job/ employment increases skill level 5. Behavioural changes are prompted	1. Existing labour demand for employment services beneficiaries 2. Correct barriers and constraints for youth on the labour market are addressed 3. Adequate economic, social, institutional and administrative conditions established 4. No stigmatizing effects

Mentoring programmes are also provided to youth who are not currently unemployed but are in education or have just entered the labour market (post-placement support). Accordingly, in some circumstances, mentors encourage mentees to stay in education or in on-the-job training. In many countries, employment agencies adopt a case-management approach (identifying barriers to employment, designing individual action plans, referring jobseekers to appropriate interventions and monitoring job-search activity), which has been argued to be the most effective method of providing these services (Walther & Pohl, 2005).

While in some countries public employment agencies continue to be the main providers of employment services, other countries have moved into subcontracting, opening an important role for private employment agencies to address mismatches and information failures in the labour market. Box 4 illustrates a subcontracting model applied by a French public employment agency to facilitate counselling and job-placement for educated youth.

Box 4: Employment services: Counselling and job placement for young graduate jobseekers in France

In France, the government agency Pôle Emploi matches jobseekers with potential employers and provides benefits and job counselling to the unemployed. In 2007, the French Government decided to experiment with subcontracting employment services for young graduates who had been unemployed for at least six months to private providers. The jobseeker assistance programme aimed to help jobseekers find work and to support the former jobseeker in retaining that job or finding a new job. For the first six months of the programme, the private employment agency counselled the jobseeker and helped to find a job with a contract duration of at least six months. During the first six months of employment, the client continued to be supported and advised by the agency.

A randomized experiment measured the direct and indirect (displacement) impacts of job-placement assistance on the labour market outcomes of young people. The evaluation found that the reinforced counselling programme had a positive impact on the employment status of young jobseekers eight months after assignment to the treatment group, compared to untreated jobseekers. However, these positive effects appeared to have come partly at the expense of eligible workers who did not benefit from the programme, particularly in labour markets where they were competing mainly with other educated workers and in weak labour markets.

Sources: based on information available at: www.youth-employment-inventory.org [19 Feb. 2016]; Crépon, Duflo, Gurgand, Rathelot and Zamora, 2013.

There are indications that involvement in employment services (and in ALMPs in general) has a stigmatizing effect on participants (Boone & van Ours, 2004; Kluve, Lehmann & Schmidt, 1999). Addressing this adverse effect is a prior condition for successful implementation. To this end, job-placement and job-search assistance programmes are often connected to financial incentives for jobseekers and/or employers. For example, such schemes may involve the imposition of sanctions on

the unemployed for failure to comply with the terms of the intervention. Similarly, marketing of unemployed youth may be combined with the offer of short-term subsidies to employers.

1.3.4 Subsidized employment

Insufficient labour demand is one of the main constraints faced by young job market entrants – particularly in developing economies. Subsidized employment interventions comprise two main areas: wage subsidies and labour-intensive public employment programmes (Table 5), both of which are designed to increase the job and training opportunities available to unemployed youth. The main aim of both types of intervention is to ensure that individuals who do not find a job on the regular labour market remain integrated and connected to economic and social life. To that end, such programmes offer short-term interventions but primarily work towards longer term labour market impacts.

Wage subsidies are transfers to employers or employees in order to fully or partially cover eligible individuals' wage or non-wage employment costs. Most often, the measures aim to incentivize employers to hire members of a specific target group. Wage subsidies come in numerous forms and can be offered through various mechanisms, ranging from direct transfers to firms or workers to reductions in social security contributions or payroll taxes or tax credits.

Employer-side subsidies reduce the financial costs or risks associated with not knowing the productivity of the person to be employed. As with employment services, this is a scheme which is particularly relevant to youth entering the labour market for the first time, and whose (perceived) marginal productivity may be below market wages. Employer-side subsidies may also serve to lower the costs to employers of providing on-the-job youth training. Such training subsidies offer the possibility of expanding the number of work-based training places for disadvantaged young people.

Employee-side subsidies promote labour supply through increasing the returns from employment and hence increasing incentives to seek and retain employment. While it is believed that employer-side subsidies may also encourage more active job-search (because youths believe they will be able to find work), providing employee-side earning supplements may permit more effective targeting of specific socio-demographic groups. Furthermore, whereas employer-side subsidies tackle a lack of labour demand, employee-side subsidies may be more appropriate in countries that face labour supply constraints, for example due to reservation wages.

It is important to acknowledge the limited use and evidence of wage subsidies in developing countries. Almeida et al. (2014) detail the results of experimental and quasi-experimental impact evaluations around the world. Most evidence comes from the United States with rather mixed results concerning the effectiveness of wage subsidies as tools for fostering job creation.

Table 5: Subsidized employment interventions: Results chain

Inputs	Activities		Outputs	Outcomes
1. Budget 2. Staff 3. Local counterparts 4. Trainers 5. Partnerships 6. Facilities 7. Equipment 8. Supplies 9. Technical expertise 10. Curricula	Wage subsidy is offered and transferred through payroll tax cuts or direct payments to young people or employers	Intervention offers: 1. A job (of short of long duration) 2. A job with work-based training 3. A job plus work-based training and/or job search assistance	1. Direct job creation (at least for the duration of the subsidy) 2. Participants (re)gain labour market contact 3. Participants increase (or demonstrate increased) productivity 4. Skills formation or increased job skills (technical and non-technical) through on-the-job training and exposure to the work environment: a. Development of a work ethic and work habits 5. More positive attitudes towards employment/increased incentives to apply or to work 6. Participants integrate into networks 7. Incentives to continue education	1. Increased probability of (formal, well-regulated) employment beyond the programme duration 2. Reduced time to find future job/ shorter unemployment spells in the future/ more efficient job search 3. Increased ability to retain a job/longer job duration (hours worked) 4. Better quality of employment (contract type, a job conducive to human capital development, a salary) 5. Increased earnings or consumption 6. Increased returns from employment, including long-lasting human capital accumulation
	Public employment programmes in infrastructure development projects, social development, community works and services projects	Assigning target group to public employment programmes – placing youth in jobs	1. As above 2. Improved sense of contribution to community development 3. Improved social skills	As above

Inputs	Activities		Outputs	Outcomes
Assumptions				
		Complete information about the programme available for both employers and youth Target group (first time jobseekers, disadvantaged/low-skilled youth, unemployed youth and employers) participates in programme	1. Participants are motivated to work and sufficiently qualified (adequate profiling) 2. Participants learn on the job (i.e., experience increased skill levels) 3. Programme induces (positive/expected) behavioural changes/no adverse behavioural changes 4. Subsidies are not exploited by firms or conditionalities are in place to avoid unintended behaviours by employers 5. Appropriate targeting (to avoid windfall for the firm and deadweight for society)[5]	1. Correct barriers and ensure constraints for youth attempting to access the labour market are addressed 2. Work experience adequately signals higher skills and employability 3. Acquired skills/work experience match labour market demands 4. No stigmatizing effects[6] 5. No windfall, deadweight, substitution or displacement effects. 6. Danger of programme/welfare dependency among participants taken into account[7]

[5] Targeting is critical to avoid misuse of subsidized employment programmes.

[6] Programmes may create a stigmatizing effect on participants, particularly the most highly educated. In France, participation in the workfare programme carried a stigma that hindered participants in their transition to better and more durable jobs (Bonnal, Fougere & Sérandon, 1997; Brodaty, Crépon & Fougère, 2000). Adequate programme marketing and publicizing is important to address stigma issues.

[7] Programmes may create higher dependency among participants, hindering the transition into unsubsidized employment. Evidence from public works programmes in Poland indicates that the effect of the programmes on reemployment gradually diminishes after the fifteenth month of registering as unemployed (O'Leary, 1998). Addressing this concern, Galasso, Ravallion and Salvia (2001) conducted a randomized experiment that aimed to provide comprehensive services to workfare participants in Argentina in order to promote their transition out of welfare. The experiment, called Proempleo, offered wage subsidies and specialized training to programme participants and reported positive cost-effective impacts on their employment prospects.

Evidence on the impact of youth-targeted wage subsidies in developing countries is limited and results are mixed. Evaluations looking into wage subsidies in Jordan (Groh, Krishnan, McKenzie & Vishwanath, 2012) and South Africa (Levinsohn, Rankin, Roberts & Schoer, 2014) show positive though rather short-lived effects and a narrow participation from firms. Details of the Jordan New Opportunities for Women pilot are shown in Box 5: Subsidized employment: Jordan New Opportunities for Women (Jordan NOW). A recent review of wage subsidies for youth argues that, if well targeted, the interventions can be effective in improving employment outcomes of disadvantaged youth (Bördős, Csillag & Scharle, 2016).

Box 5: Subsidized employment: Jordan New Opportunities for Women (Jordan NOW)

The Jordan New Opportunities for Women (Jordan NOW) pilot aims to increase employment of female community college graduates in Jordan by offering wage subsidies and training to graduating students. Groh, Krishnan, McKenzie and Vishwanath (2012) examined the impact of the pilot in a randomized experiment. Female graduating students were randomly allocated into four groups: a treatment group which received a job voucher; a treatment group which was invited to attend an employability skills training course designed to provide key soft skills demanded by employers; a treatment group which received both the voucher and the training; and a comparison group.

The pilot targeted young female graduates who could take the job voucher to a firm while searching for jobs. The job voucher paid the employer an amount equal to the mandatory minimum monthly wage of 150JD (US$210) per month for a maximum of six months within an 11-month period, if they hired the worker, thereby acting as a wage subsidy.

The analysis finds that the job voucher led to an increase in employment in the short term, but that most of this employment was not in the formal sector, and the average effect was much smaller and no longer statistically significant four months after the voucher period had ended. The voucher does appear to have had persistent impacts outside the capital, where it almost doubled the employment rate of graduates. However, the analysis suggests that employment gains may have resulted from displacement effects.

Source: Groh, Krishnan, McKenzie and Vishwanath, 2012. Note: The description above focuses on the wage subsidy intervention of the pilot.

The second type of labour market intervention analysed in this category is *labour-intensive public employment programmes*, also known as public works. These programmes are commonly used to increase aggregate demand for labour in contexts where markets are unable to create productive employment on the required scale. In addition to their ability to create direct jobs, public employment programmes also generate income and deliver public assets and services. Despite the strong association of these programmes with infrastructure and construction works, they can be quite versatile, with works and projects in the social sector, environmental services and multi-sectoral, community-driven programmes (Lieuw-Kie-Song, Philip, Tsukamoto & Van Imschoot, 2010; Lieuw-Kie-Song, Puerto & Tsukamoto, forthcoming).

In this type of intervention, basic social income recipients are recruited for public jobs and receive a small earning supplement to their unemployment assistance. Programmes usually target unskilled, disadvantaged or long-term unemployed workers with the aim of keeping them in contact with the labour market and mitigating the depreciation of human capital during periods of unemployment.

While public employment programmes have often been recommended as a measure in times of crises (such as seasonal shocks or economic recession),[8] they are increasingly used as a regular component of wider employment policies (Lieuw-Kie-Song et al., 2010). In addition, they have become popular as a mechanism for addressing youth unemployment (Grosh, del Ninno, Tesliuc & Ouerghi, 2008), serving both as an introduction to the world of employment and as a tool to maintain social integration. This is particularly relevant for youth service programmes, in which youth can "play an active role in community and national development while learning new skills, increasing their employability, and contributing to their overall personal development" (Cunningham, McGinnis, Verdú, Tesliuc & Verner, 2008).

Most wage subsidies and public employment programmes are designed to support employment only in the short or medium term. A positive effect on final outcomes is only attainable if the work experience and training received during the period of subsidized work also improves the longer term employment prospects of participants. For this reason, (i) wage subsidies are often granted to firms that agree to provide additional training to subsidized employees (i.e., in connection with apprenticeship schemes)[9] and (ii) public employment programmes are often paired with exit strategies, such as skills training or entrepreneurship.

1.4 WHY THE REVIEW IS NEEDED

Policymakers and practitioners are seeking answers to the youth employment challenge; looking for ideas and evidence on what works best and why, in order to improve the labour market conditions of young people. Youth employment interventions, such as entrepreneurship promotion, training and skills development, employment services, mentoring and subsidized employment are considered common measures to improve youth labour market outcomes. Yet few overview and cross-country studies have reviewed and analysed their impact on these outcomes and what determines success among youth. Even though the number of studies contributing to rigorous evidence on the effectiveness of ALMPs has increased over

[8] The programmes' potential to yield stabilization benefits is higher when they are implemented at the right time. Some programmes – particularly in South Asia – are implemented seasonally to ensure that employment is available during the agricultural slack seasons. Others, such as Argentina's Trabajar Program, are implemented during sharp economic crises as a means of increasing the incomes of poor families and those badly affected by recessions.

[9] The ability to retain work following the expiration of the wage subsidy period also serves as a signal of the acquisition of certain work-related behavioural skills to potential future employers.

the past decade, many fundamental questions remain unanswered, particularly with regard to context, programme type, design features and target groups.

- *The role of context*: Evidence on youth employment programmes is most common among developed countries and is particularly scarce in Africa, the Middle East and North Africa, Asia and sub-Saharan Africa. While contextual variables, such as levels of income and development, seem to play a role in shaping the probability of positive outcomes from youth ALMPs (Betcherman et al., 2007) more information is needed to understand how similar intervention models may affect youth differently in developed as opposed to developing contexts. Moreover, further evidence is required on the interventions and design features that are better suited to rural than to urban contexts, informal rather than formal settings, and in post-conflict and fragile-state environments.

- *The question of programme focus*: The majority of evaluations focus on the area of training and skills development, while evidence on other types of youth employment interventions, such as subsidized employment, employment services and entrepreneurship promotion, is relatively scarce. There is a significant knowledge gap regarding the effectiveness of combining different types of programme; for example, bundling up skills training, job-search assistance and mentoring.

- *The efficacy of various design features*: Little is known about the effectiveness of programme alternatives. There are several areas where policy choices can make a significant difference: design of the interventions; targeting mechanisms; length of exposure to the interventions; pedagogy; governance, management and administration; delivery channel (public, private, partnerships); delivery setting (classroom, on-the-job); and contracting, auditing and payment systems to providers of services. More evidence needs to be gathered on these design aspects.

- *The range of beneficiaries*: More evidence is needed to provide clarity on how different types of programmes affect individuals differently by age cohort, gender, level of education, ethnicity and socio-economic background.

Focusing on youth employment and understanding what works in terms of improving the labour market outcomes of youth is therefore of significant practical relevance. With the aim of impacting policymaking and programming with informed recommendations, this systematic review takes stock of the available evidence and examines changes in labour market outcomes prompted by labour market interventions for youth.

Assessing the impact of ALMPs has been a major focus of social welfare policies for decades, particularly in developed economies. It has also become a regular feature of recent public programmes in developing and transitional economies, given the increased budget constraints and need for policy decisions that are based on rigorous evidence of programme benefits and losses.

Such assessments have been regularly undertaken through social experiments that allow the estimation of programme impact by comparing observed changes in outcomes against what would have happened in the absence of a programme. In these experiments, random assignment is used to allocate the intervention among members of an eligible population. Differences in outcomes between the programme participants and their comparison group counterparts can be attributed solely to the programme since, according to the design parameters, there should be no correlation between participant characteristics and the outcome (3ie, 2013).[10]

Experimental evaluation evidence is growing in the field of youth employment. Most available evidence relies on quasi- or non-experimental methods. The Youth Employment Inventory (YEI)[11], an online global repository of information on labour market programmes for youth, offers records of impact evaluation studies of youth employment interventions worldwide. While rigour varies between studies, there is a clearly observed transition towards randomized experiments and stylized methods of evaluating impact.

This systematic review examined experimental and quasi-experimental evaluations of ALMPs that target youth. It looked at the available evidence in order to fill the knowledge gap on the impact and effectiveness of these interventions in a systematic and rigorous manner. Section 3 provides further information on the methodology adopted for the review's analysis.

Other reviews have looked at impact evaluations of youth employment programmes from different angles and at varied levels of depth. Table 6 presents the available evidence on completed reviews, identifying key differences between them and this review and summarizing its added value.

While some previous studies synthesize the evidence-base on the effectiveness of ALMPs (e.g., Card, Kluve & Weber, 2010 and 2015), very few reviews specifically focus on programmes and outcomes for youth. The most relevant review of labour market interventions for youth to date, Betcherman et al. (2007), has served as the basis for technical assistance and policy advice worldwide. Since then, a vast amount of research has been published, using experimental or quasi-experimental methods to determine the impact of new and innovative employment programmes. While some recent reviews cover this new evidence, these do not synthesize the existing empirical evidence using empirical methods such as meta-analysis (J-PAL 2013) or they only look at (potentially selective) subsets of the available evidence (IEG, 2012; Eichhorst & Rinne 2015). Other studies only include specific types of intervention or

[10] In a review of evaluation methods used in ALMPs, Heckman, La Londe and Smith (1999) identified a number of methodological lessons, ranging from recognition of the multiplicity of parameters and heterogeneous impacts intrinsic to ALMPs to the need for appropriate comparison groups and the importance of addressing selection bias. Experimental evaluations can effectively learn from these lessons by providing a framework that relies on credible comparison groups and minimizes selection bias.

[11] Available at: www.youth-employment-inventory.org [20 Feb. 2016].

outcomes (Tripney et al., 2013; Grimm & Paffhausen, 2015; Piza et al., 2016; Valerio et al., 2014), with the implication that some of the evaluations included in these studies were also included in this systematic review.

To the best of the review team's knowledge, this is the first systematic review of the impact of employment interventions on youth labour market outcomes to collate global evidence from youth ALMPs, examine the most relevant outcomes along the causal chain and identify study effect sizes through a rigorous meta-analysis.

Table 6: Existing reviews

Betcherman et al. (2007)	In 2007, the World Bank produced a review of labour market interventions for youth based on the information gathered by the YEI. The review covered 289 studies, of which only one-quarter had estimates of net impact and just one in ten offered evidence on cost effectiveness. Most evaluation evidence came from developed countries. A meta-analysis of the studies looked into factors that increase the probability of positive effects on employment or earnings of young people. The World Bank's review, translated into French and Spanish, has served as the basis for technical assistance and policy advice worldwide and is still cited in the current debate about youth employment policies and programmes. (See, for example: World Bank's Independent Evaluation Group Report on Youth Employment Programmes (IEG, 2013); review of training for young people published by the UK Department for Business, Innovation and Skills (Wilson, 2013).) The study's main limitations include insufficient systematic search and risk of bias assessment as well as a lack of reflection on effect sizes. The quantitative analysis focused on the determinants of positive main labour market outcomes in the framework of a probability model. The systematic review builds on the work of Betcherman et al., utilizing a much more structured analytical model that will provide information on the magnitude of programme impacts on youth labour markets.
J-PAL (2013)	A 2013 review paper produced by the Abdul Latif Jameel Poverty Action Lab covered an array of youth interventions from education and health to labour market programmes. The paper discusses existing knowledge about and gaps in policies focused on youth. It identifies unanswered questions and sets a research agenda that will be updated periodically. In the area of ALMPs for youth, the review considers open questions on the effectiveness of employment services, training, subsidized employment and public works programmes. There is very limited information about the search methodology behind the review but it is clear that it builds on results from cross-country reviews and impact evaluations to identify and discuss knowledge gaps. The review does not rely on a statistical meta-analysis or study effect sizes.

Eichhorst and Rinne (2015)	A 2015 research report by the Institute for the Study of Labor (IZA) provides an assessment of youth employment interventions based on the information gathered by the YEI and following the approach established by Betcherman et al. (2007). The review is based on information about the programme design, implementation and results of 730 projects in 110 countries and includes a quantitative analysis. While the review relies on a larger number of observations than the Betcherman et al. review, there is not enough evidence to make an assessment for the majority of interventions in the YEI. The review's meta-analysis is complemented by a qualitative analysis as a large number of interventions in the YEI have not been rigorously evaluated with respect to their impacts.
IEG (2013)	In 2012, the World Bank's Independent Evaluation Group carried out a meta-review of evaluations of the World Bank Group's youth employment projects. The review built on the YEI as well as on evidence from recent impact evaluations and lessons focused mainly on how to improve the Bank and International Finance Corporation's (IFC) performance and delivery in the youth employment field. There is limited information about the search process, no risk of bias assessment or measure of effect sizes.
Tripney et al. (2013)	Campbell Collaboration Group published a systematic review of Technical and Vocational Education and Training (TVET). The review aimed to summarize the available evidence on the effects of TVET interventions for young people in developing countries to inform policy, practice and research. The review built on evidence from 26 studies of 20 TVET interventions with a rigorous search process, risk of bias assessment and the statistical analysis of effect sizes. Tripney et al.'s review offers a slight overlap with the type of interventions and sample of studies that will be covered in this review, specifically with regard to ALMPs in the areas of vocational training, on-the-job training and apprenticeship training in developing countries. In contrast to Tripney et al.'s review, this review covers countries from all levels of development and will disregard any training programme that is delivered in a formal education setting.
Reviews looking at ALMPs in general, not focused on youth	There is a series of cross-country studies that reviewed the impact of ALMPs with specific findings from youth employment programmes, including: Betcherman et al. (2004), Dar and Tzannatos (1999) and Card et al., (2015). The sample of programmes specifically targeting youth is limited, as are the findings. Only Card et al. (2015) offer a relatively rigorous search and quantitative analysis of impact based on study significance. Other studies with similar limitations include those which looked at programmes implemented in Organisation for Economic Co-operation and Development (OECD) countries only, e.g., Heckman et al. (1999), Kluve and Schmidt (2002), and Kluve (2006 and 2010).

Reviews of entrepreneurship interventions, not focused on youth	Cho and Honorati (2013) synthesized evidence from interventions aimed at promoting the development of micro-, small- and medium-sized enterprises in developing countries. While the review, and corresponding meta-analysis, does not focus on youth, it does provide some insights into the effectiveness of the programmes when targeted at vulnerable populations, such as youth and women. The review does not rely on a risk of bias assessment and the search strategy was limited to some central indexes, such as Ideas and Google Scholar, as well as snowball search.
	Grimm and Paffhausen (2015) assess the direct or indirect effectiveness of interventions aimed at creating employment in micro-, small and medium-sized enterprises in low- and middle-income countries. The review explores both policies and programmes with a credible link to job creation and the context, environment and circumstances that influence such impacts. The review relies on a rigorous search strategy and provides effect sizes in selected outcomes, namely number of employees and/or growth rate.
	Valerio et al. (2014) examine a global sample of evaluations of entrepreneurship education and training programmes that aim to provide individuals with the entrepreneurial mindsets and skills to enable them to participate in entrepreneurial activities. In addition to evaluations based on experimental and quasi-experimental designs, the study also includes tracer studies as well as monitoring and evaluation reports that rely largely on administrative data. The study does not include a statistical meta-analysis or study effect sizes.
	Piza et al. (2016) review the impacts of business support services for small and medium-sized enterprises (SMEs) on firm performance indicators, employment generation and labour productivity in low- and middle-income countries. The review examines interventions including tax simplification, boosting exports and facilitating access to external markets; support for innovation policies; support for local production systems; training and technical assistance; and SME financing and credit guarantee programmes. The review relies on a systematic search strategy and provides statistical meta-analysis.
Reviews looking at labour market regulations, not focusing on youth	A review by Nataraj, Perez-Arce, Kumar and Srinivasan (2013) looks at the impact of regulations (such as minimum wages, regulations covering dismissal and various aggregate measures) on employment outcomes. The review does not consider programmes and, while it discusses the inconclusive evidence of the impact of minimum wage policies on youth, it provides no insights into the impact and effectiveness of interventions specifically serving young people.

2 Objectives

Addressing the youth employment challenge ranks high on the development agenda and it is essential to gather rigorous evidence to enhance the ability of governments – as well as civil society, the private sector and multilaterals – to diagnose and address the problems facing youth in accessing wage- or self-employment. The current political and economic context makes even more imperative the need to provide this evidence on what does work, what does not work, why, and how to improve labour market outcomes of young men and women.

Policymakers are constantly searching for solutions to the challenge of effectively supporting youth in the labour market. However, their objectives require a solid evidence base. During the 2012 International Labour Conference, governments and social partners recognized the need for more rigorous evaluation of youth employment interventions in order to review their effectiveness and, in particular, asked the International Labour Office to strengthen the evidence base on youth entrepreneurship interventions (ILO, 2012). Similar requests for information, technical and financial assistance are often made to the World Bank by client countries. Non-governmental organizations (NGOs), donors and employment practitioners in general are also intent on identifying success factors to support youth.

In this context, this systematic review aimed to provide policymakers and practitioners with evidence-based recommendations on what works to effectively support youth in the labour market by summarizing and integrating empirical research to investigate the impact of labour market interventions on labour market outcomes of young people. The review also examined whether the evidence supports the underlying assumptions about what active labour market policies (ALMPs) for youth are designed to achieve.

The following research questions framed the analysis to establish what constitutes effective measures, which will ultimately help decision-makers in the allocation of their resources and determining their investment level and portfolio on youth employment:

1. What is the impact of youth employment interventions on labour market outcomes of youth? In particular, the review investigates skills training, entrepreneurship promotion, employment services and subsidized employment interventions.

2. Which of these interventions are the most effective?

By synthesizing the evidence on the relative effectiveness of different labour market interventions for youth, this systematic review has contributed to closing the knowledge gap in this field, which will have a real impact on the 73 million young men and women who are currently actively looking for a job.

3 Methods

3.1 TITLE REGISTRATION AND PROTOCOL OF THE SYSTEMATIC REVIEW

The title registration for this systematic review was published in The Campbell Collaboration Library of Systematic Reviews on 1 November 2013. The protocol of this review (Kluve et al., 2014) was published on 3 November 2014. Both the title registration and protocol are available at:
http://www.campbellcollaboration.org/lib/project/306/

3.2 INCLUSION CRITERIA

This systematic review focused on studies that investigated the impact of interventions on labour market outcomes of young people and met the following inclusion criteria:

1. The studies were reports from evaluations of the youth employment interventions listed in the research question or combinations of these interventions

2. The studies reflected completed experimental and quasi-experimental evaluations of the abovementioned interventions and were included if, and only if,

3. Such studies measured impacts on the labour market outcomes presented in Section 3.2.4.

The selection of studies was based on the inclusion criteria outlined by the screening questionnaire presented in Section 8.6.2 of the Appendix. The review's PICOS[12] is detailed in the following subsections.

[12] In a systematic review, PICOS stands for Population (or sub-groups) to be included; the Intervention(s) in question; the Comparison groups or Counterfactuals to be considered; the Outcomes to be investigated; and the Study designs eligible.

3.2.1 Population and context

The review was global in coverage and considered interventions from all countries, regardless of their level of development. Studies investigated active labour market policies (ALMPs) that were designed for – or targeted primarily – young women and men aged 15 to 35, in consideration of varying national definitions of youth. The ALMPs examined in the study (i) targeted the unemployed or those with low levels of skills or limited work experience or who were generally disadvantaged in the labour market and (ii) aimed to promote employment and/or earnings/wage growth among the target population, rather than simply providing income support (Heckman et al., 1999).

3.2.2 Intervention

Eligible studies evaluated an ALMP that provided at least one of the following categories of intervention (also shown in Section 1.3): training and skills development, entrepreneurship promotion, employment services and/or subsidized employment. An overview of the categories of intervention is presented in Table 7.

Table 7: Youth employment programme interventions

Interventions	Description
Training and skills development	Comprised programmes outside the formal education system (and therefore did not consider Technical and Vocational Education (TVE) programmes) that offered skills training to young people in order to improve their employability and facilitate their transition into the labour market.
Entrepreneurship promotion	Aimed to provide entrepreneurial skills as well as physical, financial and social capital for youth becoming self-employed and starting a business and for those seeking to expand and grow their businesses.
Employment services	Delivered job counselling, job-search assistance and/or mentoring services, which were often complemented by job placements and technical or financial assistance.
Subsidized employment	Considered mainly those programmes which provided wage subsidies or interventions that aimed to reduce labour costs for employers taking on young workers as well as labour-intensive programmes or public works which provided short-term employment to youth in infrastructure or social development and community projects.

As discussed in more detail in the protocol (Kluve et al., 2014), this review made an important distinction between programmes, interventions and components of an intervention: A youth employment programme was considered to be a single entity that might consist of one or several interventions. In addition, each of these

interventions could have different components: It was possible to find a comprehensive intervention that offered, for instance, both skills training and employment services (to the same participant). Some examples of such multi-component interventions included the Job Corps programme in the United States, the Economic Empowerment for Adolescent Girls programme in Liberia, the Projoven programme in Peru and the Employment Fund in Nepal.

Interventions were therefore specific tracks or sub-programmes of an overall programme that were offered to different samples of participants. They were defined based on their characteristics, such as the category of intervention or the population targeted. For example, if a programme had a training track and an employment services track and participants took one or the other, they were considered to be two interventions within the same programme. Note that, according to this definition of track, it was assumed that each intervention within a programme had separate groups of participants which did not overlap. In order to provide evidence on which interventions and combinations were shown to work best, these different types were evaluated separately in the meta-analysis in the empirical Section 4.3 on Synthesis of results.

Additional consideration was given to identifying primary intervention types among multi-component designs. The review defined "main category of intervention" as the largest and predominant intervention type within a programme. If several intervention types were equally distributed across the target population (i.e., an individual was exposed to more than one intervention type with the same level of intensity), the main category of intervention was classified as unspecified.

3.2.3 Comparison

The systematic review included studies that measured change in at least one outcome of interest among intervention participants and relative to non-intervention participants based on a counterfactual analysis. Eligible comparison groups (counterfactual) included those which received no intervention or were due to receive the intervention in a pipeline or waitlist study. Note that the comparison group of some studies might have been exposed to interventions other than the evaluated intervention.

3.2.4 Outcomes

Eligible studies reported at least one selected outcome variable measuring the following primary outcomes of interest presented in Table 8: Employment outcomes, earnings outcomes and business performance outcomes. The review also captured outcomes which were measured conditional on other outcomes and excluded studies that focused solely on intermediary outcomes without measuring impact on the abovementioned primary outcomes.

Table 8: Outcome categories

Outcome category	Outcome for which effect size is measured
Employment	1 = Employment probability 2 = Unemployment probability 3 = Participation rate 4 = Hours worked 5 = Unemployment duration 6 = Quality of employment (e.g., contract, fixed term, benefits)
Earnings	7 = Earnings/income 8 = Household income 9 = Consumption 10 = Salary/wage
Business performance	11 = Profits 12 = Sales 13 = Number of employees/jobs created 14 = Capital & investment 15 = Business creation 16 = Business survival

3.2.5 Study designs

The review focused on completed experimental and quasi-experimental evaluations and considered the following research design categories and impact evaluation methods to estimate quantitatively the causal effect of the intervention on the outcome it intended to influence: (i) randomized experiments, (ii) methods for causal inference under unconfoundedness (classical regression methods, statistical matching, propensity score matching) and (iii) selection on unobservables (instrumental variables, regression discontinuity design, difference-in-differences).

1. *Randomized experiments*: The most straightforward case for analysis occurred when assignment to treatment was randomized (in controlled conditions) and, therefore, independent of covariates $\boldsymbol{X}$ as well as the potential outcomes $\boldsymbol{Y}$. In such classical randomized control trials (RCTs) it was relatively easy to obtain estimators for the average effect of the treatment using, for example, the simple difference-in-means by treatment status. Randomized experiments have been used in the evaluation of labour market programmes since the 1970s (starting in the United States), with an increasing trend over the past decade. The descriptive analysis by research design included in this review in Figure 3 confirmed that RCTs have increasingly been used to assess the impact of youth employment interventions in recent years.

2. *Methods for causal inference under unconfoundedness*: In this case, researchers analysed data from non-experimental (also called "observational") studies. Non-experimental data generally created challenges in estimating causal effects but, in one important special case, variously referred to as unconfoundedness, exogeneity, ignorability or selection on observables, questions regarding identification and estimation of the policy effects were fairly well understood (Imbens & Wooldridge, 2009). All these labels referred to some variant of the

assumption that adjusting treatment and comparison groups for differences in observed covariates $\boldsymbol{X}$ (i.e., pretreatment variables) removed all biases in comparisons between treated and comparison units (Imbens & Wooldridge, 2009). This case was of great practical relevance, with many impact evaluation studies relying on some form of this assumption: specifically, this category comprised classical **regression methods** (e.g., adjusting for covariates in a **linear regression**). Another method that was based on the unconfoundedness assumption and has been applied with increasing frequency is **statistical matching**, generating balanced samples in $\boldsymbol{X}$ of treated and comparison units and thus mimicking an experiment ex post. In practice, in recent years the most frequently used version of a selection-on-observables design has been **propensity score matching**, adjusting for a scalar, the **(estimated) conditional** probability of receiving the treatment given the covariate vector $\boldsymbol{X}$.

3. *Selection on unobservables*: Without unconfoundedness, there is no general approach to estimating treatment effects, although various methods have been proposed for special cases (see Imbens & Wooldridge, 2009) and three of them were important for this systematic review. One such method is the **instrumental variables (IV)** approach that relies on the presence of additional "treatments", the so-called instruments, which satisfy specific exogeneity assumptions. Essentially, in the case in which treatment assignment is endogenous (i.e., confounded with the potential outcomes), researchers look for instrumental variables that satisfy two assumptions. First, the instrument is correlated with the treatment (testable assumption) and, second, the instrument does not exert a direct impact on observed outcomes, but only through the treatment (maintained hypothesis). A second method is the **regression discontinuity (RD)** design that applies to settings in which (in its pure form, the so-called "sharp" RD) overlap is completely absent because the assignment is a deterministic function of one or more covariates, but causal comparisons can be made exploiting continuity of average outcomes as a function of the covariates. (In the "fuzzy" RD design, the assignment probability does not switch from 0 to 1 as in the sharp design, but only requires a (sufficiently large) discontinuity in the probability of treatment assignment at the threshold determined by the forcing covariate(s).) Regression discontinuity methods have received increasing attention in the economic impact evaluation literature in recent years. Finally, the third method, **difference-in-differences (DiD)**, relies on the presence of additional data in the form of samples of treated and comparison units before and after the treatment (these can be panel data or repeated cross-sections). In the simplest setting, outcomes are observed for units in one of two groups, in one of two time periods. Then the average gain over time in the comparison group is subtracted from the gain over time in the treatment group. This double differencing removes biases in second-period comparisons between the treatment and comparison group resulting from permanent differences between the groups, as well as biases from comparisons over time in

the treatment group resulting from time trends unrelated to the treatment. The intuitive way in which the DiD design can remove important biases, coupled with its broad applicability in many different contexts, has made this method one of the most frequently applied designs for estimating causal effects. Nonetheless, in practical applications attention must be paid to challenges to the design (e.g., sensitivity of estimates to the timing of measuring outcomes; time trends differentially affecting treatment and comparison groups; etc.). Finally, note that the approaches presented in this third category are often associated with the concept of "natural experiments", in which policy changes (or other "exogenous shocks") can be used to effectively define (randomly assigned, though not in a controlled way) treatment and comparison groups.

The form of publication of eligible studies included peer-reviewed journal, working paper, mimeo, book, policy or position paper, evaluation or technical report and dissertation or thesis. Eligible studies could be published in any language as long as they met all other eligibility criteria. The date of publication or reporting of the study had to fall between 1990 and 2014.

3.3 SEARCH METHODS

The search for relevant literature was based on a variety of sources in order to ensure that published and unpublished studies ("grey literature") relevant to the research question were included in the search process. The search process included (i) a primary search – searching of a wide range of general and specialized databases – and (ii) a complementary search – hand-searching of relevant websites, searching of dissertations, theses and grey literature databases, citation tracking, screening of reference lists and contacting authors and experts. The search included search terms in English, Spanish, French, German and Portuguese, but no language restrictions were applied in the selection process. Country restrictions were not applied to the search and selection process. The search and selection process was restricted to the period 1990 to 2014 with regard to date of publication or reporting of the study. Detailed information about the search methods can be found in the protocol of the systematic review (Kluve et al., 2014).

3.3.1 Scoping search

Prior to implementing the primary and complementary search, the review team conducted a scoping search of potentially relevant sources to determine their relevance and to develop customized search strategies which would yield relevant results. The scoping search entailed an iterative process of testing and documenting several search strategies and identifying one or more preferred search strategies and search strings for each source in order to yield a comprehensive and precise set of potentially relevant results. The relevance of sources was determined by screening the results obtained from implementing each customized search strategy.

Based on a review of preferred search strategies and the results obtained during the scoping search, selected databases and websites were not included in the final search strategy if the review team did not have access to the source (e.g., SocIndex), the results obtained from the source were of low relevance (e.g., African Economic Outlook) or the source was covered by another source (e.g., ILO working papers are included in Labordoc). The final primary and complementary search strategy covered more than 70 sources, which included general databases, specialized databases, institutional websites, conference websites, dissertations and theses databases and grey literature databases. Section 8.6.1 in the Appendix presents the list of sources included in the final primary and complementary search.

3.3.2 Primary search

The primary search included 11 general databases and 12 databases that specialize in literature relevant to development economics and labour market issues. The search terms used for the primary search were based on the inclusion criteria and tried to strike a balance between sensitivity (e.g., finding all available articles in a topic area) and specificity (e.g., finding only relevant articles). For electronic databases with advanced search functions, the preferred search was based on a search of exposure, outcome and subject terms using Boolean operators in title and abstract from 2000 onwards. Highly relevant databases, such as EconLit, were searched for studies published since 1990 in order to include potentially relevant studies between 1990 and 2000. The search terms for electronic databases and examples for RePEc/IDEAS, EconLit and ERIC are presented in Sections 8.6.3 and 8.6.4 of the Appendix. The search strategy was modified according to the specifications of each database. Wherever possible, synonyms as well as wildcards and truncation symbols were applied as appropriate. The use of synonyms also accounted for British or American English spelling. To account for terminology differences across disciplines, database thesauri were consulted to ensure that all appropriate synonyms were included. Where available, the team also relied on the database's index terms and/or free-text terms. For databases or websites with basic search functions, the review team adjusted the search terms to accommodate the limited functionality of search functions and adapted these customized search strategies to relevant keyword searches and/or topic/theme searches based on the test results of keyword combinations of search terms. The search of electronic databases was completed in February 2014. From November 2014 to January 2015, the review team contacted experts and authors of included studies, screened reference lists of included studies and conducted citation tracking in order to identify additional studies. The search dates for each source used during the primary and complementary search process are presented in Section 8.6.1 in the Appendix.

3.3.3 Complementary search

The primary search was complemented by hand-searching and screening of 35 websites, such as institutional and conference websites, five dissertation, thesis and grey literature databases, nine other reviews and meta-analyses and literature

snowballing as well as contacting experts and relevant institutions. The hand-searching strategy was customized for each relevant institutional website. Search terms were used for websites that included a search facility. Otherwise, relevant sections (for example, “documents” or “publications”) were searched. Websites of conferences that were deemed relevant to the research question were searched for potentially relevant studies. To include potentially relevant dissertations and theses that were not indexed in bibliographic databases, the review team searched national and international dissertation and thesis databases. The review team also conducted citation tracking and screened reference lists of included studies and relevant existing reviews and meta-analyses to identify further studies for inclusion. The review team contacted authors of previous reviews and included studies, as well as experts and individuals coordinating youth employment related topics in relevant institutions, to ask whether they knew of any studies that might be applicable in addition to the studies that were included after the full text review of full reports. Ongoing and unpublished studies within the grey literature were identified through the screening and hand-searching of relevant websites/gateways and conference websites, citation tracking and contacting experts and relevant institutions. In addition, a keyword search was undertaken for the grey literature databases.

3.4 DATA COLLECTION AND ANALYSIS

3.4.1 Data extraction

Relevant information from included studies was systematically extracted using a coding tool and coding manual. The coding tool, which is presented in Section 8.7 of the Appendix, included information about variables related to study methods, the characteristics of the intervention and its implementation, the characteristics of the subject samples of analysis, the outcome variables and statistical findings, and contextual features.

At effect size level, the coding tool captured sub-group analysis of employment, earnings and business performance outcomes and estimated treatment effects by age cohorts, gender, educational level, income level and location, among other dimensions. Types of outcomes were further disaggregated by occupation category (dependent vs. self-employment), status of occupation (formal vs. informal) and conditional on other outcomes. To describe the data and empirical methods, the coding tool included information about the research design, statistical methodology, type of significance test, type and method of measurement, date of data measurement and data source. The coding tool also captured the form and year of publication.

For each category of intervention (i.e., skills training, entrepreneurship promotion, employment services and subsidized employment), the coding tool extracted information about the type of intervention, targeting and delivery mechanism, payment system and provider, duration of specific interventions, selection of

participants and conditionality of eligibility. General programme characteristics recorded details of the target group by age, gender, educational level, income level, location and employment status as well as the type of organizations involved in designing, financing and implementing the programme. The coding tool kept a record of region, country, scale and average duration of the programme. In addition to any awareness-raising efforts and gender considerations integrated into programme design and implementation, it also captured the incentives, monitoring mechanisms and sanctions for non-compliance connected with the programme.[13]

A separate section of the coding tool was used to record information when the study reported intermediary outcomes or outcomes other than the ones considered in this review. This section also captured additional sub-group analyses, relative treatment effects, general equilibrium effects, costs of the programme or cost-benefit analysis, as well as any implementation problems or empirical identification problems described by the author.

A coding manual provided detailed instructions for coders in order to ensure consistency in extracting and interpreting relevant information, in particular with regard to the selection of appropriate treatment effect estimates. Guidelines were provided to identify the treatment effect estimates with the lowest risk of bias when studies reported multiple estimates for the same types of outcomes. Coders selected the preferred method of estimating the effect and their choices were verified by a second reviewer. For example, estimates based on experimental designs were considered to provide the lowest risk of bias followed by natural experiments and quasi-experimental designs. Other considerations outlined in the manual to mitigate the effects of potential bias included the use of covariates, the type of data used and the statistical methodology applied for the estimation.

Information extracted from included studies was discussed with a second reviewer and coding decisions involving assumptions were documented by each researcher. Further information about the selection of studies and data extraction can be found in the protocol of the systematic review (Kluve et al., 2014).

3.4.2 Standardizing effect size estimates

To compare estimated treatment effects across studies, the standardized mean difference (SMD) was computed for both continuous outcome variables (e.g., income) and dichotomous outcome variables (e.g., employment probability) reported in the primary studies. In addition, researchers computed a binary variable holding the value of one if a treatment effect was positive and statistically significant

[13] To minimize the number of missing values in programme-related variables considered relevant for the analysis, additional information was gathered from sources beyond the study (which is the core unit of analysis), including project reports and project websites. The variables coded from these sources were: monitoring mechanisms, participant profiling, incentives to participants (for programme participation and/or performance), and incentives to service providers (payments conditional on outcomes of programme participants), and are presented in Section 8.7 of the Appendix.

(PSS).[14] The main results were based on the SMD measure, while the PSS indicators served for descriptive purposes and sensitivity analysis.

Box 6: Understanding effect sizes

Effect size is a generic term used to describe the estimated treatment effect for a study. This treatment effect is the observed relationship between an intervention and an outcome. In order to compare effect sizes across studies and outcome constructs, this systematic review used a meta-analysis to synthesize the data extracted from primary studies.

The SMD was used as a summary statistic in meta-analysis to combine results from studies which used different ways of measuring the same outcome (e.g., income). The SMD is a dimensionless measure of the relative magnitude of the treatment effect, which allowed estimated treatment effects to be compared across studies and different outcome constructs. The direction and magnitude of the effect of the intervention on reported outcomes of interest were essential data elements in assessing the effectiveness of active labour market programmes for youth.

For the analysis in this systematic review, estimated treatment effects were extracted from the primary studies and SMDs were computed. An SMD of zero indicates that the intervention, on average, resulted in an equivalent effect for the treatment group and the (comparison) group which did not receive the treatment; whereas an SMD greater than zero indicates the degree to which, on average, the treatment group had a better outcome.
Source: Authors, adapted from *Cochrane handbook for systematic reviews of interventions* (Version 5.1.0). Available at http://handbook.cochrane.org/ [24 Mar. 2016] and Glossary of Cochrane terms at http://community-archive.cochrane.org/glossary [24 Mar. 2016].

The SMD captured the relative magnitude of the treatment effect in a way that is dimensionless and hence comparable across outcomes and studies. It was the ratio of the treatment effect for a specific outcome relative to the standard deviation of that outcome within the evaluation sample used to estimate the treatment effect. Most studies reported either matching- or regression-based estimates of the treatment effect (even for RCT-based designs).[15] Hence, SMDs in most cases were computed using the formulae given by Waddington et al. (2004, p. 372f), namely:

For studies using parallel group or matching-based strategies Hedges' g and its standard error SE_g were computed as

$$g = \frac{\bar{Y}_t - \bar{Y}_c}{S_p} * \left[1 - \frac{3}{4 * (n_t + n_c - 2) - 1}\right]$$

$$SE_g = \sqrt{\left[\frac{n_t + n_c}{n_c * n_t} + \frac{g^2}{2*(n_c + n_t)}\right]}.$$

[14] A treatment effect was considered to be statistically significant if it had a p-value from a two-tailed test of less than 0.05.

[15] This is in line with experiences documented by previous systematic reviews in related fields, such as Baird, Ferreira, Özler and Woolcock (2013) or Tripney et al. (2013).

Where $\bar{Y}_t$ and $\bar{Y}_c$ are the mean outcome in the treatment group and comparison group, respectively. Similarly, n_t and n_c are the respective sample sizes. The term in brackets represents the small-sample correction procedure developed by Hedges and Olkin (1985). The numerator of g represents the causal raw impact of the programme on the outcome. In matching-based studies, $\bar{Y}_t - \bar{Y}_c$ is reflected by the average treatment effect on the treated (ATET). S_p is the pooled standard deviation of the outcome after treatment and is computed as:

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

With S_t and S_c as the standard deviation in the treatment and comparison group respectively (Hedges' approach). If either the comparison or treatment group's standard deviation was not reported, the standard deviation of the total sample S_T or the comparison group standard deviation was used to compute g. In the case of dichotomous outcome variables, the S_t and S_c were computed based on the number of observations and the proportion in the respective group, if available.

For partial effect sizes estimated using multivariate analysis, g and its standard error were estimated based on formula described in Keef and Roberts (2004):

$$g = \frac{\hat{\beta}}{\hat{\sigma}}$$

$$SE_g = \sqrt{\frac{g^2}{v-2} * (\frac{v}{t^2} + v * [c(v)]^2 - v + 2)} \quad \text{, where } \frac{1}{c(v)} = \sqrt{\frac{v}{2}} * \frac{\Gamma(\frac{v}{2}-\frac{1}{2})}{\Gamma(\frac{v}{2})}$$

Where $\hat{\beta}$ refers to the coefficient of the treatment variable in the regression, $\hat{\sigma}$ is the pooled standard deviation of the outcome, v is n-k degrees of freedom and $\Gamma()$ is the gamma function.[16]

There are two approaches for the calculation of the pooled standard deviation from regression-based studies. In Hedges' approach, $\hat{\sigma}$ is the standard deviation of the error term in the regression. As this was rarely reported, the team followed Cohen's approach and computed $\hat{\sigma}$ from the standard deviation of the dependent variable across all observations (S_T) (cf. Lipsey & Wilson, 2001):

$$\hat{\sigma} = \sqrt{\frac{S_T * (n_t - 1) - (\beta^2 * (n_c * n_t)/(n_c + n_t))}{n_t - 1}}.$$

If information for calculating SE_g was not available, it was approximated by

$$SE_g = \frac{g}{t}$$

where t is the t-value associated with a t-test on the treatment effect of a regression.

[16] For studies with large n, $c(v)$ was considered equal to 1.

If none of the values for S_p, S_T or S_c could be obtained from the report (or by contacting the authors), the standard deviation of the outcome variable was approximated using the formula from Borenstein, Cooper, Hedges and Valentine (2009)

$$S_p = SE * \sqrt{\frac{n_c * n_t}{n_c + n_t}}$$

where SE in the standard error of a means test (e.g., regression coefficient). Since this formula is technically only correct for bivariate effect sizes, a sensitivity analysis was performed on the sample without these imputations.

For some studies, the review team transformed reported effect size statistics (often *t*, *F*, *p* or *z*-values) prior to calculating effect sizes following the procedures suggested in Lipsey and Wilson (2001).

Prior to synthesizing computed effect sizes, checks were made for outliers which could have been a result of erroneous coding or misleading assumption in the computation of SMD. In cases where SMDs or their standard errors seemed implausibly large, the original reports were revisited to check whether these were in accordance with the findings stated by the authors. In cases where the effect sizes where correctly coded and computed but still appeared implausible, the authors were contacted for clarification. As it was not possible to solve all outlier issues following this approach, the data from remaining outliers was censored, initially by winsorizing the data and then finally by dropping any remaining outliers. Winsorizing is a method of censoring data by limiting extreme values in the statistical data to reduce the effect of possibly spurious outliers. Winsorizing refers to a method where all outliers are set to a specific percentile. In this case, the top 1 per cent and bottom 1 per cent of observations were set to the value of the 1st and 99th percentile, respectively. The robustness of the results was tested with respect to the level of winsorizing and the cut-off ranges for trimming outliers.

3.4.3 Unit of analysis issues

Originally, it had been planned to correct the standard errors for a possible unit of analysis error by adjusting the standard errors (SE_g) according to the formula suggested in Higgins and Green (2011, p. 502ff). A unit of analysis error typically arises if the study conducts analysis and programme placement at different levels and the analysis does not adequately account for this clustering (e.g., by using cluster robust standard errors or variance components analysis). In such cases, the analysis would yield narrower confidence intervals than the true confidence intervals, increasing the risk of Type-I error. This can be a problem in cluster randomized trials or in quasi-experimental studies in which treatment allocation is clustered. However, no studies were identified where there was a suspicion that the unit of analysis was not adequately addressed in the statistical analysis.

3.4.4 Dealing with missing data

In several instances, primary reports did not supply sufficient information to compute standardized effect sizes from reported treatment effect estimates. Most often, post-intervention mean and/or standard deviation of the outcome variable could not be obtained. The frequency with which missing information was encountered indicates that better reporting standards are required for impact evaluations studies.

As a first step, authors of included papers were contacted to provide missing information and to clarify discrepancies. This was an important and time-consuming measure, carried out via standardized letters and missing information forms, into which authors or research assistants could easily insert the results and data requested.[17]

Initially, the review team reached out to the authors of 100 included reports (note that this number represents almost the entire sample of included reports in the systematic review), requesting additional information to facilitate the computation of the effect sizes or to achieve clarity on the quantitative results or interventions details.[18] In the event that an author did not reply, the same request was sent two more times. In total, the authors of 34 papers replied, while no response was received from the authors of 63 reports. In the remaining cases, no valid, up-to-date email addresses could be found for the authors.

In several instances, information was missing about, for instance, standard deviations, sample sizes or average outcomes in the comparison group follow-up data collection. In these cases, the missing data were imputed from available information based on specific assumptions. For instance, when the overall sample size was provided but not the sample sizes for the treatment and comparison groups separately, an assumption of equal sample sizes was made (splitting the overall sample size in half). The same assumption was applied in cases in which only the treatment or comparison group sample size was reported. Results from a meta-analysis were reported based on the more conservative sample (without imputing missing information) as part of the sensitivity analysis.

In cases where the information necessary to compute an effect size (e.g., sample size, mean outcomes and/or standard deviation) could not be derived from the available information, the effect size was excluded from the analysis.[19]

---

[17] In order to increase response rates, however, the review team did not ask authors for missing information on intervention characteristics.

[18] The difference between the total number of included studies and the total number of missing responses is mainly due to studies for which no additional information was required.

[19] In some cases, the required information could be obtained for the overall sample, but not for specific sub-group analysis. In those cases only the specific effect size was included but not the entire study.

3.4.5 Dealing with dependent effect sizes

In a meta-analysis, the unit of analysis is the study. Section 3.2.2 clarified that a single *programme* could include more than one *intervention*, which was regarded as the review's primary unit of interest (instead of the overall impact of one programme, the team was interested in the impact of each specific intervention). Each intervention may have been evaluated by more than one *study* (e.g., evaluation), each of which may have been published in multiple *reports* (e.g., working papers, technical reports or journal publications). Two reports were treated as part of the same study if they were based on the same data and hence could not be treated as independent, even if they were written by different authors. Therefore, an intervention population (all participants) might be different from the study population (all in one data set), which might itself differ from the sample population for a specific treatment effect estimate on a specific outcome construct.

Estimated treatment effects may be regarded as independent from each other when the underlying data were derived from different sample populations. To maintain the independence assumption, it was important that only one effect size per outcome construct and study was included in the analysis (Borenstein, Hedges, Higgins and Rothstein, 2009). However, each report might present different treatment effect estimates for the same outcome construct and the same sample population – for example for different sub-group analyses or employing different statistical methods. This implied that different estimates within each study (sometimes across reports) had to be combined into one effect size per sub-group.

Creating effect size aggregates and summary effect sizes (e.g., at the intervention or study level or across different sub-groups as part of the moderator analysis) required careful estimation to avoid the situation where a single group of participants influenced the summary effect size disproportionately. For example, a treatment effect might be reported in a study for the entire (pooled) sample and subsequently reported for sub-groups of the same sample, such as males and females.[20] The median number of treatment effect estimates per study in the sample was 12, with some reports providing more than 100 estimates. In such instances, a multitude of treatment effects could be reported for the same group where there was no a priori reason to give preference to one measure over another.

In these scenarios it was possible to mitigate the disproportionate influence on the aggregate effect sizes by applying the following steps. First, by identifying a set of effect sizes that were derived from the same independent group of participants and then, where applicable, selecting the effect sizes for this group where it was possible to establish a preference. (For example, keeping only pooled estimates and discarding sub-group estimates except when needed in the analysis.) By dropping some of the effect sizes derived from the sample this redundancy was removed from

[20] No studies were encountered in the sample which assessed different treatments using the same group of individuals as the comparison group (multi-arm studies with pooled comparison).

the analysis as far as possible.[21] This method provided a better approach to the data than averaging effect sizes across all overlapping sub-groups. [22]

Second, in cases where multiple effect sizes were reported for each independent group without clear justification for dropping some rather than others (e.g., the where same outcomes were reported at several points in time for the same group), the aggregate ("synthetic") effect sizes were estimated for each independent group. Based on the method for combining effect sizes from the same independent population suggested by Borenstein, Hedges, Higgins and Rothstein (2009), the approach was as follows: Let g_{ij} and SE_{ij} be the ith effect size, where $i = (1, \ldots, \mathrm{m})$, and its standard error, respectively, for the sample population identified by j. To arrive at a single combined (aggregate) effect size for group j the team took the simple average:

$$g_{\mathrm{j}} = \frac{1}{\mathrm{m}} \sum_{\mathrm{i}=1}^{\mathrm{m}} g_{\mathrm{ij}}$$

and the standard error of g_{j} given by

$$SE_{\mathrm{j}} = \sqrt{\left(\frac{1}{\mathrm{m}}\right)^2 \left(\sum_{\mathrm{i}=1}^{\mathrm{m}} g_{\mathrm{ij}}^2 + \sum_{\mathrm{i} \neq \mathrm{k}} \rho_{\mathrm{ikj}} \sigma_{\mathrm{ij}} \sigma_{\mathrm{kj}}\right)}$$

where ρ_{ikj} is the correlation coefficient between g_{ij} and g_{kj} in study j.[23]

Hence, the independent group aggregates were assembled at the relevant unit of analysis, such as at the intervention or study level (depending on the assumed correlation addressed in the procedure). Then the random-effects meta-analysis was applied to the aggregated data and estimated summary effect sizes.

3.4.6 Synthesis methods

Summary effect sizes are provided for the three outcome categories: (1) employment outcomes, (2) earnings outcomes and (3) business performance outcomes. The summary effect sizes were estimated via a random-effects meta-analysis based on

[21] Here, redundancy indicates providing additional information about a group that is not needed for the desired level of aggregation. For example, if the goal is to create programme aggregates for all participants, then male and female sub-group estimates may be dropped. On the other hand, if the goal is to create an aggregate for females for each programme, then pooled estimates would be dropped.

[22] For the purpose of brevity the guidelines used to drop effect sizes within each group are not included here. This information is available upon request.

[23] The first best option is to attempt to estimate ρ_{ik} from the data. However, in cases where there was an insufficient number of observations, then some assumption about ρ_{ik} had to be be made. Assuming that $\rho_{ik} = 0$ is likely to overestimate precision, and assuming that $\rho_{ik} = 1$ is likely to underestimate precision, the more conservative assumption was adopted, that $\rho_{ik} = 1 \; \forall \; (i, j)$ where $i \neq k$.

the intervention-outcome level aggregates using the –metan– command in Stata.[24] Random-effects meta-analysis is recommended in settings which present significant contextual heterogeneity in terms of study population, intervention and implementation. To account for differences in individual studies' sample sizes, effect sizes were averaged across studies by using inverse-variance weighting of the individual effect sizes. This weighting resulted in the individual effect sizes from studies with larger sample size being given more weight in the combined effect size. The summary effect sizes generated in this manner are presented alongside the 95 per cent confidence intervals in forest plots (Section 4.3). In addition to the aggregate effect size, these forest plots display the weight each intervention carries towards the summary effect size.

Heterogeneity tests were used to examine whether the variation in effect size estimates within outcome categories was larger than expected from sampling error alone (Deeks, Altman & Bradburn, 2001). To test for heterogeneity, the team employed I^2 statistics and Q-statistics, as well as the τ^2 statistic. These statistics tested whether the percentage of variability in effect estimates was due to heterogeneity rather than by chance. A significant Q (p-value <0.05) and an I^2 value of at least 50 per cent were considered to be indicators of heterogeneity.

3.4.7 Moderator analysis

Moderator analyses were performed when there was evidence of heterogeneity. The analyses tested hypotheses about whether variation in the (average) effect sizes reported in studies was associated with differences in study, participant and intervention characteristics (moderators). These moderator analyses also served as a test for correlations of effect size magnitude with specific characteristics of interventions and population groups. They therefore formed the basis for the answers to the research questions regarding factors of intervention effectiveness.

In a first step, a univariate approach was implemented, analogous to an analysis of variance (ANOVA) analysis, again via a random-effects meta-analysis based on the intervention-outcome level aggregates. Specifically, the review team investigated heterogeneity within outcome categories by (i) main intervention category, (ii) country income level, (iii) gender, (iv) participant income status and (v) time elapsed after programme completion. Results from these models are presented in the form of forest plots in Section 4.3.3.

Ideally, moderator analysis should be conducted with a minimum of ten studies for each individual moderator variable (Borenstein, Hedges, Higgins & Rothstein, 2009). A decision was made to present forest plots for sub-groups (e.g., intervention types) that had at least four individual interventions. The number of effect size estimates and individual interventions for each sub-group are displayed in the

[24] A detailed elaboration of random-effects meta-analysis estimation is given in Borenstein, Hedges, Higgins and Rothstein (2009).

respective forest plots to provide the reader with an indication of the size of the evidence base.

In a second step, a multivariate random-effects meta-regression model was estimated using the –metareg– command in Stata. This allowed the team to test which factors correlated with the magnitude in the effect size estimate (SMD) while controlling for other potentially moderating factors, such as research design. As discussed above, multivariate meta-regression is particularly well suited to drawing inferences from literature that reports diverse estimates and where there is heterogeneity in the combination of intervention characteristics and research designs.

The regression models fitted to the data were random-effects meta-regression models of the type:

$$g_{\mathrm{ij}} = \beta_0 + \sum \beta_k Z_{\mathrm{jk}} + u_{\mathrm{j}} + \varepsilon_{\mathrm{ij}}$$

where g_{ij} is the ith effect size estimate (SMD) from study j. Z_{ik} represents K (k = 1, …, K) moderator variables representing study-level or intervention-level heterogeneity. $uj \sim (0,2)$ is the study-level error term with τ_{j}^2 as the between study variance in true effects, assumed equal across estimates within each study. $\varepsilon_{\mathrm{ij}} \sim N(0, \sigma_{\mathrm{ij}}^2)$ is the random-effects error term and σ_{ij}^2 the standard error corresponding to effect size estimate g_{ij}.

The team estimated the random-effects meta-regression model using method of moments (DerSimonian and Laird 1986) and Knapp and Hartung (2003) adjusted standard errors, which can be considered more conservative and rigorous than other regression approaches. Note, however, that the multivariate meta-regression model was estimated on the effect size level (i.e., prior to aggregating effect sizes as described in Section 3.4.5). Consequently, this multivariate analysis does not explicitly take into account dependencies across effect size estimates as was the case for the bivariate analysis. The main reason behind this decision was that aggregating effect sizes prior to the multivariate analysis caused methodological difficulties in estimating the effect of various covariates that often vary within studies (e.g., gender). The robustness of the results was tested under a robust-variance estimation (Hedges, Tipton & Johnson, 2010) in a restricted specification and sample. The team did not find that this altered regression results substantially.

A large array of study-level, intervention-level and contextual variables were identified and coded which it was assumed could be correlated with the reported effect size. The code description in Section 8.7 in the Appendix provides an overview of all variables that are included in the multivariate meta-regression. In addition to these, tests were conducted for the influence of various other moderator variables but the decision was made to exclude any that were deemed non-significant.

3.4.8 Supporting interpretation of effect sizes: The percentage change

In addition to the SMD, the review team computed the simple percentage change of the intervention over the control group mean as a more intuitive indicator of the intervention's impact. The percentage change was calculated by dividing the raw effect size (i.e., the mean difference between treatment and comparison group) by the mean value of the outcome variable for the comparison group. As a consequence, the percentage change indicates the direction of change for the treatment groups, with negative values meaning that the treatment group's outcome was lower than the comparison group's. This percentage change was then averaged by independent effect size group (i.e., by a grouped combination of intervention and study level). Subsequently, the team weighted the group-wise percentage changes using the inverse-variance weights (as throughout the analysis) and computed the final percentage changes.

3.4.9 Sensitivity analysis

A range of sensitivity checks were conducted to test the robustness of the results. Sensitivity analysis was carried out by restricting the meta-analysis to a subset of all studies included in the original meta-analysis. First, following guidance from the Campbell Collaboration (2014, p. 9), an examination was carried out to establish whether findings were influenced by the rigour of the evidence. Specifically, the team tested heterogeneity across study design (randomized vs. quasi-experimental) and publication status (published vs. unpublished studies). Second, the sensitivity of the results was tested with regard to the assumptions made for computing the SMD effect size in the presence of missing information. Third, the team tested the validity of the method of dealing with statistical outliers (dropping observations vs. winsorizing the data).

3.4.10 Risk of bias and study design assessment

During the research and coding process, the team found that impact studies often lacked important details that would allow a confident appraisal of the plausibility of the identifying assumptions on which the empirical analyses were based. This lack of detailed reporting in many publications limited the extent to which a full risk of bias assessment, for example, based on Waddington and Hombrados (2012), was possible and would have been informative. As a consequence, an alternative framework was adopted (proposed in Duvendack, Palmer-Jones, Copestake, Hooper, Loke & Rao (2011) and Duvendack, Hombrados, Palmer-Jones & Waddington (2012)) in order to assess the statistical rigour of primary studies. This approach combined an assessment of both research design and the method of statistical analysis. In addition to the original approach, the assessment was further disaggregated by the statistical method (DiD, statistical matching, etc.) used for addressing potential confounders of the original research design (randomized experiment, natural experiment, etc.). By placing RCTs at one end of the spectrum and cross-section designs at the other, the tool aimed to reflect the potential capacity

of different empirical identification strategies to control for possible biases.[25] In addition to the sensitivity analysis, the team therefore tested whether different research designs and empirical approaches yielded different effect sizes on average.

3.4.1.1 Assessment of reporting bias

Publication bias or "file drawer effects" refers to the underreporting of studies which establish a negative or mixed evaluation finding (Franco, Malhotra & Simonovits, 2014). The review team assessed the danger of publication bias in the sample of included studies by several means. First, by testing the influence of study design and publication status as part of the sensitivity analysis. Second, by performing standard tests for publication bias: plotting the effect size against standard errors (funnel plots) using the –metafunnel– and –metacum– commands in Stata. Moreover, the team also implemented Egger, Davey Smith, Schneider and Minder's (1997) meta-regression test using the –metabias– command in Stata. The idea underlying the small-sample assessment to detect publication bias is that "researchers who have small samples and low precision will be forced to search more intensely across model specifications, data, and econometric techniques until they find larger estimates" hence "such considerations suggest that the magnitude of the reported estimate will depend on its standard error" (Doucouliagos & Stanley, 2012).

Tests were also made to establish whether there were observable differences in reported effect sizes between peer-reviewed and unpublished studies. For example, it was possible that estimates reported in the top journals might contain greater selection bias because these journals are more selective (Stanley, 2013).

3.5 DEVIATIONS FROM THE PROTOCOL

The protocol of the systematic review was published in November 2014 and was followed by the implementation of the search and selection process outlined in the protocol. The primary and complementary search process benefitted from the extensive scoping search and development of tailored search strategies for each source prior to the publication of the protocol, allowing the review team to follow the planned search process closely. The main search in electronic databases was completed in February 2014. The systematic search resulted in a high number of studies to be screened, classified and coded in 2014. While the selection and data extraction process was ongoing, the review team decided to consider additional sources that were made available in 2014 (e.g., studies presented at the Doha Evidence Symposium in March 2014). Following the selection and data extraction process, the review team contacted experts and authors of included studies, screened reference lists of included studies and conducted citation tracking in order to identify additional studies from November 2014 to January 2015.

[25] As the authors emphasize, this framework should not be taken to endorse a universal "hierarchy of methods" but rather as providing an objective and efficient framework for assessing the potential risk of bias in randomized and quasi-randomized studies.

During the process of data collection and synthesis, the team made changes to the coding tool and empirical methodology which represent deviations from the protocol published by the Campbell Collaboration Group, i.e., Kluve et al., 2014.

- In addition to the variables proposed in the protocol, three additional intervention-level variables were coded. The variables relate to the design of the intervention and were deemed relevant and apriori strongly correlated to reported effect sizes:
    - Participant profiling for services provided: The variable captured whether the intervention (i) identified individual factors or characteristics that implied a risk in the labour market and (ii) relied on such information to assign youth to specific services. Examples include caseworker discretion, screening or specific eligibility rules.
    - Incentives to participants: Capture whether participants received payments conditional on (monitored) programme participation or success. This also included participants' eligibility to welfare or unemployment benefits.
    - Incentives to service providers: This variable captured whether payments (or bonuses) to the implementing agency were conditional on outcomes of intervention participants.

- The protocol had proposed to review specific cases were evaluations measured general equilibrium or spillover effects. However, the frequency of such analyses and measures was so low that the review team focused on studies looking into partial equilibrium effects.

- Given its relevance in policymaking, the protocol had proposed the coding and analysis of Intention-to-Treat (ITT) estimates. The plan was to approximate ITT estimates from studies which reported only Average Treatment Effect on the Treated (ATET) estimates, using the formula suggested in Bloom (2006). However, of those studies estimating the ATET, very few reported the share of individuals who were originally assigned to the treatment group but did not take up treatment (i.e., non-compliers, defiers or no-shows). The approximation proved to be especially difficult for quasi-experimental studies, as the distinction between ITT and ATET estimates was not always clear. Instead of converting treatment effect estimates, the team decided to test differences between ITT and ATET estimates as part of the sensitivity analysis.

4 Results

4.1 DESCRIPTION OF STUDIES AND INTERVENTIONS

4.1.1 Search results and selection of studies

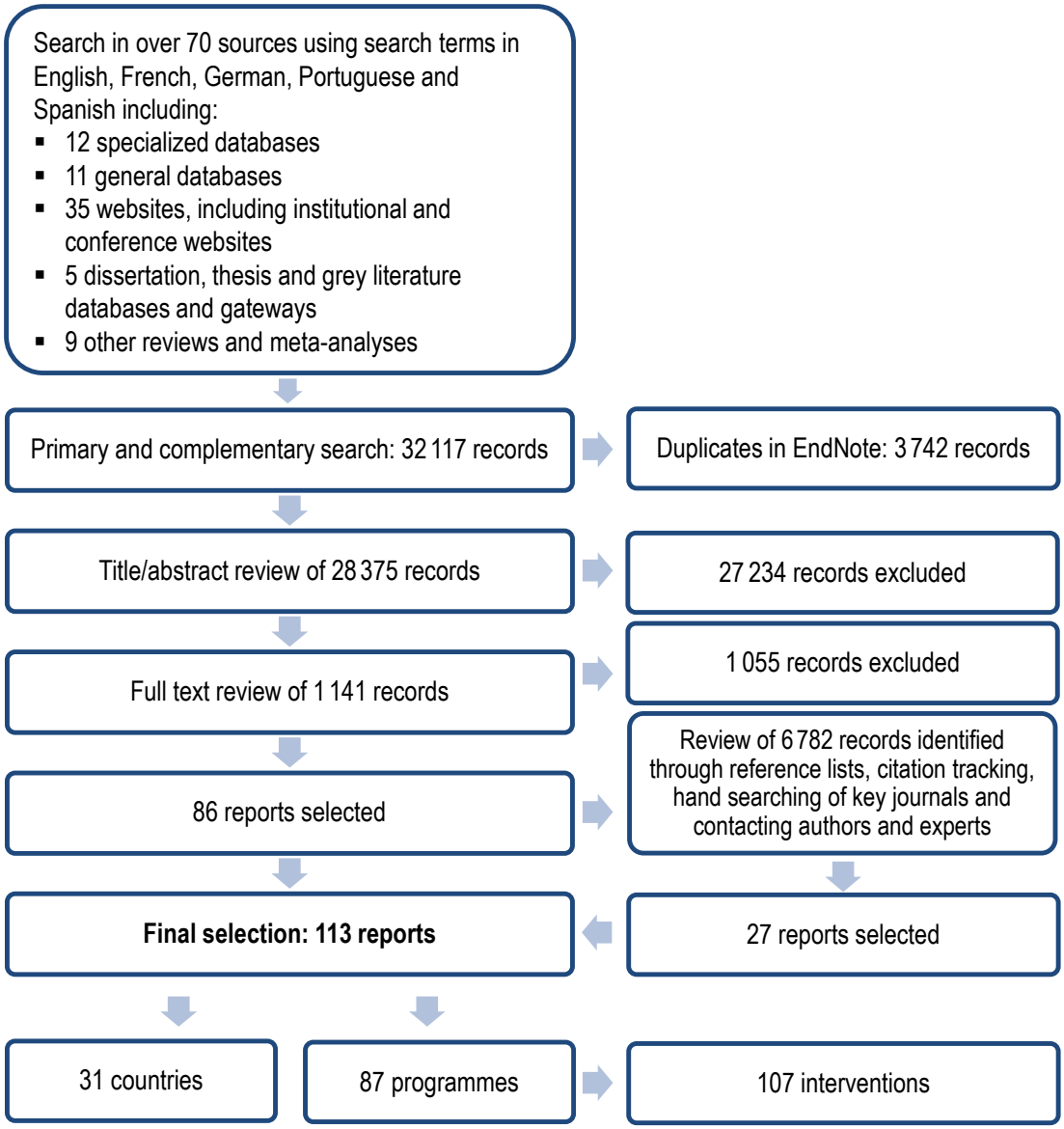
The primary and complementary search identified 32,117 records, based on a search of over 70 sources, including 12 specialized databases, 11 general databases, 35 websites, such as institutional and conference websites, five dissertation, thesis and grey literature databases, and nine other reviews and meta-analyses. The search in electronic databases was completed in February 2014. From November 2014 to January 2015, the review team contacted experts and authors of included studies, screened reference lists of included studies and conducted citation tracking in order to identify additional studies. The list of included sources as well as the search dates for each source used during the primary and complementary search process are presented in Section 8.6.1 in the Appendix. After removing duplicates in the reference management software EndNote, screening of 28,375 records by title and abstract was carried out by individual reviewers, applying the inclusion criteria of the screening questionnaire (see Section 8.6.2 of the Appendix). A total of 1,141 records were identified for full text screening.

In order to minimize bias, included and excluded results were cross-checked by a second researcher and discrepancies were resolved by both researchers. This systematic screening process led to the identification of 86 reports which were considered to be of adequate content and methodological rigour to inform the systematic review. The main reasons for excluding reports at full-text stage were the following criteria: study design, target group, intervention. In addition, several reports were excluded because a more recent or updated version of the same report was available, the report only focused on relative effects, the impact evaluation study was ongoing or the report did not examine any of the outcomes of interest considered in this review. Examples of excluded reports and the reasons for excluding them are presented in Section 6.2.

After extracting data from the preliminary set of 86 included reports, the review team screened 6,782 additional records that were identified through reference lists and citation tracking of included studies, hand searching of key journals in which a large number of included studies were found and contacting authors and experts.

This search process led to the selection of 27 additional reports. Overall, this comprehensive search and selection process identified 113 reports which were considered eligible for inclusion in this review. The search and screening process is illustrated in Figure 2.

Figure 2: Search results



4.1.2 Characteristics of included reports

The systematic screening process led to the identification of 113 reports that met all criteria for inclusion (Section 3.2).[26] As shown in Table 9, panel A, more than half of the impact evaluations of youth employment interventions were conducted in high-income countries where there is an established practice of results measurement,

[26] Note: Each intervention may have been evaluated by more than one *study* (e.g., evaluation), each of which may have been published in multiple *reports* (e.g., working papers, technical reports or journal publications). Further information about the relation between *study* and *report* is provided in Section 3.4.5.

particularly with regard to government employment measures. The large share of reports from high-income countries in this systematic review (65 out of 113 reports representing eleven of the 31 countries in the sample) is an important feature that justifiably suggests that some caution should be exercised when interpreting the results in global terms.

Table 9: Characteristics of included reports

	n	%		*n*	%
(A) Country income level			**(E) Timing of evaluation follow-up**		
High-income country	65	58	Less than or equal to one year	58	51
Low- and middle-income country	48	42	Longer than one year	71	63
(B) Year of publication			**(F) Sub-group analysis in addition to the overall analysis**		
1991–2000	14	12	Gender disaggregated	56	50
2001–2005	20	18	Low-income participants	4	4
2006–2010	27	24	Education level of participants	13	12
2011–2014	52	46			
			(G) Outcome category		
(C) Type of publication			Employment	98	87
Peer-reviewed journal	41	36	Earnings	91	81
Working paper	28	25	Business performance	10	9
Evaluation/technical report	30	27			
Other (book/dissertation)	14	12	**(H) Main intervention**		
			Skills training	74	65
(D) Evaluation design			Entrepreneurship promotion	12	11
Experimental	53	47	Employment services	11	10
Natural experiment	11	10	Subsidized employment	17	15
Quasi-experimental	50	44	Unspecified	9	8

Note: *n* = 113. Reports may not be exclusive across the different typologies in this table, e.g., one study may estimate multiple outcomes or examine more than one intervention type.

The number of reports assessing the impact of youth employment interventions has increased steadily over the past few years (panel B), with nearly half of the sample published after 2010 and 21 reports published in 2014 alone.[27] Interestingly, this surge in evaluation has benefitted developing countries by providing a greater quantity of better quality evidence about what works to support youth in the labour market. There were 48 reports of interventions implemented in low- and middle-income countries, with a particular prevalence of impact evaluations in Latin America and the Caribbean.

It appears the extensive search effort was successful in identifying a variety of publications from the grey literature (panel C). Only around one-third of the reports

[27] In contrast, the 2007 synthesis of the Youth Employment Inventory reported 73 studies with a counterfactual-based impact evaluation of youth employment programmes implemented between 1950 and 2006 (Betcherman et al., 2007). Notably, most impact evaluations recorded in the inventory and implemented prior to 1990 took place in high-income countries (mainly the United Kingdom and the United States).

come from peer-reviewed journals, with the remainder split between working papers, technical reports from implementing organizations and others, such as books or dissertations. Most of the reports published in 2014 were working papers, identified through the complementary search process, making the detection of grey literature one of the strengths of this review.

While the review focused on counterfactual impact evaluations, the search process uncovered a large variety of different evaluation designs, namely experimental designs, natural experiments and quasi-experimental designs (as discussed in Section 3.2.5). In contrast to other systematic reviews, this review contained a significant share of randomized experiments (53 reports, as shown in Table 9, panel D). Many of the results from these randomized control trials (RCTs) have been published recently (66 per cent after 2010) and hence were not included in previous reviews. Figure 3 shows the recent surge in rigorous evidence. Prior to 2011 most RCTs in the sample were conducted in high-income countries (Figure 4), while the past five years have seen a remarkable increase in RCTs in developing countries. Most notably, in 2014, 12 out of 15 RCTs included in this review were from low- and middle-income countries; seven of them evaluating youth employment programmes in Africa (Box 7).

Quasi-experimental designs, such as panel and cross-sectional evaluations were the second most common study design (50 reports), frequently relying on propensity score matching (PSM) and difference-in-difference (DiD) for causal inference under unconfoundedness. Quasi-experimental designs have been more widely employed over the past decade, with approximately 40 reports published after 2004 in Latin America and the Caribbean and in OECD countries. Finally, the review included 11 natural experiments, all of which were implemented in high-income and upper middle-income countries between 2004 and 2014.

In relation to the evaluation features, 39 reports provided impact estimates at multiple time points. In addition, 71 reports measured changes in outcomes of interest over 12 months after treatment exposure (panel E). These longer term effects were estimated primarily across skills training interventions. Few studies provided a sub-group analysis in addition to the overall analysis (panel F). In particular, only half of the reports in the sample provided separate results for males and females (excluding those that evaluated gender-targeted programmes). Very few reports in the sample provided separate treatment effects for disadvantaged, low-income or low-educated youth.

Table 9 also provides an overview of the types of outcomes measured across the included reports (panel G). Three-quarters of the reports in the sample reported results for more than one type of outcome. Employment and earnings outcomes were extensively reported. Employment probability was by far the most commonly measured and reported outcome within the set of reports. More than 88 reports provided an estimate of the programme impact on employment probability. Another 35 reports estimated the effect of an intervention on hours worked.

Figure 3: Total number of reports and reports relying on RCTs by year of publication

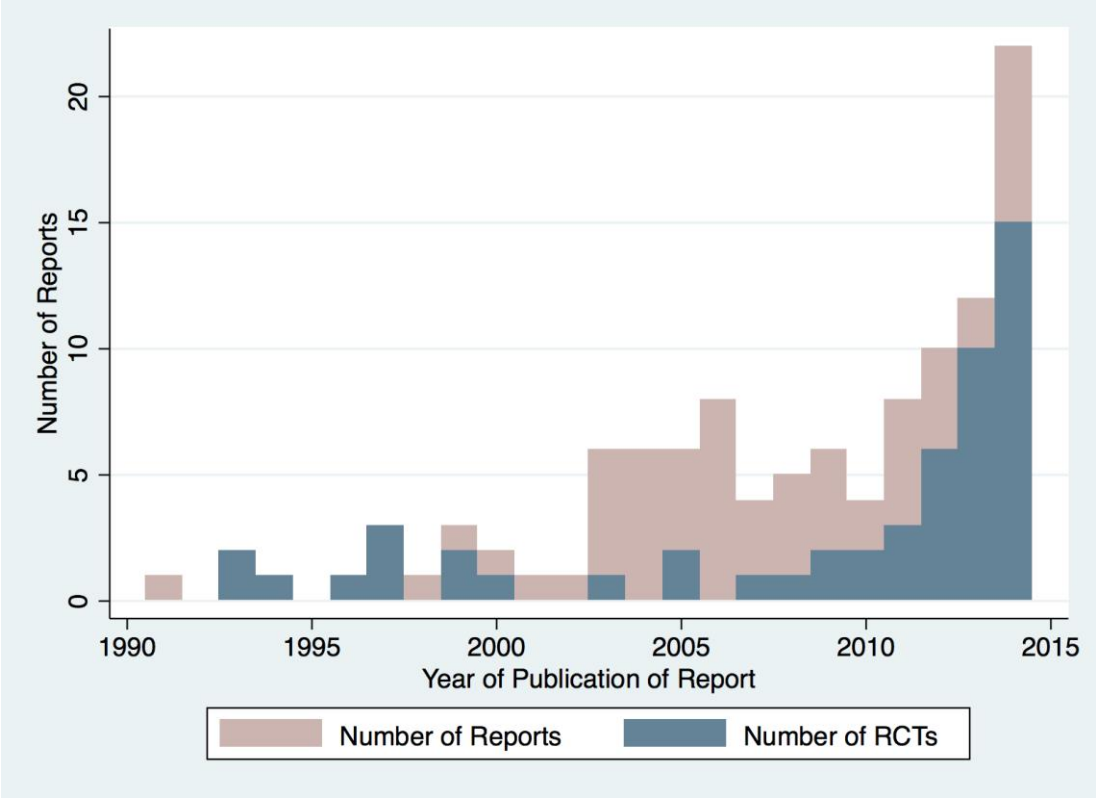
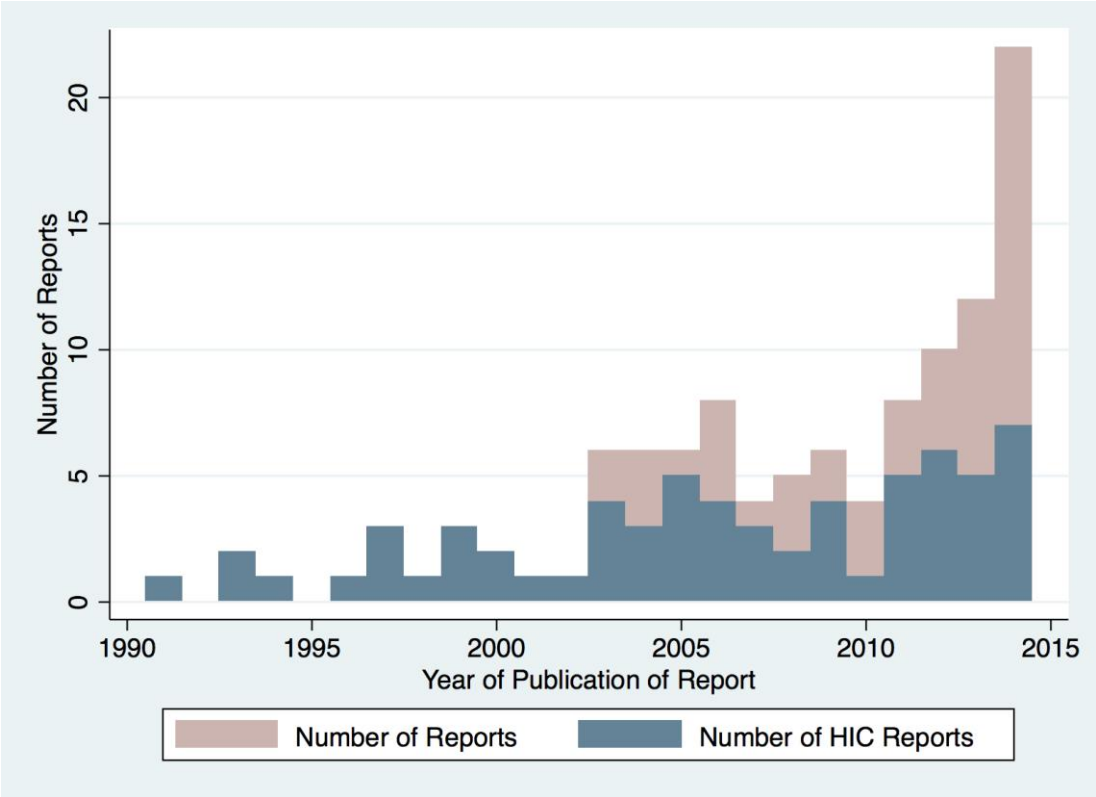


Figure 4: Total number of reports and reports from high-income countries (HICs) by year of publication



Box 7: Studies of youth employment interventions in Africa

This review included 13 reports of impact evaluations carried out in African countries. None of these 13 reports predated 2010. Most (nine studies) were published in working papers with only two reports published in peer-reviewed journals (by January 2015). With only one exception, all quantitative results came from RCTs, which often reported the intention-to-treat estimator as well as the effect of the intervention on the average participants who completed the programme – this was due to compliance problems which are common across evaluated interventions in the region.

Only six reports measured changes in outcomes of interest over a year after the young person's exposure to the intervention. This s an important aspect, as labour market impacts often materialize only over the long term.

Studies focused mainly on assessing changes in employment (13 reports) and earnings outcomes (12 reports), and to a lesser extent on understanding changes in business performance, survival or expansion (six reports). A sizable number of entrepreneurship promotion interventions were implemented in Africa and included in the review (eight out of 17).

Source: Based on a background report on African studies (Pasali, 2015).

Table 9 also displays the limited number of reports (ten out of 113) measuring changes in business performance outcomes. Nine of these related to RCTs. They were most commonly found among interventions aiming to promote entrepreneurship among young people.

4.1.3 Characteristics of evaluated interventions

As shown in Figure 2, the search process led to 113 reports that assessed impacts of 87 youth employment programmes. The review drew a key conceptual distinction between programmes, interventions and components (Section 3.2.2). Youth employment programmes can consist of one or more interventions. These are exclusive tracks offered to discrete samples of participants. For example, in the New Deal for Young People programme, implemented in the United Kingdom and described in Box 8, youth had to choose one of four different tracks, namely, (i) education or training; (ii) a job with a voluntary sector employer; (iii) a job on the environmental task force; or (iv) employment in a wage subsidy programme.

Interventions, on the other hand, have one or several components, which were classified as skills training, entrepreneurship promotion, employment services or subsidized employment measures. Table 10 provides an overview of the 107 interventions in the review. Main category (panel A) refers to the interventions where it was possible to identify a primary component. In line with previous reviews (Kluve, 2010; Betcherman et al., 2007), skills training proved to be the most common type of main intervention category, followed by subsidized employment, entrepreneurship promotion and employment services.

Table 10: Characteristics of included interventions

	n	%		n	%
A) Main category			**D) Country income level**		
Skills Training	55	51	High-income country	60	56
Entrepreneurship Promotion	15	14	Low and middle-income country	47	44
Employment Services	10	9			
Subsidized Employment	21	20	**E) Intervention Region**		
Unspecified	6	6	Sub-Sahara Africa	15	14
			Europe and Central Asia	4	4
B) Has Component			Latin America and Caribbean	22	21
Skills Training	68	64	Middle East and North Africa	6	6
Entrepreneurship Promotion	17	16	South Asia	4	4
Employment Services	40	37			
Subsidized Employment	25	23	**F) Scale of Intervention**		
			National	59	55
C) Combinations			Regional	21	20
Skills Training Only	32	30	Local or pilot	30	28
Entrepreneurship Promotion Only	14	13			
Employment Services Only	9	8	**G) Intervention features**		
Subsidized Employment Only	12	11	**Target group:**		
Skills Training & Entrepreneurship Promotion	1	1	Women only	16	15
Skills Training & Employment Services	24	22	Unemployed at intervention start	48	45
Skills Training & Subsidized Employment	8	7	Low-Income/Disadvantaged Youth	45	42
Entrepreneurship Promotion & Employment Services	1	1	**Implementated with participation of:**		
Employment Services & Subsidized Employment	3	3	Government	75	70
Skills Training & Employment Services & more	3	3	Private Sector	63	59
			NGO/Non-profit	37	35
			Mulitlateral organisation	11	10

Note: *n* = 107.

There were six interventions for which no main category of intervention could be identified, and these were therefore classified as unspecified. Their components were bundled in such a way that made it impossible to identify one type of intervention as being predominant over the others. They were truly multi-dimensional in nature and formed part of the following programmes: Active labour market programme for disadvantaged youth in Germany (study by Ehlert, Kluve & Schaffner, 2012); the National Guard Youth Challenge Programme in the United States (study by Millenky, Bloom, Muller-Ravett & Broadus, 2011); the New Chance Programme in the United States (studies by Chang, Huston, Crosby & Gennetian, 2007 and Quint, Bos & Polit, 1997); the New Deal for Young People in the United Kingdom (studies by Blundell et al., 2004, De Giorgi, 2005 and Wilkinson, 2003); the Teenage Parent Demonstration in the United States (study by Maynard, Nicholson & Rangarajan, 1993); and the Youth Opportunity Grant Initiative in the United States (study by Jackson et al., 2007). Details of the New Deal for Young People in the United Kingdom are presented in Box 8.

Box 8: New Deal for Young People (NDYP) in the United Kingdom

The New Deal for Young People (NDYP) was introduced in the United Kingdom in 1998 and aimed to help the young unemployed into work and to increase their employability by combining different types of interventions, especially job-search assistance and subsidized employment. Participation was mandatory for all people aged 18–24 who had claimed unemployment benefit (Jobseeker's Allowance) for a period of six months or more. Participants entered a "gateway" period of intensive job-search under the supervision of a personal adviser, intended to last no longer than four months. Those who were still receiving the Jobseeker's Allowance at the end of the gateway period were obliged to take one of four options: (i) entry into full-time education or training for those without basic qualifications; (ii) a job with a voluntary sector employer; (iii) a job on the environmental task force; (iv) employment in a wage subsidy programme. In addition, under the terms of the scheme, employers were obliged to offer education or training on at least one day per week.

Evaluations showed that the programme appeared to have generated an increase in the probability of young men (who had been unemployed for six months) finding a job within the next four months (Blundell, Costa Dias, Meghir & Van Reenen, 2004) and suggested that a period of subsidized employment was a more effective means of exiting unemployment and securing unsubsidized employment than the other options available under NDYP (Dorsett, 2006).

Sources: based on information available at: www.youth-employment-inventory.org [20 Feb. 2016].

While the remaining interventions had one main component to address the labour market constraints of youth, more than one-third extended the intervention's scope with one or more additional measures. As panel B shows, some 64 per cent of interventions in the review incorporated a skills training component; but almost half of these combined skills training with some other measure. The most common combination was skills training and employment services, observed in 27 interventions.

Entrepreneurship promotion interventions that focused on youth were comparatively scarce. Entrepreneurship-related components were only reported in 17 interventions, and these components often seemed to be delivered in a way that was disconnected from other active labour market measures. It is important to highlight that the results chain for entrepreneurship promotion (Table 3) already incorporates the delivery of training services in relation to entrepreneurial and business development and management skills, avoiding potential overlaps between skills training and entrepreneurship promotion categories.[28]

As discussed above, the majority of the reports included in this review assessed impacts of youth employment programmes implemented in high-income countries, which translated into a sample of 60 interventions (panel D). There were 56

[28] While both skills training and entrepreneurship promotion interventions comprised training activities and skills development goals, they were differentiated by their overall objective and, often, different target groups.

interventions (52 per cent) from OECD countries alone (panel E), a proportion comparable to those seen in previous reviews (e.g., Card et al., 2010 and 2015; Betcherman et al., 2007).

The second largest share of impact evaluations stemmed from interventions in Latin America and the Caribbean, where many countries have experimented with active labour market policies (ALMPs) since the early 1990s – particularly through quasi-experimental designs embedded in the Jóvenes Programmes, a series of skills training interventions implemented throughout the region[29] (see Box 2 for an example).

The review captured 17 interventions evaluated in Africa (15 in sub-Saharan Africa and two in North Africa, panel E), all of which were covered in recent impact evaluations published after 2011. In contrast, there was a relatively small number of evaluated interventions from other developing and emerging regions: Four in Europe and Central Asia, South Asia and the Middle East, respectively. There was no evidence from interventions implemented in East Asia and the Pacific.

With regard to scale (panel F), most interventions had a national coverage. In 30 cases the evaluations examined localized interventions implemented as pilots. The disaggregation across urban and rural areas demonstrated a significant lack of evidence about what works to support rural youth. The review's sample included only six evaluated interventions in rural areas, 33 in urban areas, and 62 interventions at a national scale with (imputed) coverage at both urban and rural level.

A close examination of programme targeting (panel G) led to the identification of 16 interventions (15 per cent) designed to serve only young women, 48 interventions (45 per cent) targeting youth who were unemployed prior to joining the intervention and 45 (42 per cent) that focused exclusively on low-income and disadvantaged youth.

Public and private sector actors were the most common implementing entities. Their implementing role was more prevalent among high-income countries, while evaluated interventions with an implementation role for non-governmental organizations (NGOs) and non-profit organizations tended to be more common in low-income countries.

Detailed descriptions of the intervention features and overall treatment effects are presented in the Appendix in Sections 8.1 to 8.5.

[29] An interesting learning curve in evaluation methods in Latin America and the Caribbean is shown in Figure 48, in Section 9 of the Appendix.

4.2 ASSESSMENT OF INCLUDED STUDY DESIGNS

Impact studies often lacked important details that would allow a confident assessment of the plausibility of the identifying assumptions on which the empirical analyses were based. In order to assess the rigour of the designs in the included primary studies, the review team used the framework proposed in Duvendack et al. (2011, 2012). The approach combined an assessment of both the research design and the method of statistical analysis leading to an implicit hierarchy of study designs, with RCTs as the most rigorous design and cross-section designs at the bottom end of the expected accuracy of results scale. Given the study design, the rigour of the statistical analysis was also a function of the statistical methods, ranging from more advanced methods, such as DiD, PSM, instrumental variables (IV), or regression discontinuity designs (RDD), to multivariate regressions and simple (means) tabulations. In addition to the sensitivity analysis, the team tested whether different research designs and empirical approaches yielded different effect sizes on average.

Table 11 shows the classification of evaluation reports that were included in the systematic review. Almost half of the reports (67 reports or 47 per cent of cases, see Table 9) were conducted as RCTs, meaning that the accuracy of the results associated with them is naturally high. Twelve RCTs did, however, rely only on simple tabulation methods, slightly reducing the expected rigour of their analysis. Figures in Table 11 count the number of cases when a particular report relied on a particular statistical method. It was possible, for example, for the same RCT to rely on more than one method, which explains why the total number of RCTs in Table 11 surpasses that reported in Table 9.

Table 11: Number of reports for study design assessment

	Statistical methods of analysis										
Research design	Di D	Regresion-adjusted DiD	Matching and DiD	Matching	IV	RDD	Multivariate		Tabulation	Other	Total
							Linear	Non-linear			
RCT	2	2	2	0	10	0	29	4	12	6	67
Natural experiment	1	1	3	1	0	2	0	0	0	4	12
Pipeline	1	0	1	0	0	0	0	0	0	0	2
Only panel	3	5	14	5	0	0	1	0	0	1	29
Only cross-section	3	1	0	12	1	0	4	5	3	0	29
Total	**10**	**9**	**20**	**18**	**11**	**2**	**34**	**9**	**15**	**11**	**139**

Notes: Based on Duvendack et al. (2012). One research design could rely on more than one statistical method of analysis. "Other" includes 11 cases (nine reports) that could not be readily classified within the other statistical methods of analysis. They comprised non-parametric statistical approaches (three reports), a combination of matching and IV (two reports) and principal stratification approaches (two reports). Given that these are rather sophisticated methods (more than a simple tabulation of means), their occurrence with RCTs or natural experiments was considered of potential low risk of bias.

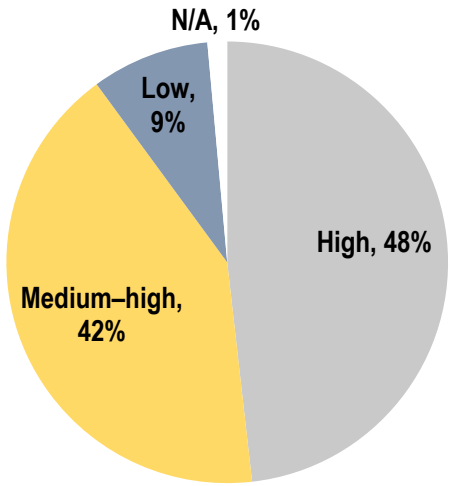
A further 11 per cent of reports (12 cases reported in Table 11) were based on natural experiments, combined with sophisticated statistical methods that went beyond

simple tabulation of means. Accordingly, these studies can also be categorized as having a high level of rigour in their analysis.

There were a total of 60 reports with pipeline, only panel or only cross-section designs (Table 9, under quasi-experimental designs). In 46 cases they used and/or combined DiD, PSM, IV or RDD methods, associated with a high to medium level of accuracy. There were only 12 instances (11 per cent) of low statistical rigour, when the above-mentioned designs relied on multivariate analysis or tabulation methods. There was only one unclassified report that combined panel and multivariate analysis.

In summary, the analysis showed that the included reports were generally very rigorous pieces of work, with almost 48 per cent of cases presenting a high level of rigour, 42 per cent a high–medium level, and only 9 per cent with a low level (Figure 5). This finding somewhat alleviated concerns of prevalent biases to the internal validity of included reports – which might in turn have rendered the results invalid. However, it was clear that the design approach could only provide a first approximation of potential factors affecting the internal validity of empirical research designs. Section 4.3.5 assesses the robustness of the results as part of the review's sensitivity analysis by testing whether studies classified as having a low level of statistical and analytical rigour contained statistically significant different effect sizes in comparison to more rigorous studies.

Figure 5: Share of included reports by study design rigour



4.3 SYNTHESIS OF RESULTS

The following set of tables displays the main results of the systematic review and meta-analysis, which will be further explained throughout the Section.

Table 12: Summary of results by outcome category

Parameters of interest	Employment outcomes	Earnings outcomes	Business performance outcomes
Standardized Means Difference I^2	0.04	0.05	0.03
Standard errors of the effect size	0.01	0.01	0.04
95% confidence interval	0.03 0.06	0.03 0.06	-0.05 0.12
I Squared	63.66	81.64	48.83
Number of SMDs	1,330	670	169
Number of interventions	105	92	14
Sample size	38,219,046	12,696,812	62,905
Mean difference	4.67	2,084.35	-109.75
Control outcome	30.65	10,355.27	499.91
Treatment outcome	29.14	11,210.02	349.60
Percentage change	0.07	0.09	-0.06

Table 13: Summary of results for employment outcomes by main category of intervention

Parameters of interest	Skills training	Entrepreneurship promotion	Employment services	Subsidized employment	Unspecified
Standardized Means Difference I^2	0.05	0.16	0.01	0.02	0.03
Standard errors of the effect size	0.01	0.05	0.01	0.02	0.04
95% confidence interval	0.02 0.07	0.06 0.26	-0.02 0.04	-0.01 0.06	-0.04 0.1
I Squared	64.77	71.41	0	50	0
Number of SMDs	904	43	104	193	86
Number of interventions	67	7	10	16	5
Sample size	3,439,703	61,502	2,340,789	32,198,189	178,863
Mean difference	7.78	3.11	-1.09	0.51	-2.65
Control outcome	38.74	6.14	19.16	9.76	53.56
Treatment outcome	38.10	7.61	14.41	6.40	48.89
Percentage change	0.08	0.24	0.02	0	0.1

Table 14: Summary of results for earnings outcomes by main category of intervention

Parameters of interest	Skills training	Entrepreneurship promotion	Employment services	Subsidized employment	Unspecified
Standardized Means Difference I^2	0.07	0.09	0.01	-0.01	Dropped from analysis
Standard errors of the effect size	0.01	0.04	0.01	0.02	
95% confidence interval	0.05 0.08	0.01 0.18	0 0.02	-0.05 0.03	
I Squared	85.7	63.81	0	61.24	
Number of SMDs	495	50	36	57	
Number of interventions	60	12	8	9	
Sample size	2,045,960	42,530	194,713	10,358,155	
Mean difference	3,781.16	2,346.78	-16.53	-1524.22	
Control outcome	15488.55	2,744.07	226.15	7,057.33	
Treatment outcome	17274.87	3,639.95	186.38	5,535.96	
Percentage change	0.11	0.22	0	-0.01	

Note: Dropped from the analysis indicates the number of studies was insufficient (lower than four) to allow for the computation of the effect sizes.

Table 15: Summary of results for business performance outcomes by main category of intervention

Parameters of interest	Skills training	Entrepreneurship promotion
Standardized Means Difference I^2	Dropped from analysis	0.1
Standard errors of the effect size		0.05
95% confidence interval		0 0.19
I Squared		39.09
Number of SMDs		162
Number of interventions		10
Sample size		58,519
Mean difference		47.72
Control outcome		371.77
Treatment outcome		395.75
Percentage change		0.15

Notes: Dropped from the analysis indicates the number of studies was insufficient (lower than four) to allow for the computation of the effect sizes. The table does not show report on all categories of intervention as related studies did not measure changes in business performance outcomes.

4.3.1 Descriptive analysis of effect size estimates

To synthesize the results of the 113 empirical reports of youth employment interventions, the review relied on the reported treatment effect as a measure of impact. The search and screening process led to the identification and coding of 3,629 treatment effects. Based on the reported (or acquired) information, it was possible to compute the direction and statistical significance for 3,105 treatment effect estimates. The computation of standardized mean difference (SMD) required further information (the minimum requirement being the number of observations in treatment and/or comparison groups). Even after using the methods of imputing missing information described above (Section 3.4.4), it was only possible to compute the SMD from 2,259 reported treatment effect estimates, as shown in the third column of Table 16.

Table 16: Sample size of treatment effect estimates

				SMD and standard error		
		Coded treatment effects	Computed statistical significance and direction	Computed SMD (Hedges' *g*) and standard error		No. of independent studies
				Total	of which imputed	
(A)	**Type of outcome**					
	Employment	2 410	1 983	1 403	466	107
	Earnings	1 045	949	682	377	94
	Business performance	174	173	174	14	15
(B)	**Evaluation follow-up timing**					
	Less than or equal to one year	1 125	1 058	727	240	69
	Longer than one year	1 657	1 435	977	303	75
(C)	**Main Intervention category**					
	Skills training	2 182	2 036	1 488	555	75
	Entrepreneurship promotion	264	262	260	46	15
	Employment services	205	171	140	50	10
	Subsidized employment	570	461	251	135	16
	Unspecified	408	175	120	71	5
	Total	**3 629**	**3 105**	**2 259**	**857**	**121**

It was possible to compute a substantially higher number of effect sizes than in other systematic reviews thanks to an intensive and demanding effort to acquire missing

information from authors. There were 121 independent samples to account for dependencies within studies due to overlap of the study population across effect estimates, as described in Section 3.4.5 Dealing with dependent effect sizes.

Among all effect sizes in the data set, the proportion that were significant is similar to other meta-analyses in the social sciences. About one-third of the reported impact estimates were positive and statistically significant at the 5 per cent significance level and only 7.6 per cent of effect sizes were negative and statistically significant. In comparison, Card et al. (2010) found 39 per cent positively and significant effect sizes, while Cho and Honorati (2013) reported 28 per cent to be positive significant. Grimm and Paffhausen (2014) reported that 46 per cent of impact estimates were positive and statistically significant at the 10 per cent level.

4.3.2 Univariate random-effects meta-analysis

The following sections discuss results from the univariate meta-analysis approach to explore the differences in average effect size estimates across interventions in the sample. The analysis built on forest plots, which are commonly used to graphically describe the results of a meta-analysis. Forest plots are based on an inverse-variance weighted least squares random-effects meta-analysis model (see Box 9).

To improve the readability of this report, only "summary" forest plots are included in the main text. These provide the summary estimate for each sub-group in the respective analysis and, where appropriate, the respective overall summary SMD.[30] Fifteen "disaggregated" forest plots, with study-level SMDs for each outcome category and main intervention category, are provided in the Appendix.[31] Results presented in these forest plots were based on the sample using all imputations available and winsorizing the top 1 per cent of statistical outliers. Results from the restricted sample and/or obtained under different assumptions regarding outliers are presented as part of the sensitivity analysis in Section 3.4.9.

Synthesis of the overall evidence by outcome

Figure 6 presents the overall summary effect sizes for each selected outcome category of interest – namely, employment outcomes, earnings or income outcomes and business performance outcomes.[32] Note that an individual study may have contributed to multiple outcome categories and hence the individual sub-groups may not be independent (in other words, the same sample of participants may have provided an estimate for earnings and employment outcomes, in which case the two estimates are not independent. Consequently, the overall effect size in the figure should not be interpreted.

---

[30] The overall summary SMD is not displayed in cases where a single study may provide estimates for multiple categories in the sub-group analysis, as may be the case for different outcome measures or gender, for example.

[31] Disaggregated forest plots for each moderator analysis in the following sub-sections can be obtained from the authors on request.

[32] See the corresponding full forest plot in Appendix Section 9.1, Figure 49.

Box 9: Reading a forest plot

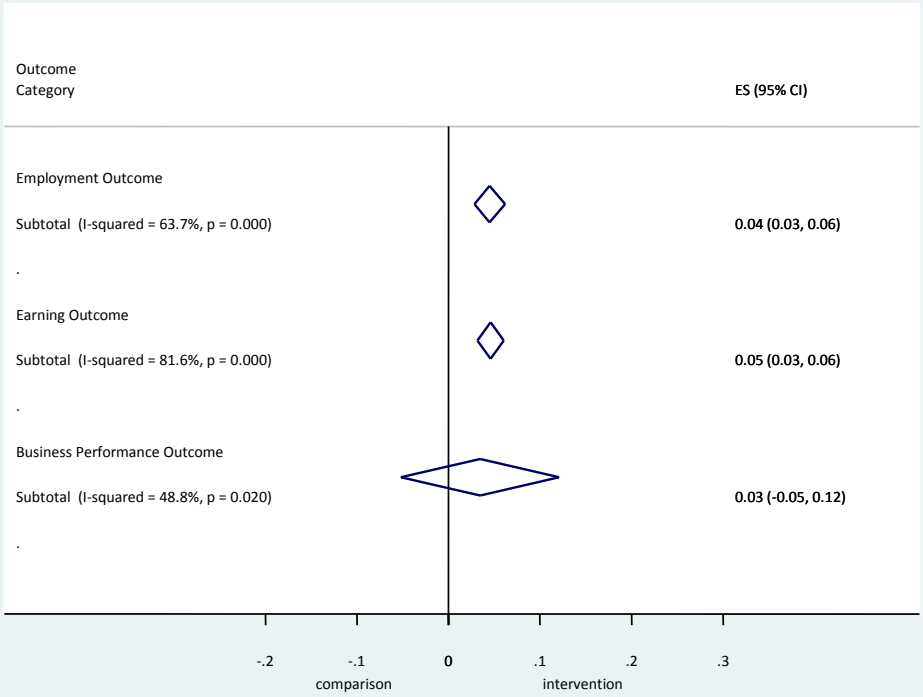
This review presents effect size estimates and confidence intervals for the respective outcomes of interest of an intervention. This information is displayed in forest plots, which can be read as follows:

- Each sub-group (for summary plots) or intervention (for full plots) is represented by one line in the plot.
- The SMD is reported under effect size (ES), along with its corresponding confidence interval. The same information is represented graphically through the diamonds. An SMD greater than zero indicates that, on average, the treatment group had a better outcome than the (comparison) group, which did not receive the treatment. This is considered a positive effect.
- The vertical, unbroken line represents no effect from the interventions on the outcomes of interest.
- The edges of the diamonds represent the confidence interval (CI). For instance, in the summary forest plots shown below, the size of the diamonds represents the confidence interval per sub-group analysed in the respective plot.
- The weight is the inverse of the variance of that particular sub-group or intervention. It shows the contribution or strength that each particular sub-group (for summary plots) or intervention (for full plots) gives to the overall summary effect size.
- The overall effect estimate is reported at the bottom of the plot. The SMD value is further marked by a vertical dotted line, making it easier to compare where sub-group SMDs fall in relation to the overall SMD.
- The level of heterogeneity is captured in the I^2 statistic.
- Notes below each aggregate forest plot provide the number of SMDs and the number of independent studies that form the basis for each computed summary SMD.

Employment and earnings outcomes were the largest contributors to the overall meta-analysis: 105 of 119 independent studies estimated an employment outcome and 92 estimated an earnings outcome.[33] The overall effect on earnings outcomes across all intervention categories was 0.05 SMDs (CI = 0.03, 0.06; I^2 = 82 per cent; number of interventions = 92) and statistically significant at the 5 per cent level. The summary effect on employment outcomes was similar and also statistically significant (0.04 SMD; CI = 0.03, 0.06; I^2 = 64 per cent; number of interventions = 105). Only impact estimates from studies that measured business performance outcomes exhibited a relatively large confidence interval and the summary effect was not statistically significant (0.03 SMD; CI = -0.05, 0.12; I^2 = 49 per cent; number of interventions = 14). At the same time, the plot also exposed high heterogeneity (represented by the I^2 statistic) within each outcome category, suggesting that a large share of the variation in effect sizes is explained by inter-study heterogeneity. Earnings outcomes displayed the highest I^2 value at 82 per cent, suggesting that more than three-quarters of the variation in the effect sizes is not by chance and rather due to heterogeneity between interventions.

[33] Discrepancies with Table 16 are due to the treatment of outliers prior to analysis.

Figure 6: Summary forest plot of all outcomes (full sample) by outcome category



No. of SMDs/Studies: Total: 2169/119
Employment Outcome: 1330/105. Earning Outcome: 670/92. Business Performance Outcome: 169/14.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The WINGS programme in Uganda (Blattman et al., 2014) (0.82 SMD; 95% CI = 0.13, 1.50), the PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18), and the Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceed the SMDs limit of 0.75 and were therefore not included in the analysis for the forest plot above.

In order to explore the factors driving such differences, the remainder of the report explored the effect sizes within each outcome category through moderator and sensitivity analyses. Since the number of independent studies that measured specific outcomes for certain moderators was small for some moderators, the review team only assessed those outcomes where at least four interventions were obtained.

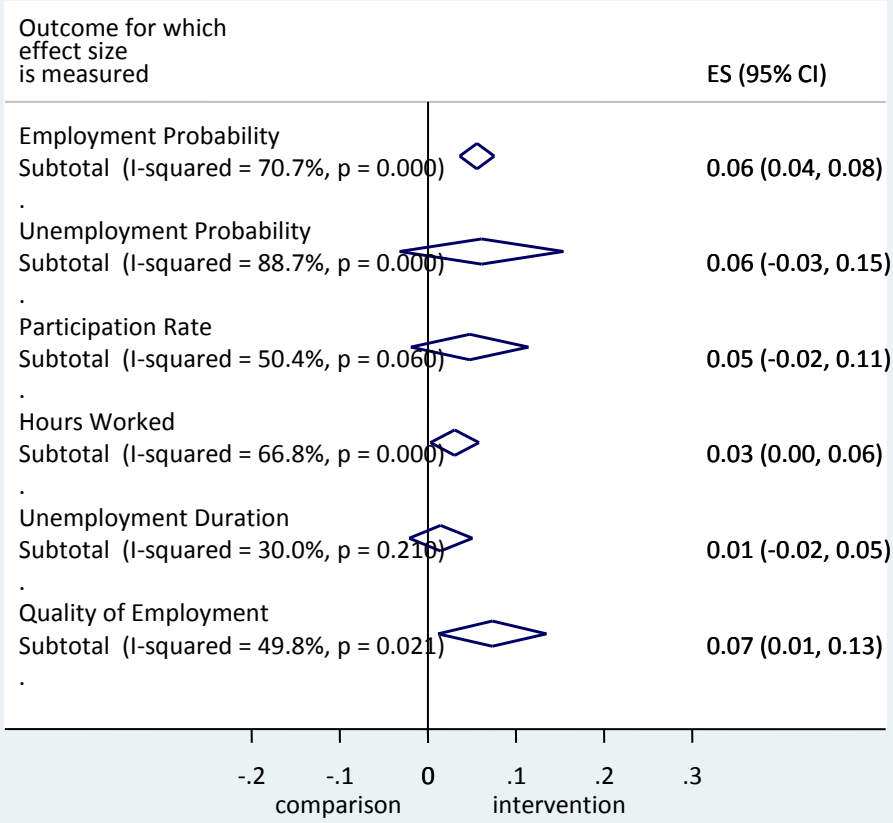
4.3.3 Univariate moderator analysis

As a first step, the team tested whether summarizing effect sizes within the three outcome categories presented a viable procedure or whether significant heterogeneity was already detectable across outcome constructs in each outcome category. Following this, tests for heterogeneity were carried out by investigating the influence of several factors as part of the moderator analysis: (i) main intervention type; (ii) country income level; (iii) time after exposure to treatment; (iv) study-level summaries of participant characteristics, including gender and participant's income status; (v) programme characteristics, including scale of the programme and implementing organization. The moderator analyses generally provided results that were stratified by main category of intervention in order to avoid "comparing the incomparable".

4.3.3.1 Outcome measure

To factor in the diverse nature of each outcome, the team assessed the effect size of each outcome measured separately by outcome category (see Figure 7, Figure 8 and Figure 9). Employment probability represented the largest share of effect sizes and carried the greatest weight across employment measures, with an SMD of 0.06 (CI = 0.04, 0.08; I^2 = 71 per cent; number of interventions = 98) followed by hours worked, with an effect size of 0.03 SMD (CI = 0.00, 0.06; I^2 = 67 per cent; number of interventions = 35). See as well Table 17 for further details.

Figure 7: Summary forest plot of employment outcomes (full sample) by outcome measure

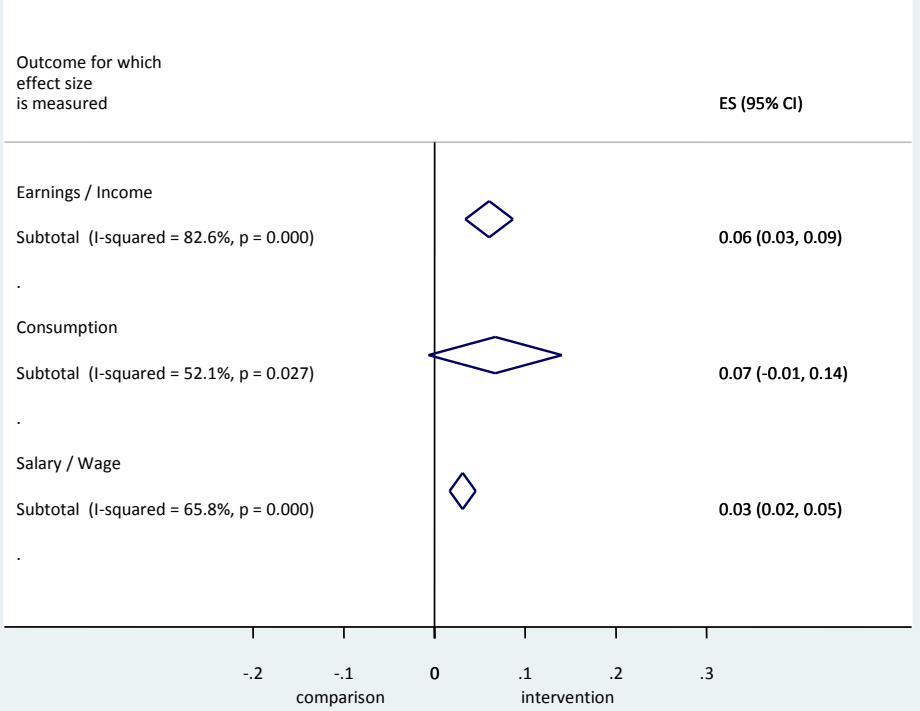


No. of SMDs/Studies: Total: 1312/105
Employment Probability: 1029/98. Unemployment Probability: 50/13. Participation Rate: 18/7. Hours Worked: 137/35. Unemployment Duration: 30/6. Quality of Employment: 48/13.
Sub-group dropped (<4 independent studies).
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Within earnings-related outcomes, wages and reported earnings drove the overall effect size, with individual effect sizes of 0.03 SMD (CI = 0.02, 0.05; I^2 = 66 per cent; number of interventions = 36) and 0.06 SMD (CI = 0.03, 0.09; I^2 = 83 per cent; number of interventions = 66), respectively. The significantly smaller sample of effect sizes for business performance outcomes presented greater variability, with overall negative effects on profits (-0.02 SMD; CI = -0.09, 0.05; I^2 = 0 per cent; number of interventions = 7) and sales (-0.06 SMD; CI = -0.16, 0.04; I^2 = 0 per cent; number of interventions = 5) and a large effect among capital and investment

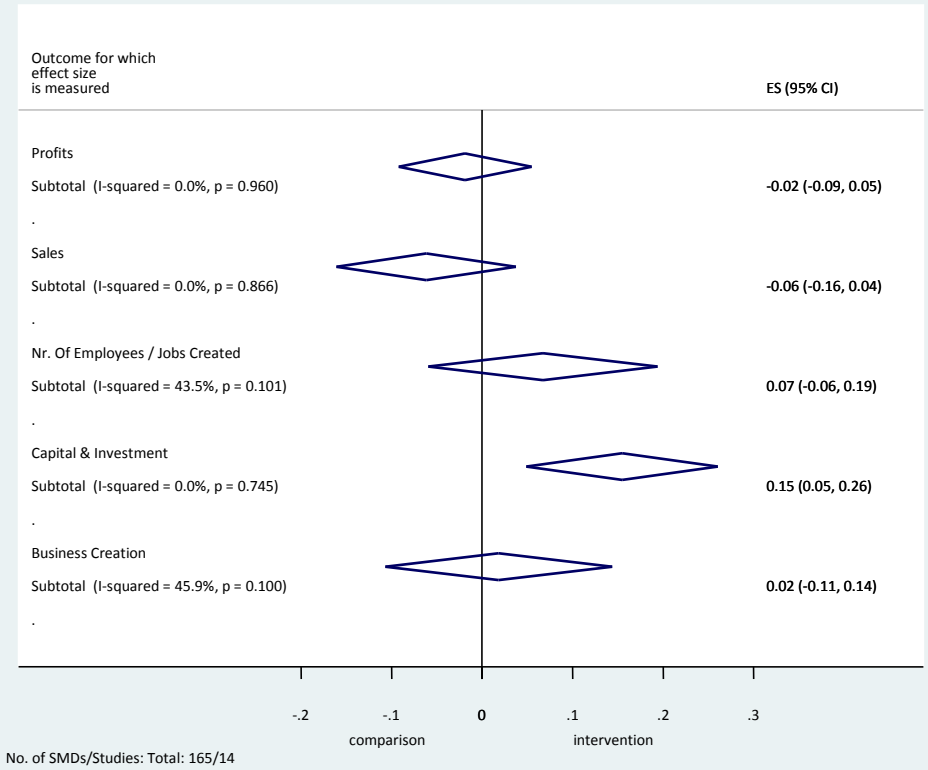
reported outcomes (0.15 SMD, CI = 0.05, 0.26; I^2 = 0 per cent; number of interventions = 6). Additional parameters are reported in Table 18 and Table 19.

Figure 8: Summary forest plot of earnings outcomes (full sample) by outcome measure



No. of SMDs/Studies: Total: 661/90
Earnings / Income: 473/66. Consumption: 29/10. Salary / Wage: 159/36.
Sub-group dropped (<4 independent studies): Household Income.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 9: Summary forest plot of business performance outcomes (full sample) by outcome measure



No. of SMDs/Studies: Total: 165/14
Profits: 59/7. Sales: 36/5. Nr. Of Employees / Jobs Created: 35/7. Capital & Investment: 26/6. Business Creation: 9/6.
Sub-group dropped (<4 independent studies): Business Survival.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The WINGS programme in Uganda (Blattman et al., 2014) (0.82 SMD; 95% CI = 0.13, 1.50) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Table 17: Summary of results for measures within employment outcomes

Parameters of interest	Employment probability	Unemployment probability	Participation rate	Hours worked	Unemployment duration	Quality of employment
Standardized Means Difference I^2	0.06	0.06	0.05	0.03	0.01	0.07
Standard errors of the effect size	0.01	0.05	0.03	0.01	0.02	0.03
95% confidence interval	0.04	-0.03	-0.02	0	-0.02	0.01
	0.08	0.15	0.11	0.06	0.05	0.13
I Squared	70.7	88.66	50.39	66.79	29.97	49.85
Number of SMDs	1029	50	18	137	30	48
Number of interventions	98	13	7	35	6	13
Sample size	29,200 313	3,889 055	2,120 112	1,798 941	579 513	223 329
Mean difference	2.73	0.04	0.05	-24.20	-1.60	0.01
Control outcome	0.31	0.23	0.51	299.27	21.11	1.14
Treatment outcome	0.31	0.23	0.57	289.76	16.92	1.09

Parameters of interest	Employment probability	Unemployment probability	Participation rate	Hours worked	Unemployment duration	Quality of employment
Percentage change	0.1	-0.19	0.01	0.03	0.06	0.47

Table 18: Summary of results for measures within earnings outcomes

Parameters of interest	Earnings/Income	Household income	Consumption	Salary/wage
Standardized Means Difference I^2	0.06	Dropped from analysis	0.07	0.03
Standard errors of the effect size	0.01		0.04	0.01
95% confidence interval	0.03 0.09		-0.01 0.14	0.02 0.05
I Squared	82.58		52.13	65.85
Number of SMDs	473		29	159
Number of interventions	66		10	36
Sample size	2,860 231		22 837	9,801 683
Mean difference	4,517.90		3,334.44	474.79
Control outcome	19159.10		10199.94	3,595.88
Treatment outcome	20740.19		13236.30	4,023.22
Percentage change	0.11		0.33	0.06

Note: Dropped from the analysis indicates the number of studies was insufficient (lower than four) to allow for the computation of the effect sizes.

While there was heterogeneity across the different outcome measures within each outcome category, this was not statistically significant based on the random-effects meta-analysis model within each outcome category (the 95 per cent confidence interval of the sub-group average included the overall mean represented by the dotted red line), except for the case of unemployment duration and capital and investment measures). Based on these results, the team was confident that it was viable to pool results across outcome measures in the subsequent analysis.

Table 19: Summary of results for measures within business performance outcomes

Parameters of interest	Profits	Sales	No. of employees/Jobs created	Capital and investment	Business creation	Business survival
Standardized Means Difference I^2	-0.02	-0.06	0.07	0.15	0.02	Dropped from analysis
Standard errors of the effect size	0.04	0.05	0.06	0.05	0.06	
95% confidence interval	-0.09 0.05	-0.16 0.04	-0.06 0.19	0.05 0.26	-0.11 0.14	

Parameters of interest	Profits	Sales	No. of employees/Jobs created	Capital and investment	Business creation	Business survival
I Squared	0	0	43.51	0	45.94	
Number of SMDs	59	36	35	26	9	
Number of interventions	7	5	7	6	6	
Sample size	21 972	11 373	9 938	12 384	4 867	
Mean difference	-463.92	-141.49	75.54	268.39	0.00	
Control outcome	917.73	1 021.32	142.91	1 115.14	0.25	
Treatment outcome	356.31	920.79	203.94	1 198.96	0.25	
Percentage change	-0.35	-0.18	0.06	0.65	-0.1	

Note: Dropped from the analysis indicates the number of studies was insufficient (lower than four) to allow for the computation of the effect sizes.

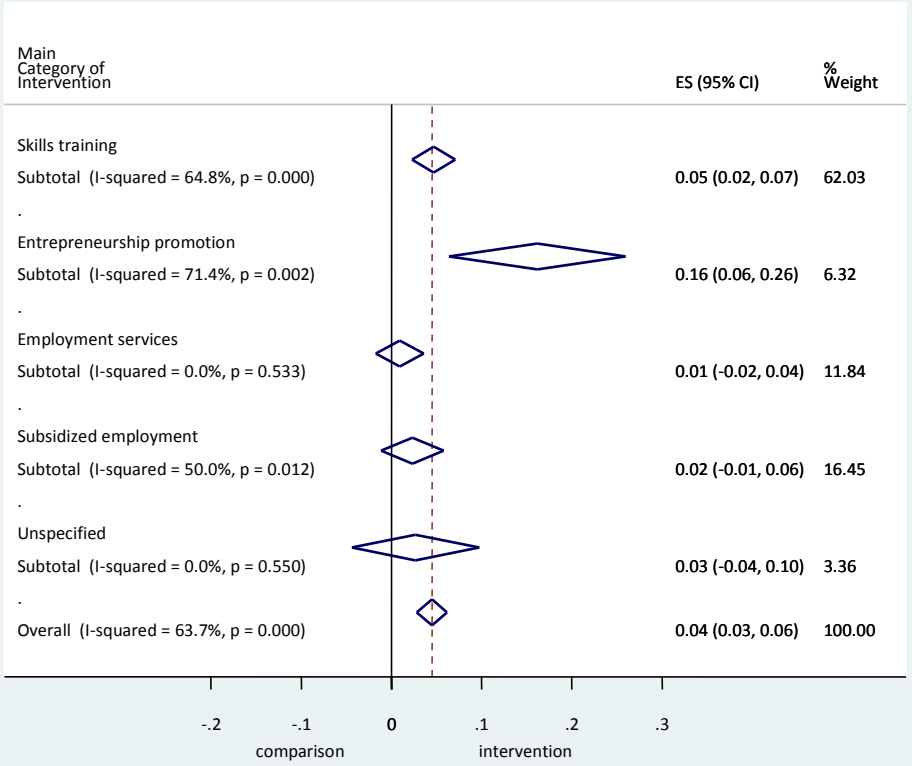
4.3.3.2 Main category of intervention

After restricting the analysis to cases where **employment outcomes** were reported (Figure 10), both entrepreneurship promotion (0.16 SMD; CI = 0.06, 0.26; I^2 = 71 per cent; number of interventions = 7) and skills training interventions (0.05 SMD; CI = 0.02, 0.07; I^2 = 65 per cent; number of interventions = 67) exposed larger than average effect sizes. It is important to note the evidence base behind this result. While there are 67 interventions supporting the magnitude of impact skills training measures, there are only seven related to entrepreneurship promotion. Detailed results are displayed in Table 13.

Interventions providing mainly employment services to youth were the least successful (0.01 SMD; CI = -0.02, 0.04; I^2 = 0 per cent; number of interventions = 10). In agreement with the descriptive analysis of interventions, interventions with skills training as the main category had the greatest weight within the overall employment-related effect size.

In most cases, confidence intervals overlapped with the overall mean SMD, suggesting that there were no significant differences in average effect size across types of interventions. The I^2 tests, however, reported statistically significant heterogeneity within the sub-groups for skills training, entrepreneurship promotion and subsidized employment interventions.

Figure 10: Summary forest plot of employment outcomes (full sample) by main category of intervention

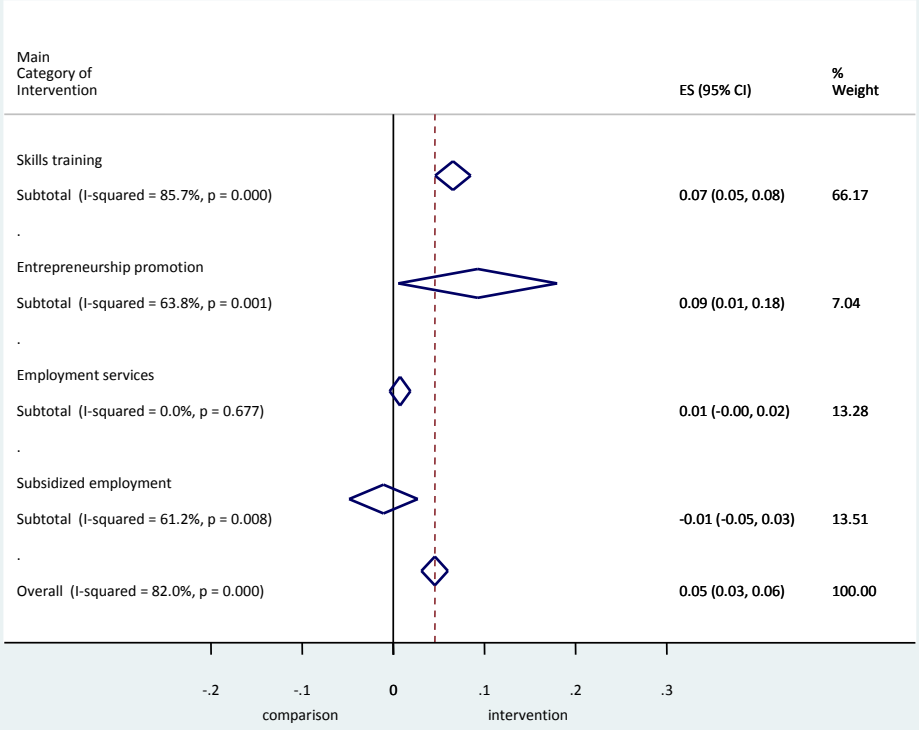


No. of SMDs/Studies: Total: 1330/105
Skills training: 904/67. Entrepreneurship promotion: 43/7. Employment services: 104/10. Subsidized employment: 193/16. Unspecified: 86/5.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

There was no evidence of heterogeneity across cases where it was not possible to identify a main category of intervention (i.e., in the unspecified category). Such cases reported an SMD of 0.03 (C I= -0.04, 0.10; I^2 = 0 per cent; number of interventions = 5) on employment outcomes. The category was dropped from the earnings outcome analysis due to insufficient sample size.

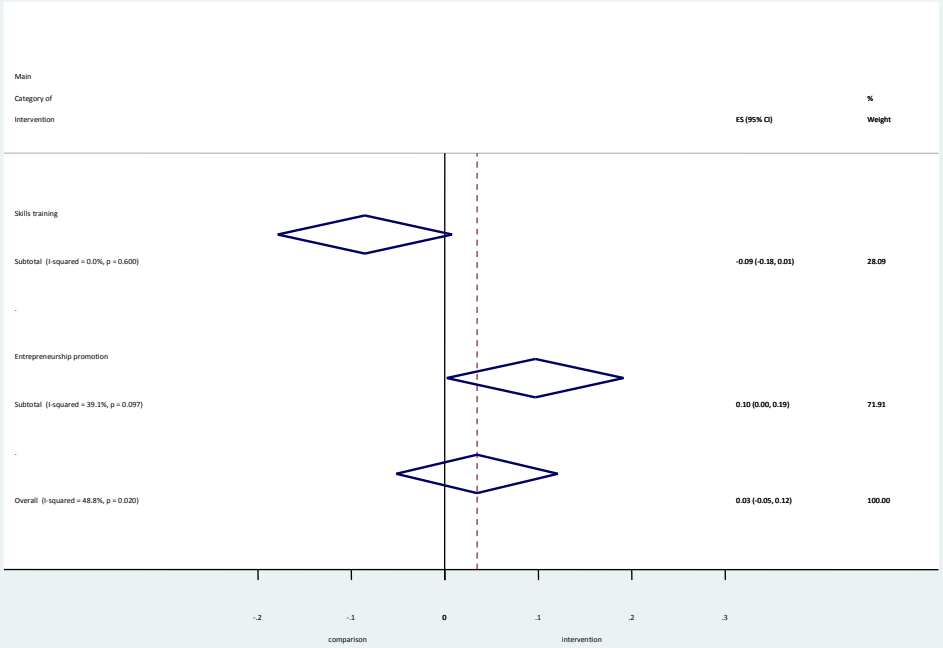
Qualitatively, the results of effect sizes from **earnings- or income-related outcomes**, across main intervention types, mimicked those from employment outcomes, though in this case subsidized employment interventions offered the lowest (and negative) effect size (-0.01 SMD; CI = -0.05, 0.03; I^2 = 61 per cent; number of interventions = 9).

Figure 11: Summary forest plot of earnings outcomes (full sample) by main category of intervention



No. of SMDs/Studies: Total: 638/89
Skills training: 495/60. Entrepreneurship promotion: 50/12. Employment services: 36/8. Subsidized employment: 57/9.
Sub-group dropped (<4 independent studies): Unspecified.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 12: Summary forest plot of business performance outcomes (full sample) by main category of intervention



No. of SMDs/Studies: Total: 169/14
Skills training: 7/4. Entrepreneurship promotion: 162/10.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The WINGS programme in Uganda (Blattman et al., 2014) (0.82 SMD; 95% CI = 0.13, 1.50) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

The computed effect sizes (displayed in Figure 11) suggested that skills training (0.07 SMD; CI = 0.05, 0.08; I^2 = 86 per cent; number of interventions = 60) and entrepreneurship interventions (0.09 SMD; CI = 0.01, 0.18; I^2 = 64 per cent; number of interventions = 12) positively and consistently impacted both the employment and earnings prospects of young people, while evidence from other intervention types showed rather lower impacts on both outcome categories. However, significant heterogeneity was detected within all categories of intervention except for employment services (0.01 SMD; CI = 0.00, 0.02; I^2 = 0 per cent; number of interventions = 8). See

Table 14 for further information on the results and parameters.

The summary forest plot for **business performance outcomes** (Figure 12 and Table 15) relies on a sample of 169 effect sizes, computed from treatment effects reported in 14 studies. Notably, the impact of skills training interventions, which measured impacts on business performance outcomes (four cases), was negative with an average SMD of -0.09 (CI = -0.18, 0.01; I^2 = 0 per cent; number of interventions = 4). On the other hand, the magnitude of impact from entrepreneurship interventions on business outcomes was comparatively large (0.10 SMD; CI = 0.00, 0.19; I^2 = 39 per cent; number of interventions = 10) and just statistically significant at 10 per cent significance level.

4.3.3.3 Country income level

This section explores differential impacts across country income levels. The analysis recognized (i) the differences in labour market barriers facing youth in the context of different country income levels (Robalino, Margolis, Rother, Newhouse & Lundberg, 2013); (ii) the role of context on the ability of youth employment interventions to shape labour market outcomes of youth (Betcherman et al., 2007); and (iii) the intrinsic and differentiated characteristics of labour markets and institutions across middle- and low-income countries in comparison to high-income countries (Fields, 2011; Cho, Margolis, Newhouse & Robalino, 2012).

The analysis capitalized on the sizable number of studies under each country income group. There were 65 and 48 reports of interventions implemented in high-income countries and low- and middle-income countries, respectively.

Interventions in high-income countries were typically national programmes, implemented and designed by government agencies. Evidence from local or pilot interventions was scarce (only 15 per cent of the total sample). In low- and middle-income countries, more than 40 per cent of the evidence was generated from small-scale local programmes. These programmes often targeted specific groups, such as young women. While only 5 per cent of interventions in high-income countries targeted young women, they were the focus of 27 per cent of the interventions evaluated in low-income countries.

Evaluated interventions also varied across country income levels.[34] While, in high-income countries, evaluations of employment services, subsidized employment and skills training were common, only a negligible number of entrepreneurship promotion interventions were evaluated. In contrast, both entrepreneurship and skills training interventions were relatively frequently reported in countries outside the high-income economies, but there were few cases of evaluated interventions providing mainly employment services or subsidized employment interventions.

Research designs also varied across country income groups. A significant proportion (>50 per cent) of the recent evidence from middle- and low-income countries had been generated from relatively small-scale experimental evaluation designs. In contrast, quasi-experimental approaches using administrative data made up a large share (60 per cent) of the studies from high-income countries.

The review team also observed that many of the interventions in high-income countries were designed and implemented with the participation of government agencies. However, in some cases other stakeholders were involved, in particular the private sector (for example, in the form of private firms providing training or employment services).

---

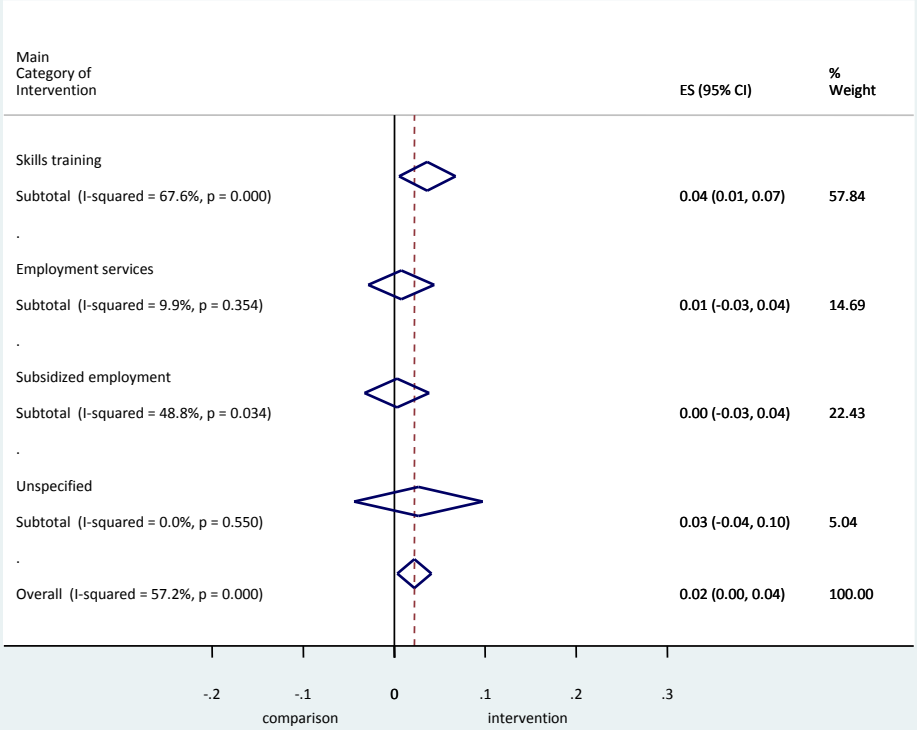
[34] Results did not necessarily reflect the intervention types which were predominately implemented as these may not have been evaluated.

Summary forest plots are provided for high-income and low- and middle-income countries.[35] Effect sizes reported on both employment and earnings outcomes were generally higher among low- and middle-income countries (see Figure 14, Table 21, Figure 16, and Table 23) with overall effect sizes of 0.08 SMD (CI = 0.04, 0.11; I^2 = 64 per cent; number of interventions = 48) and 0.12 SMD (CI = 0.08, 0.15; I^2 = 86 per cent; number of interventions = 53) for each outcome, respectively. In comparison, reported effect sizes for high-income countries' employment and earnings outcomes were 0.02 SMD (CI = 0.00, 0.04; I^2 = 57 per cent; number of interventions = 52) and 0.01 SMD (CI = -0.01, 0.02; I^2 = 70 per cent; number of interventions = 31), respectively. This suggested that active labour market measures had a greater impact among youth in developing countries compared to youth in advanced economies. The result coincided with Betcherman et al. (2007), which demonstrated that the probability that a programme has a positive impact on labour market outcomes declines as the country's income level rises.

Skills training interventions generated the highest magnitude of impacts in high-income countries, reportedly 0.04 SMD (CI = 0.01, 0.07; I^2 = 68 per cent; number of interventions = 29) for employment (Figure 13, Table 20) and 0.02 SMD (CI = 0.00, 0.04; I^2 = 72 per cent; number of interventions = 21) for earnings (Figure 15, Table 22) outcomes, respectively. In contrast, skills training interventions in low- and middle-income countries showed lower effect size (0.06 SMD; CI = 0.02, 0.10; I^2 = 63 per cent; number of interventions = 38) among employment outcomes, compared to a 0.18 SMD (CI = 0.06, 0.29; I^2 = 68 per cent; number of interventions = 5) for **entrepreneurship promotion** interventions.

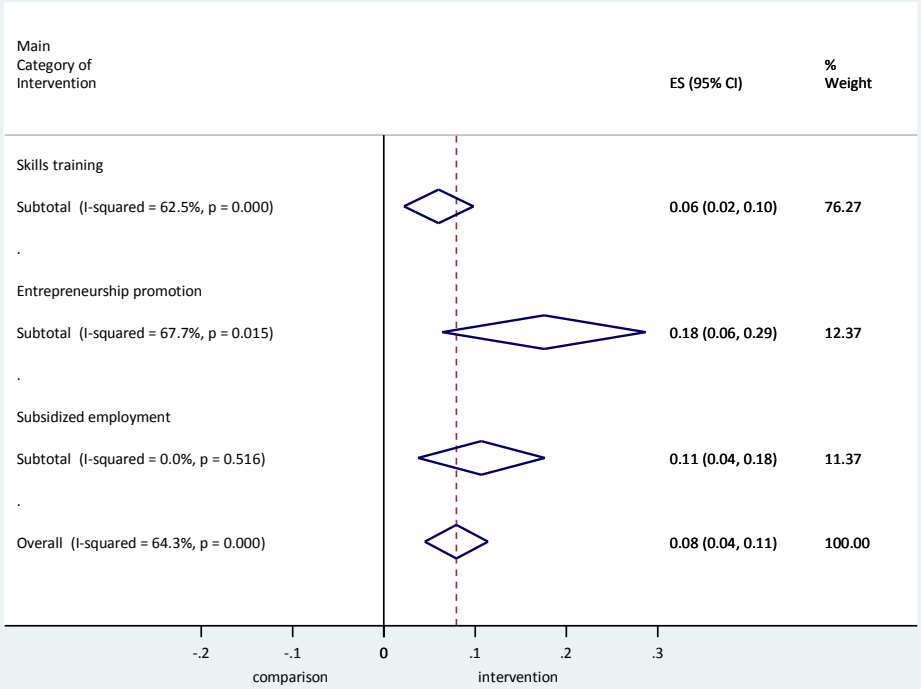
[35] Due to limited data availability, the analysis did not differentiate between impacts in low- and middle-income countries.

Figure 13: Summary forest plot of employment outcomes by main category of intervention for high-income countries



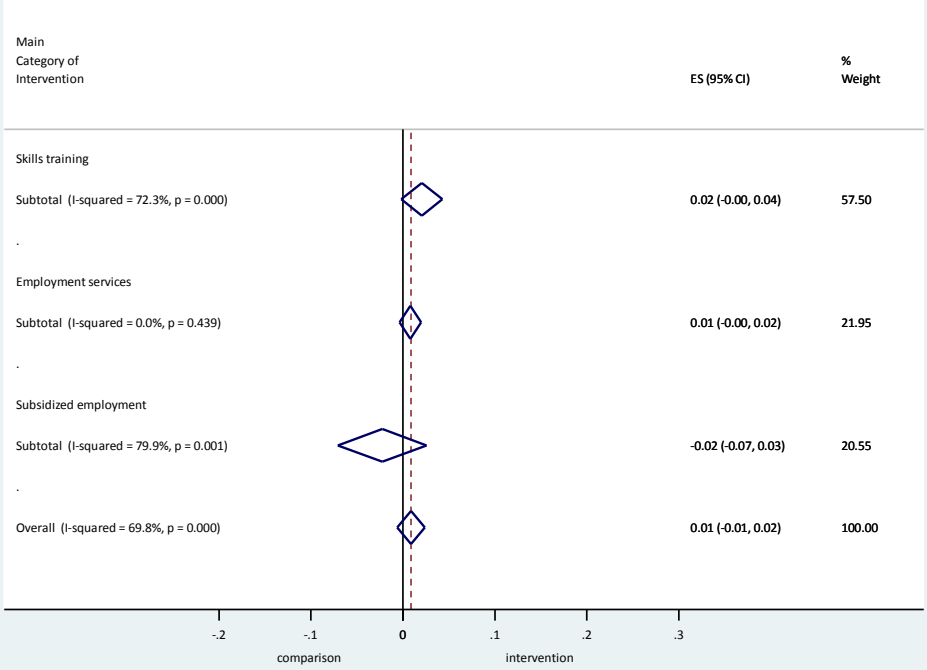
No. of SMDs/Studies: Total: 867/52
Skills training: 534/29. Employment services: 87/7. Subsidized employment: 160/11. Unspecified: 86/5.
Sub-group dropped (<4 independent studies): Entrepreneurship promotion.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 14: Summary forest plot of employment outcomes by main category of intervention for low- and middle-income countries



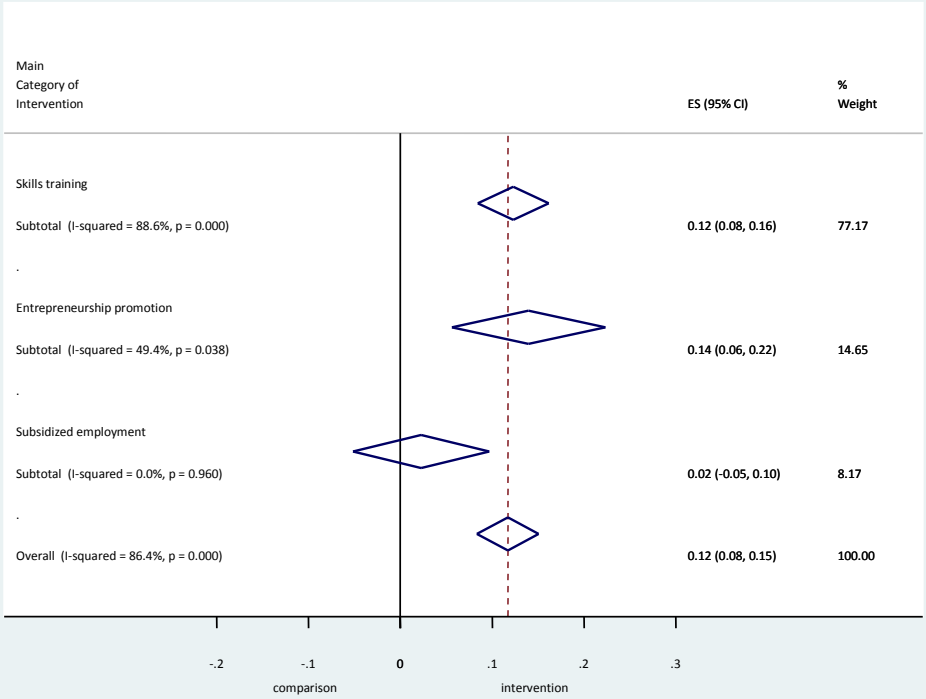
No. of SMDs/Studies: Total: 438/48
Skills training: 370/38. Entrepreneurship promotion: 35/5. Subsidized employment: 33/5.
Sub-group dropped (<4 independent studies): Employment services.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75

Figure 15: Summary forest plot of earnings outcomes by main category of intervention for high-income countries



No. of SMDs/Studies: Total: 381/31
Skills training: 303/21. Employment services: 32/5. Subsidized employment: 46/5.
Sub-group dropped (<4 independent studies): Entrepreneurship promotion Unspecified.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75

Figure 16: Summary forest plot of earnings outcomes by main category of intervention for low- and middle-income countries



No. of SMDs/Studies: Total: 242/53
Skills training: 192/39. Entrepreneurship promotion: 39/10. Subsidized employment: 11/4.
Sub-group dropped (<4 independent studies): Employment services.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

While the effect size displayed wide variance, entrepreneurship promotion interventions offered positive prospects for stimulating the labour market outcomes of youth in the developing world. The limited number of effect sizes from entrepreneurship interventions in high-income countries caused the category to drop out of the analysis.

Evidence from **subsidized employment** interventions in low- and middle-income countries was rather limited but still presented positive impacts (on average) and non-heterogeneity within the sub-group for both outcome types. For high-income countries, the interpretation of the findings was less encouraging, as the studies reported no impact on employment and negative impacts on earnings.

Employment services interventions held an important position in terms of the effect of youth-targeted ALMPs in high-income countries. While their impact on employment was negligible, they tended to yield positive income gains among participating youth. The category dropped out of the low- and middle-income countries analysis due to its reduced sample size.

In conclusion, observing both outcome categories, the team ruled out statistically significant differences between intervention types for both high-income and low- and middle-income countries.

Table 20: Summary of results on employment outcomes across main categories of intervention in high-income countries

Parameters of interest	Skills training	Entrepreneurship promotion	Employment services	Subsidized employment	Unspecified
Standardized Means Difference I^2	0.04	Dropped from analysis	0.01	0	0.03
Standard errors of the effect size	0.02		0.02	0.02	0.04
95% confidence interval	0.01		-0.03	-0.03	-0.04
	0.07		0.04	0.04	0.1
I Squared	67.55		9.89	48.82	0
Number of SMDs	534		87	160	86
Number of interventions	29		7	11	5
Sample size	2,394,204		2,326,518	32,031,060	178,863
Mean difference	1.86		-1.57	0.54	-2.65
Control outcome	33.90		27.80	11.09	53.56
Treatment outcome	33.44		20.89	7.20	48.89
Percentage change	0.03		0.02	0.01	0.1

Note: Dropped from the analysis indicates the number of studies was insufficient (lower than four) to allow for the computation of the effect sizes.

Table 21: Summary of results on employment outcomes across main categories of intervention in low- and middle-income countries

Parameters of interest	Skills training	Entrepreneurship promotion	Employment services	Subsidized employment
Standardized Means Difference I^2	0.06	0.18	Dropped from analysis	0.11
Standard errors of the effect size	0.02	0.06		0.04
95% confidence interval	0.02	0.06		0.04
	0.1	0.29		0.18
I Squared	62.54	67.66		0
Number of SMDs	370	35		33
Number of interventions	38	5		5
Sample size	1,045,500	54,205		167,129
Mean difference	14.92	4.19		0.16
Control outcome	40.40	7.83		0.89
Treatment outcome	39.71	9.98		0.94
Percentage change	0.14	0.36		-0.04

Note: Dropped from the analysis indicates the number of studies was insufficient (lower than four) to allow for the computation of the effect sizes.

Table 22: Summary of results on earnings outcomes across main categories of intervention in high-income countries

Parameters of interest	Skills training	Entrepreneurship promotion	Employment services	Subsidized employment	Unspecified
Standardized Means Difference I²	0.02	Dropped from analysis	0.01	-0.02	Dropped from analysis
Standard errors of the effect size	0.01		0.01	0.02	
95% confidence interval	0		0	-0.07	
	0.04		0.02	0.03	
I Squared	72.32		0	79.91	
Number of SMDs	303		32	46	
Number of interventions	21		5	5	
Sample size	1,163,479		190,770	10,347,125	
Mean difference	4,800.83		-16.51	-2106.18	
Control outcome	18537.80		215.34	9,585.32	
Treatment outcome	20514.93		174.91	7,477.58	
Percentage change	0.02		0	-0.03	

Note: Dropped from the analysis indicates the number of studies was insufficient (lower than four) to allow for the computation of the effect sizes.

Table 23: Summary of results on earnings outcomes across main categories of intervention in low- and middle-income countries

Parameters of interest	Skills training	Entrepreneurship promotion	Employment services	Subsidized employment
Standardized Means Difference I²	0.12	0.14	Dropped from analysis	0.02
Standard errors of the effect size	0.02	0.04		0.04
95% confidence interval	0.08	0.06		-0.05
	0.16	0.22		0.1
I Squared	88.62	49.45		0
Number of SMDs	192	39		11
Number of interventions	39	10		4
Sample size	882,481	34,542		11,030
Mean difference	726.10	3,013.40		16.09
Control outcome	5,520.71	3,286.44		551.43
Treatment outcome	6,198.67	4,447.00		588.67
Percentage change	0.18	0.3		0.08

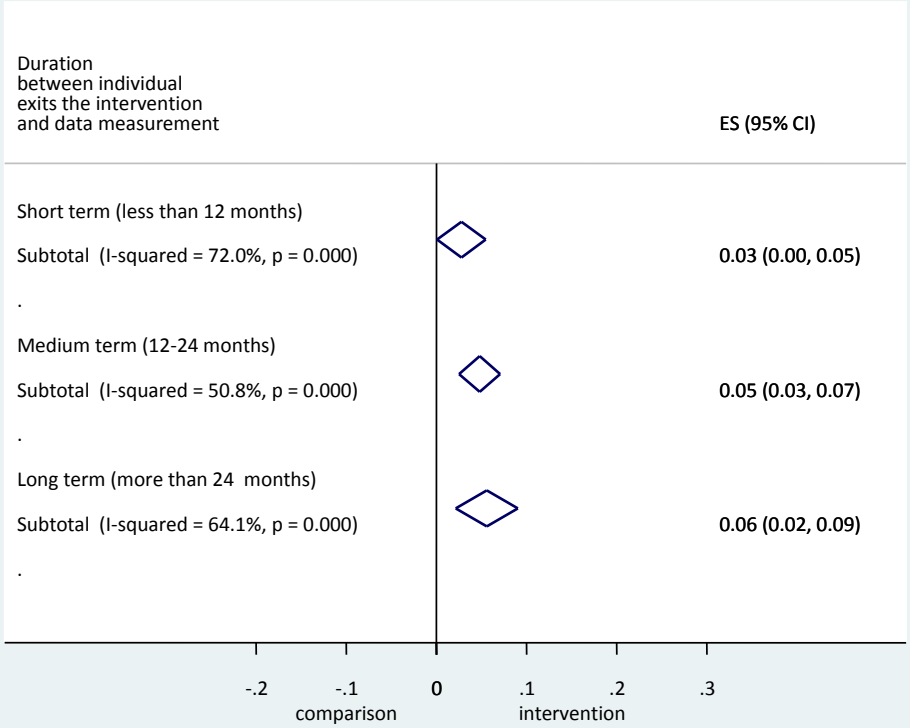
Note: Dropped from the analysis indicates the number of studies was insufficient (lower than four) to allow for the computation of the effect sizes.

4.3.3.4 Duration after treatment

Not all research studies reported information on the time lag between exposure to treatment and measurement of changes in outcomes. After imputations, only 72 per cent of the SMDs could be classified according to study timing after treatment into short- (data collected less than 12 months after the end of the treatment), medium- (12–24 months) and long-term studies (more than 24 months). Longer term outcomes were most common in evaluations from high-income country.

In the restricted sample, the overall effect size of employment and earnings outcomes was roughly the same (employment outcomes: 0.04 SMD; CI = 0.02, 0.06; I^2 = 66 per cent; number of interventions = 85; earnings outcomes: 0.05 SMD; CI = 0.04, 0.07; I^2 = 81 per cent; number of interventions = 79). In both cases, short- and medium-term studies had a similar weight on the overall effect size of the entire meta-analysis (around 45 per cent and 35 per cent, respectively). While there was unaccounted heterogeneity within the different duration terms, it appeared that effect size estimates from longer term evaluations (>1 year) were relatively larger than shorter and medium term estimates in the case of studies measuring employment outcomes. As displayed in Figure 17 and Table 24, effect sizes for medium and long term were 0.05 SMD (CI = 0.03, 0.07; I^2 = 51 per cent; number of interventions = 43) and 0.06 SMD (CI = 0.02, 0.09; I^2 = 64 per cent; number of interventions = 21), respectively. This suggested a certain time lag before outcomes materialize.

Figure 17: Summary forest plot of employment outcomes by duration of period between individual exiting the intervention and data measurement (short, medium and long term)



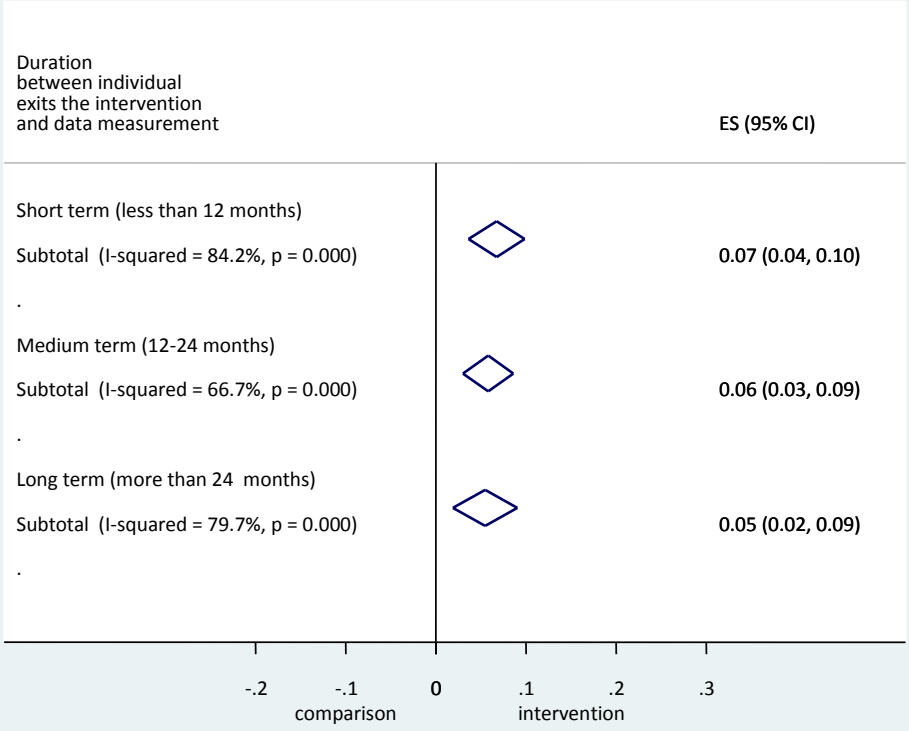
No. of SMDs/Studies: Total: 905/85
Short term (less than 12 months): 348/57. Medium term (12-24 months): 377/43. Long term (more than 24 months): 180/21.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Table 24: Summary of results on employment outcomes by duration

Parameters of interest	Short term (less than 12 months)	Medium term (12–24 months)	Long term (more than 24 months)
Standardized Means Difference I^2	0.03	0.05	0.06
Standard errors of the effect size	0.01	0.01	0.02
95% confidence interval	0	0.03	0.02
	0.05	0.07	0.09
I Squared	71.98	50.81	64.08
Number of SMDs	348	377	180
Number of interventions	57	43	21
Sample size	1,790,980	1,566,096	1,897,714
Mean difference	3.39	6.42	3.58
Control outcome	18.36	25.97	112.76
Treatment outcome	15.97	25.09	109.56
Percentage change	0.07	0.1	0.11

Earnings outcomes showed a reversed pattern. Figure 18 (Table 25) shows impacts that decrease as duration between exposure to treatment and measurement increases. Medium- and long-term effect sizes were 0.06 SMD (CI = 0.03, 0.09; I^2 = 67 per cent; number of interventions = 38) and 0.05 SMD (CI = 0.02, 0.09; I^2 = 80 per cent; number of interventions = 20), respectively. While the effect size for short-term duration was 0.07 SMD (CI = 0.04, 0.1; I^2 = 84 per cent; number of interventions = 54).

Figure 18: Summary forest plot of earnings outcomes by duration of period between individual exiting the intervention and data measurement (short, medium and long term)



No. of SMDs/Studies: Total: 543/79
Short term (less than 12 months): 207/54. Medium term (12-24 months): 167/38. Long term (more than 24 months): 169/20.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Table 25: Summary of results on earnings outcome by duration

Parameters of interest	Short term (less than 12 months)	Medium term (12–24 months)	Long term (more than 24 months)
Standardized Means Difference I^2	0.07	0.06	0.05
Standard errors of the effect size	0.02	0.01	0.02
95% confidence interval	0.04 0.1	0.03 0.09	0.02 0.09
I Squared	84.23	66.74	79.74
Number of SMDs	207	167	169
Number of interventions	54	38	20

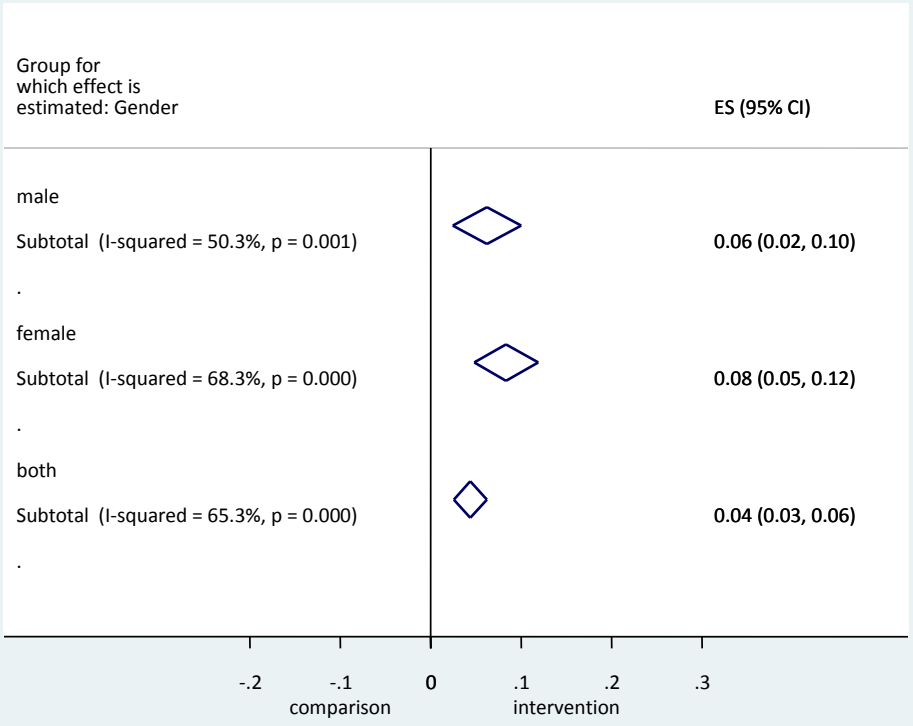
Parameters of interest	Short term (less than 12 months)	Medium term (12–24 months)	Long term (more than 24 months)
Sample size	1,297,776	501,005	912,958
Mean difference	-654.38	7,715.05	1,359.53
Control outcome	4,361.13	25,542.80	9,785.48
Treatment outcome	3,530.38	29,645.58	11,207.11
Percentage change	0.08	0.12	0.2

At the same time, for both outcome types, the confidence intervals for each sub-group fell within the mean overall SMD and, hence, differences were not statistically significant. As the team suspected that other study-level characteristics might have confounded the analysis, this question was explored in more detail through multivariate meta-regression.

4.3.3.5 Gender

The analysis by gender relied on whether an effect size was reported for female or male participants only. Pooled results (meaning those estimated on data that could not be disaggregated by gender) did not form part of this sub-group analysis.

Figure 19: Summary forest plot of employment outcomes (full sample) by gender



No. of SMDs/Studies: Total: 1330/105
male: 198/28. female: 340/42. both: 792/84.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

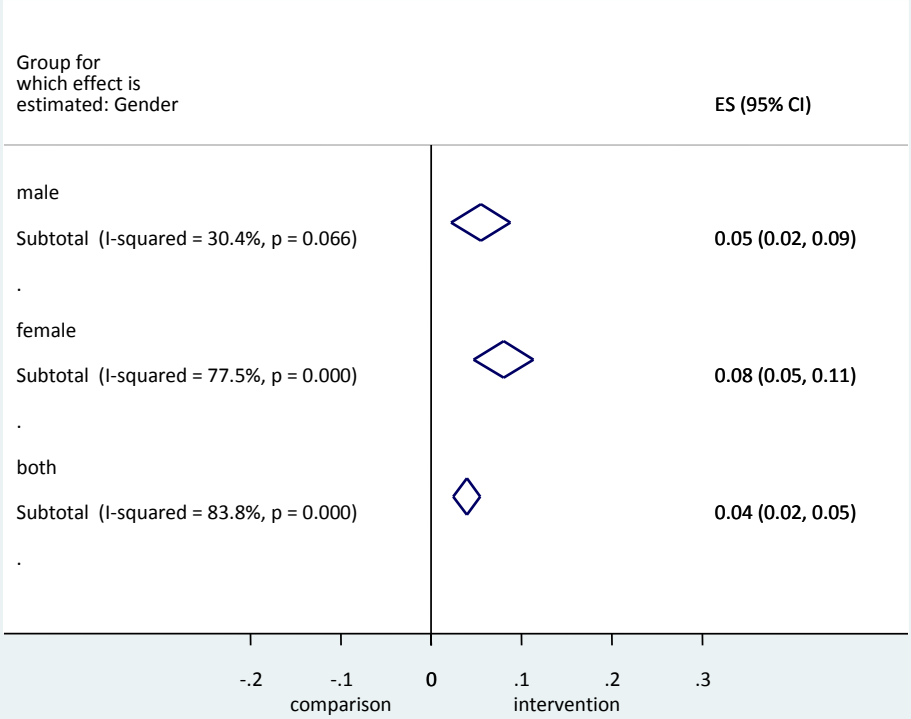
A large body of literature focuses on differences in the effectiveness of labour market interventions by gender. These differences were reflected in the interventions and studies in the sample. There were 39 interventions that reported male-only outcomes, of which more than half were located in high-income countries. Conversely, of the 54 interventions reporting outcomes separately for females, only 44 per cent were in high-income economies. More than one-quarter of these 54 interventions (15) specifically targeted only female participants (the vast majority of which (12) were located in low- and middle-income countries). In contrast, many other intervention characteristics were distributed relatively evenly between interventions that reported male-only and/or female-only estimates, in particular the main category of the intervention, age, education and income status of the target population, or programme implementers.

Summary forest plots showed greater effect sizes for young women compared to young men across employment and earnings outcomes. This suggested that interventions which specifically measured changes in outcomes by gender tended to have higher returns for women.

Table 26: Summary of results on employment outcomes reported by a specific gender or both

Parameters of interest	Male	Female	Both
Standardized Means Difference I^2	0.06	0.08	0.04
Standard errors of the effect size	0.02	0.02	0.01
95% confidence interval	0.02	0.05	0.03
	0.1	0.12	0.06
I Squared	50.28	68.3	65.27
Number of SMDs	198	340	792
Number of interventions	28	42	84
Sample size	2,646,312	2,755,581	32,817,154
Mean difference	14.37	14.55	1.55
Control outcome	57.02	10.47	35.28
Treatment outcome	52.49	9.62	33.85
Percentage change	0.06	0.18	0.1

Figure 20: Summary forest plot of earnings outcomes (full sample) by gender



No. of SMDs/Studies: Total: 670/92
male: 134/28. female: 190/43. both: 346/72.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Table 27: Summary of results on earnings outcomes reported by a specific gender or both

Parameters of interest	Male	Female	Both
Standardized Means Difference I^2	0.05	0.08	0.04
Standard errors of the effect size	0.02	0.02	0.01
95% confidence interval	0.02 0.09	0.05 0.11	0.02 0.05
I Squared	30.38	77.51	83.76
Number of SMDs	134	190	346
Number of interventions	28	43	72
Sample size	1,166,483	1,228,642	10,301,687
Mean difference	2,157.98	1,368.85	2,162.41
Control outcome	16,718.45	8,816.04	10,010.51
Treatment outcome	18,855.68	11,016.96	10,828.26
Percentage change	0.07	0.2	0.07

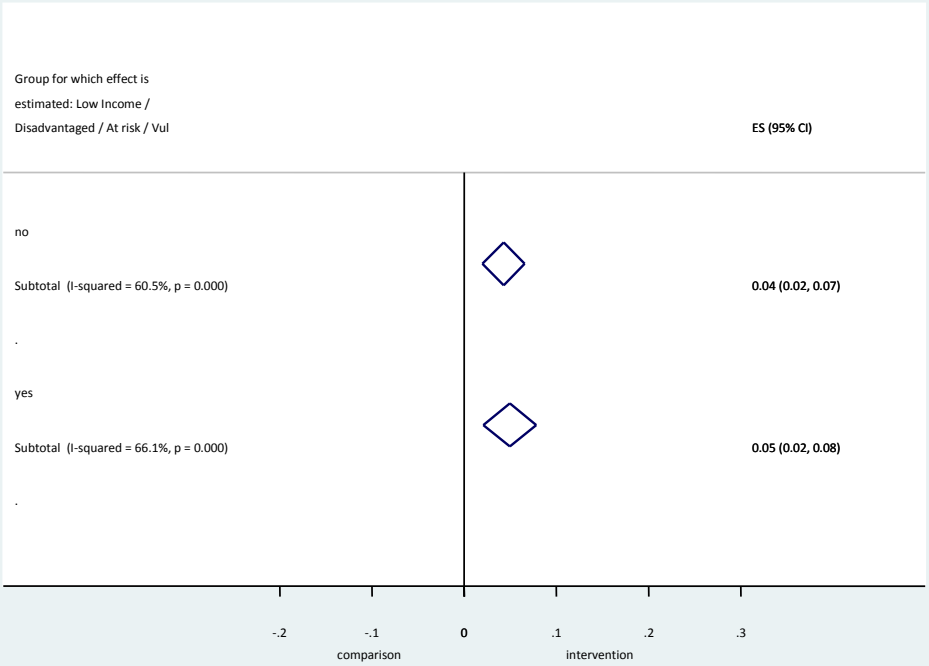
4.3.3.6 Participant income status

This subsection looks at differential effects by participant sub-group, focusing on low-income/disadvantaged/at risk/vulnerable youth. Estimates for this sub-group (labelled "disadvantaged youth" for ease of reference) existed for almost half of the interventions (47 per cent) or more than half of the programmes (51 per cent), equally distributed across gender. The share of interventions containing separate estimates for the disadvantaged youth sub-group was considerably higher in low- and middle-income countries (57 per cent) than in high-income countries (38 per cent).

Figure 21 (Table 28) and Figure 22 (Table 29) display the results on employment and earnings outcomes respectively by participant income status. It appears that interventions had a greater impact on earnings outcomes of disadvantaged youth (0.12 SMD; CI = 0.08, 0.16; I^2 = 81 per cent; number of interventions = 53) compared to non-disadvantaged youth (0.02 SMD; CI = 0, 0.03; I^2 = 73 per cent; number of interventions = 37). Results on earnings for disadvantaged youth were larger than results across employment outcomes in the same target group (0.05 SMD; CI = 0.02, 0.08; I^2 = 66 per cent; number of interventions = 57).

Across intervention types, entrepreneurship promotion increased employment and earnings impacts among disadvantaged youth. Computed SMDs were large in magnitude and variance. Total overall effect size was larger for earnings outcomes (0.13 SMD; CI = 0.09, 0.18; I^2 = 82 per cent; number of interventions = 48) than for employment outcomes (0.06 SMD; CI = 0.02, 0.09; I^2 = 66 per cent; number of interventions = 54). These results contrasted quite sharply with the analyses of interventions which did not specifically target disadvantaged youth. For that sub-group, overall effect sizes were substantially smaller (employment outcomes: 0.03 SMD; CI = 0.01, 0.06; I^2 = 56 per cent; number of interventions = 40; earnings outcomes: 0.02 SMD; CI = 0.00, 0.03; I^2 = 73 per cent; number of interventions = 37) with skills training yielding the highest effects.

Figure 21: Summary forest plot of employment outcomes by participant income group (where yes is low-income, disadvantaged, at risk or vulnerable youth)

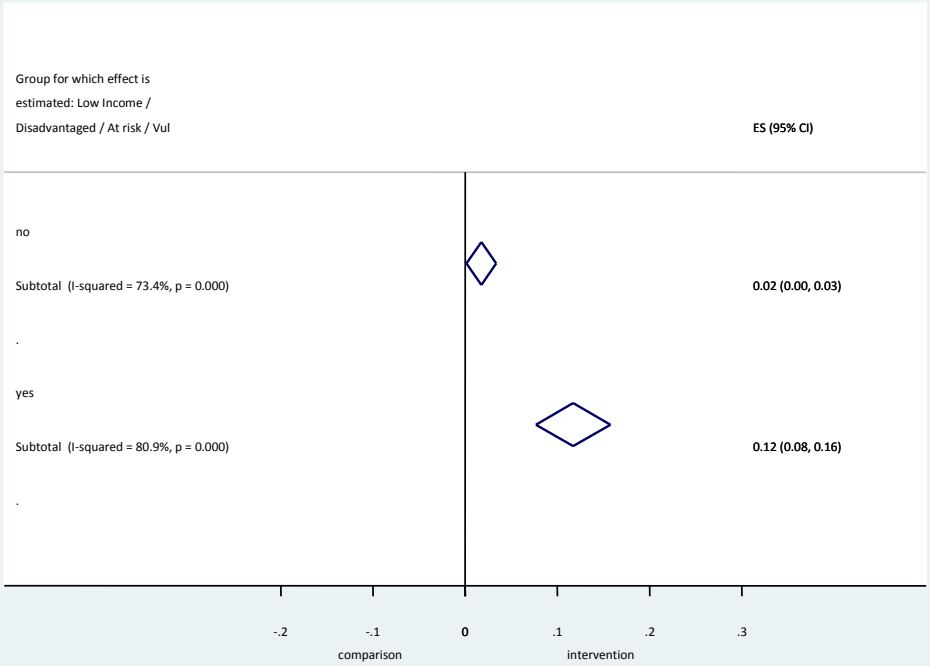


No. of SMDs/Studies: Total: 1321/101
no: 471/44. yes: 850/57.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Table 28: Summary of results on employment outcomes by participant income group (where yes is low-income, disadvantaged, at risk or vulnerable youth)

Parameters of interest	No	Yes
Standardized Means Difference I^2	0.04	0.05
Standard errors of the effect size	0.01	0.01
95% confidence interval	0.02 0.07	0.02 0.08
I Squared	60.52	66.08
Number of SMDs	471	850
Number of interventions	44	57
Sample size	35,835,136	1,933,130
Mean difference	0.93	9.30
Control outcome	39.70	25.38
Treatment outcome	36.82	25.04
Percentage change	0.01	0.13

Figure 22: Summary forest plot of earnings outcomes by participant income group (where yes is low-income, disadvantaged, at risk or vulnerable youth)

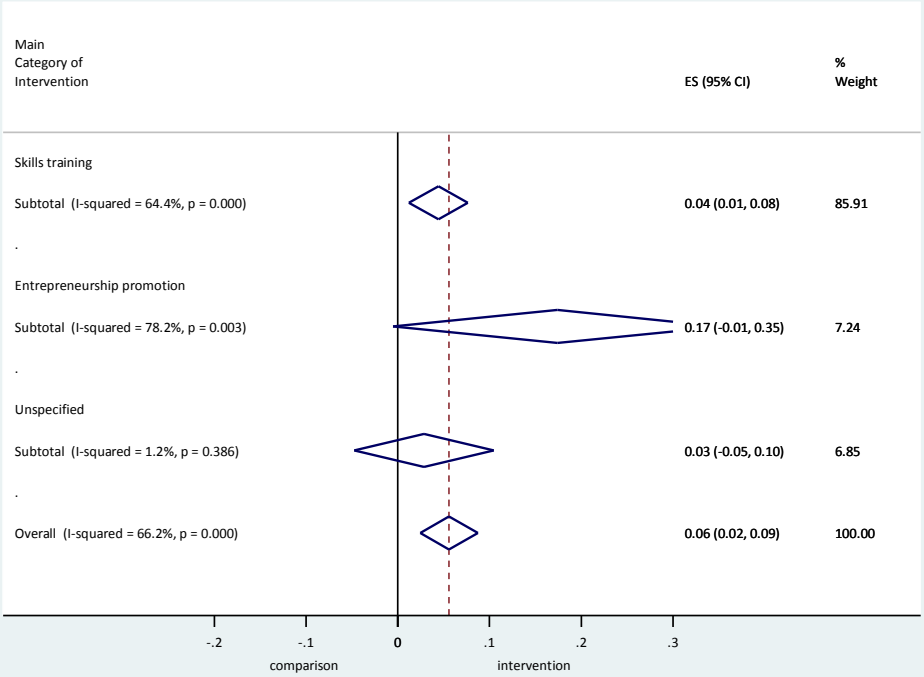


No. of SMDs/Studies: Total: 665/90
no: 163/37. yes: 502/53.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Table 29: Summary of results on employment outcomes by participant income group (where yes is low-income, disadvantaged, at risk or vulnerable youth)

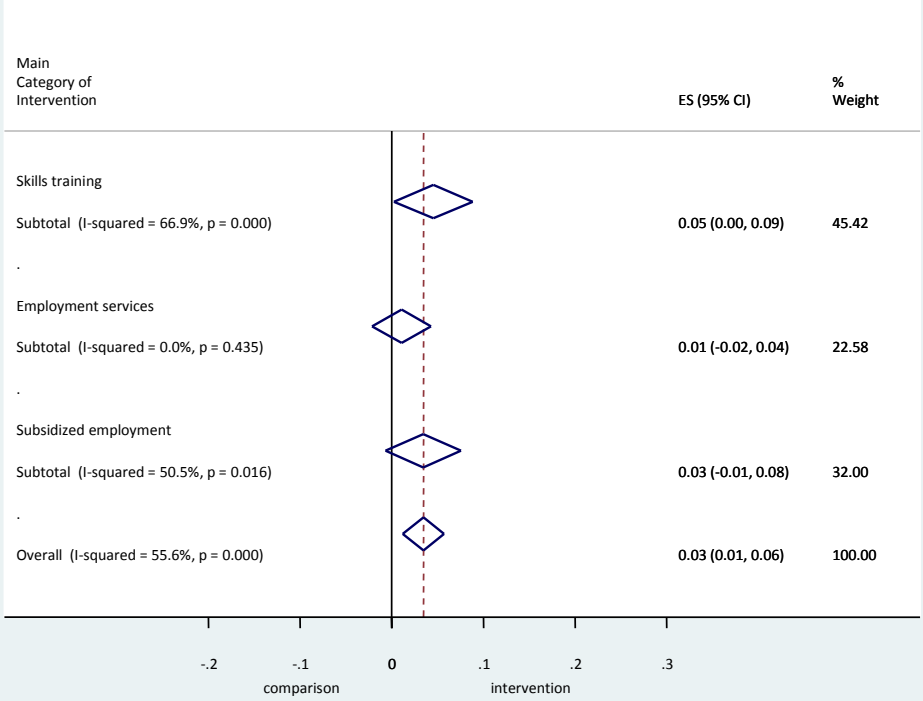
Parameters of interest	No	Yes
Standardized Means Difference I^2	0.02	0.12
Standard errors of the effect size	0.01	0.02
95% confidence interval	0 0.03	0.08 0.16
I Squared	73.35	80.91
Number of SMDs	163	502
Number of interventions	37	53
Sample size	10,822,670	1,872,850
Mean difference	-577.28	9,001.62
Control outcome	3,138.52	27,440.73
Treatment outcome	2,477.55	31,707.80
Percentage change	0.07	0.16

Figure 23: Summary forest plot of employment outcomes by main category of intervention for low-income and disadvantaged participants



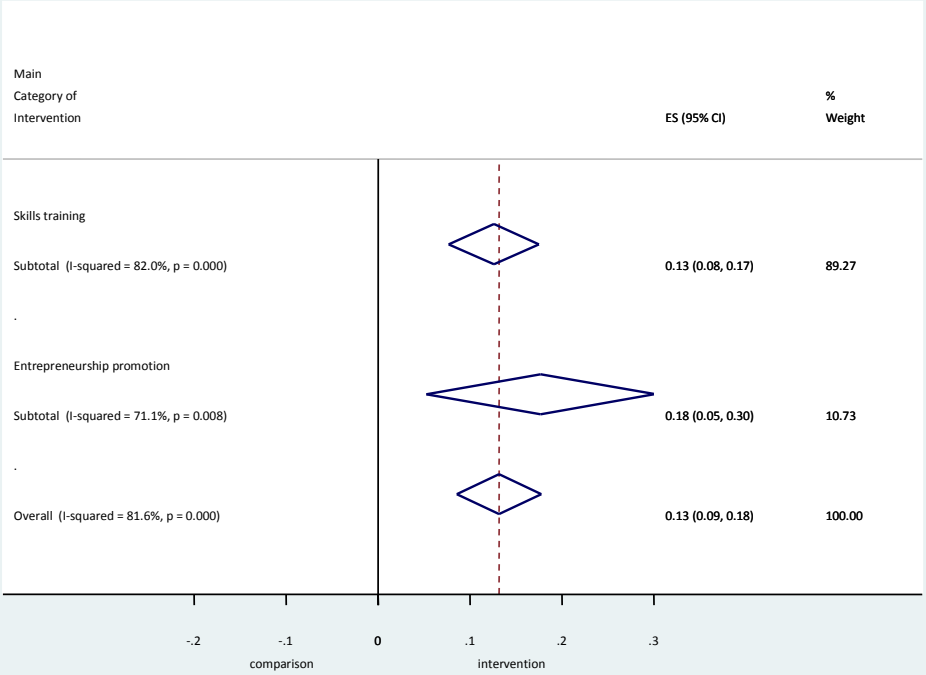
No. of SMDs/Studies: Total: 829/54
Skills training: 743/46. Entrepreneurship promotion: 19/4. Unspecified: 67/4.
Sub-group dropped (<4 independent studies): Employment services Subsidized employment.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75

Figure 24: Summary forest plot of employment outcomes by main category of intervention for non-low-income/non-disadvantaged participants



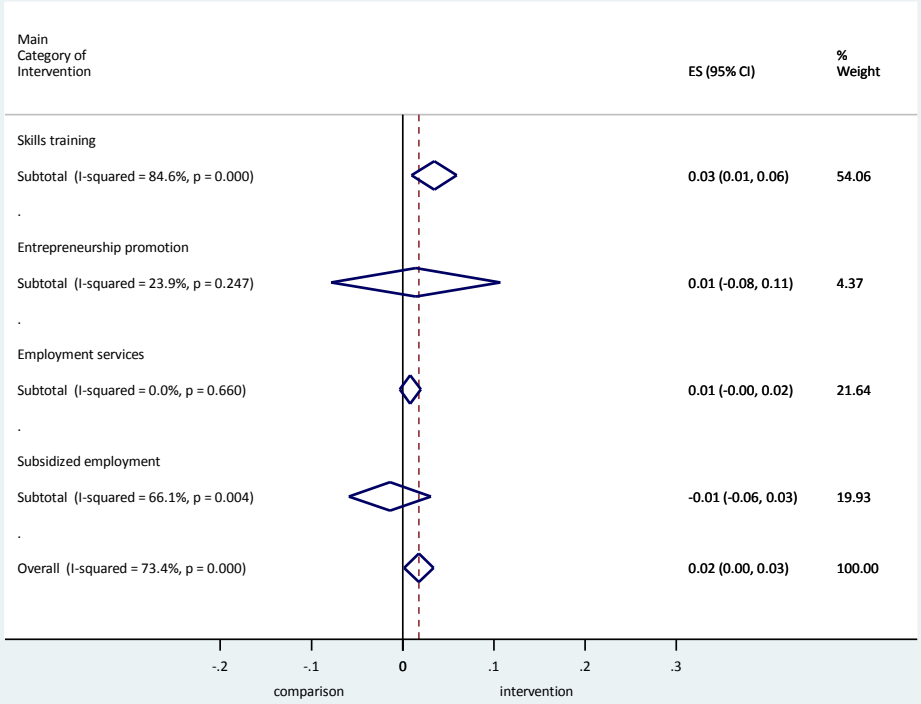
No. of SMDs/Studies: Total: 428/40
Skills training: 152/17. Employment services: 102/9. Subsidized employment: 174/14.
Sub-group dropped (<4 independent studies): Entrepreneurship promotion Unspecified.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 25: Summary forest plot of earnings outcomes by main category of intervention for low-income/disadvantaged participants



No. of SMDs/Studies: Total: 445/48
Skills training: 431/43. Entrepreneurship promotion: 14/5.
Sub-group dropped (<4 independent studies): Employment services Subsidized employment Unspecified.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 26: Summary forest plot of income outcomes by main category of intervention for non-low-income/non-disadvantaged participants



No. of SMDs/Studies: Total: 163/37
Skills training: 59/15. Entrepreneurship promotion: 36/7. Employment services: 35/7. Subsidized employment: 33/8.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75

Table 30: Summary of results on employment outcomes by participant income group (low-income participants includes also disadvantaged, at risk or vulnerable youth)

Parameters of interest	Low-income participants	Non-low-income participants		
	Skills training	Skills training	Employment services	Subsidized employment
Standardized Means Difference I^2	0.04	0.05	0.01	0.03
Standard errors of the effect size	0.02	0.02	0.02	0.02
95% confidence interval	0.01 0.08	0 0.09	-0.02 0.04	-0.01 0.08
I Squared	64.38	66.87	0	50.54
Number of SMDs	743	152	102	174
Number of interventions	46	17	9	14
Sample size	1,314,345	1,674,578	2,337,042	31,707,837
Mean difference	11.49	2.24	-1.60	0.69
Control outcome	27.83	66.40	28.14	12.08
Treatment outcome	27.54	64.96	21.14	7.92
Percentage change	0.11	0.01	0.02	-0.01

Notes: Entrepreneurship promotion, employment services, subsidized employment and unspecified categories were dropped from the analysis for the group of low-income participants due to the small number of independent studies. Entrepreneurship promotion and unspecified categories were dropped from the analysis for the group of non-low-income participants due to the small number of independent studies.

Table 31: Summary of results on earnings outcomes by participant income group (low-income participants includes also disadvantaged, at risk or vulnerable youth)

Parameters of interest	Low-income participants		Non-low-income participants			
	Skills training	Entrepreneur-ship promotion	Skills training	Entrepreneur-ship promotion	Employment services	Subsidized employment
Standardized Means Difference I^2	0.13	0.18	0.03	0.01	0.01	-0.01
Standard errors of the effect size	0.03	0.06	0.01	0.05	0.01	0.02
95% confidence interval	0.08 0.17	0.05 0.3	0.01 0.06	-0.08 0.11	0 0.02	-0.06 0.03
I Squared	82.04	71.08	84.62	23.86	0	66.08
Number of SMDs	431	14	59	36	35	33
Number of interventions	43	5	15	7	7	8
Sample size	1,360,610	15,331	684,058	27,199	192,159	9,919,254
Mean difference	11,202.67	4,535.91	-515.53	-1.00	-16.42	-1,898.11

Parameters of interest	Low-income participants		Non-low-income participants			
	Skills training	Entrepreneur-ship promotion	Skills training	Entrepreneur-ship promotion	Employment services	Subsidized employment
Control outcome	34,760.16	5,106.97	2,868.19	162.03	215.01	8,792.60
Treatment outcome	40,082.61	6,863.33	2,245.09	145.55	174.83	6,899.06
Percentage change	0.16	0.23	0.1	0.26	0	-0.01

Notes: Employment services, subsidized employment and unspecified categories were dropped from the analysis for the group of low-income participants due to the small number of independent studies. The unspecified category was dropped from the analysis for the group of non-low-income participants due to the small number of independent studies.

4.3.3.7 Programme characteristics

This section analyses the presence of effect size heterogeneity across studies that evaluated programmes on a different scale or implemented by different actors. In the sample these two characteristics did not differ for interventions within the same programme, they were therefore referred to as programme-level characteristics.

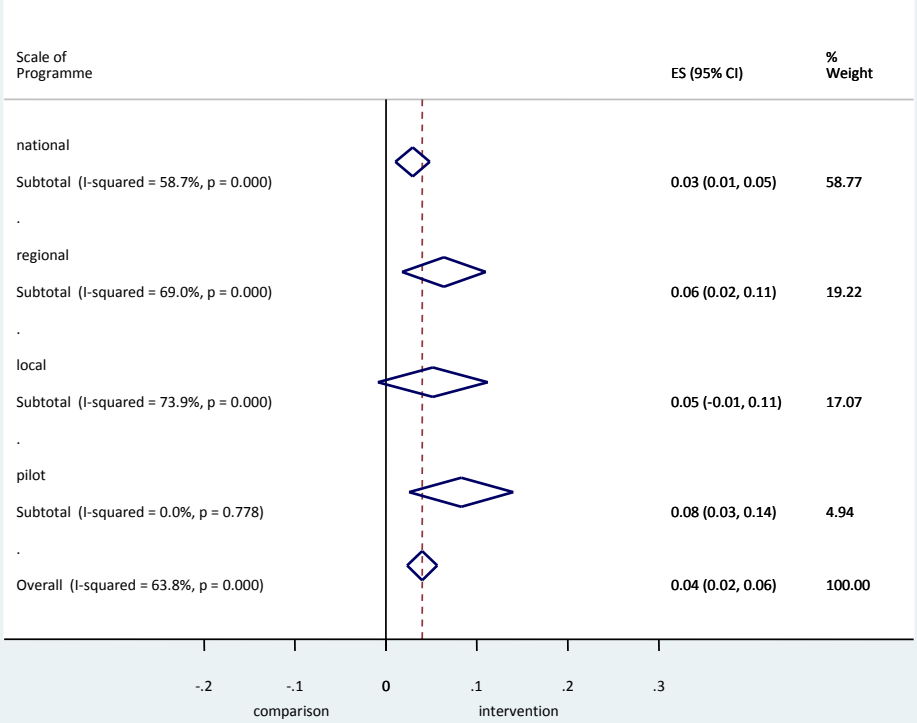
The review team coded the scale of the programme using four categories, which generally referred to the level on which the programme was implemented, namely:

1. National level, which comprised programmes that were implemented across several regions in a country.
2. Regional level, referring to programmes that had clear geographical targeting on selected administrative regions.
3. Local level, when multiple areas in the entire country were selected (e.g., cities).
4. Pilot level, capturing programmes that were implemented as a trial, with relative low scope and the expectation of future scale-up.

Note that the variable was coded on the intervention rather than the study (sample) level: That is, the classification did not reflect whether the evaluation was conducted for a sub-sample of the entire programme, but rather the main objective was to test the difference between (small-scale) local or pilot programmes and national-level policies.

Results are presented in Figure 27 (Table 32) and Figure 28 (Table 33). Studies of national-level programmes generally reported somewhat smaller effect sizes for both earnings (0.03 SMD; CI = 0.02, 0.05; I^2 = 76 per cent; number of interventions = 47) and employment outcomes (0.03 SMD; CI = 0.01, 0.05; I^2 = 59 per cent; number of interventions = 55). However, the difference in terms of smaller scale programmes is not statistically significant. In addition, there was large unexplained heterogeneity within all sub-groups except the sample of pilot programmes.

Figure 27: Summary forest plot of employment outcomes by scale of the programme

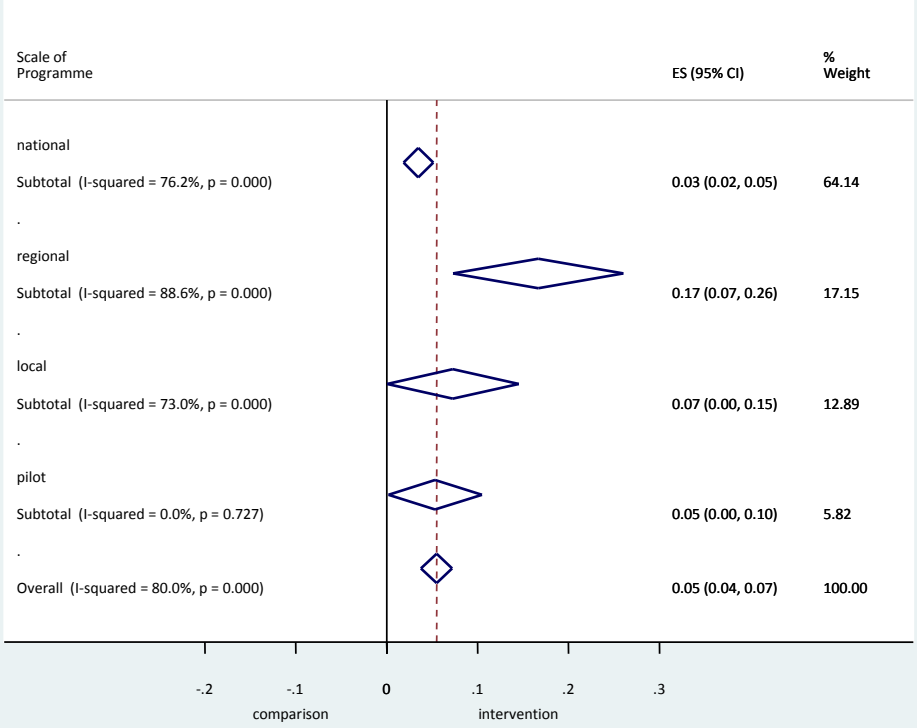


No. of SMDs/Studies: Total: 1293/102
national: 853/55. regional: 130/17. local: 135/24. pilot: 175/7.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Table 32: Summary of results on employment outcomes by scale of the programme

Parameters of interest	National	Regional	Local	Pilot
Standardized Means Difference I^2	0.03	0.06	0.05	0.08
Standard errors of the effect size	0.01	0.02	0.03	0.03
95% confidence interval	0.01	0.02	-0.01	0.03
	0.05	0.11	0.11	0.14
I Squared	58.69	69.02	73.93	0
Number of SMDs	853	130	135	175
Number of interventions	55	17	24	7
Sample size	37,218,469	700,150	170,586	107,989
Mean difference	2.13	3.35	11.70	13.72
Control outcome	47.60	18.54	1.62	23.10
Treatment outcome	44.46	19.15	2.45	35.67
Percentage change	0.04	0.08	0.14	0.33

Figure 28: Summary forest plot of earnings outcomes by scale of the programme



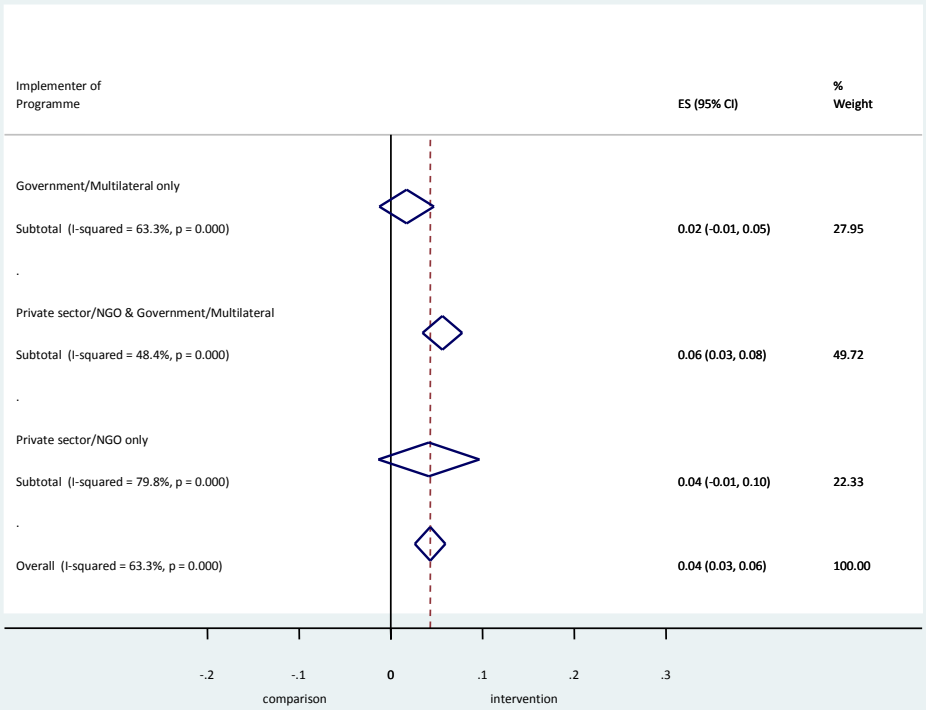
No. of SMDs/Studies: Total: 656/89
national: 417/47. regional: 108/21. local: 32/16. pilot: 99/6.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Table 33: Summary of results on earnings outcomes by scale of the programme

Parameters of interest	National	Regional	Local	Pilot
Standardized Means Difference I^2	0.03	0.17	0.07	0.05
Standard errors of the effect size	0.01	0.05	0.04	0.03
95% confidence interval	0.02	0.07	0	0
	0.05	0.26	0.15	0.1
I Squared	76.15	88.6	72.97	0
Number of SMDs	417	108	32	99
Number of interventions	47	21	16	6
Sample size	12,169,329	417,652	42,854	59,592
Mean difference	2,462.12	117.42	961.50	4,248.82
Control outcome	11,814.07	835.94	11,762.48	8,349.14
Treatment outcome	12,725.79	863.83	12,794.23	12,254.15
Percentage change	0.08	0.11	0.19	0.32

Figure 29 (and Table 34) and Figure 30 (and Table 35) provide summary SMDs for studies that analysed programmes implemented by different agencies. Implementers were categorized into public institutions, i.e., governments or multilateral organizations, and private entities, which were private sector firms or NGOs. In the analysis, the review team looked at the differential impact of programmes implemented by (i) governments and/or multilaterals, (ii) private sector firms and/or NGOs, or (iii) a combination of public and private sector (i.e., governments and/or multilaterals combined with private sector firms and/or NGOs). Any programmes that were not classified according to these three groups were called "other"; for example, when the implementing agency remained unknown to the reviewers.

Figure 29: Summary forest plot of employment outcomes by implementer



No. of SMDs/Studies: Total: 1315/103
Government/Multilateral only: 225/21. Private sector/NGO & Government/Multilateral: 825/59. Private sector/NGO only: 265/23.
Sub-group dropped (<4 independent studies): Other.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

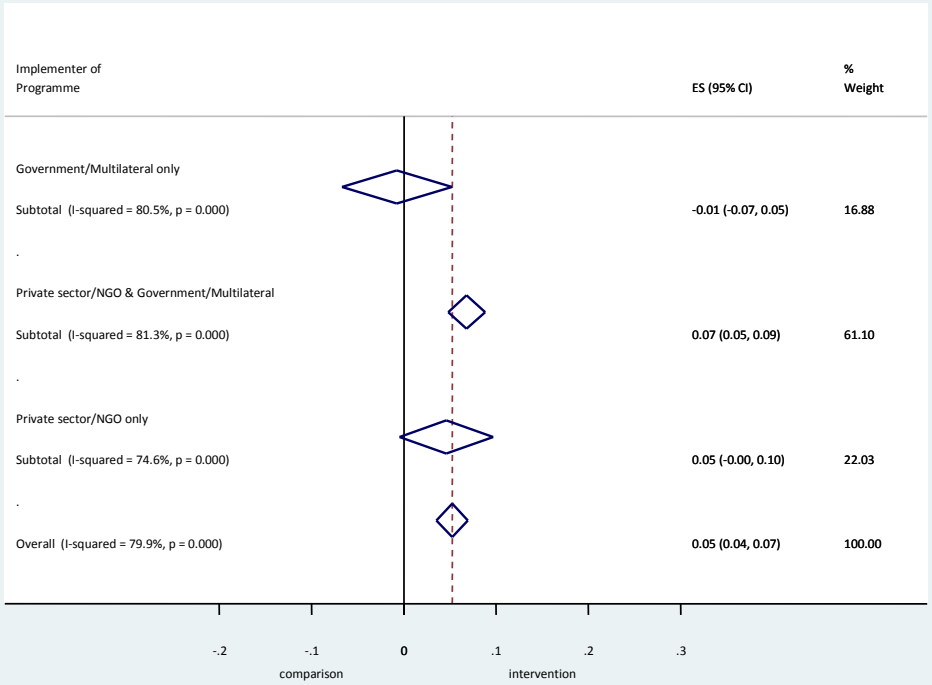
The team found that, for employment and earnings outcomes, a combination of public and private sector implementation led to the highest SMDs of around 0.06 (CI = 0.03, 0.08; I^2 = 48 per cent; number of interventions = 59) (employment) or 0.07 (CI = 0.05, 0.09; I^2 = 81 per cent; number of interventions = 57) (income) and significantly different from zero.

Private sector only implemented programmes (i.e., implemented by private sector firms and/or NGOs) led to moderate gains for both employment and income of around 0.04 SMDs (CI = -0.01, 0.10; I^2 = 80 per cent; number of interventions = 23) (employment) or 0.05 (CI = 0.00, 0.10; I^2 = 75 per cent; number of interventions = 21) (earnings) and with the summary SMD barely reaching significance at the 5 per cent level.

Table 34: Summary of results on employment outcomes by implementer

Parameters of interest	Government/Multilateral only	Other	Private sector/NGO & Government/ Multilateral	Private sector/ NGO only
Standardized Means Difference I^2	0.02	Dropped from analysis	0.06	0.04
Standard errors of the effect size	0.02		0.01	0.03
95% confidence interval	-0.01		0.03	-0.01
	0.05		0.08	0.1
I Squared	63.32		48.38	79.82
Number of SMDs	225		825	265
Number of interventions	21		59	23
Sample size	12,132,632		25,620,955	455,068
Mean difference	2.39		7.35	0.96
Control outcome	19.69		54.38	1.67
Treatment outcome	16.76		52.65	2.09
Percentage change	0.02		0.06	0.15

Figure 30: Summary forest plot of earnings outcomes by implementer



No. of SMDs/Studies: Total: 668/90
Government/Multilateral only: 86/12. Private sector/NGO & Government/Multilateral: 427/57. Private sector/NGO only: 155/21.
Sub-group dropped (<4 independent studies): Other.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

The summary SMD of studies of public sector only implemented programmes (i.e., government and/or multilateral agency as implementers) was statistically insignificant for both employment and earnings outcomes.

However, as it was possible that the analysis could have been confounded with other intervention- or study-level characteristics that were correlated with programme scale (e.g., country income level), this difference was explored in more detail as part of the multivariate meta-analysis.

Table 35: Summary of results on earnings outcomes by implementer

Parameters of interest	Government/Multilateral only	Other	Private Sector/NGO & Government/ Multilateral	Private Sector/ NGO only
Standardized Means Difference I^2	-0.01	Dropped from analysis	0.07	0.05
Standard errors of the effect size	0.03		0.01	0.03
95% confidence interval	-0.07		0.05	0
	0.05		0.09	0.1
I Squared	80.54		81.26	74.56
Number of SMDs	86		427	155
Number of interventions	12		57	21
Sample size	3,524,836		8,919,289	251,413
Mean difference	-1,526.65		3,859.60	1,296.64
Control outcome	7,180.16		15,993.60	3,242.02
Treatment outcome	5,285.78		17,770.68	4,003.11
Percentage change	0.06		0.09	0.16

4.3.4 Multivariate meta-regression analysis

This section presents empirical estimates of the correlates of programme effectiveness from meta-regressions, as described in Section 3.4.6. Table 36 to Table 38 display these regression results. The bottom part of each table reports key (sub-) sample characteristics, including the number of observations, number of interventions and studies and programme type distribution for each specification.

First, in Table 36, the main results for the full sample are reported: the first panel on the left-hand side contains estimates for the pooled sample, i.e., comprising all countries and all outcome measures. The middle panel includes impact estimates for employment outcomes only, while the right-hand panel includes impact estimates for earnings outcomes. Each of the panels contains four different specifications in column (1) to (4) that introduce additional covariates in a stepwise manner.

Table 37 and Table 38 follow the same structure as Table 36, but consider the two main country groups separately: Table 37 focuses on impact estimates from high-income countries, while Table 38 includes impact estimates from low- and middle-income countries. The analysis by country income level recognized the role of context and the significant differences in labour markets across country type, from labour market barriers to the level of investment in youth employment and degree of enforcement of labour market regulations and institutions.

Table 36: Meta-regression results (full sample)

	All outcomes pooled				Employment Outcomes				Earnings/Income Outcomes			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Type of intervention												
Base category: skills training												
Main category: entrepreneurship promotion	0.018	0.013	0.054***	0.069***	0.032	0.025	0.090***	0.142***	0.018	0.02	0.049*	0.008
	[1.26]	[0.76]	[3.08]	[2.71]	[1.20]	[0.93]	[3.16]	[3.98]	[0.82]	[0.88]	[1.82]	[0.17]
Main category: employment services	-0.029***	-0.004	0.090***	0.050*	-0.041**	0.01	0.092***	0.053	-0.042***	-0.015	0.100***	0.007
	[3.16]	[0.40]	[5.54]	[1.69]	[2.58]	[0.59]	[4.10]	[1.35]	[3.09]	[1.07]	[3.24]	[0.14]
Main category: subsidized employment	-0.037***	-0.009	0.012	0.01	-0.041***	-0.014	0.014	0.011	-0.038***	-0.016	0.032	-0.033
	[4.70]	[1.02]	[0.82]	[0.40]	[3.06]	[1.03]	[0.63]	[0.32]	[3.90]	[1.53]	[1.30]	[0.76]
Main category: unspecified	-0.038***	-0.029**	-0.057***	-0.082***	-0.041**	-0.005	-0.058**	-0.059*	0.006	0	-0.056*	-0.130***
	[3.01]	[2.16]	[3.36]	[3.13]	[2.03]	[0.22]	[2.48]	[1.73]	[0.31]	[0.02]	[1.77]	[2.84]
= 1 if has an additional component	-0.019***	-0.004	0.043***	0.079***	-0.028***	-0.026**	0.031***	0.073***	-0.020***	0.039***	0.077***	0.144***
	[2.97]	[0.54]	[4.79]	[4.82]	[2.59]	[2.21]	[2.62]	[3.75]	[2.67]	[3.68]	[4.53]	[4.39]
Study characteristics												
Standard error of SMD		0.873***	0.895***	0.374***		1.075***	1.266***	0.792***		0.924***	0.925***	0.844***
		[9.64]	[8.72]	[2.88]		[8.05]	[9.06]	[4.00]		[7.20]	[4.82]	[3.45]
Publication peer-reviewed		0.013*	-0.006	0.001		-0.006	-0.025**	-0.035***		0.029***	0.015	0.044***
		[1.93]	[0.81]	[0.10]		[0.51]	[2.18]	[2.59]		[3.26]	[1.21]	[2.99]
Evaluation design: RCT		-0.01	-0.019**	-0.076***		-0.016	-0.005	-0.070***		-0.057***	-0.054***	-0.116***
		[1.25]	[2.09]	[5.93]		[1.37]	[0.41]	[4.22]		[4.66]	[3.09]	[5.21]
High-income country		-0.036***	-0.068***	-0.148***		-0.045***	-0.040***	-0.098***		-0.024***	-0.101***	-0.154***
		[4.86]	[7.53]	[9.53]		[3.79]	[3.31]	[4.55]		[2.61]	[5.91]	[5.25]
Outcome measure												
Base category: business outcomes												
Employment outcome		0.077***	0.062***	0.03								
		[3.07]	[2.61]	[1.10]								
Earnings/income outcome		0.084***	0.068***	0.042								
		[3.32]	[2.87]	[1.60]								
Estimated unadjusted difference in means		0.017*	-0.022**	-0.011		0.035*	-0.031*	-0.039		0.016	-0.005	0.025
		[1.79]	[2.06]	[0.67]		[1.87]	[1.84]	[1.60]		[1.42]	[0.37]	[1.22]
Follow-up over one year later			0.044***	0.059***			0.037***	0.057***			0.053***	0.056***
			[6.18]	[6.43]			[3.54]	[4.59]			[4.84]	[3.99]
Evaluation sample												
Base category: pooled sample												
Low-income/disadvantaged youth			0.073***	0.011			0.059***	0.042			0.087***	-0.002
			[6.64]	[0.46]			[3.67]	[1.36]			[4.73]	[0.04]
Male participants			-0.004	0.011			-0.008	0.007			-0.005	0.002
			[0.44]	[1.10]			[0.63]	[0.53]			[0.37]	[0.10]
Female participants			-0.008	0.003			-0.004	0.004			-0.005	0.008
			[1.08]	[0.35]			[0.35]	[0.27]			[0.44]	[0.56]
Younger participants			0.001	0.022			-0.003	-0.01			0.01	0.064**
			[0.16]	[1.29]			[0.29]	[0.44]			[0.93]	[2.56]
Intervention characteristics												
Design includes paticipant profiling				0.051**				0.075***				-0.019
				[2.55]				[3.02]				[0.47]
Participant engagement mechanism (supervision or incentives)				0.072***				0.066***				0.069**
				[4.86]				[3.52]				[2.14]
Design includes incentives for service providers				0.043***				0.044**				0.012
				[2.87]				[2.18]				[0.46]
Implementer of programme: government				-0.048***				-0.084***				-0.062**
				[2.88]				[3.80]				[2.07]
Implementer of programme: NGO/non-profit				0.030***				0.008				0.036*
				[2.68]				[0.54]				[1.79]
Implementer of programme: private sector				-0.009				0.032*				-0.061*
				[0.63]				[1.67]				[1.80]
Programme has soft skills training				0.016				-0.024				0.033
				[0.87]				[0.97]				[1.12]
Younger participants and programme has soft skills training				-0.02				0.004				-0.064**
				[1.02]				[0.15]				[2.18]
Constant	0.052***	-0.049*	-0.107***	-0.025	0.080***	0.039***	-0.061***	-0.017	0.040***	0.025***	-0.048**	0.072
	[10.37]	[1.87]	[4.24]	[0.62]	[10.34]	[2.59]	[3.23]	[0.33]	[6.27]	[3.39]	[2.44]	[0.98]
n	2,213	2,213	1,428	1,021	1,377	1,377	796	504	670	670	469	365
No. of interventions/studies	100/100	100/100	62/75	36/51	91/88	91/88	56/66	31/44	77/75	77/75	52/59	31/43
No. of Interventions by main category of intervention	ST:54. EP:15. ES:10. SE:16. UN:5	ST:54. EP:15. ES:10. SE:16. UN:5.	ST:37. EP:11. ES:5. SE:6. UN:3.	ST:17. EP:8. ES:5. SE:5. UN:1.	ST:53. EP:7. ES:10. SE:16. UN:5.	ST:53. EP:7. ES:10. SE:16. UN:5.	ST:37. EP:5. ES:5. SE:6. UN:3.	ST:17. EP:3. ES:5. SE:5. UN:1.	ST:45. EP:12. ES:8. SE:9. UN:3.	ST:45. EP:12. ES:8. SE:9. UN:3.	ST:33. EP:10. ES:4. SE:3. UN:2.	ST:16. EP:7. ES:4. SE:3. UN:1.

Notes: All columns represent results from a random-effects meta-regression using the restricted maximum likelihood (REML) estimator. *n* refers to the number of effect sizes (before dropping sub-group estimates and aggregating at the study level). The main category of intervention is given by the following abbreviations: ST = skills training, EP = entrepreneurship promotion, ES = employment services, SE = subsidized employment, UN = unspecified. *t*-test statistics are shown in square brackets below coefficient estimates. Asterisks indicate statistical significance: *, ** and *** denote statistical significance at the 10 per cent, 5 per cent and 1 per cent level of significance, respectively.

Table 37: Meta-regression results (high-income countries sample)

	All outcomes pooled				Employment Outcomes				Earnings/Income Outcomes			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Type of intervention												
Base category: skills training												
Main category: entrepreneurship promotion	-0.054	-0.081*	-0.109***		-0.025	-0.055	-0.105*		-0.083*	-0.093*	-0.08	
	[1.49]	[1.81]	[2.99]		[0.37]	[0.82]	[1.84]		[1.74]	[1.87]	[1.48]	
Main category: employment services	-0.022**	-0.013	0.106***	0.089	-0.045**	0	0.132***	0.12	-0.031**	-0.027*	0.107	
	[1.97]	[1.10]	[4.01]	[1.44]	[2.29]	[0.02]	[3.46]	[1.31]	[2.29]	[1.87]	[1.63]	
Main category: subsidized employment	-0.034***	-0.019*	0.022	-0.154***	-0.048***	-0.031*	0.018	-0.147*	-0.030***	-0.029***	0.019	-0.193**
	[3.53]	[1.85]	[1.23]	[2.70]	[2.86]	[1.76]	[0.60]	[1.72]	[2.83]	[2.76]	[0.62]	[2.08]
Main category: unspecified	-0.026*	-0.022	-0.027	0.145**	-0.006	0.024	-0.012	0.197**	0.018	0.006	-0.045	
	[1.95]	[1.49]	[1.42]	[2.39]	[0.27]	[0.99]	[0.39]	[2.47]	[1.01]	[0.28]	[1.41]	
= 1 if has an additional component												
	-0.023***	-0.028**	0.019	-0.180***	-0.061***	-0.060***	0.017	-0.184***	-0.023**	0.009	0.038	0.026
Study characteristics	[2.71]	[2.57]	[1.49]	[4.32]	[4.20]	[3.61]	[0.88]	[4.07]	[2.48]	[0.51]	[1.29]	[0.56]
Standard error of SMD		0.784***	1.151***	0.912***		1.145***	1.652***	0.728***		0.531***	0.665**	1.167***
		[5.84]	[6.80]	[4.09]		[6.01]	[7.15]	[2.63]		[3.40]	[2.51]	[3.09]
Publication peer-reviewed		0.007	-0.003	-0.016		-0.001	-0.023	-0.021		0.016*	0.004	-0.013
		[0.76]	[0.26]	[1.06]		[0.04]	[1.36]	[0.90]		[1.68]	[0.31]	[0.64]
Evaluation design: RCT		0	-0.023*	-0.098***		-0.029*	-0.029	-0.087**		-0.035*	-0.051	-0.166***
		[0.01]	[1.72]	[3.38]		[1.69]	[1.60]	[2.29]		[1.94]	[1.53]	[3.17]
Outcome measure												
Base category: bsuiness outcomes												
Employment outcome		0.037	-0.001	-0.005								
		[0.50]	[0.01]	[0.58]								
Earnings/income outcome		0.039	0.001									
		[0.53]	[0.01]									
Estimated unadjusted difference in means		0.022**	0.003	-0.026		0.066***	0.005	-0.012		0.006	-0.007	-0.028
		[2.00]	[0.23]	[1.59]		[2.81]	[0.25]	[0.46]		[0.49]	[0.47]	[1.43]
Follow-up over one year later			0.060***	0.060***			0.055***	0.049***			0.079***	0.080***
			[6.51]	[5.97]			[3.51]	[3.63]			[6.29]	[5.38]
Evaluation sample												
Base category: pooled sample												
Low-income/disadvantaged youth			0.087***	0.121**			0.067**	0.09			0.119**	
			[3.79]	[2.07]			[2.02]	[1.23]			[2.32]	
Male participants			0.002	0.019*			-0.011	0.019			0.003	0.017
			[0.22]	[1.76]			[0.70]	[1.44]			[0.21]	[0.98]
Female participants			-0.027***	-0.020*			-0.039**	-0.022			-0.019	-0.018
			[2.66]	[1.76]			[2.43]	[1.54]			[1.36]	[0.95]
Younger participants			-0.003	0.035			0.009	0.028			0	0.327***
			[0.36]	[0.86]			[0.59]	[0.67]			[0.01]	[3.62]
Intervention characteristics												
Design includes paticipant profiling				0.259***				0.250***				0.132***
				[6.97]				[5.95]				[2.99]
Design includes incentives for service providers				0.156***				0.178***				0.002
				[4.55]				[4.85]				[0.05]
Implementer of programme: government				0.231**				0.231**				-0.016
				[2.55]				[2.32]				[0.11]
Implementer of programme: NGO/non-profit				0.032				-0.007				0.11
				[0.62]				[0.12]				[0.80]
Implementer of programme: private sector				0.227***				0.239***				0.120***
				[4.46]				[3.90]				[3.59]
Programme has soft skills training				-0.172**				-0.116				0.063
				[1.98]				[1.03]				[0.45]
Younger participants and programme has soft skills training				-0.048				-0.044				-0.342***
				[1.17]				[0.99]				[3.73]
Constant	0.043***	-0.029	-0.132**	-0.419***	0.083***	0.008	-0.129***	-0.424**	0.031***	0.021**	-0.149***	-0.284
	[5.71]	[0.39]	[2.16]	[3.13]	[7.34]	[0.47]	[3.33]	[2.60]	[3.55]	[2.31]	[2.75]	[1.62]
n	1,306	1,306	628	471	873	873	369	271	424	424	250	200
No. of interventions/studies	53/54	53/54	27/38	15/26	52/47	52/47	26/32	15/22	35/36	35/36	20/28	Nov-20
No. of interventions by main category of intervention	ST:28. EP:2. ES:7. SE:11. UN:5.	ST:28. EP:2. ES:7. SE:11. UN:5.	ST:15. EP:2. ES:3. SE:4. UN:3.	ST:7. EP:0. ES:3. SE:4. UN:1.	ST:27. EP:2. ES:7. SE:11. UN:5.	ST:27. EP:2. ES:7. SE:11. UN:5.	ST:15. EP:1. ES:3. SE:4. UN:3.	ST:7. EP:0. ES:3. SE:4. UN:1.	ST:20. EP:2. ES:5. SE:5. UN:3.	ST:20. EP:2. ES:5. SE:5. UN:3.	ST:12. EP:2. ES:2. SE:2. UN:2.	ST:6. EP:0. ES:2. SE:2. UN:1.

Notes: All columns represent results from a random-effects meta-regression using the restricted maximum likelihood (REML) estimator. *n* refers to the number of effect sizes (before dropping sub-group estimates and aggregating at the study level). The main category of intervention is given by the following abbreviations: ST = skills training, EP = entrepreneurship promotion, ES = employment services, SE = subsidized employment, UN = unspecified. *t*-test statistics are shown in square brackets below coefficient estimates. Asterisks indicate statistical significance: *, ** and *** denote statistical significance at the 10 per cent, 5 per cent and 1 per cent level of significance, respectively.

The empirical results from the meta-regression suggested that, overall, **programme effectiveness differed by intervention type**, and that certain contextual and design factors (in terms of both the intervention and the evaluation of it) mattered in determining the estimated impacts that a programme could attain. Table 36 (upper part labelled "Type of intervention") presents coefficients by intervention type.

- The first most notable result was that *entrepreneurship promotion interventions were correlated to larger effect sizes, relative to the base category (skills training), and that they appeared to be the most successful* interventions. The middle panel of Table 36 (Employment outcomes) and the panel on the right-hand side (Earnings and income outcomes) indicate that these positive correlations of the entrepreneurship programmes and the SMD were driven by employment outcomes – probably, above all, the impact on self-employment – but not by earnings outcomes. Looking across country groups, Table 38 shows that the positive impacts of entrepreneurship promotion in low- and middle-income countries drove the reported pooled effects.

- The second result concerning the intervention type was that *employment services also appeared to be more successful than skills training, specifically once the time of the follow-up impact measurement and participant profiles capturing gender and age* were accounted for (i.e., the third of the four specifications reported in each panel). The evidence for this came mainly from high-income countries (Table 37) and may have reflected efficiency gains from effective job matching and case management. For the low- and middle-income country sample, all coefficients for the employment services intervention type were not statistically different from zero (Table 38).

- The third result is that *multi-pronged interventions worked* (Table 36). While this finding had already been reported by other reviews (e.g., Betcherman et al., 2007), this systematic review captured the role of country income level in this result. The success of multi-component interventions was largely determined by country group heterogeneity and the evaluations of such programmes from low- and middle income countries. For the low- and middle-income country group, the evidence suggested that combining multiple components generated more successful programmes: The coefficients on the indicator variable specifying whether an intervention had more than one component ("= 1 if has an additional component") in Table 38 were positive and significant in most specifications. At the same time, in the high-income country sample the meta-regression estimates suggested that interventions which focused on one main category were more effective.

Table 38: Meta-regression results (low- and medium-income countries sample)

	All outcomes pooled				Employment Outcomes				Earnings/Income Outcomes			
	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)	(1)	(2)	(3)	(4)
Type of intervention												
Base category: skills training												
Main category: entrepreneurship promotion	0.005	0.040**	0.099***	-0.01	0.040*	0.046**	0.139***	0.019	0.002	0.029	0.078**	0.013
	[0.40]	[2.35]	[4.56]	[0.27]	[1.72]	[1.97]	[4.34]	[0.35]	[0.08]	[1.04]	[2.31]	[0.20]
Main category: employment services	-0.021	-0.015	0.024	-0.07	0	-0.007	0.029	-0.163	-0.097	-0.09	-0.116	-0.115
	[0.68]	[0.48]	[0.65]	[0.95]	[0.01]	[0.19]	[0.73]	[1.14]	[1.24]	[1.22]	[1.30]	[0.99]
Main category: subsidized employment	0.004	-0.004	0.059	0.178***	0.024	0.027	0.072*	0.265***	-0.077	-0.090*	0.057	0.114
	[0.20]	[0.20]	[1.51]	[3.03]	[0.93]	[1.08]	[1.70]	[3.21]	[1.57]	[1.92]	[0.62]	[1.18]
= 1 if has an additional component												
	0.048***	0.055***	0.074***	0.056	0.049***	0.027*	0.024	0.102	0.060***	0.083***	0.130***	0.031
Study characteristics	[4.02]	[4.36]	[5.10]	[0.93]	[3.23]	[1.68]	[1.39]	[0.65]	[2.65]	[3.88]	[4.64]	[0.28]
Standard error of SMD		0.811***	0.608***	0.082		1.093***	0.991***	0.464		1.329***	1.066***	0.653**
		[6.57]	[4.25]	[0.48]		[6.51]	[4.98]	[1.38]		[5.36]	[3.81]	[2.11]
Publication peer-reviewed		0.045***	0.023	0.044*		0.011	-0.034*	0.037		0.087***	0.111***	0.140***
		[3.76]	[1.56]	[1.92]		[0.70]	[1.69]	[0.80]		[4.24]	[4.34]	[4.20]
Evaluation design: RCT		-0.022*	-0.02	-0.066**		0.002	0.005	-0.021		-0.071***	-0.096***	-0.105***
		[1.93]	[1.32]	[2.40]		[0.17]	[0.27]	[0.40]		[3.41]	[3.61]	[2.67]
Outcome measure												
Base category: business outcomes												
Employment outcome		0.100***	0.079***	-0.001		-0.054	-0.087			0.034	0.102	0.141
		[4.28]	[2.75]	[0.03]		[0.69]	[1.12]			[0.30]	[0.89]	[1.28]
Earnings/income outcome		0.111***	0.100***	0.018								
		[4.66]	[3.56]	[0.59]								
Estimated unadjusted difference in means		-0.027	-0.043	0.095								
		[0.47]	[0.73]	[0.83]								
Follow-up over one year later			0.004	0.005			0.015	0.004			-0.038*	-0.056*
			[0.32]	[0.26]			[0.88]	[0.14]			[1.76]	[1.95]
Evaluation sample												
Base category: pooled sample												
Low-income/disadvantaged youth			0.078***	0.044			0.080***	0.054			0.049*	0.025
			[4.86]	[1.31]			[3.65]	[1.07]			[1.95]	[0.47]
Male participants			-0.040**	0			-0.032	0.01			-0.061**	-0.049*
			[2.55]	[0.02]			[1.43]	[0.35]			[2.35]	[1.91]
Female participants			0.003	0.018			0.022	0.039*			0.021	0.016
			[0.27]	[1.16]			[1.37]	[1.69]			[1.01]	[0.73]
Younger participants			0.008	0.055**			-0.009	0.088*			0.015	0.01
			[0.55]	[2.07]			[0.48]	[1.86]			[0.65]	[0.31]
Intervention characteristics												
Design includes paticipant profiling				0.02				0.068				-0.019
				[0.42]				[0.90]				[0.25]
Participant engagement mechanism (supervision or incentives)				-0.016				-0.057				-0.011
				[0.50]				[1.00]				[0.20]
Design includes incentives for service providers				-0.036				-0.087				0.008
				[1.01]				[1.27]				[0.18]
Implementer of programme: government				-0.026				-0.034				-0.026
				[1.07]				[0.52]				[0.67]
Implementer of programme: NGO/non-profit				0.003				-0.025				0.067
				[0.11]				[0.50]				[1.39]
Implementer of programme: private sector				-0.174***				-0.175**				-0.158**
				[4.89]				[2.47]				[2.48]
Programme has soft skills training				0.001				0				-0.038
				[0.04]				[0.00]				[0.87]
Younger participants and programme has soft skills training				-0.019				-0.130*				0.221**
				[0.36]				[1.87]				[2.47]
Constant	0.074***	-0.082***	-0.109***	0.219***	0.075***	0.001	-0.051**	0.199	0.085***	0.009	0.005	0.173
	[11.62]	[3.07]	[3.58]	[3.19]	[8.50]	[0.07]	[2.14]	[1.39]	[7.77]	[0.51]	[0.17]	[1.45]
n	907	907	800	550	504	504	427	233	246	246	219	165
No. of interventions/studies	47/46	47/46	35/37	21/25	39/41	39/41	30/34	16/22	42/39	42/39	32/31	20/23
Np. of interventions by main category of intervention	ST:26. EP:13. ES:3. SE:5. UN:0.	ST:26. EP:13. ES:3. SE:5. UN:0.	ST:22. EP:9. ES:2. SE:2. UN:0.	ST:10. EP:8. ES:2. SE:1. UN:0.	ST:26. EP:5. ES:3. SE:5. UN:0.	ST:26. EP:5. ES:3. SE:5. UN:0.	ST:22. EP:4. ES:2. SE:2. UN:0.	ST:10. EP:3. ES:2. SE:1. UN:0.	ST:25. EP:10. ES:3. SE:4. UN:0.	ST:25. EP:10. ES:3. SE:4. UN:0.	ST:21. EP:8. ES:2. SE:1. UN:0.	ST:10. EP:7. ES:2. SE:1. UN:0.

Notes: All columns represent results from a random-effects meta-regression using the restricted maximum likelihood (REML) estimator. *n* refers to the number of effect sizes (before dropping sub-group estimates and aggregating at the study level). The main category of intervention is given by the following abbreviations: ST = skills training, EP = entrepreneurship promotion, ES = employment services, SE = subsidized employment, UN = unspecified. *t*-test statistics are shown in square brackets below coefficient estimates. Asterisks indicate statistical significance: *, ** and *** denote statistical significance at the 10 per cent, 5 per cent and 1 per cent level of significance, respectively.

Looking at design factors of the impact evaluation, the regression results suggested that **active labour market measures can positively impact employment and earnings of youth but the change does not take effect immediately: improvements in outcomes take time to materialize.**

- First, it is important to note that, as captured by the "Outcome measure" set of covariates, there was some evidence that *effects on employment and earnings outcomes were larger than those on business performance indicators* (the base category). This can be seen in the pooled – by outcome – regressions in the panel on the left-hand side in each of Table 36, Table 37 and Table 38. The overall effect is mainly driven by the evidence from low- and middle-income countries.

- Second, Table 36 and Table 37 show that – overall, and in particular for high-income countries – *intervention impacts were significantly larger if measured at a later stage*, as captured by the indicator variable showing whether follow-up data were collected at least one year after an individual had exited the intervention ("Follow-up over one year later"). This finding was in line with recent results from the meta-analysis in Card et al. (2015), whose sample consisted mainly of ALMPs that were not specifically targeted at youth. Their estimation results pointed to a pronounced increasing pattern of programme impacts over time post-programme, which was particularly strong in the case of skills training programmes.

Other study factors were also correlated with the SMD of the estimated programme impact; for instance:

- *The research design mattered*. Studies based on RCTs consistently reported a significantly smaller magnitude of the effect size than non-experimental studies.

- In addition, there was some indication that *peer-reviewed studies provided smaller effect sizes* (which were also less frequently positive and significant); this finding would, if anything, appear to refute the suggestion that publication bias could systematically affect the data used here. This result is discussed in more detail as part of the analysis of publication bias (Section 4.3.6).

- One important consideration in this regard is that the review team *did not find differential treatment effects by gender, or by the age group* considered in the original impact study (the latter referring to an indicator variable that mapped whether the sample looked at the younger or older half of the youth age distribution).

The next set of covariates introduced in the meta-regressions captured features of the intervention design as well as characteristics of the participants for whom impacts were estimated. The results implied that **certain design and**

implementation features determine programme success.

- Concurrently, programme design features, such as *participant profiling, participant engagement mechanisms (through supervision or incentives) and specific incentives for service providers were all positively correlated with larger effect sizes*; this overall result was driven by the corresponding positive coefficients in the estimations for the high-income country sample (Table 37).

- Other programme features also appeared to be important. Specifically, the type of programme implementer correlated with the dependent variable in a differential way by country group. *In the high-income country sample, programmes implemented by the Government and the private sector were correlated with a larger SMD*. At the same time, *in the low- and middle-income country sample, private sector participation was correlated with a smaller effect size*.

- Finally, for both country groups the indicator which captured whether or not a programme contained a soft skills training component showed no significant correlation with larger effect sizes.

Overall, the evidence from the meta-regressions indicated that programme type and design features played an important role in determining effect sizes. In addition, many of these factors seemed to differ in important ways between the two country samples of high-income and low- and middle-income countries. For instance, programme evaluations generally showed larger effect sizes in low- and middle-income countries than in high-income countries. Also – as discussed above – entrepreneurship promotion and skills training were the most effective programmes in low- and middle-income countries. In high-income countries, the evidence was somewhat mixed, but, besides skills training, it was mainly employment services interventions that constituted the more effective youth labour market programmes.

4.3.5 Sensitivity analysis

In this section, the robustness of the results from both the univariate analysis and the multivariate random-effects meta-regression (Sections 4.3.2, 4.3.3, and 4.3.4) are tested. For the sake of brevity, this section discusses the sensitivity of the results from the overall synthesis of the evidence (pooled sample) and the moderator analysis by main intervention category. Hence, the robustness of each moderator analysis is not discussed, as these generally reflected findings in the pooled sample.

The following section focuses on three types of decisions which may have affected the overall results: First, different assumptions in computing (or imputing) effect sizes are tested. Second, the robustness of some of the decisions in the data synthesis (e.g., regarding outliers) is checked. Third, the question of whether the variance in effect sizes might be caused by factors related to the applied evaluation design (i.e., study type, risk of bias) is investigated.

For the univariate analysis, the respective summary forest plots are again included in the main text while, for this section, forest plots showing each intervention SMD separately have not been appended.[36] Table 39 tests whether results based on the two main model specifications (columns (2) and (4) in each table respectively) in Section 4.3.4 hold when restricting the sample along the different dimensions of the sensitivity analysis discussed below.

In addition to the sensitivity analysis discussed in this section, the review team performed various other checks on the analysis (e.g., using Cohen's *d* instead of Hedges' *g*; testing for differences between Intention-to-Treat (ITT) and Average Treatment Effect on the Treated (ATT) estimates; additional methods of data imputation). Since none of these checks significantly altered the main results, they are not discussed in detail.

[36] The full results are available from the authors upon request.

Table 39: Sensitivity analysis of meta-regression (full sample)

	Pooled SMD (full imputation)	Pooled SMD (no imputation)	Pooled SMD (full imputation, only RCT)	Pooled PSS probit (full imputation)	Pooled SMD (full imputation)	Pooled SMD (no imputation)	Pooled SMD (full imputation, only RCT)	Pooled PSS probit (full imputation)
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Type of Intervention								
Base category: skills training								
Main category: entrepreneurship promotion	0.013	0.052***	0.052***	0.111	0.069***	0.077**	0.025	0.006
	[0.76]	[3.30]	[3.94]	[1.11]	[2.71]	[2.42]	[0.99]	[0.08]
Main category: employment services	-0.004	-0.008	0.054***	-0.145	0.050*	0.033	0.048	-0.347**
	[0.40]	[1.02]	[4.57]	[1.45]	[1.69]	[1.06]	[1.17]	[1.97]
Main category: subsidized employment	-0.009	0.007	0.016	-0.148*	0.01	0.055	0.094**	-0.262
	[1.02]	[1.19]	[0.59]	[1.87]	[0.40]	[1.30]	[2.01]	[1.55]
Main category: unspecified	-0.029**	-0.003	-0.006	0.034	-0.082***	-0.119***	-0.148***	0
	[2.16]	[0.16]	[0.51]	[0.30]	[3.13]	[3.23]	[5.24]	
= 1 if has an additional component	-0.004	0.018**	0.072***		0.079***		0.140***	
	[0.54]	[2.16]	[7.35]	0.016	[4.82]	0.126***	[5.45]	0.153
Study characteristics				[0.23]		[5.35]		[1.48]
Standard Error of SMD (log evaluation sample size for PSS)	0.873***	0.803***	0.212**	0.050***	0.374***	0.041	0.215	0.050*
	[9.64]	[11.62]	[2.41]	[2.66]	[2.88]	[0.27]	[1.48]	[1.72]
Publication peer-reviewed	0.013*	0.026***	0.007	-0.114	0.001	0.01	0.012	0.018
	[1.93]	[4.25]	[1.05]	[1.64]	[0.10]	[0.67]	[0.92]	[0.27]
Evaluation design: RCT	-0.01	-0.019***		-0.053	-0.076***	-0.080***		-0.008
	[1.25]	[2.65]		[0.78]	[5.93]	[3.86]		[0.11]
High-income country	-0.036***	0.003	-0.082***	-0.067	-0.148***	-0.151***	-0.104***	-0.234**
	[4.86]	[0.47]	[8.72]	[0.90]	[9.53]	[6.90]	[4.63]	[2.13]
Outcome measure								
Base category: business outcomes								
Employment outcome	0.077***	0.037**	0.028*	0.094	0.03	0.015	0.027	0.059
	[3.07]	[2.17]	[1.78]	[0.76]	[1.10]	[0.57]	[1.08]	[0.62]
Earnings/income outcome	0.084***	0.036**	0.018	0.083	0.042	0.03	0.028	0.147
	[3.32]	[2.15]	[1.16]	[0.66]	[1.60]	[1.17]	[1.14]	[1.64]
Estimated unadjusted difference in means	0.017*	0.005	-0.001	-0.023	-0.011	-0.025	-0.002	-0.079
	[1.79]	[0.70]	[0.15]	[0.16]	[0.67]	[1.04]	[0.15]	[0.53]
Follow-up over one year later					0.059***	0.025**	0.056***	0.133*
					[6.43]	[2.09]	[5.86]	[1.81]
Evaluation sample								
Base category: pooled sample								
Low-income/disadvantaged youth					0.011	0.041	0.013	0.144
					[0.46]	[1.25]	[0.46]	[1.02]
Male participants					0.011	0.025*	0.022**	-0.057
					[1.10]	[1.88]	[2.08]	[1.21]
Female participants					0.003	-0.001	0.012	-0.012
					[0.35]	[0.09]	[1.11]	[0.20]
Younger participants					0.022	0.071**	0.021	0.062
					[1.29]	[1.98]	[0.64]	[0.50]
Design includes paticipant profiling							0.021	
					0.051**	0.059**	[0.83]	0.332***
Intervention Characteristics					[2.55]	[2.14]		[2.72]
Participant engagement mechanism (supervision or incentives)					0.072***	0.097***	0.051***	0.327***
					[4.86]	[3.99]	[2.90]	[2.72]
Design includes incentives for service providers					0.043***	0.028	0.003	0.162
					[2.87]	[1.48]	[0.11]	[1.48]
Implementer of programme: government					-0.048***	-0.091***	-0.064***	-0.319***
					[2.88]	[4.88]	[2.94]	[3.03]
Implementer of programme: NGO/non-profit					0.030***	0.039**	0	-0.006
					[2.68]	[2.20]	[0.00]	[0.08]
Implementer of programme: private sector					-0.009	-0.058***	-0.078***	-0.125
					[0.63]	[2.76]	[4.01]	[1.51]
Programme has soft skills training					0.016	-0.004	-0.035	-0.376***
					[0.87]	[0.19]	[1.37]	[3.03]
Younger participants and programme has soft skills training					-0.02	-0.083**	-0.034	0.053
					[1.02]	[2.06]	[1.01]	[0.36]
Constant	-0.049*	-0.040**	0.005		-0.025	0.063	0.02	
	[1.87]	[2.33]	[0.25]		[0.62]	[1.20]	[0.42]	
n	2,213	1,166	1,266	2,081	1,021	582	742	968
No. of interventions/studies	100/100	76/69	42/49	99/92	36/51	29/36	24/32	36/47
No. of interventions by main category of intervention	ST:54. EP:15. ES:10. SE:16. UN:5.	ST:43. EP:10. ES:8. SE:10. UN:5.	ST:21. EP:11. ES:4. SE:3. UN:3.	ST:54. EP:14. ES:10. SE:16. UN:5.	ST:17. EP:8. ES:5. SE:5. UN:1.	ST:15. EP:7. ES:4. SE:2. UN:1.	ST:11. EP:8. ES:3. SE:1. UN:1.	ST:37. EP:11. ES:5. SE:6. UN:3.

Notes: Columns (1)–(4) represent model specification (2) and columns (5)–(8) represent model specification (4) from the full sample meta-regression results in Table 43. With the exception of columns (4) and (8), all models represent estimates from a random-effects meta-regression using the restricted maximum likelihood (REML) estimator. Columns (4) and (8) represent marginal effects from probit regression on the indicator variable whether the treatment effect was positive and statistically significant at the 5 per cent level. Probit regressions are weighted by the inverse of the number of observations of the respective study. Standard errors are clustered at the study level. Asterisks indicate statistical significance: *, ** and *** denote statistical significance at the 10 per cent, 5 per cent and 1 per cent level of significance, respectively. *n* refers to the number of effect sizes (before dropping sub-group estimates and aggregating at the study level).

4.3.5.1 Imputation of missing information

As discussed in Section 3.4.4, in some cases it was necessary to impute information and make assumptions in order to compute SMDs for specific studies.

- First, the review team had to make certain assumptions regarding the sample size of the treatment and/or comparison group if either of these was not reported.
- Second, the team approximated SMDs using the formula provided by Borenstein, Cooper, Hedges and Valentine (2009) to approximate SMDs where information on the pooled standard deviation could not be obtained otherwise.

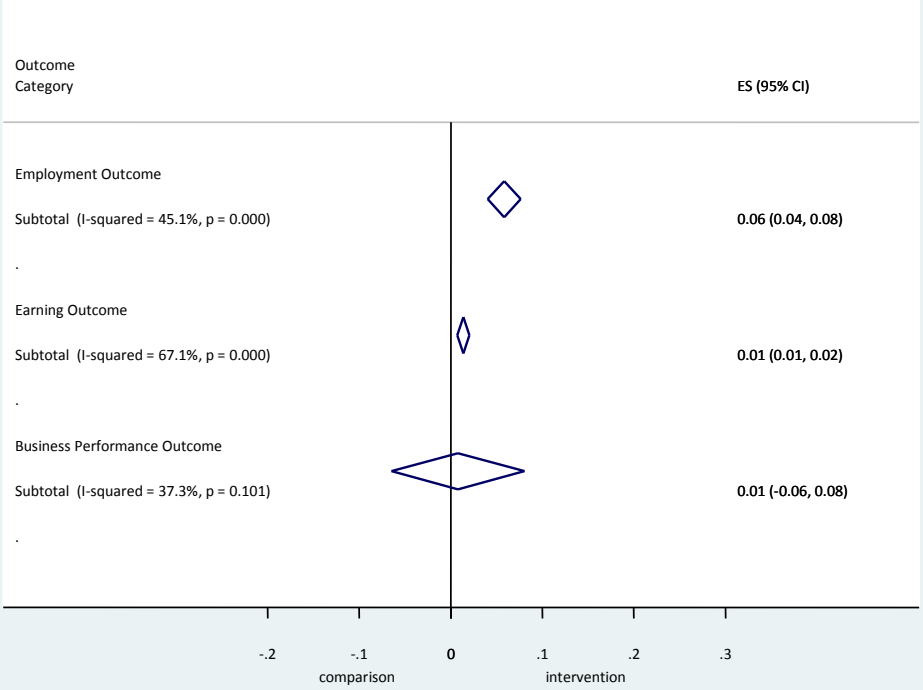
This section compares results using the entire sample of studies (including all imputed values) with a restricted sample excluding all studies where SMDs (or their standard error) could not be computed without these assumptions.[37]

Figure 31 replicates the forest plot displayed in Figure 6, displaying the summary SMDs by outcome category. Not imputing any missing information reduced the overall sample from 2,169 SMDs and 119 studies by almost half, to 1,116 SMDs (82 studies). The average SMD for employment outcomes increased, though the increase was not statistically significant. At the same time, the summary effect size (i.e., SMDs) for earnings outcomes was significantly reduced to 0.01; leading to an (insignificant) reduction in the overall SMD of youth employment interventions across outcomes.

Figure 32, Figure 33 and Figure 34 replicate the forest plots in the moderator analysis by main intervention category (Section 4.3.3.2) in the limited non-imputation sample. The sample size dropped from 1,312 effect sizes (105 studies) to 682 effect sizes (69 studies) for employment outcomes and from 661 effect sizes (90 studies) to 279 effect sizes (4 studies) for earnings outcomes. The number of business performance outcomes was reduced to 153 effect sizes (11 studies).

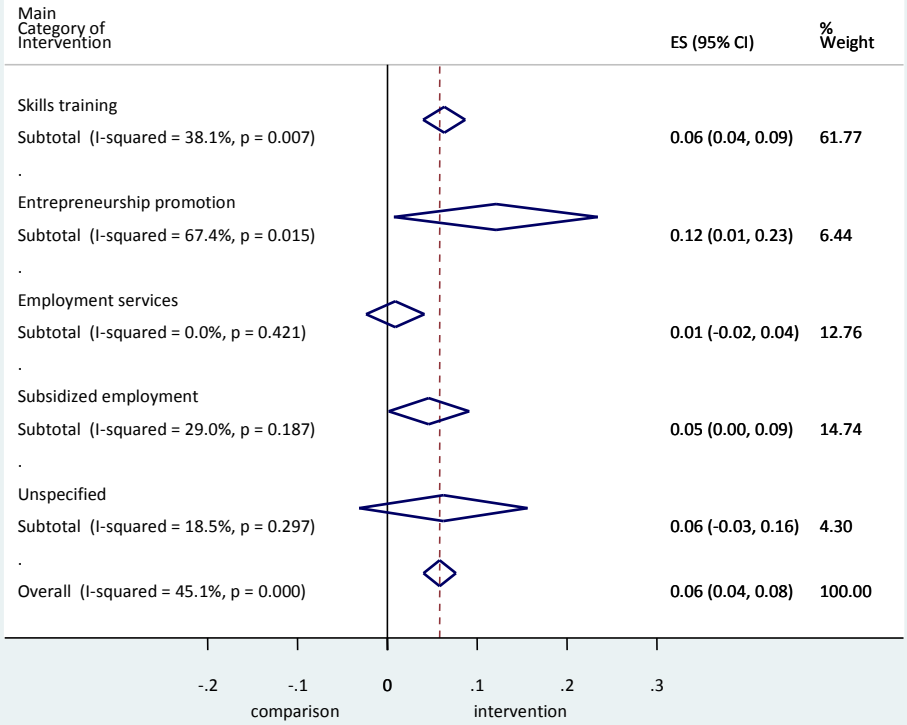
[37] In addition, the team performed a within-study check by comparing SMDs computed using the Borenstein, Hedges, Higgins and Rothstein (2009) formula to alternative methods of computation for studies where sufficient data was available. Since effect sizes were very similar under various computation methods, that the Borenstein formula appeared to deliver an adequate approximation of the true SMD in other studies as well.

Figure 31: Forest plot of all outcomes by outcome category without imputations



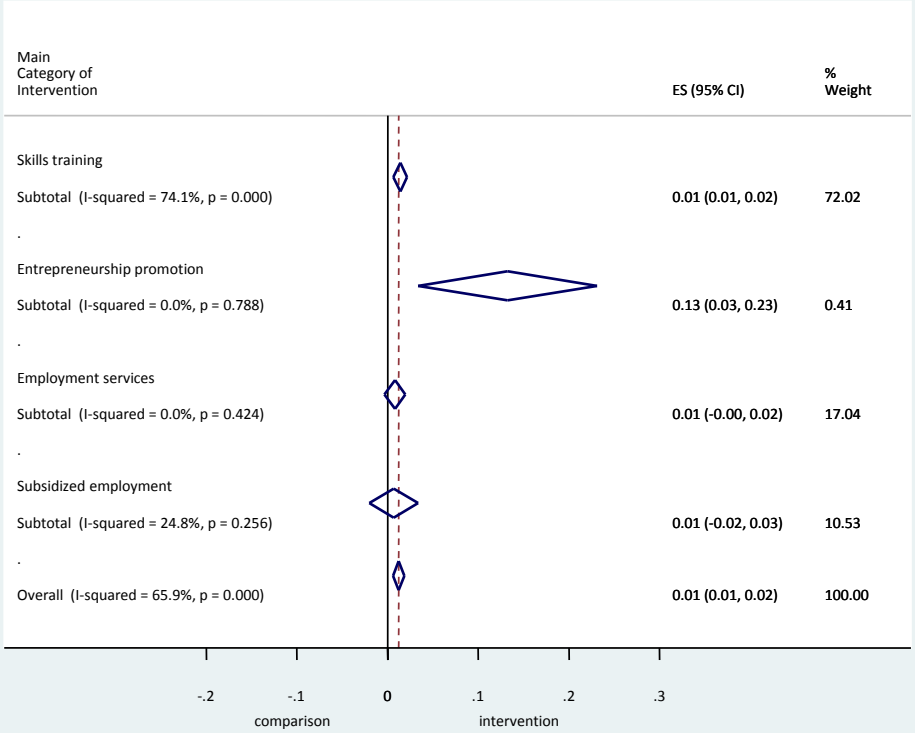
No. of SMDs/Studies: Total: 1116/82
Employment Outcome: 682/69. Earning Outcome: 281/48. Business Performance Outcome: 153/11.
Sub-group dropped (<4 independent studies): .
Note: Imputation: none, SMD limit = .75, SMD S.E. limit = .75
Note: The WINGS programme in Uganda (Blattman et al., 2014) (0.82 SMD; 95% CI = 0.13, 1.50), the PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18), and the Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceed the SMDs limit of 0.75 and were therefore not included in the analysis for the forest plot above.

Figure 32: Summary forest plot of employment outcomes by main category of intervention without imputations



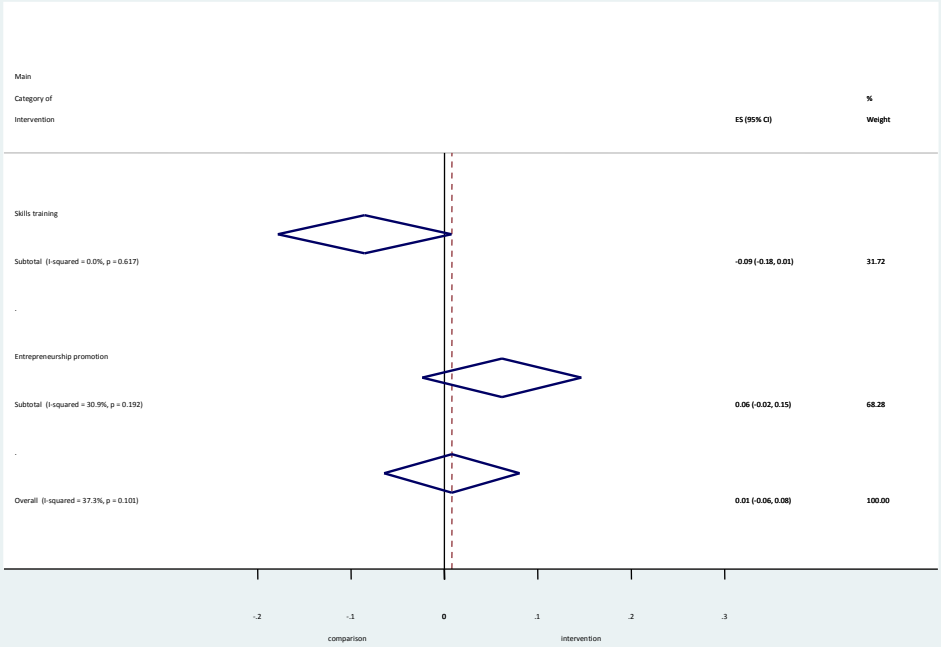
No. of SMDs/Studies: Total: 682/69
Skills training: 467/44. Entrepreneurship promotion: 28/5. Employment services: 63/6. Subsidized employment: 83/9. Unspecified: 41/5.
Sub-group dropped (<4 independent studies): .
Note: Imputation: none, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 33: Summary forest plot of earnings outcomes by main category of intervention without imputations



No. of SMDs/Studies: Total: 279/46
Skills training: 196/30. Entrepreneurship promotion: 29/6. Employment services: 26/5. Subsidized employment: 28/5.
Sub-group dropped (<4 independent studies): Unspecified.
Note: Imputation: none, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 34: Summary forest plot of business performance outcomes by main category of intervention without imputations



No. of SMDs/Studies: Total: 153/11
Skills training: 7/4. Entrepreneurship promotion: 146/7.
Sub-group dropped (<4 independent studies): .
Note: Imputation: none, SMD limit = .75, SMD S.E. limit = .75
Note: The WINGS programme in Uganda (Blattman et al., 2014) (0.82 SMD; 95% CI = 0.13, 1.50) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Despite reducing the sample size significantly, the basic results regarding the effectiveness of different intervention types held in the smaller sample. In fact, results and average effect sizes for employment and business performance outcomes were very similar to the main results that included all imputed values. Only in the case of earnings/income outcomes, was the average impact of skills training significantly reduced (the confidence intervals for skills training in the upper and lower panel did not overlap). Also, the precision of the estimate for entrepreneurship promotion intervention was somewhat reduced (i.e., had a larger confidence interval).

Similarly, results of the meta-regression analysis appeared robust against the method of imputation. Columns (2) and (6) in Table 39show specifications (2) and (4) from the tables in Section 4.3.4, estimated using the non-imputed samples. Again, the number of observations was only about half of the imputed sample size for both model specifications. Similarly, the number of studies was reduced from 100 studies (51 studies in the extended specification 4) to 69 studies (36 studies). Nonetheless, the direction and significance of most correlates appeared stable. Interestingly, some point estimates actually became significant only in the restricted sample (e.g., the indicator variable for male and younger participants). Since changes in the correlation of the estimates appeared largely at random, it was clear that reporting quality was not systematically related to any of the covariates (which could have introduced a bias into the estimates).

4.3.5.2 Assumptions in the analysis

In the main meta-analysis, the team applied a procedure to remove implausibly large or influential estimates (cf. Section 3.4.2). Although meta-analysts would not want to erroneously exclude relevant estimates, balance demands that no single estimate, especially one among hundreds, should determine how an entire research literature is viewed or understood (Stanley & Doucouliagos, 2015).

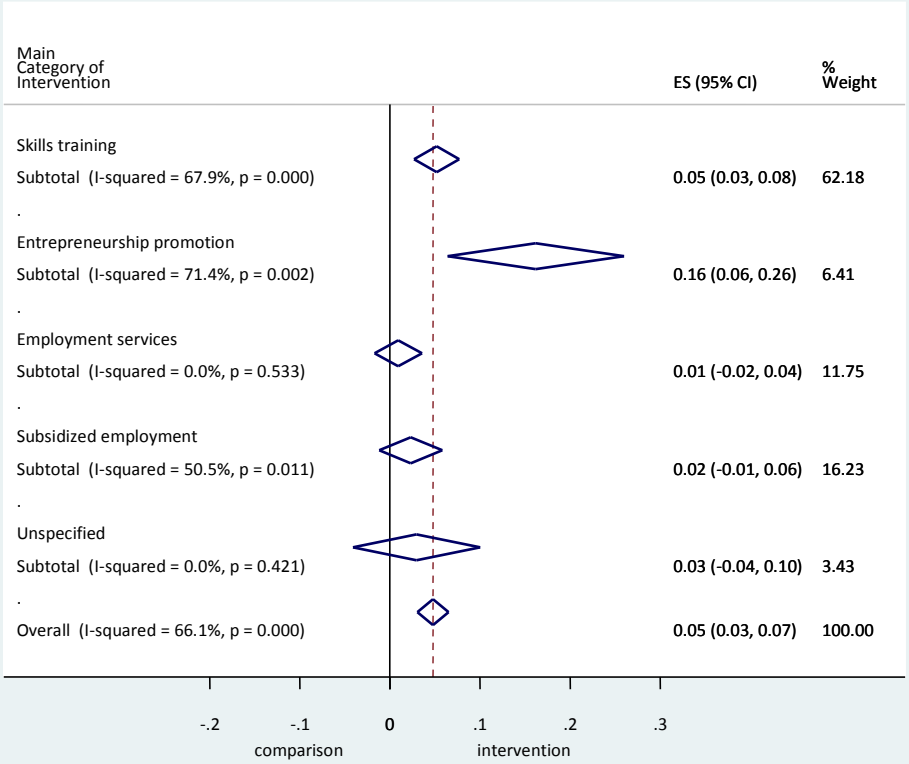
As described in the respective sections, the team first winsorized the highest and lowest 1 per cent of coded SMD effect sizes estimates. Generally, this affected the (unweighted) mean SMD and its standard deviation only marginally. Subsequently any observations with an SMD or an SMD standard error of more than 0.75 were dropped. In the full sample, roughly 30 SMDs from 4 studies were excluded; most of these stemming from sub-group analysis and therefore all but one study was retained in the sample. This section tests whether the results are robust against these assumptions. The full results for all tests in the report are not displayed but those that appeared to be of major importance are highlighted. For example, the team also tested whether winsorizing at the 5 per cent level (instead of 1 per cent) altered the results, but could not find any definitive evidence and therefore this issue is not discussed.

The upper panel of Figure 35, Figure 36 and Figure 37 again replicates the results of the main moderator analysis regarding the main category of intervention but this time without winsorizing the data and only dropping outliers with an SMD or standard error above three (effectively not dropping any observations).

In fact, since very few observations were dropped, the results changed marginally and even confidence intervals did not increase as much as one might have expected. Correspondingly, the team found that results from the meta-regression model were not affected by the decisions to censor the specified data and so this robustness check is not discussed further in the current report.

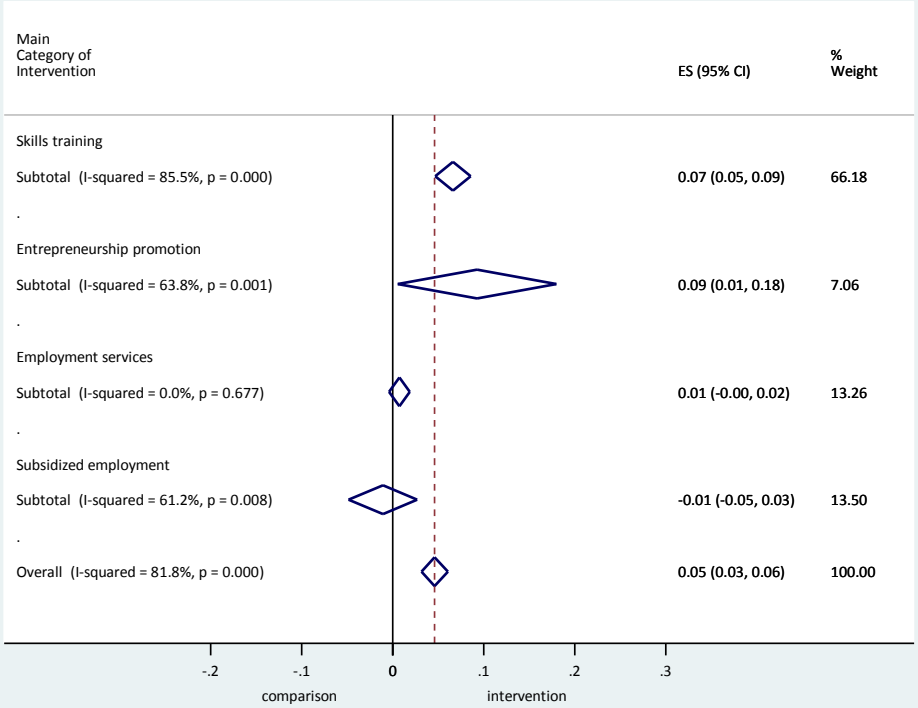
Tests were also conducted on whether the results were robust to the level of aggregation of effect sizes before synthesizing results based on the random-effects meta-analysis. Specifically, the team checked whether aggregating effect size across all studies of one intervention, rather than only aggregating studies using the same data set, made a difference. The latter increased the number of observations in the full analysis (based on all outcome variables) from 100 interventions to 120 individual studies. Similarly, the reviewers tested whether the level of cluster in the meta-regression model significantly affected results but found no evidence that the level of first-step aggregation or cluster significantly altered results.

Figure 35: Robustness check: Summary forest plot of employment outcomes by main category of intervention, including outliers



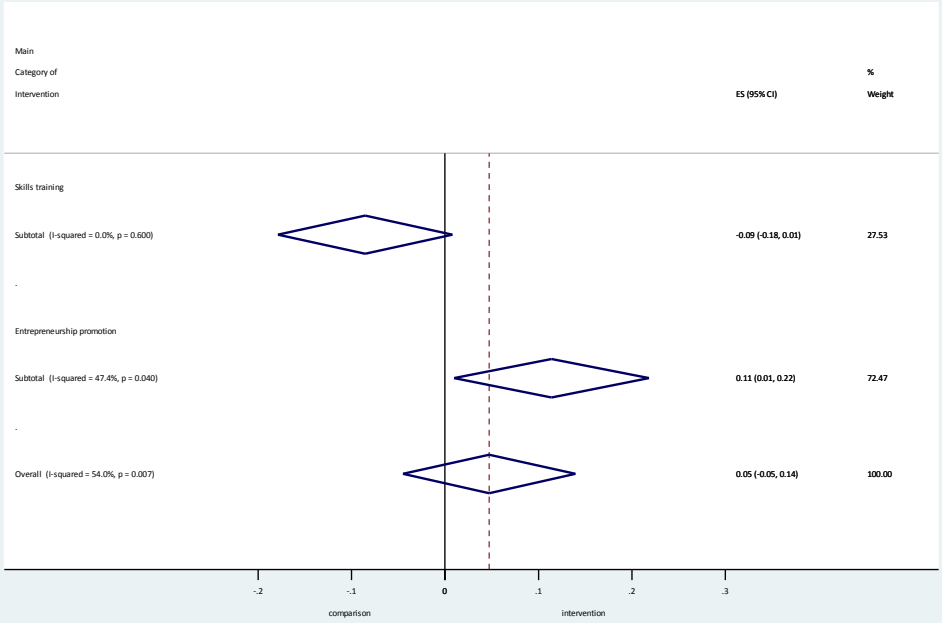
No. of SMDs/Studies: Total: 1350/106
Skills training: 922/68. Entrepreneurship promotion: 43/7. Employment services: 104/10. Subsidized employment: 194/16. Unspecified: 87/5.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = 2, SMD S.E. limit = 2

Figure 36: Robustness check: Summary forest plot of earnings outcomes by main category of intervention, including outliers



No. of SMDs/Studies: Total: 643/91
Skills training: 500/62. Entrepreneurship promotion: 50/12. Employment services: 36/8. Subsidized employment: 57/9.
Sub-group dropped (<4 independent studies): Unspecified.
Note: Imputation: full, SMD limit = 2, SMD S.E. limit = 2

Figure 37: Robustness check: Summary forest plot of business performance outcomes by main category of intervention, including outliers



No. of SMDs/Studies: Total: 174/15
Skills training: 7/4. Entrepreneurship promotion: 167/11.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = 2, SMD S.E. limit = 2

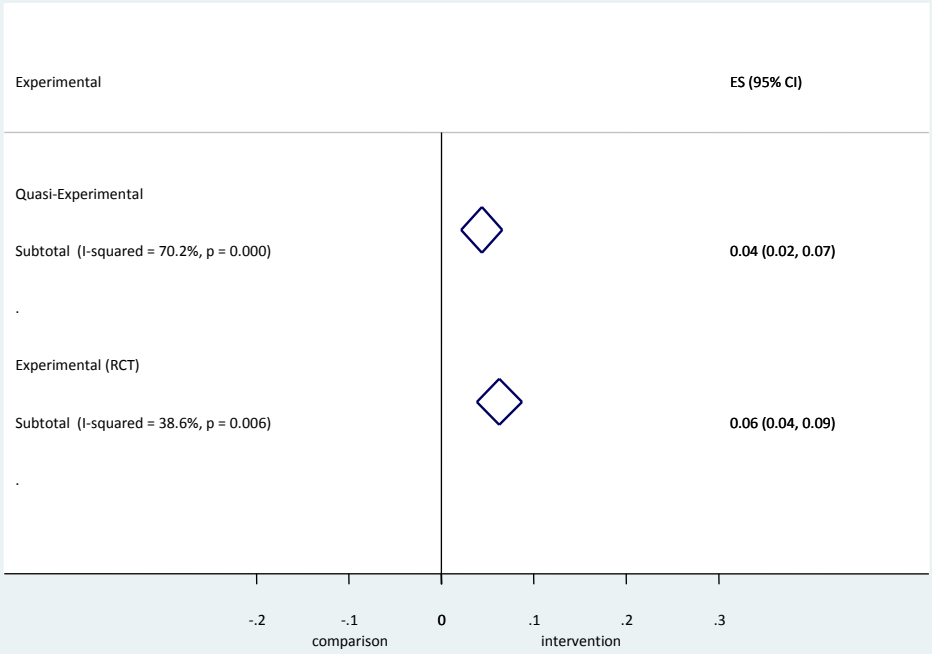
In summary, the team tested the robustness of the results towards different decisions made in the process of compiling and analysing the data, such as the imputation of missing information or the handling of statistical outliers. Since all the sensitivity analysis yielded similar results to the main analysis, the team can be quite confident that the findings reported in sections 4.3.2 to 4.3.4 were not influenced by the method of analysis.

4.3.5.3 Research design

This section tests whether the results depended on the applied evaluation design. In the combined meta-analysis, studies of randomized control trials and quasi-experimental evaluation approaches were pooled but more rigorous evaluation designs might have systematically yielded different effect sizes than less robust evaluation designs.

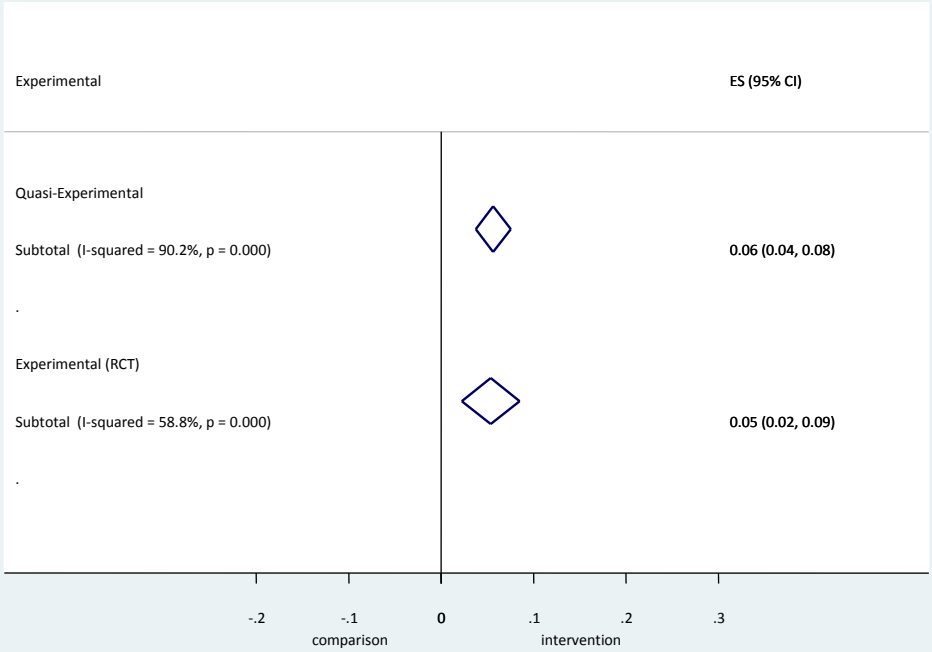
Figure 38 and Figure 39 provide a moderator analysis by research design for both employment and earnings outcomes. In contrast to expectations, experimental studies actually produced larger effect sizes in both cases, but the difference between experimental and quasi-experimental was not statistically significant at the 5 per cent level in either case. Based on the one-way random effects ANOVA model, the team was able to rule out a systematic difference on average between effect sizes generated from experimental and quasi-experimental studies.

Figure 38: Summary forest plot of employment outcomes by research design



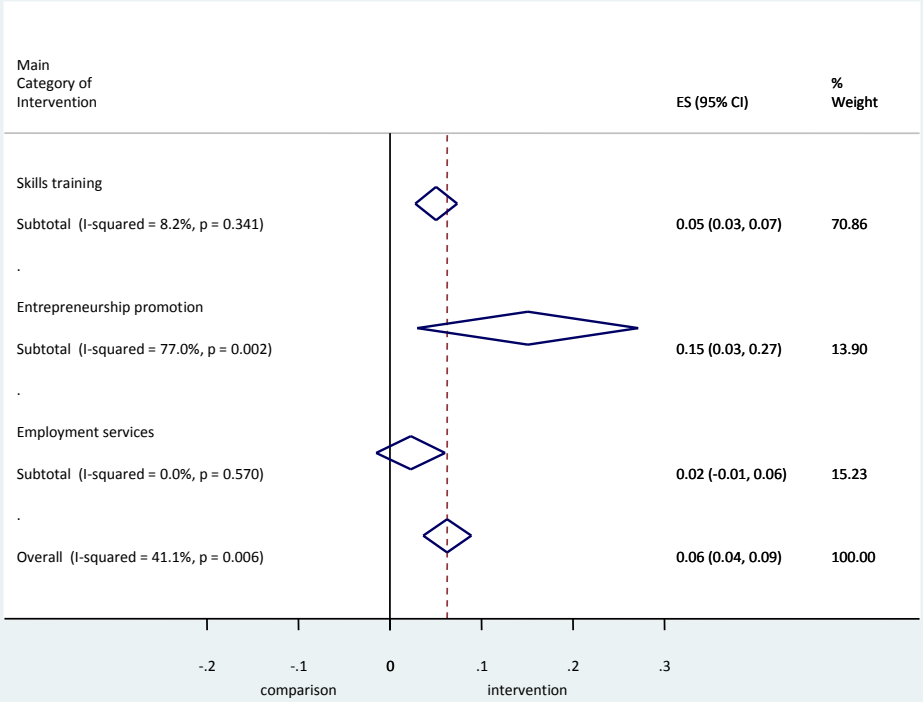
No. of SMDs/Studies: Total: 1378/109
Quasi-Experimental: 702/67. Experimental (RCT): 676/43.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 39: Summary forest plot of earnings outcomes by research design



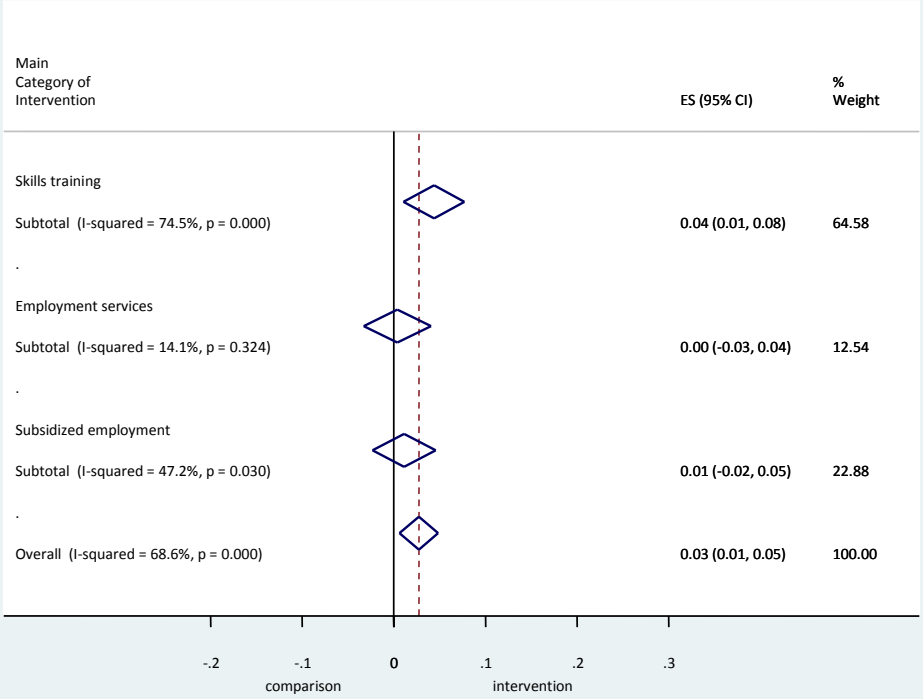
No. of SMDs/Studies: Total: 671/92
Quasi-Experimental: 246/51. Experimental (RCT): 425/42.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 40: Summary forest plot of employment outcomes by main category of intervention for experiments



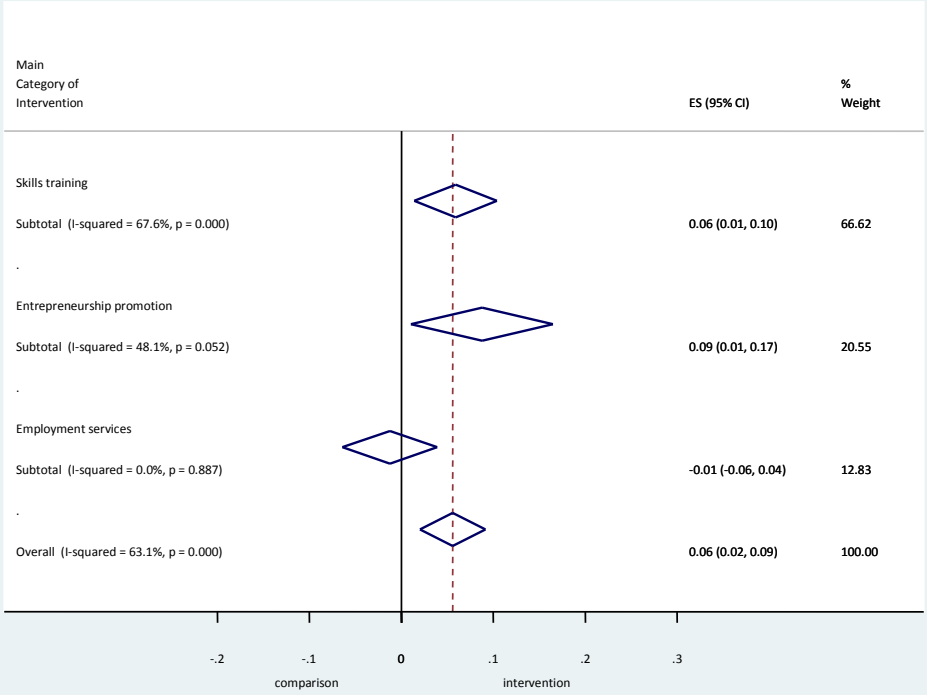
No. of SMDs/Studies: Total: 597/37
Skills training: 506/28. Entrepreneurship promotion: 38/5. Employment services: 53/4.
Sub-group dropped (<4 independent studies): Subsidized employment Unspecified.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75

Figure 41: Summary forest plot of employment outcomes by main category of intervention for quasi-experiments



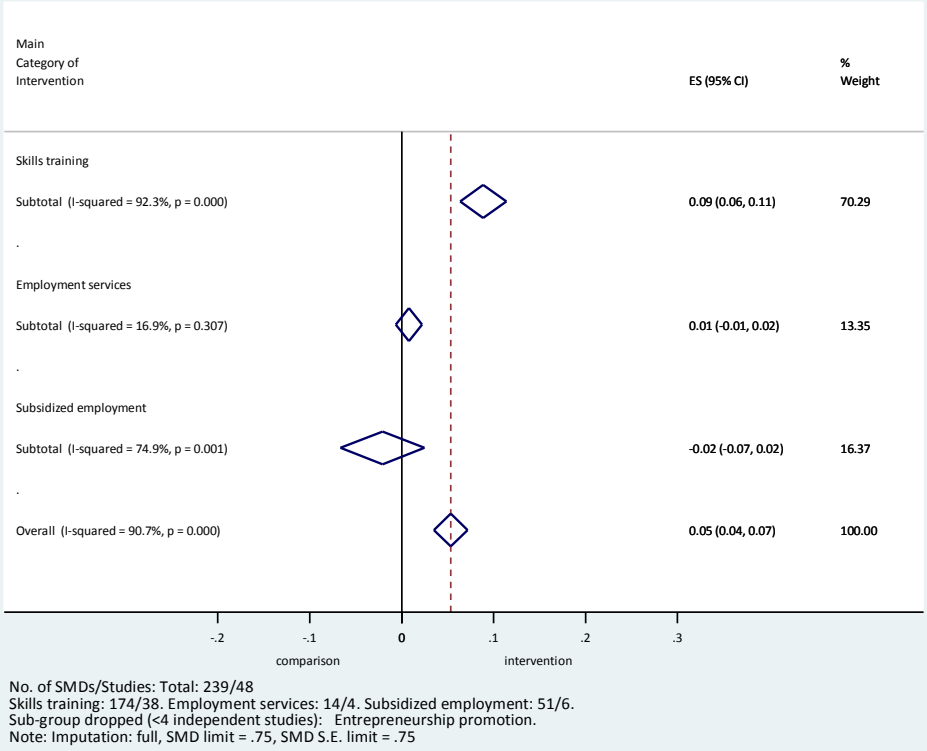
No. of SMDs/Studies: Total: 625/59
Skills training: 398/40. Employment services: 51/6. Subsidized employment: 176/13.
Sub-group dropped (<4 independent studies): Entrepreneurship promotion Unspecified.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 42: Summary forest plot of earnings outcomes by main category of intervention for experiments



No. of SMDs/Studies: Total: 387/36
Skills training: 321/23. Entrepreneurship promotion: 44/9. Employment services: 22/4.
Sub-group dropped (<4 independent studies): Subsidized employment Unspecified.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 43: Summary forest plot of income outcomes by main category of intervention for quasi-experiments



No. of SMDs/Studies: Total: 239/48
Skills training: 174/38. Employment services: 14/4. Subsidized employment: 51/6.
Sub-group dropped (<4 independent studies): Entrepreneurship promotion.
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75

However, this may differ according to the category of intervention. It is, for example, plausible that a particular intervention type consistently displays higher results when evaluated through experimental studies than through quasi-experimental ones. Therefore, the team tested whether the basic results on the effectiveness of different intervention types held when considering only evidence from experimental studies, which was arguably more reliable than quasi-experimental results. In a first descriptive step, Figure 40, Figure 41, Figure 42 and Figure 43 replicate the moderator analysis by main category of intervention for employment and earnings outcomes separately for experimental and quasi-experimental studies.

The evidence for entrepreneurship interventions was entirely based on empirical research and, hence, the review's results pertain. In contrast, all studies of subsidized employment interventions were derived from quasi-experimental approaches. To some degree, this correlation between intervention type and research design may have confounded the analysis. The spectrum of research designs employed for evaluating skills training interventions was more mixed. But regardless of evaluation design, skills training interventions appear to have been the most successful intervention type, along with entrepreneurship programmes. (The difference in SMDs between entrepreneurship and skills training interventions was still not statistically significant.) Unfortunately, studies of interventions classified as unspecified were dropped from the analysis, since fewer than four interventions were found which could have been classed as falling within either sub-group (experimental vs. quasi-experimental).

As with all univariate analysis, one issue was that the difference between experimental and quasi-experimental studies observed in forest plots might also have been driven by other factors (such as the fact that the majority of experiments were conducted in low- and middle-income countries, generally yielding larger effect sizes). Based on the univariate analysis, it is not possible to state with certainty that the aggregate effect size was actually downward biased by including evidence from quasi-experimental studies. But this, in fact, appears to have been the case: Controlling for other factors, which may be correlated with the average SMD in the multivariate meta-regression, the review team found that reported effect sizes from experimental studies were, on average, significantly smaller in magnitude (Table 36).

Given these caveats, the robustness of the meta-regression model was examined by removing effect sizes from quasi-experimental research from the data set. The results are reported in columns (3) and (7) of Table 39. Dropping all estimates from quasi-experimental studies still left roughly 60–70 per cent of effect sizes from the overall sample. Comparing point estimates to the full model in columns (1) and (5) showed that almost all of the meta-analysis findings were supported when considering only evidence from experimental studies. Even in the fully specified model, which suffered from the loss in statistical power (columns (5) and (7)), the direction of the covariate correlates was stable and only a few point estimates became insignificant.

Contrary to expectations, experimental studies actually reported larger effect sizes, on average. However, this seemed to contradict other factors that were correlated with the research design, most notably the type of intervention and country income status. After controlling for these aspects in the multivariate analysis, the review team found that experimental research in general leads to smaller effect size estimates.

Nonetheless, the meta-analysis results appeared robust to the type of evaluation design and the main findings were corroborated by rigorous evidence from experimental studies.

4.3.6 Analysis of small-sample bias and publication bias

This section uses funnel plots and Egger's tests to check whether there was any indication of publication bias in the sample of studies. Figure 44 and Figure 45 present funnel plots for the entire sample (including all outcomes and all sub-groups). The figure displays plots of the effect size (SMD) on the horizontal axis and the standard error of the effect size (SE SMD) on the vertical axis. In Figure 44, effect sizes are aggregated at the study level, and each dot represents an individual study. In Figure 45, the data is entirely disaggregated, meaning that each dot represents one effect size estimate. The solid line crosses the horizontal axis at the overall average fixed effect estimate.

Figure 44: Funnel plot of all outcomes and sub-groups, aggregated at study level

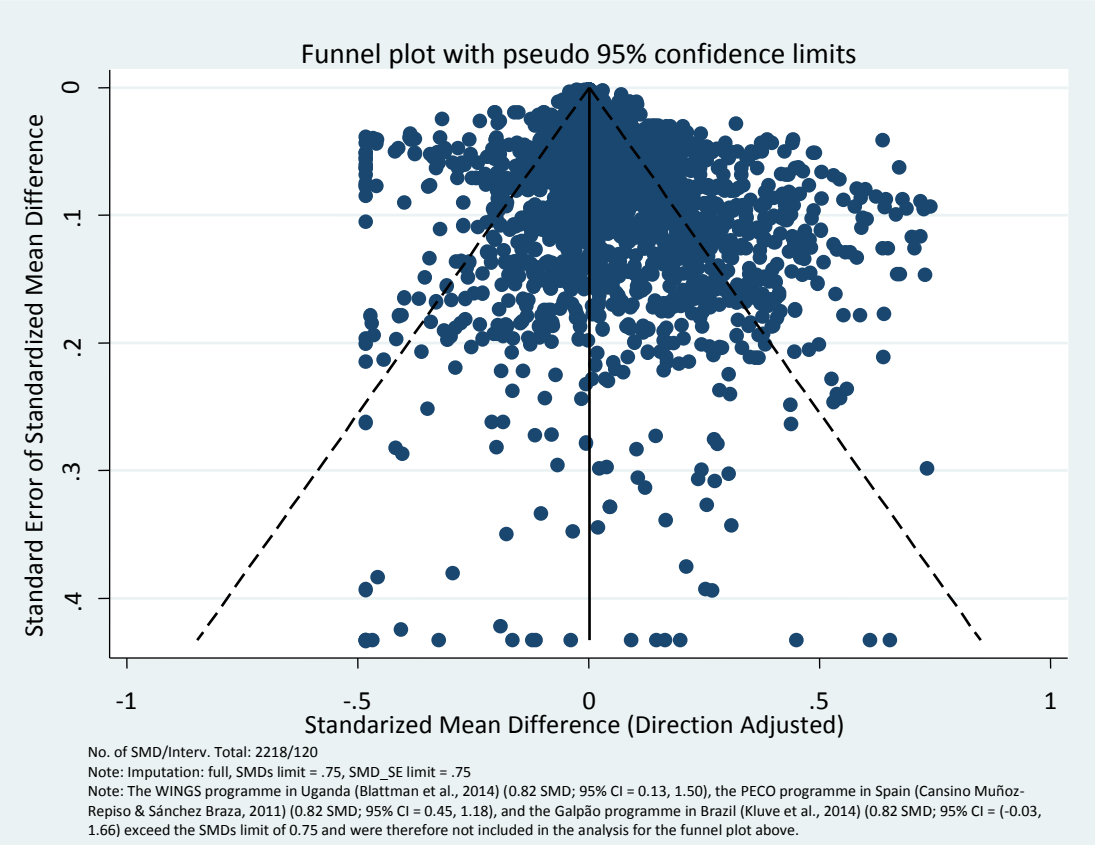
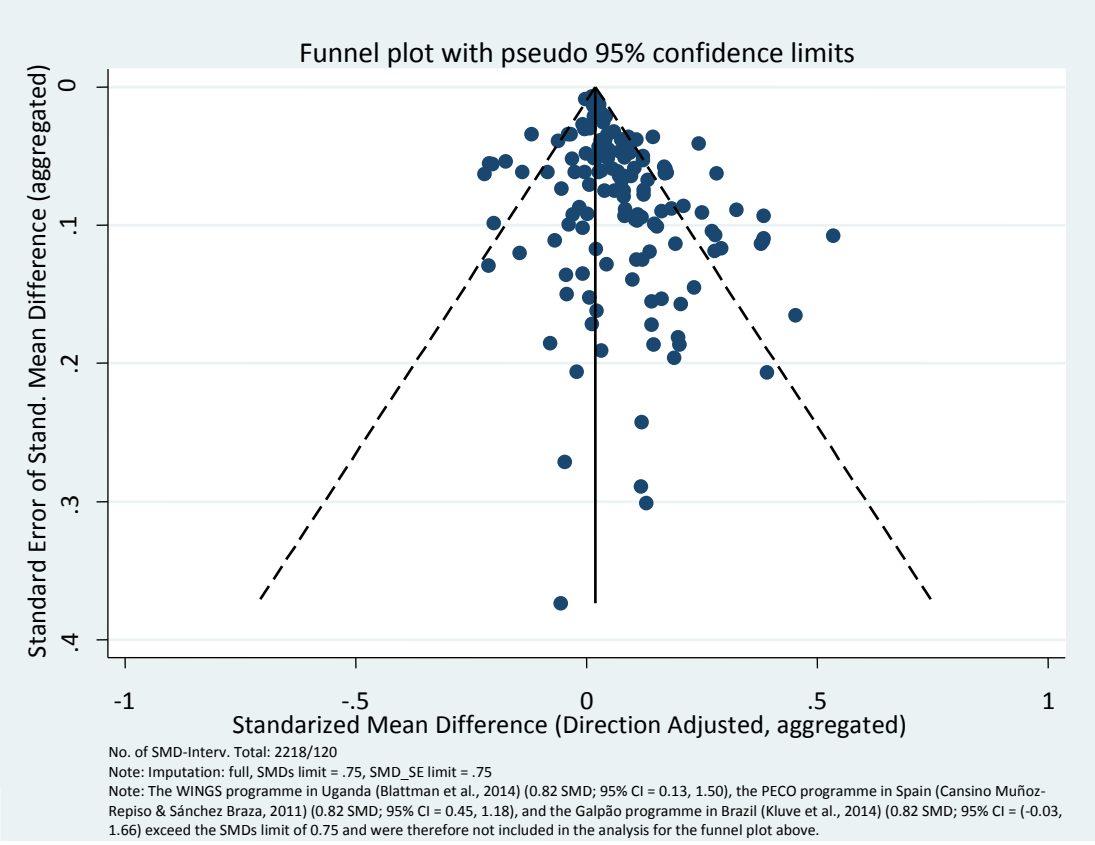


Figure 45: Funnel plot of all outcomes and sub-groups, disaggregated (on effect size estimate level)



Although most of the dots (studies) are spread around the solid line and within the triangular area (indicating the 95 per cent confidence interval), a degree of tendency

towards the right is observable. These represent studies that reported positive effects (with a medium level of precision, as measured by the standard error). This slight asymmetry may be an indicator of publication bias.

Results presented in Table 40 from Egger's test for publication bias confirmed the visual indication: The coefficient of the variable bias was positive and statistically significant at the 5 per cent level.

Table 40: Egger's test for small-sample bias

		Outcomes		
	Full sample	Employment	Earnings/income	Business performance
	(1)	(2)	(3)	(4)
Slope	-0.002***	-0.007***	-0.001***	0.000
Standard error	(0.001)	(0.002)	(0.001)	(0.001)
Bias	0.868***	0.990***	0.982**	0.102
Standard error	(0.058)	(0.083)	(0.097)	(0.116)
n	2219	1379	671	169

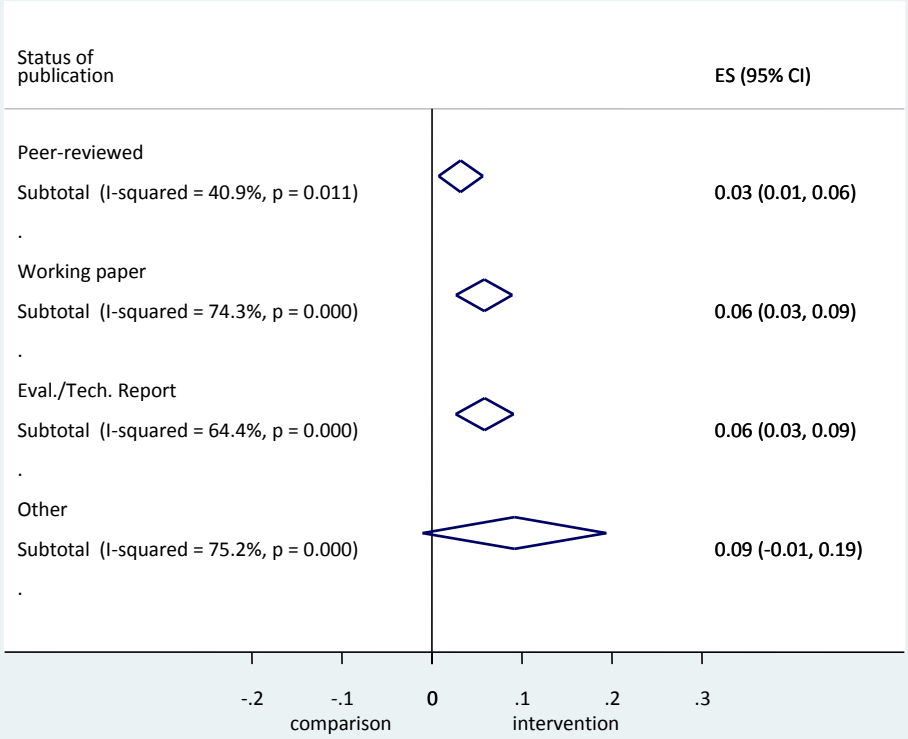
Note: *, ** and *** denote statistical significance at 10 per cent, 5 per cent and 1 per cent level of significance respectively.

As in the previous sections, the analysis is further disaggregated by outcome categories. Figure 61, Figure 62 and Figure 63 in Section 9.2 of the Appendix show funnel plots (aggregated at the study level) for the three outcome variables separately.[38] As can be seen from these figures, effect sizes of earnings and income outcomes appeared most strongly skewed to the right. This was also confirmed when performing Egger's test for each outcome category separately. For business performance outcomes, Egger's test was not significant at the 5 per cent level.

This potential publication bias was accounted for in the multivariate meta-regression model using the procedure described in Doucouliagos and Stanley's study in 2009: The authors argued that including the standard error of the SMD in the random-effect model would account for the potential effect of publication bias and the resulting coefficient estimate would provide an indication of the magnitude (and significance) of the effect. Following this approach, the team found a clear indication of selection for statistically positive results: In Table 36, Table 37 and Table 38, the point estimate was consistently positive and statistically significant. In addition, the summary effect estimate (represented by the constant) in the model, which pools all outcomes, turned non-significant when accounting for publication bias using this approach. This seemed to be largely driven by the negative (insignificant) results on business performance outcomes, while the summary effect on employment and earnings was still significant even accounting for publication bias.

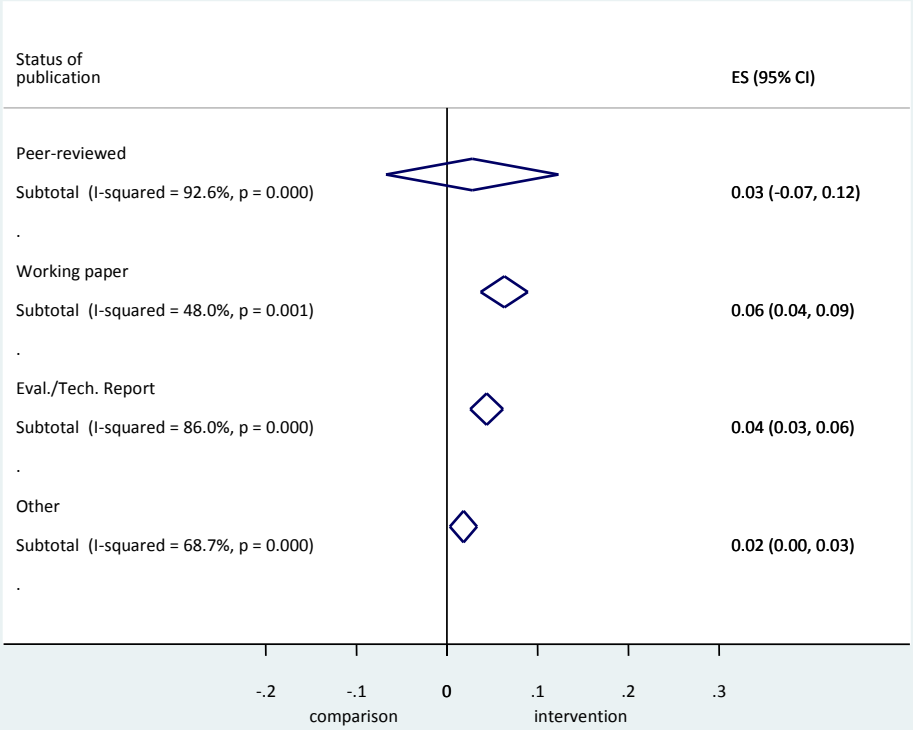
[38] Note that some dots are not reflected in the overall forest plots for all outcome variables since effect sizes were aggregated across outcome categories within studies before plotting them.

Figure 46: Summary forest plot of employment outcomes by publication status



No. of SMDs/Studies: Total: 1378/109
Peer-reviewed: 230/29. Working paper: 494/41. Eval./Tech. Report: 492/39. Other: 162/11.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 47: Summary forest plot of earnings outcomes by publication status



No. of SMDs/Studies: Total: 671/92
Peer-reviewed: 145/15. Working paper: 171/34. Eval./Tech. Report: 270/35. Other: 85/15.
Sub-group dropped (<4 independent studies): .
Note: Imputation: full, SMD limit = .75, SMD S.E. limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

In addition to the above test for publication bias, the team also tested whether reported effect sizes differed between peer-reviewed articles, working papers (some of which were unpublished at the time of publication search), evaluation reports/technical reports and other types of reports (such as books and dissertations). No statistically significant differences in average effect sizes by publication status were found, as can be seen by the forest plot shown in Figure 46. These results held when the analysis was disaggregated by intervention type or outcome category (not reported). Similarly, the dummy for publication status (peer-reviewed) in the multivariate results did not provide a clear picture.

Funnel plots and Egger's test indicated some publication bias towards studies showing positive effects of youth employment interventions on labour market outcomes. Using the procedure proposed in Doucouliagos and Stanley (2009), the reviewers accounted for publication bias in the multivariate meta-regression model. While the overall effect was significantly reduced, youth employment interventions still showed a significant positive effect on employment and earnings outcomes. Nonetheless, the team concluded that the summary effect size of youth employment outcomes probably represented an upper bound for the true impact of these interventions. At the same time, no correlation of reported effect sizes with publication status was detected.

5 Discussion

5.1 SUMMARY OF MAIN RESULTS

5.1.1 The why, the what and the how of the systematic review

The 2030 Agenda for Sustainable Development has placed the importance and urgency of achieving full and productive employment and decent work for all squarely at the centre of the new development vision, with youth as a top priority.

Why are measures to achieve these objectives so urgently needed? The youth employment crisis has had long-lasting effects on labour markets and economies across the globe. Of the estimated 200 million unemployed people today, about 37 per cent – more than 73 million – are between the ages of 15 and 24. Youth unemployment stands at a much higher level than the average unemployment rate for adults, in some cases three times as high, while two out of five young people in the labour force are either working but poor or unemployed (ILO, 2015a).

High levels of youth unemployment and underemployment have fuelled discontent and increased detachment from the labour market. Youth face today longer and more insecure pathways to stable employment. ILO school-to-work transition surveys conducted in 28 low- and middle-income countries show that almost eight out of ten young workers work in the informal economy, either as wage earners or necessity-driven self-employed (ILO, 2015a).

To achieve full and productive employment and decent work for youth, it will be necessary for governments, the private sector, social partners, development partners, non-governmental organisations (NGOs), and all youth employment practitioners to coordinate and deliver effective interventions – rooted in solid and rigorous evidence about what works. Resources are limited, so decision makers must ask themselves not only whether they are implementing a youth employment programme right but also whether they are implementing the right programme vis-à-vis the challenge, the context and the means.

To support informed decision making, the systematic review examined the existing evidence on the effectiveness of interventions that aimed to improve the labour market outcomes of youth. The review relied on a structured and comprehensive search that allowed the identification and assessment of all relevant impact evaluation studies carried out worldwide between 1990 and 2014 across the following intervention types:

- **Training and skills development**, which comprises programmes outside the formal education system (and therefore does not consider Technical and Vocational Education (TVE) programmes) that offer skills training to young people in order to improve their employability and facilitate their transition into the labour market.

- **Entrepreneurship promotion**, aiming to provide entrepreneurial skills as well as physical, financial and social capital for youth becoming self-employed and starting a business and for those seeking to expand and grow their businesses.

- **Employment services**, delivering job counselling, job-search assistance and/or mentoring services, which are often complemented by job placement and technical or financial assistance.

- **Subsidized employment**, which are government efforts to boost labour demand and incentivize hiring and meaningful work experience for young women and men. This type of interventions include wage subsidy programmes and labour-intensive public employment programmes.

The key labour market outcomes considered were the post-treatment measures of employment, earnings, and business performance.

- **Employment outcomes** include employment and/or unemployment probabilities, participation rates, hours worked, unemployment duration and quality of employment.

- **Earnings outcomes** include reported earnings and income, household income, consumption, and salary and/or wage.

- **Business performance outcomes** include profits, sales, number of employees and jobs created, capital and investment, business creation and business survival.

The key research questions were:

1. What impact do youth employment interventions have on labour market outcomes of youth?

2. Which of these interventions are most effective?

In the process of understanding what works, the review also focused on the way in which interventions work, relying on prior theories of change for the selected intervention types as well as on observed programme characteristics reported in the studies.

A strict search and screening process led to the identification of 113 reports that were considered of adequate content and methodological rigour to be included in the analysis. The 113 reports represented 107 interventions (Table 41). The evidence base spanned 31 countries and covered 55 skills training interventions, 15 entrepreneurship promotion interventions, ten employment services interventions, and 21 subsidized employment interventions. There were six interventions for which no clear main category of intervention could be established. A large share of the evidence derived from recent publications, with nearly half of the sample produced after 2010. Evaluation designs vary, with 47 per cent of reports relying on experimental designs, 10 per cent in natural experiments, and 44 per cent in quasi-experimental evaluation. Many of the most recent studies were experimental evaluations of interventions implemented in low- and middle-income countries, notably from Africa, Latin America, and the Caribbean.

Table 41: Summary table of main characteristics of studies and interventions included in the systematic review

		Reports	%			Interventions	%
Total number of reports		**113**		**Total number of interventions**		**107**	
(A)	**Country income level**			**(E)**	**Country income level**		
	High-income country	65	58		High-income country	60	56
	Low- and middle-income country	48	42		Low- and middle-income country	47	44
(B)	**Outcome category**			**(F)**	**Intervention region**		
	Employment	98	87		OECD	56	52
	Earnings	91	81		Latin America and Caribbean	22	21
	Business performance	10	9		Sub-Sahara Africa	15	14
					Middle East and North Africa	6	6
(C)	**Evaluation design**				Europe and Central Asia	4	4
	Experimental	53	47		South Asia	4	4
	Natural experiment	11	10				
	Quasi-experimental	50	44	**(G)**	**Main category**		
					Skills training	55	51
(D)	**Main intervention**				Entrepreneurship promotion	15	14
	Skills training	74	65		Employment services	10	9
	Entrepreneurship promotion	12	11		Subsidized employment	21	20
	Employment services	11	10		Unspecified	6	6
	Subsidized employment	17	15				
	Unspecified	9	8				

Note: Reports may not be exclusive across the different typologies in this table, e.g., one study may estimate multiple outcomes or look into more than one intervention type.

The comprehensive systematic search led to the identification and coding of a total of 3,629 treatment effect estimates. These estimates along with further information reported and/or retrieved from authors of the primary studies allowed the computation of 1,402 standardized mean difference (SMD) effect sizes.[39] The SMD captures the relative magnitude of the treatment effect in a way that is dimensionless and hence comparable across outcomes and studies. It is the ratio of the treatment effect for a specific outcome relative to the standard deviation of that outcome within the evaluation sample used to estimate the treatment effect. If the information required to compute SMDs could not be obtained from the reports (or the authors), standard deviations were approximated using the formulae suggested in Borenstein et al. (2009), leading to a total of 2,259 SMDs for the basis of the analysis.

SMDs were computed for both continuous and dichotomous outcome variables (e.g., income and employment probability, respectively), compared across different outcome constructs, and summarized within and across reports to one effect size per outcome for each study. Random-effects meta-analysis methods were then employed to synthesize and compare effect sizes reported in the primary studies. Subsequently, multivariate meta-regression models were estimated and information about intervention-level, study-level and country-level characteristics were included to assess factors associated with the magnitude of reported effect size estimates.

5.1.2 Main results

1. **The systematic review showed that investing in youth through active labour market measures does pay off.** Interventions to support young women and men in the labour market lead to positive outcomes, increasing their chances of finding or staying in employment and improving their income. The positive effect on employment and earnings was statistically significant with effect sizes measured by 0.04 and 0.05 SMDs, respectively, demonstrating the responsiveness of these outcomes to youth's exposure to active labour market programmes (ALMPs). With substantially less evidence, the effect on business performance outcomes was not statistically significant at 0.03 SMD; however, when only entrepreneurship promotion interventions were considered, its impact was larger and significant, at 0.10 SMD (see SMD figures in Table 42).

[39] Effect size is a generic term used to describe the estimated treatment effect for a study. This treatment effect is the observed relationship between an intervention and an outcome. In order to compare effect sizes across studies and outcome constructs, this systematic review used a meta-analysis to synthesize the data extracted from primary studies. The SMD is used as a summary statistic in meta-analysis to combine results from studies which used different ways of measuring the same outcome. An SMD of zero indicates that the intervention, on average, resulted in an equivalent effect for the treatment group and the (comparison) group which did not receive the treatment; whereas an SMD greater than zero indicates the degree to which, on average, the treatment group had a better outcome than the comparison group.

Table 42: Summary of results: SMDs and 95 per cent confidence intervals across outcomes and intervention types

Intervention types	Employment outcomes	Earnings outcomes	Business outcomes
Skills training	0.05 (0.02,0.07)	0.07 (0.05,0.08)	n/a
Entrepreneurship promotion	0.16 (0.06,0.26)	0.09 (0.01,0.18)	0.10 (0.00,0.19)
Employment services	0.01 (-0.02,0.04)	0.01 (-0.00,0.02)	n/a
Subsidized employment	0.02 (-0.01,0.06)	-0.01 (-0.05,0.03)	n/a
Unspecified	0.03 (-0.04,0.10)	n/a	n/a
Overall	**0.04 (0.03,0.06)**	**0.05 (0.03,0.06)**	**0.03 (-0.05,0.12)**

Note: n/a implies the study sample was too small (under four) to compute the effect size or there were no studies measuring the selected outcomes across the intervention types.

2. **It is 'however' an investment that needs time to grow.** The multivariate regression showed that there appears to be a time lag before impacts can be captured (Table 36). This is, changes in youth outcomes do not happen immediately after exposure to the intervention, they take time to materialize. The result applied in particular to studies from interventions in high-income countries and highlighted the importance of regarding youth employment programming as a long term investment. Moreover, the finding underscored the fact that changes in the labour market outcomes of youth do not happen overnight and it is therefore important for evaluations to measure changes at different points in time.

3. **Programme impacts conceal major contextual differences.** The systematic review and meta-analysis aimed to capture the overall effects of youth employment interventions based on global evidence from 1990 to 2014. Even after factoring in differences across interventions, their effects on labour market outcomes of youth were not consistent across studies. The review assessed numerous factors that correlated with reported effect sizes to different intervention results, from country context to programme and participants' characteristics. Tests for heterogeneity showed substantial variation in the effect size magnitude due to country income level, the design and implementation of the interventions and the profile of programme beneficiaries. These elements are further explored below.

4. **The underlying evidence base varies by country income level.** Intervention characteristics and research designs differ significantly between high-income and low- or middle-income countries. A large proportion of the evidence from high-income countries derived from quasi-experimental studies

of national programmes, implemented in collaboration with government organizations. In contrast, the evidence from low- and middle-income countries was predominantly based on experimental impact evaluations of rather small-scale, targeted interventions, which were often implemented by NGOs or international organizations.

5. **Impact is higher in low- or middle-income countries than in high-income countries.** Evaluation studies from low- or middle-income countries produced larger effect size estimates than studies conducted in high-income countries. The result holds for employment and earnings outcomes (Table 43) and after controlling for differences in research design and intervention characteristics. The studies pointed to a factual difference across country contexts: Being unemployed or unskilled in a high-income country – where labour demand is skill intensive – puts youth at a highly disadvantaged position vis-à-vis a cohort that is on average well educated. While ALMPs help these youth to reconnect to the labour market, they do not fully compensate for knowledge or skills not acquired earlier, in the education system. In lower income countries, with large cohorts of disadvantaged youth, marginal investments in skills and employment opportunities lead to larger changes in outcomes. This finding coincides with earlier reviews by Betcherman et al. (2007) and Fares and Puerto (2009).

Table 43: Summary of results: SMDs and 95 per cent confidence intervals across outcomes, intervention types, and country income level

Intervention types	Employment outcomes		Earnings outcomes		Business outcomes	
	HICs	LMICs	HICs	LMICs	HICs	LMICs
Skills training	0.04 (0.01,0.07)	0.06 (0.02,0.1)	0.02 (0,0.04)	0.12 (0.08,0.16)	n/a	-0.09 (-0.18,0.01)
Entrepreneurship promotion	n/a	0.18 (0.06,0.29)	n/a	0.14 (0.06,0.22)	n/a	0.15 (0.07,0.23)
Employment services	0.01 (-0.03,0.04)	n/a	0.01 (0,0.02)	n/a	n/a	n/a
Subsidized employment	0.00 (-0.03,0.04)	0.11 (0.04,0.18)	-0.02 (-0.07,0.03)	0.02 (-0.05,0.1)	n/a	n/a
Unspecified	0.03 (-0.04,0.1)	n/a	n/a	n/a	n/a	n/a

Notes: (i) HICs: High-income countries, LMIC: Low- and middle-income countries; (ii) n/a implies the study sample was too small (under four) to compute the effect size or there were no studies measuring the selected outcomes across the intervention types.

6. ***In low- and middle-income countries* entrepreneurship and skills training interventions offer the greatest impacts.** The evidence from low- and middle-income countries showed that youth employment interventions lead to a meaningful impact on both employment and earnings of youth. In particular, entrepreneurship and skills training interventions yield positive results, on average, especially in terms of income gains. This is an

important finding, which points to the merits of combining both supply- and demand-side interventions to support youth. It also provides tangible evidence about the effect of human capital investment. The effect of entrepreneurship promotion interventions should be interpreted with care because despite the large magnitude of impact, this intervention category also reports the largest confidence intervals across outcome maeasures. More primary studies are therefore needed to increase the accuracy of the finding.

7. ***In low-and middle-income countries* appropriate design features (e.g. Participant profiling, monitored programme Participation, and incentives of Participants and service providers) can boost the impact of subsidized employment measures.** Most evaluated subsidized employment programmes come from high-income countries. There was no evidence from either wage subsidies or public employment programmes implemented in low-income countries and limited evidence of wage subsidies in middle-income countries. With this caveat, the multivariate meta-regression analysis showed that once design features such as appropriate profiling, supervision and incentives for participation of youth and further incentives for services providers were accounted for, wage subsidy programmes led to significant positive effects on employment outcomes, even greater than skills training interventions.[40]

8. ***In low-and middle-income countries* comprehensive measures combining different programme components lead to better outcomes.** There is evidence that interventions combining several different measures were more often successful than single-component interventions. The same explanatory power and impact direction was not seen across programmes in high-income countries, where having an additional component tended to reduce programme effects on employment outcomes.

9. ***In high-income countries*, the role of intervention type is less tangible.** Based on the random-effects meta-regression, no single type of intervention provided clear evidence of a significant effect on the employment or earnings of youth in high-income countries. Skills training appeared slightly more likely to effectuate some (albeit small) impact on employment or earnings, but the difference in comparison to other intervention types was generally not significant. As indicated before, multi-component interventions in high-income countries did not seem to translate into better outcomes for

[40] Participant profiling takes places when an intervention (i) identifies individual factors or characteristics that implied a risk in the labour market for a particular young person and (ii) then relies on this information to assign him/her to specific services. Incentives to participants comprise design features that offer participating youth benefits (e.g., payments, unemployment benefits) conditional on (monitored) programme participation or compliance with programme rules. Incentives to service providers involve payments (or bonuses) to the implementing agency conditional on the achievement of certain outcomes among participating youth (e.g., placement rate).

youth. Longer term employment and income estimates in high-income countries were higher than estimates that considered only the short term (less than one year after treatment exposure). In other words, changes in labour market outcomes take time.

10. **Programme design tends to drive results more strongly than the type of intervention: participant profiling, monitored programme participation, and incentives of participants and service providers are key determinants of success**. In other words, the "how" seemed to be more important than the "what". This was particularly pronounced in high-income countries where the presence of a system to profile participants, mechanisms to engage participants, or incentives for service providers to perform and achieve targets yielded larger impacts.

11. **Programmes lead to better outcomes when they target low-income and disadvantaged youth.** Across measures of targeting, a focus on low-income or youth with low levels of education triggers higher employment and earnings for youth across all country income levels. The analysis by gender is less conclusive. While the overall effect size for employment and earnings appears to be larger for young women than for young men, we find no strong patterns in the multivariate regression analysis to suggest that targeting women only will lead to better outcomes.

12. **The evidence on soft skills is mixed and deserves further impact research.** Given the recent focus on the importance of soft skills for youth among employers, the meta-regression tested their role in determining programme impacts and found no evidence to conclusively suggest that programmes with a soft skills component were more or less successful, even when controlling by age.

13. **There is no clear indication about the impact that public, private or civil society implementers bring to the equation.** While the involvement of public and private entities in the implementation of a youth employment programme led to positive impacts in high-income countries, their role in low- and middle-income countries was non-significant or negative. More impact research is needed to account for implementation agents and mechanisms.

14. **The results appear robust in terms of the quality of the underlying evidence**, as well as across different assumptions and model specifications. Most importantly, they held up under a restricted sample of experimental impact evaluations. While there was some evidence of publication (or small-sample) bias, it was controlled for in the meta-regression analysis, demonstrating that the results were not driven by sample composition bias.

The findings of the systematic review provided important guidelines about how to support youth in the labour market, in a crucial time when governments, donors, international organizations and practitioners are seeking ways to enhance youth's integration into the labour market and fine-tune their actions for more and better results. These findings must nevertheless be contextualized in order to capture local and national specificities as well as programme design and delivery conditions.

The systematic review also supported the identification of important evidence gaps where further quality evidence and work is needed:

- More and better evidence is needed on employment services, wage subsidies and public employment programmes for youth, particularly in low- and middle-income countries. The evidence base on public employment programmes is particularly limited, with only two studies from high-income countries in the sample of included reports.

- It is important to note as well that despite the large and significant magnitude of effect of entrepreneurship promotion interventions, the evidence base is still limited and exhibits high variance. Therefore, more evidence will enhance the accuracy of the synthetized results.

- While the review highlighted a growing evaluation evidence from youth employment programmes implemented in sub-Saharan Africa, it also reported very limited information from the Middle East and North Africa, South Asia and East Asia and the Pacific. These are regions where more targeted action to expand the evidence base should be considered.

- Similarly, more research is needed on intermediate outcomes. This is linked to the importance of improving research reporting standards and expanding the scope of outcomes of interest in order to better synthesize evidence about how interventions impact knowledge, skills, attitudes and behaviours. More and better information on these intermediate outcomes will improve overall understanding about the causality and pathways of change between the intervention and the final outcomes.

- Last, despite the rigorousness of the systematic review, findings will still be incomplete without the availability of cost information. More information is needed on programme costs as well as systematic comparisons against programme effects. What may look highly effective may in fact be too expensive to replicate or scale up.

5.1.3 Unpacking the causal chain across youth employment interventions

Section 1.3 proposed a series of causal chains connecting youth-targeted ALMPs to expected outputs such as direct job creation or changes in skills, knowledge, attitudes, or behaviours, and ultimately linking programme delivery to projected labour market outcomes as well as other closely correlated outcomes such as accumulation of human capital.

Some of these anticipated connections were confirmed by the results of the systematic review, shedding light on the impacts of skills training, entrepreneurship, employment services and subsidized employment on labour market outcomes of youth. The main effect size estimates (i.e., SMDs) are given in Table 43 (above).

This section re-examines the proposed result chains, reflecting on the transmission channels that lead from activities to outcomes. It relies on the findings from the meta-analysis and digs deeper into the individual studies, unpacking features of programme design and implementation that triggered success in the achievement of intermediate and final outcomes.

5.1.3.1 Skills training programmes

Education and training are key determinants of success in the labour market and strong predictors of non-vulnerable jobs among youth (Sparreboom & Staneva, 2014). While time spent on education and training certainly pays off, returns are far more likely to be realized if there are strong, explicit links between education and training policies and the world of work.

Youth training programmes seek to develop skills that enhance human capital and lead to long-term gains in employment. A simplified results chain depicted in Table 44 draws a road map of how exposure to a training programme and the skills acquired through it can lead to improvements in employment, earnings and business performance. The causal hypothesis relies on a series of assumptions and the achievement of some intermediate results, such as positive changes in knowledge, skills, attitudes and behaviours, which are expected to occur in the short term and lead to changes in labour market outcomes such as the probability of employment after programme participation.

The road map is complex, as there are a number of parameters to consider in the design and delivery of training, including (i) the curriculum; (ii) the skills or combination of skills embedded in the curriculum (technical, soft); (iii) training provider's experience and quality; (iv) participation of employers (as well as workers' associations) in programme design and implementation; (v) the setting (in-classroom, on-the-job, mixed); (vi) financial and non-financial incentives for participation of both youth and employers; (vii) targeting mechanisms; (viii) mechanisms for the selection of training providers; (ix) monitoring and reporting; (x) alignment with other ALMPs.

Table 44: Simplified results chain for interventions offering skills training

Activities	Outputs	Outcomes
Technical and business skills training; Literacy or numeracy skills training; and Behavioural, life skills or soft skills training Offered through: 1. Provision of skills training (e.g., distance or classroom training) 2. Placement in workplace training (e.g., internships, on-the-job training schemes) 3. Placement in apprenticeship schemes 4. Provision of financial incentives to young apprentices and employers providing apprenticeship training	1. Improved technical competencies in a specific trade 2. Improved management skills and understanding of business mechanisms 3. Improved financial literacy 4. Improved reading, writing, and mathematical skills 5. Improved psychosocial, decision-making, communication, and teamwork skills 6. Increased self-management and self-esteem 7. Improved physical and mental health	1. Increased probability of employment 2. Reduced time to find job/ shorter unemployment duration/ greater efficiency in the job search 3. Increased ability to retain job/longer job duration (hours worked) 4. Better quality of employment (contract type, job type) 5. Increased earnings or consumption 6. Increased business performance (efficiency, profits, investments, output of entrepreneurs)[41]
Assumptions		
1. Target group participates in training (there is awareness about the programme's existence) 2. Contracted training institutions conduct training and link participants to employers (if conditionality exists) 3. Employers train young people on-the-job and offer placement (if conditionality exists)	1. Participants attend and complete the training 2. Training addresses participants' constraints (e.g., existing skill shortages) as well as the constraints of the labour market 3. Participants learn in training/training increases skill level/training is well matched to interests and abilities of participants 4. Training induces expected behavioural and attitudinal change	1. Existing labour demand for skilled labour 2. Learned skills match labour market needs/demand 3. No stigmatizing effects 4. Training completion and related certificate signals acquisition of increased level of skill and higher (expected/observed) productivity 5. Employers value certified training 6. Participants gain recognized and valued qualifications 7. Adequate economic, social, institutional and administrative conditions are in place

Skills training interventions are the most widely used youth employment intervention worldwide and are increasingly combined with other measures to boost employability (Betcherman et al., 2007; Fares & Puerto, 2009). A total of 55 out of the 107 evaluated interventions (51 per cent, as shown in Table 10) examined by this review fell within the main category of skills training interventions, with 53 per cent of these being conducted in high-income countries, 35 per cent in middle-income countries and 13 per cent in low-income countries.[42]

[41] Additional entrepreneurship-related outcomes are listed in Section 5.1.3.2.

[42] Figures do not total 100 per cent due to rounding.

On average, skills training interventions improved employment outcomes among young women and men by 0.05 SMDs (CI = 0.02, 0.07; I^2 = 65 per cent; number of interventions = 67, as shown in Table 13) and also led to higher earnings (0.07 SMDs; CI = 0.05, 0.08; I^2 = 86 per cent; number of interventions = 60, as shown in Table 14). Some key results emerged from the meta-review:

1) **Skills training programmes lead to positive changes in labour market outcomes.** With a sizable evidence base, effect size estimates across all country income types were positive (Table 45). The result supports the economics of active labour market training programmes which aim to help youth enter the labour market and accumulate the necessary skills to compete for jobs and improve their productivity – with subsequent positive impacts on wages provided that there is no depreciation in skills (Heckman, Lochner & Cossa, 2002).

Table 45: Main results from skills training interventions

Parameters of interest	Employment outcomes		Earnings outcomes	
	HICs	LMICs	HICs	LMICs
Standardized Means Difference I^2	0.04	0.06	0.02	0.12
Standard errors of the effect size	0.02	0.02	0.01	0.02
95% confidence interval	0.01	0.02	0	0.08
	0.07	0.1	0.04	0.16
I Squared	67.55	62.54	72.32	88.62
Number of SMDs	534	370	303	192
Number of interventions	29	38	21	39
Sample size	2,394 204	1,045 500	1,163 479	882 481
Mean difference	1.86	14.92	4800.83	726.10
Control outcome	33.90	40.40	18537.80	5520.71
Treatment outcome	33.44	39.71	20514.93	6198.67
Percentage change	0.03	0.14	0.02	0.18

Notes: HICs: High-income countries, LMIC: Low- and middle-income countries.

2) **The effect of training is higher for youth in low- and middle-income countries compared to youth in high-income countries** (Table 45). The result echoed the findings from Betcherman et al. (2007) and highlighted the role of contextual variables, such as access to basic and technical vocational education and training, and to social protection systems, and suggested that, while training programmes led to positive outcomes in high-income countries, they were unable to compensate for skills that were not acquired at school.

The multifaceted nature and evolution of skills training interventions was also observed in the evidence from single studies:

3) **Comprehensive, multi-service training interventions were more prevalent and worked best in low- and middle-income countries.** Skills training interventions have evolved into holistic measures (Fares & Puerto, 2009). Some 36 out of the 107 interventions (34 per cent) examined by this review combined skills training with one or more additional intervention types: 24 interventions combined training with employment services only, eight interventions with subsidized employment only, and one intervention with entrepreneurship promotion only. There were three cases in which skills training interventions were combined with more than one intervention type.

   The combination of skills training and entrepreneurship promotion (and potentially further intervention types) was particularly prevalent in low- and middle-income countries, emphasizing youth's scant opportunities in (formal) employment and the limited ability of public and private sectors to absorb the growing youth labour force. Some examples of evaluated interventions that consisted, at a minimum, of a skills training and an entrepreneurship intervention component included the Employment and Livelihood for Adolescents (ELA) in Uganda, the Economic Empowerment of Adolescent Girls (EPAG) Programme in Liberia and the Livelihoods Training for Adolescent Living Programme in India.

4) **Recent evidence points to the relevance of incentives and profiling mechanisms within the design of the interventions.** Incentives and profiling measures were correlated with better employment and earnings outcomes, as discussed in Section 4.3.4. The Adolescent Girls Employment Initiative (AGEI) of the Employment Fund in Nepal provided technical and life skills training with a comprehensive incentive scheme. Training providers, who were selected through a competitive bidding process, were offered a bonus payment based on the number of trainees that had obtained "gainful" employment six months after completing the training and a second bonus for the share of participants that met pre-specified vulnerability criteria and were successfully placed in employment (Ahmed, Chakravarty, Lundberg & Nikolov, 2014).

5) **Despite the growing awareness and demand for soft skills, aggregated results did not imply that they systematically led to better outcomes.** While most interventions covered by this review offered technical skills, soft or non-technical skills were increasingly embedded in training packages (28 out of 55 skills training interventions), reflecting employers' demand for these abilities (Cunningham et al., 2010; Youth Employment Network & International Youth Foundation, 2009).

   The meta-regression results did not suggest a significant correlation of the inclusion of a soft skills component with larger effect size estimates. In fact,

when restricting the sample to high-income countries, the availability of soft skills in the programme curriculum was correlated with lower employment effects, particularly among the younger cohort. An example from a high-income country is the JOBSTART programme from the United States. The programme applied an intensive exposure model that combined basic education, occupational skills training, training-related support services and jobs development and placement assistance – which included work-readiness, life and communication skills – for school dropouts and economically disadvantaged youth. While there is no disaggregation of impacts by skills set delivered, the evaluation showed overall meagre impacts on employment outcomes (Cave, Bos, Doolittle & Toussaint, 1993).

Evidence from single studies in low-income countries offered more promising results. The combination of life and vocational skills provided to adolescent girls by the ELA Programme in Uganda led to large and significant changes in behaviours and an increased probability of employment and self-employment (Bandiera, Buehren, Burgess, Goldstein, Gulesci, Rasul & Sulaiman, 2014). These mixed results called for further investigation about the role of soft skills in the causal chain from intervention to final outcomes.

6) **Multi-setting approaches enhanced the acquisition of relevant skills and led to better labour market outcomes.** Skills training interventions expanded the exposure of trainees to different environments, particularly by combining in-classroom with on-the-job training (Fares & Puerto, 2009). This combination was prevalent in almost half of the evaluated skills training interventions (25 out of 55). When not combined, classroom training alone was more frequently observed (in 45 out of 55 skills training interventions, compared to 32 where training was given at the work place).

The Jóvenes Programmes in Latin America and the Caribbean were well-represented in this systematic review, with (often several) impact evaluation studies for programmes implemented in Argentina, Chile, Colombia, Dominican Republic, Panama and Peru. The model, piloted in the 1990s, combined in-classroom and on-the-job training in a demand-driven fashion. On the one hand, the design of the programme ensured private sector involvement in the definition of training content, securing the correspondence between the skills taught and those demanded by the productive sector. On the other hand, implementation was demand driven through stringent, competitive bidding processes for the selection of training providers, and incentive payment schemes were based on trainees' outcomes. The first and most successful of the Jóvenes Programmes in terms of impact on employment was Chile Jóven, with an effect size for employment outcomes of 0.35 SMD (CI = 0.13, 0.58) and for income outcomes of 0.23

SMD (measured with less precision, CI = -0.16, 0.60). The employment effect sizes of other Jóvenes Programmes were lower but still positive and close to the sample mean for skills training interventions (which was SMD 0.05; CI = 0.02, 0.07).

It is important to note that, while it was not possible for the systematic review to assess treatment effects on intermediate outcomes, such as knowledge, skills acquisition, attitudes and behaviours, some single studies did find (i) positive impacts of youth employment programmes on educational outcomes (in the United States) and (ii) noticeable changes in behaviours, expectations and non-cognitive skills (in Dominican Republic).

In conclusion, Table 46 provides an evidence check against the expected outcomes for skills training interventions outlined in the results chain.

Table 46: Evidence checks for skills training interventions

Expected outcomes	Evidence checks
1. Increased probability of employment	There was ample evidence demonstrating the ability of skills training to increase the probability of employment among youth after programme exposure. The evidence applied to all country income levels and across wage employment (France's Contrat de Qualification) and self-employment (Uganda's ELA programme).
2. Reduced time to find job/ shorter unemployment duration/ greater efficiency in the job search	Few studies reported on the job search or time looking for a job after the programme. Measurements of unemployment duration or unemployment probability were less common and offered mixed results. For example, young men that benefited from the programme Juventud y Empleo in the Dominican Republic saw an increase in formalization (written contract) coupled with an increase in duration (weeks) of unemployment and hours spent job-seeking on last working day (Ibarrarán et al., 2014). Comprehensive measures that combined training with counselling and job search assistance offer potential to impact the job search. However, more evidence is needed to support this proposed causality.
3. Increased ability to retain job/longer job duration (hours worked)	While the evidence was clear about the positive impact of training on the probability of employment, it was less so about its impact on employment duration. Furthermore, employment probabilities and hours of work did not necessarily react in the same way to the same intervention; e.g. the evaluation of Galpao in Brazil reported positive and negative average SMDs for employment probability and hours worked, respectively (Calero et al., 2014). In contrast, programmes in Nepal and India, reported both high employment probabilities and high hours of work (Ahmed et al., 2014 and Maitra & Mani, 2014)
4. Better quality of employment (contract type, job type)	Skills training increased job quality. The evidence was more common among programmes in low- and middle-income countries and it was correlated with better wages or earnings; e.g. Colombia's Formación Técnica y Tecnológica and Jóvenes en Acción Programmes, the Ninaweza Youth Empowerment Programme in Kenya, Procajoven in Panama, and Projoven in Peru.
5. Increased earnings	Skills training interventions led to higher earnings among youth, supporting the argument that investments in human capital lead to higher wages and therefore better employment outcomes in the long

Expected outcomes	Evidence checks
	term. A review of impact evaluations of the Jóvenes Programmes in Latin America and the Caribbean showed positive short term impacts on earnings, slightly larger for young women (in Colombia and Panama) than young men. The data was however less reliable as retrospective evaluations had to rely on retrospective income data (Ibarrarán & Rosas-Shady, 2009).
6. Increased business performance (efficiency, profits, investments, output of entrepreneurs)	No clear evidence that demonstrated positive changes in business outcomes. While training programmes increasingly incorporate management skills and business courses in the training content (Chile Joven, Chile's Formación en Oficios para Jóvenes de Escasos Recursos Programme, Liberia's EPAG, Apprenticeship Training Programme and Entrepreneurial Support for Vulnerable Youth in Malawi, and Nepal's Employment Fund), studies did not show systematic measurement of changes in business performance outcomes. Some studies that measured business creation found negative impacts (Alvares de Azevedo, Davis & Charles, 2013 for Ninaweza in Kenya and Cho et al., 2013 for the Apprenticeship programme in Malawi). However, it was important to emphasize that starting a business was not always the primary goal of those interventions.

5.1.3.2 *Entrepreneurship promotion interventions*

Entrepreneurship promotion interventions are designed to address the individual and external constraints that young people encounter in starting or growing a business by providing entrepreneurial skills and facilitating access to capital for self-employment – including physical, financial and social capital.

The systematic review examined 15 entrepreneurship interventions that offered mainly business skills training, business advisory services and/or access to credit or grants. Table 47 presents a simplified version of the results chain in Section 1.3.2 to outline the outcomes expected from entrepreneurship interventions, including (i) employment outcomes such as increased probability of employment, (ii) earnings outcomes and (iii) business performance outcomes, such as increased sales.

Table 47: Simplified results chain for entrepreneurship promotion

Activities	Outputs	Outcomes
1. Business and management training 2. Business advisory services, mentoring and coaching 3. Access to markets and value chains 4. Credit or access to credit 5. Grants (monetary or in-kind) 6. Microfranchising	1. Increased entrepreneurial impetus, business and management skills and financial literacy 2. Improved understanding of business practices, laws and regulations 3. Increased access to markets and networks	1. Increased employment probability or number of hours worked 2. Increased earnings or consumption among young entrepreneurs 3. Business started 4. Increased business investment, performance and competitiveness (e.g. profits, sales, capital and investment, business survival) 5. Additional jobs created

Activities	Outputs	Outcomes
	4. Increased access to capital or financial services 5. Increased incentives to start own business	
Assumptions		
1. Content, intensity and delivery of services is tailored to the needs of the target group and to the programme objective 2. Correct group is interested in the intervention and is targeted 3. Target group participates in programme and completes entire programme cycle	1. Participants learn from training and advisory service 2. Training and advice prompted expected behavioural change 3. Credit or grant is used for enterprise 4. Credit agency/franchisor does not exploit entrepreneur	1. Created and supported businesses meet existing consumer demand 2. Adequate regulatory and business environment 3. Fertile macroeconomic environment 4. Adequate economic, social, institutional and administrative conditions 5. Start-ups benefit from additional investment/credit/networks 6. Credit or grant is used for productive investments

Some important results emerged from the analysis and review of single studies:

1) **On average, entrepreneurship promotion interventions lead to positive effects on employment outcomes** (0.16 SMD; CI = 0.06, 0.26; I^2 = 71 per cent; number of interventions = 7), **earnings outcomes** (0.09 SMD; CI = 0.01, 0.18; I^2 = 64 per cent; number of interventions = 12) **and business performance outcomes** (0.10 SMD; CI = 0.00, 0.19; I^2 = 39 per cent; number of interventions = 10). Similar to the analysis of Cho and Honorati (2013), this review observed a wide variation of effects depending on the services provided within the intervention package and the context.

2) **Most of the evidence originated from interventions set in low-income countries (Liberia and Uganda) and middle-income countries (Bosnia and Herzegovina, Colombia, Peru and Tunisia)**, and their evidence was notably recent. Ten of the 15 entrepreneurship interventions were evaluated between 2012 and 2014, with evidence predominantly coming from Africa. Only two interventions were implemented in high-income countries (France and United Kingdom), which implies they were dropped from the analysis due to insufficient sample. Table 48 therefore presents the effects of entrepreneurship interventions in low- and middle-income countries. Detailed characteristics of these entrepreneurship interventions are presented in Section 8.2.

3) **The effects were intensified in low- and middle-income countries where entrepreneurship interventions reported larger effects on employment outcomes** (0.18 SMD; CI = 0.06, 0.29; I^2 = 68 per cent; number of interventions = 5), **earnings outcomes** (0.14 SMD; CI = 0.06, 0.22; I^2 = 49 per cent; number of interventions = 10) **and business performance outcomes** (0.15 SMD; CI = 0.07, 0.23; I^2 = 0 per cent; number of interventions = 9). Entrepreneurship interventions appeared to work well if they address specific constraints: In Uganda, the evaluation of the Youth Opportunities Programme (YOP) showed that grants for non-agricultural vocational training and business start-up had substantial economic impacts on **earnings** for young people in the capital-constrained environment of a conflict-affected region. This finding is in line with Section 4.3.3.6, which highlighted similar positive effects of entrepreneurship interventions for disadvantaged youth.

Table 48: Main results from entrepreneurship interventions

Parameters of interest	Employment outcomes		Earnings outcomes		Business performance outcomes	
	HICs	LMICs	HICs	LMICs	HICs	LMICs
Standardized Means Difference I^2	Dropped from analysis	0.18	Dropped from analysis	0.14	Dropped from analysis	0.15
Standard errors of the effect size		0.06		0.04		0.04
95% confidence interval		0.06		0.06		0.07
		0.29		0.22		0.23
I Squared		67.66		49.45		0
Number of SMDs		35		39		150
Number of interventions		5		10		9
Sample size		54,205		34,542		48,101
Mean difference		4.19		3,013.40		112.36
Control outcome		7.83		3,286.44		360.83
Treatment outcome		9.98		4,447.00		456.85
Percentage change		0.36		0.3		0.34

Note: (i) HICs: High-income countries, LMIC: Low- and middle-income countries; (ii) Results for employment, earnings and business performance outcomes of entrepreneurship interventions in HICs were dropped from the analysis due to an insufficient number of observations.

4) **Entrepreneurship interventions followed a trend towards multi-component services.** About two-thirds of the evaluated interventions offered a combination of business skills training, business advisory services (including mentoring) and/or access to finance. An intervention which adopted this multi-pronged approach was the Women's Income Generation Support (WINGS) programme in Uganda with the largest effect size for employment outcomes across all evaluated interventions examined by the review. The programme combined business skills training, cash grants and

follow-up support to young women, leading to an increase in working hours from 14 to 25 hours per week. This programme seemed to be the main driver of the overall positive impact of entrepreneurship interventions.

5) **Similarly, interventions providing both entrepreneurship training and business advisory services – irrespective of grants provision – showed strong, positive evaluation results on employment outcomes in low- and middle-income countries**. For example, the Economic Empowerment of Adolescent Girls (EPAG) programme in Liberia provided classroom-based training followed by six months of follow-up support and reported a 47 per cent increase in employment. In addition to changes in labour market outcomes, the evaluation showed improvements in the self-confidence of participating girls.

6) **Positive business performance outcomes (e.g., an increase in profits) were reported for interventions that provided *start-up* grants, either alone or in combination with training and advisory services.** These results were driven by interventions that specifically aimed at mitigating capital constraints for poor and vulnerable young people, as in the case of the YOP and the WINGS programmes, both implemented in northern Uganda.

7) **The evidence on grants was, however, not conclusive when it came to supporting existing young entrepreneurs in growing and expanding their businesses**. A recent randomized experiment with the Start and Improve Your Business (SIYB) programme in Uganda showed that limited access to finance was a real constraint for young business owners that could be addressed through the combination of business training and loans. This programme effect, however, only held for the subsample of young men who had expressed an interest in growing their business. Evidence suggested that, in developing countries, family pressure on women can deflect the use of grants or credit for non-business purposes (Fiala, 2014).

   The evaluation of the Partner Microcredit Foundation Experiment, a business and financial literacy programmes in Bosnia and Herzegovina, highlighted the fact that the programme led to improvements in business practices and entrepreneurial impetus, but did not directly translate into improved chances of business survival. In Peru, three entrepreneurship interventions addressing the need for a multi-component approach through business training, business advisory services and access to finance, also improved business performance outcomes of low-income youth and youth living in rural areas. The programmes Calificación de Jóvenes Creadores de Microempresas, Formación de Líderes Empresariales, and Formación

Empresarial de la Juventud relied on business plan competitions to determine eligibility for programme participation or start-up funding.

8) **Developing a business plan was a component of more than one-third of entrepreneurship interventions and was a common means of determining eligibility for participation in the programme and/or access to finance**. Additional examples included CréaJeunes in France and Turning Theses into Enterprises in Tunisia.

In conclusion, Table 49 provides an evidence check against the expected outcomes for entrepreneurship promotion interventions outlined in the results chain.

Table 49: Evidence checks for entrepreneurship promotion interventions

Expected outcomes	Evidence checks
1. Increased employment probability or number of hours worked	There was strong evidence that entrepreneurship promotion interventions in low- and middle-income countries led to increased employment probability and number of hours worked. E.g., the Economic Empowerment of Adolescent Girls (EPAG) programme in Liberia reported a large increase in employment (Adoho et al., 2014). The Women's Income Generation Support (WINGS) programme in Uganda showed the largest effect size for employment outcomes across all evaluated interventions examined by the review (Blattman et al., 2013, Blattman et al., 2014).
2. Increased earnings or consumption among young entrepreneurs	Entrepreneurship promotion interventions tend to show positive effects on earnings and consumption for young people. These effects were intensified in low- and middle-income countries where entrepreneurship interventions proved particularly effective for disadvantaged youth and in capital-constrained environments such as in the context of the Youth Opportunities Programme in Uganda (Blattman, Fiala & Martinez, 2013)
3. Business started	There was good evidence that entrepreneurship promotion is an effective approach to support business creation by young people. For example, Formación de Líderes Empresariales in Peru improved business creation by providing business training, business advisory services, business plan competitions and access to finance (Jaramillo & Parodi, 2005).
4. Increased business investment, performance and competitiveness (e.g. profits, sales, capital and investment, business survival)	Overall, the impact of entrepreneurship interventions on business performance outcomes was positive. E.g., the Youth Opportunities Programme in Uganda led to positive results on capital/investment and the WINGS programme in Uganda reported positive effects on business survival (Blattman et al., 2013, Blattman et al., 2014, (Blattman, Fiala & Martinez, 2013). However, the evidence was inconclusive on means to support existing young entrepreneurs to grow and expand their business. The Start and Improve Your Business Programme (SIYB) in Uganda showed differential impacts across gender, with young men benefiting more from combined training and loans than young women (Fiala, 2014).
5. Additional jobs created	There is no sufficient evidence to validate the causality between youth entrepreneurship promotion interventions and the creation of jobs through the newly created or expanded businesses. Blattman, Fiala & Martinez's evaluation of the Youth Opportunities Programme in Uganda offered an example of a study that captured positive effects on additional jobs created.

5.1.3.3 Employment services

Employment services generally comprise interventions focusing on labour intermediation, i.e. programmes optimizing the process that matches jobseekers with vacancies. They deliver job counselling, job-search assistance and/or mentoring services for (re)activation purposes, which are often complemented by job placements and technical or financial assistance. The basic idea behind providing employment services to youth is that young workers have difficulty signalling their skills and credentials and/or lack the networks or knowledge to effectively search for vacancies and connect with employers. Hence, these programmes often focus on improving job-seeking skills and the efficiency of the matching process (Table 50).

Table 50: Simplified results chain for employment services

Activities	Outputs	Outcomes
Job placement/ intermediation services, through: 1. Providing job placement services used by unemployed 2. Assessing and matching jobseekers and potential employers (brokering information) 3. Marketing disadvantaged jobseekers to employers 4. Matching unemployed with job vacancies	1. Improved matching of jobseekers and employers 2. Increased intensity (motivation) and efficiency of job-search	1. Increased labour-market participation 2. Increased probability of employment 3. Reduced time to find job/shorter unemployment duration 4. Increased ability to keep a job/longer job duration 5. Better quality of employment (contract type, hours worked) 6. Increased earnings or consumption
Job counselling/ job-search assistance/ mentoring, through: 1. Providing career and personal development advice 2. Providing job-search advice or training 3. Ensuring active and efficient job search	1. Participants better informed about labour market (i.e., qualifications in demand and where jobs are to be found) 2. Improved job-search skills 3. Increased intensity, motivation and efficiency of job-search 4. More informed decisions about investment in education	As above
Financial assistance for job search, through: Provision of credit or grants/stipends connected to job-search and job-acceptance (e.g., transport, childcare)	Greater ability to find and accept jobs (e.g., enhanced mobility)	As above
Assumptions		
1. Target group (unemployed and employers) takes up the service offer (there is awareness about the programme's existence)	1. Correct target group identified (participants are constrained by lack of job-search skills) 2. Participants are motivated	1. Existing labour demand for employment services beneficiaries 2. Correct barriers and constraints for youth on the labour market

Activities	Outputs	Outcomes
2. Participants complete/attend the programme 3. Participants comply with conditionalities and service requirements 4. Service matches the needs and abilities of participants	to search and take up work 3. Matched workers are able to do the job 4. Participants learn on the job/employment increases skill level 5. Behavioural changes are prompted	are addressed 3. Adequate economic, social, institutional and administrative conditions established 4. No stigmatizing effects

The review identified a sample of ten employment services interventions, a majority of which combined job counselling, job-search assistance and mentoring services. In fewer cases, the interventions provided job-placement services and/or financial assistance. The only intervention that focused solely on financial assistance for job search was a subsidized transportation experiment in Ethiopia (Franklin, 2014). Interventions were typically of short duration (three months on average) and their intensity ranged from one-off afternoon visits to job information centres for secondary students in Germany to 12 months in the Counselling and Job Placement for Young Graduate Job Seekers programme in France. Importantly, the review highlighted the increasing reliance on employment services as supplementary measures within other ALMPs, mainly training and wage subsidies.

Most evaluations took place in high-income countries (Finland, France, Germany, Portugal and the United States) where they were typically implemented by public employment agencies and operated on a national scale. In developing countries, evaluated interventions were implemented in Ethiopia, India and Jordan. They were characterized by their small scale or pilot nature and the common aim to reduce job-search costs for jobseekers, either via job screening and matching, recruiting services or transport subsidies.

On average, employment services interventions provided moderate gains in employment outcomes among young women and men by 0.01 SMDs (CI = -0.02, 0.04; I^2 = 0 per cent; number of interventions = 10, as shown in Table 13 where it is also possible to see that employment services reported the lowest SMD for the outcome category) and also led to moderate higher earnings (0.01 SMDs; CI = 0, 0.02; I^2 = 0 per cent; number of interventions = 8, as shown in Table 14).

The evidence pointed to several key patterns:

1) **Most employment services programmes tended to specialize in specific services.** In contrast to other main intervention types, employment services interventions exhibited a trend towards single-pronged approaches, mainly the provision of job counselling, job-search assistance and/or mentoring services. A relatively successful example of this monotypic intervention is the programme of mandatory visits to job information centres

for German secondary students, whereas a less effective example is the "Job Shadowing" component of the School-to-Work Opportunities Act (STWOA) in the United States.

2) **The evidence on employment services programmes from low- and middle-income countries is very thin.** This is likely in line with the fact that this programme type originates in the idea of assisting registered jobseekers within an Unemployment Insurance (UI) system in a high-income country, and is thus an uncommon main intervention type in low- and middle-income countries. In fact, the number of studies in the sample of low- and middle-income countries was too small to comply with the review's minimum requirement (four) and was therefore dropped from the effect size analysis (Table 51). An examination of the individual studies in Jordan (Groh, McKenzie, Shammout & Vishwanath, 2014), India (Jensen, 2012) and Ethiopia (Franklin, 2014) showed rather positive impacts on employment outcomes of young participants.

Table 51: Main results from employment services

Parameters of interest	Employment outcomes		Earnings outcomes	
	HICs	LMICs	HICs	LMICs
Standardized Means Difference I^2	0.01	Dropped from analysis	0.01	Dropped from analysis
Standard errors of the effect size	0.02		0.01	
95% confidence interval	-0.03 0.04		0 0.02	
I Squared	9.89		0	
Number of SMDs	87		32	
Number of interventions	7		5	
Sample size	2,326,518		190,770	
Mean difference	-1.57		-16.51	
Control outcome	27.80		215.34	
Treatment outcome	20.89		174.91	
Percentage change	0.02		0.00	

Notes: (i) HICs: High-income countries, LMIC: Low- and middle-income countries; (ii) Results for employment, earnings and business performance outcomes of entrepreneurship interventions in HICs were dropped from the analysis due to an insufficient number of observations.

3) **Aggregate empirical evidence for high-income countries showed positive effect sizes for employment and earnings outcomes**, though they were relatively smaller than the SMDs for other intervention types. Single studies in high-income countries typically found small or often non-significant effects on employment. The study by Caliendo, Künn and Schmidl (2011) was the only one that detected positive long-term effects on youth labour market outcomes from the Job Search Assistance track of the German ALMP measures.

4) **In most studies, the changes in labour market outcomes were transitory and there was no sign of a stepping-stone effect.** Evidence for this was provided, for instance, by the impact evaluations of the Counseling and Job Placement for Young Graduate Job Seekers in France (Crépon et al., 2013), the transport subsidies intervention in Addis Ababa (Franklin, 2014), and the mandatory visits to job information centres in Germany (Saniter, 2014)).

In conclusion, Table 52 provides an evidence check against the expected outcomes for employment services outlined in the results chain.

Table 52: Evidence checks for employment services

Expected outcomes	Evidence checks
1. Increased labour-market participation	There was no evidence to validate the causality between employment services for youth and this outcome construct.
2. Increased probability of employment	Current evidence is very thin in this regard an insufficient to secure the causality in the aggregate. Positive changes in employment probability were reported more often among low- and middle-income countries (Ethiopia (Franklin, 2014), India (Jensen, 2012), and Jordan (Groh et al., 2014)) than high-income ones (France (Crépon et al., 2013), Germany (Caliendo, Künn & Schmidl, 2011)). It is however important to note that positive changes were generally not accompanied by positive changes in other outcome types, such as earnings.
3. Reduced time to find job/shorter unemployment duration	The limited evidence showed that employment services increased the probability of unemployment (instead of reducing it as expected). This effect was captured in Finland (Hämäläinen, Hämäläinen & Tuomala, 2014) and Germany (Saniter, 2014). Unemployment duration was seldom measured, and when it was, results show limited gains.
4. Increased ability to keep a job/longer job duration/ increase in hours worked	There was no evidence to validate the causality between employment services for youth and this outcome construct.
5. Better quality of employment (contract type)	There was no evidence to validate the causality between employment services for youth and this outcome construct.
6. Increased earnings or consumption	Impacts on earnings were rather small with some negative reports in Jordan (Groh et al., 2014) and France (Crépon et al., 2013). Consumption changes were only measured in one study (India (Jensen, 2012)), not sufficient to support the proposed causality.

5.1.3.4 Subsidized employment interventions

Overall, subsidized employment interventions reported larger effects on employment outcomes (0.02 SMDs; CI = -0.01, 0.06; I^2 = 50 per cent; number of interventions = 105, as shown in Table 13) than on earnings (-0.01 SMDs; CI = -0.05, 0.03; I^2 = 61 per cent; number of interventions = 89, as shown in Table 14). They also appeared less successful in higher income countries (Table 53). Before

delving further into these findings, the analysis below differentiates between results and evidence from interventions delivering wage subsidies as opposed to public employment programmes – two subsidized employment measures with very distinct characteristics in design and implementation.

Table 53: Main results from subsidized employment interventions

Parameters of interest	Employment outcomes		Earnings outcomes	
	HICs	LMICs	HICs	LMICs
Standardized Means Difference I^2	0	0.11	-0.02	0.02
Standard errors of the effect size	0.02	0.04	0.02	0.04
95% confidence interval	-0.03 0.04	0.04 0.18	-0.07 0.03	-0.05 0.1
I Squared	48.82	0	79.91	0
Number of SMDs	160	33	46	11
Number of interventions	11	5	5	4
Sample size	32,031 060	167 129	10,347 125	11 030
Mean difference	0.54	0.16	-2106.18	16.09
Control outcome	11.09	0.89	9585.32	551.43
Treatment outcome	7.20	0.94	7477.58	588.67
Percentage change	0.01	-0.04	-0.03	0.08

Notes: HICs: High-income countries, LMIC: Low- and middle-income countries

5.1.3.4.1 Wage subsidy interventions

Low levels of skills, limited or no work experience, signalling barriers, or economic crises and downturns all hamper labour demand for youth. Employers may have limited scope for hiring or suspect that youth come to the market with low productivity levels – lower than the market wage for a given job. To compensate for possible low productivity and to incentivize hiring (and training) of young people, wage subsidy programmes offer a risk discount to employers that offsets certain wage and non-wage costs.

Table 54 (a shortened version of Table 5) lists (i) more and better employment outcomes (from increased probability of employment to higher job quality and more efficient job searches), (ii) higher earnings, and (iii) long-term effects on youth's human capital and employability among the expected outcomes of wage subsidy programmes.

Table 54: Simplified results chain for interventions offering wage subsidies

Activities	Outputs	Outcomes
Wage subsidy is offered and transferred through payroll tax cuts or direct payments to young people or employers.	1. Direct job creation 2. Participants (re)gain labour market contact 3. Participants increase (or demonstrate increased)	1. Increased probability of employment 2. Shorter unemployment spells in the future/more efficient job search

Intervention offers: 1. A job (of short of long duration) 2. A job with work-based training 3. A job plus work-based training and/or job-search assistance	productivity 4. Skills formation or increased job skills (technical and non-technical) through on-the-job training and exposure to the work environment 5. More positive attitudes towards employment/increased incentives to apply for jobs or to work 6. Participants integrate into networks 7. Incentives to continue education	3. Increased ability to retain a job/ longer job duration 4. Better quality of employment 5. Increased earnings or consumption 6. Increased returns from employment, including long-lasting human capital accumulation
Assumptions		
Complete information about the programme for both employers and youth Target group (first time job-seekers, disadvantaged/low-skilled youth, unemployed youth, and employers) participates in programme	1. Participants are motivated to work and appropriately qualified (adequate profiling) 2. Participants learn on the job (i.e., experience increases skills levels) 3. Programme induces (positive/expected) behavioural changes/no adverse behavioural changes 4. Subsidies are not exploited by firms or conditionalities are in place to avoid unintended behaviours by employers 5. Appropriate targeting to avoid windfall for the firm and deadweight for society	1. Correct barriers and constraints for youth attempting to access the labour market are addressed 2. Work experience adequately signals higher skills and employability 3. Acquired skills/work experience match labour market demands 4. No stigmatizing effects 5. No windfall, deadweight, substitution, or displacement effects.

The systematic review included 17 studies in which wage subsidies featured as the main category of intervention. Most of the evidence (12 out of 17 studies) came from high-income countries; namely, Australia, Canada, Chile, France, Germany, Sweden and the United States. Evaluations from middle-income countries (registered in five studies) assessed impacts of programmes implemented in Jordan, South Africa, Tunisia and Turkey[43]. The evidence distribution was important as there was noticeable heterogeneity in the results across country income types. Two key messages stemmed from the results:

1) **Wage subsidy programmes for youth performed better in middle-income countries than in high-income countries** (Table 55). Effect sizes for employment and earnings were respectively close to zero and negative in high-income countries.

[43] There is no evidence from low-income countries. In fact, all middle-income countries in this group are classified as uppermiddle-income. (Source: World Bank Country and Lending Groups 2016.)

2) **Employment outcomes were highly responsive to young people's exposure to wage subsidies**, especially in comparison to earnings outcomes.

Table 55: Main results from wage subsidy interventions

Parameters of interest	Employment outcomes		Earnings outcomes	
	HICs	LMICs	HICs	LMICs
Standardized Means Difference I^2	0	0.11	-0.04	0.02
Standard errors of the effect size	0.01	0.04	0.04	0.04
95% confidence interval	-0.03	0.04	-0.12	-0.05
	0.03	0.18	0.05	0.1
I Squared	27.14	0	84.93	0
Number of SMDs	75	33	45	11
Number of interventions	7	5	4	4
Sample size	31,227,034	167,129	10,344,230	11,030
Mean difference	0.38	0.16	-3,276.79	16.09
Control outcome	8.44	0.89	14,908.02	551.43
Treatment outcome	5.54	0.94	11,629.00	588.67
Percentage change	0.01	-0.04	-0.05	0.08

Notes: HICs: High-income countries, LMIC: Low- and middle-income countries

To explain these effects, the review pointed to the role of design features in determining programme effectiveness; echoing similar claims by Neumark and Grijalva (2013), Almeida et al. (2014) and Bördős et al. (2016). Table 36, column (4) shows that once design features such as participant profiling, supervision, and incentives were accounted for, subsidized employment interventions, heavily influenced by the wage subsidy programmes in the sample, appeared to be more successful than skills training interventions.

The design of wage subsidy programmes implied numerous decisions on: (i) targeting – general subsidies vs. hiring subsidies or the decision to focus on specific target groups; (ii) the payment vehicle – direct payment, reduction in payroll taxes or social security contributions; (iii) the payee – employer or employee; (iv) the size of the subsidy and basis for its computation; (v) the duration of the subsidy or of the intervention as a whole; (vi) the offer – a job, a job with training or a job with training and other services; (vii) conditionalities, reporting requirements and programme monitoring.

While there was no clear evidence on relative effectiveness across design options, some messages from single studies were apparent:

3) **Fine-tuning conditionalities, securing feasibility of claims and proper information and dissemination were critical to incentivize firm take up.** Conditionalities were set to curb unintended behaviours and ensure connections across the underlying theory of change. Their establishment implied appropriate monitoring, which was often linked to well-developed public employment services. Stringent conditionalities, however, have the potential to deter employers' participation, as shown in the French national programme, Contrat Jeune en Entreprise, aimed at promoting long-term contracts among disadvantaged youth. The programme offered a hiring subsidy, paid directly to the employer, and targeted youth aged 22 and younger, who had dropped out of school before passing the secondary school examination that would qualify them for entry to university. The subsidy was proportional to the part-time ratio for part-time workers and was offered in full for two years and then reduced to half during the third year. In return, employers had to commit to not dismissing a participant, except for professional misconduct, during the three-year term of the contract. The programme led to a very low take-up by employers, who argued that conditions were too strict in comparison to the perceived benefit (Roger & Zamora, 2011).

In contrast, conditionalities that were compensated with relatively high subsidies seemed to cover the employer's opportunity cost adequately and enhance their participation. The national German programme JUMP offered direct payments to employers of 40 per cent of the wage value on the hiring of unemployed youth with secondary education. The relatively generous subsidy was paired with strict conditions for no early dismissal and a guaranteed period of post-subsidy employment, equivalent in duration to half the subsidized period. An impact evaluation of the programme showed positive impacts on the probability of employment in the short and long terms, with higher effects among the more skilled youth and in regions with relatively low labour demand (Caliendo, Künn & Schmidl, 2011).

The lack of internal mechanisms at the firm level and of adequate information decreases incentives for the subsidies. A controlled experiment that provided employment vouchers to unemployed young South Africans in order to reduce the wage costs for employing firms yielded an average SMD for employment outcomes of 0.13 (CI = 0.01, 0.26). The evaluation study reported a positive probability of wage employment that reduced slightly over the longer term. However, the experiment suffered from a low take-up of the employment vouchers by eligible employers, which seemed to be partially correlated with the administrative burden of claiming the subsidy

(firms did not have internal processes in place to deal with this aspect) and the perception by employers that the vouchers were not legitimate (Levinsohn et al., 2014).

4) **Profiling was key to avoiding deadweight and substitution effects.** The Stage d'Initiation à la Vie Professionelle (SIVP) in Tunisia provided an employment subsidy for university graduates by reducing the employer's hiring costs and exempting it from social security contributions, resulting in an average programme SMD of 0.16 (CI = -0.03, 0.34). The programme decreased joblessness, increased the probability of employment in the private sector and reduced the chances of permanent contracts among young programme participants. The first-come, first-serve nature of the programme and non-reliance on profiling mechanisms is argued to have led to large deadweight effects (Broecke, 2013).

   In general, single studies hardly account for deadweight, substitution or displacement effects. This is a significant drawback that restricts the interpretation and applicability of evaluation findings (Almeida et al. 2014).

5) **What matters in the offer is the ability of programmes to enhance skills formation among youth.** A programme that only offers "a job" has the potential to lead to positive outcomes if the exposure to employment is sufficiently relevant to facilitate learning-by-doing, which will lead to higher employment in the long run (Heckman et al., 2002). Relevant exposure could imply subsidized employment of extended duration, as in the case of the JUMP wage subsidies in Germany (Caliendo, Künn & Schmidl, 2011) or exposure to a job that facilitated the acquisition of or delivered on-the-job training on new and job-relevant skills. Comprehensive designs that combined wage subsidies with skills training measures shed some light on mechanisms to boost skills gains and employability among youth. Although classified under "Unspecified main category", the New Deal for Young People programme, implemented in the United Kingdom, demonstrated the success of combining job-search assistance, a wage subsidy, on-the-job training, and sanctions to boost labour market outcomes of registered unemployed youth. The programme, introduced in the United Kingdom in 1998 to help the young unemployed into work and to increase their employability, offered multi-staged job-search assistance, followed by a menu of four tracks: training, education, wage subsidy or reinstatement in the labour market through voluntary work or environmental services. Analyses of the wage subsidy measures showed positive transitions to employment (Blundell et al., 2004) and lower probability of unemployment among programme participants (Dorsett, 2006).

In conclusion, Table 56 provides an evidence check against the expected outcomes for wage subsidy interventions outlined in the results chain.

Table 56: Evidence checks for wage subsidy interventions

Expected outcomes	Evidence checks
1. Increased probability of employment	Wage subsidy programmes did increase the probability of employment beyond the subsidy and programme duration. Design features were determinant in securing positive outcomes, particularly in middle-income countries
2. Shorter unemployment spells in the future/more efficient job search	There was also evidence to back up a consequential decrease in the probability of unemployment among youth, suggesting some efficiency gains from demonstrating high/higher productivity to employers or improvements in the job search
3. Increased ability to retain a job/longer job duration	There was no evidence to demonstrate an increased ability to retain a job or secure longer job duration after exposure to a wage subsidy programme. The few evaluations that reported on hours worked (Groh et al., 2012 and Webb, Sweetman & Warman, 2014) showed negative to no impact
4. Better quality of employment	Evidence on quality of employment is mixed. Some long-duration subsidies led to positive employment outcomes in the long-term (Caliendo, Künn & Schmidl, 2011) as well as to long-term contracts (Roger & Zamora, 2011) or fixed term contracts (Brodaty, 2007). Other schemes of shorter duration led to temporary and often unregistered jobs (e.g. Jordan NOW as reported by Groh et al., 2012)
5. Increased earnings or consumption	Overall effect sizes of earnings outcomes were smaller than those for employment, particularly among high-income countries
6. Increased returns from employment, including long-lasting human capital accumulation	There was evidence of skills formation, particularly among interventions that offered relevant jobs, sufficient exposure, or the opportunity to learn at the workplace – including through on-the-job training (Wilkinson, 2003; Blundell et al., 2004; De Giorgi, 2005)

5.1.3.4.2 Public employment interventions

Public employment programmes seek to stimulate labour demand in contexts where markets are unable to create productive employment on the required scale. In the context of youth, public employment programmes can facilitate first-time jobseekers' entry into the labour market and keep unskilled or disadvantaged youth connected to the labour market, thus mitigating skills depreciation or the negative, scarring effects of long-term unemployment.

The multi-dimensional nature of the included programmes offered scope for multiple objectives. Their connection to social protection policies also allowed the formulation of expected outcomes beyond those related to the labour market, such as consumption smoothing. Table 57 (a shortened version of Table 5), however, focuses on a list of labour market-related outcomes that included (i) more and better employment measures (probability of employment, hours worked, job quality), (ii) higher earnings and (iii) human capital accumulation (when the programme led to skills formation). Public employment programmes are complex, entailing a number of design and implementation parameters, from the selection of works and services to targeting mechanisms, wage setting, determination of benefits, work conditions and labour intensity, incentives for participation and monitoring and reporting requirements.

Table 57: Simplified results chain for interventions offering public employment programmes

Activities	Outputs	Outcomes
Public employment programmes in infrastructure development projects, social development, community works and services projects	1. Direct job creation 2. Participants (re)gain labour market contact 3. Skills formation or increased job skills (technical and non-technical) through learning by doing and exposure to the work environment a. Development of a work ethic and work habits b. Development of social skills 4. More positive attitudes towards employment/increased incentives to apply for jobs or to work 5. Improved sense of contribution to community development 6. Incentives to continue education	1. Increased probability of employment beyond the programme duration 2. Reduced time to find future job/ shorter unemployment spells 3. Increased ability to retain a job/longer job duration (hours worked) 4. Better quality of employment (contract type, a job conducive to human capital development, a salary) 5. Increased earnings or consumption 6. Increased returns from employment, including long-lasting human capital accumulation
Assumptions		
1. Complete information available about the programme for both employers and youth. 2. Target group (first time job-seekers, disadvantaged/low-skilled youth, unemployed youth and employers) participates in programme.	1. Participants are motivated to work and sufficiently qualified (adequate profiling) 2. Participants learn on the job (i.e., experience increases skill levels) 3. Programme induces (positive/expected) behavioural changes/no adverse behavioural changes 4. Appropriate targeting	1. Correct barriers and constraints for youth attempting to access the labour market are addressed 2. Work experience adequately signals higher skills and employability 3. Acquired skills/work experience match labour market demands 4. No stigmatizing effects

The evidence to support the proposed theory of change was unfortunately very sparse. During the search period, the systematic review was able to identify only two studies with public employment programmes as their main category that complied with the review's inclusion criteria. Both studies reported zero to negative treatment effects on the probability of employment after programme participation, suggesting that **public employment programmes have not effectively facilitated improvements in labour market outcomes of youth.**

Caliendo et al. (2011) assessed the impact of the German Job Creation Schemes Programme, which provided unemployed youth with secondary education with the

opportunity to work in infrastructure or social projects for a maximum of 12 months. The study found negative impacts on employment probability of young participants both in the short and long term.

Brodaty (2007) examined the French programme Travaux d'Utilité Collective (TUC), a social development and community public works project for unemployed youth. The job duration varied from three to 24 months, with contributions by both Government and employers. The study found no significant changes in employment probability compared to youth in the comparison group.

The effect sizes of these two studies fell below the overall effect size for subsidized employment interventions and also in relation to wage subsidy programmes.

Furthermore, one of the four arms during the "option" stage of the above-mentioned UK New Deal for Young People programme acted as a public employment programme. The environmental services track within the programme provided jobs for youth in housing projects, forest and park management, and reclamation of derelict or waste land. Evaluations of the New Deal showed that this particular component had limited to no impact on post-programme employment, particularly in comparison to wage subsidies, which provided a more effective means of exiting unemployment and securing unsubsidized employment (Dorsett, 2006). Similar results were found by Card et al. (2010 and 2015), where evidence that was not specifically focused on youth showed public employment programmes to be generally less successful than other types of ALMPs.

The meagre evidence on youth-targeted public employment programmes limited the discussion about what works or which design features matter most. This finding calls for further impact research on this type of intervention, particularly in low- and middle-income contexts where programme exposure may have diverse effects on youth and their families.

The search window of the systematic review missed the recording of a recent impact evaluation of a public employment programme implemented in Côte d'Ivoire by Premand, Marguerie, Crépon and Bertrand (2015). The evaluation of the Emergency Youth Employment and Skills Development project, established in 2012 to support the economic recovery following the post-electoral crisis, showed large short-term positive impacts on probability of employment and hours worked in wage occupations and positive impacts on earnings while youth were still participating in the programme, in contrast to the comparison group. While results for long-term effects are not yet available, the promising short-term results support the call for more and better evidence-gathering in developing economies.

5.2 AGREEMENTS AND DISAGREEMENTS WITH OTHER STUDIES OR REVIEWS

The effort to undertake this systematic review was initially motivated (see Section 1.4) by the statement that new evidence was needed to support decision-making on youth employment. Specifically, similar systematic reviews and studies either required urgent updating (Betcherman et al., 2007) or simply posed related, yet distinct, research questions (e.g., Card et al., 2010, 2015; Tripney et al., 2013; Grimm & Paffhausen, 2015).

The findings presented in the previous section are aligned with that motivation: the results of the empirical analysis are generally congruent with previous and related literature, but they (i) add much more depth given the rigour of the analysis and the comprehensive nature of the data, (ii) complement and carve out much more clearly the patterns indicated by previous studies, and (iii) add genuinely novel insights. So, in essence, the current study found few points of variance with related studies, but agreed on major lines, strengthening the existing knowledge base and identifying many new, detailed aspects:

1) **Agreement: The main result agreed with related studies – youth interventions are effective tools for improving labour market outcomes.** With a broader sample of target population, i.e. not only youth, Card et al., 2010 and 2015 found that ALMPs have smaller effects for older workers and youth in comparison to women and the long-term unemployed. The ability of those studies to factor in other groups offers an important insight to the results of this review: While the impact of ALMPs on youth was positive, the magnitude of reported effect sizes is smaller than ALMPs targeting all individuals, without an age target.

2) **Agreement: There was heterogeneity by programme type**, as indicated by every systematic assessment of the literature, and that entrepreneurship promotion programmes and skills training programmes were effective interventions, particularly in low- and middle-income countries. **The importance of human capital based programmes and their dynamic time horizon** – with increasing effect sizes observed over longer durations post-programme – has recently also been found in Card et al. (2015) – a pattern which was replicated in this review for youth-only interventions.

3) **Agreement:** Another pattern indicated by the previous literature (e.g., Kluve, 2010; Card et al., 2010; Betcherman et al., 2007) is that **youth labour market interventions tend to be less effective in high-income than in low-middle income countries.** Confirming this pattern was another important result of this review's meta-analysis, and it

substantially reinforced the corresponding conjectures of previous studies, by providing a basis of more comprehensive data and more profound empirical analysis.

4) **More nuanced:** While related studies also conjectured that the **more comprehensive type of interventions tended to be more successful, this result was only confirmed by the current review among low- and middle-income countries**, and not in high-income countries (agreeing, perhaps coincidentally, with Eichhorst and Rinne (2015) based on the limited information provided in the YEI data).

5) **Agreement:** Moreover, as the literature using systematic analyses of labour market programmes has grown, some indicative evidence has pointed to female participants benefitting more than male participants (echoing Card et al., 2015).

6) A **novel finding** of this meta-analysis and of the examination of the theory of change was that intervention design and implementation features tended to drive results more strongly than did the type of intervention (phrased in the Main Results as: the "how" seeming to be more important than the "what").

5.3 COMPLETENESS AND APPLICABILITY OF EVIDENCE

The evidence base on youth employment is growing and improving. While better study designs lead to lower risks of bias, limitations in the evidence still shed only partial light on what works.

1) **Insufficient consideration to spillovers and general equilibrium effects.** Exposure to ALMPs is expected to create a spillover effect among non-participants, as well as causing general equilibrium effects throughout the economy. While some of these spillovers may positively affect overall employment outcomes, in certain cases they can hamper the performance of programme non-participants. This is true of the substitution effects and windfall effects that can arise from wage subsidy programmes, which are rarely addressed in the empirical literature. Box 10 below describes a study of a wage subsidy programme in Tunisia that examined partial equilibrium effects. Accordingly, in the absence of systematic considerations of the general or partial equilibrium effects, the review's findings necessarily exhibit a degree of incompleteness and questionable external validity.

Box 10: Partial equilibrium effects: Entrepreneurship training and self-employment among university graduates in Tunisia

In Tunisia, an entrepreneurship track was introduced into the applied undergraduate (*licence appliquée*) curriculum in 2009. University students enrolled in the last year of their *licence appliquée* were invited to apply to the entrepreneurship track, which provided students with: (i) entrepreneurship courses organized by the public employment office; (ii) external private sector coaches in an industry relevant to the student's business idea; and (iii) supervision from university professors in development and finalization of the business plan. The entrepreneurship track offered students the opportunity to graduate by writing a business plan instead of a traditional undergraduate thesis. On graduation, participants were invited to submit their business plans to a competition and the competition winners became eligible to receive seed capital to establish their business.

A randomized trial aimed to identify the impact of the entrepreneurship track on beneficiaries' labour market outcomes. The study showed that the entrepreneurship track significantly increased the rate of self-employment among university graduates approximately one year after graduation, but that the effects were small in absolute terms. The employment rate among beneficiaries remained unchanged, which in partial equilibrium indicates a substitution from wage employment to self-employment. However, Almeida et al. (2012) note that the shift from wage employment into self-employment may free up job opportunities for non-participants, therefore potentially leading to higher employment overall in general equilibrium. The study design did not allow such potential general equilibrium effects to be identified.

Sources: based on information available at: www.youth-employment-inventory.org [22 Feb. 2016]; Almeida, Barouni, Brodmann, Grun and Premand, 2012.

2) **Limited reporting on the transmission channels.** The theory of change is what allows the exploration of how empirical findings in context A can be useful to decision-makers in context B. Unfortunately, studies often focus on the final outcomes and provide limited information dealing with effects on intermediate outcomes, such as changes in knowledge, skills, behaviours or attitudes. Assessing impacts on intermediate outcomes was beyond the scope of this systematic review, and coding and analysing these outcomes may be an area for future research.

3) **Insufficient consideration of cost effectiveness.** The applicability of the evidence hinges not only on its internal and external validity but also on its feasibility. Detailed analyses of costs are very limited and methods to compute net benefits and cost-benefit ratios have not yet been standardized.

5.4 QUALITY OF THE EVIDENCE

This systematic review did not undertake a full risk of bias assessment as recommended in the systematic review methodology literature. However, it relied on a framework by Duvendack et al. (2012) to assess the analytical and statistical rigour of the included studies based on the studies' design and statistical methodology. This framework correlates strongly with more sophisticated risk of bias assessments, which is why – given the results – the review team can be confident that the overall rigour of the included studies was high and, consequently, associated risks of statistical biases can be assumed to be low.

5.5 LIMITATIONS AND POTENTIAL BIASES IN THE REVIEW PROCESS

The review provides an accurate picture of the current empirical literature on labour markets impacts of youth employment interventions globally, thanks to the exhaustive search effort that allowed the identification of a number of unpublished studies during the search of grey literature.

While the review team believes that the issue of publication bias has been successfully addressed, the possibility of outcome reporting bias or file drawer effects cannot be completely excluded. In this regard, some evidence for publication bias was found (see section 4.3.6). While no correlation of reported effect sizes with publication status was found, Egger's test indicated some publication bias towards studies showing positive effects. Even though this was accounted for in the multivariate meta-regression model, the summary effect size of youth employment outcomes may represent an upper bound for the true impact of these interventions.

It is not possible to be quite as confident regarding the potential impact of missing information and reporting quality. The review team made an extensive effort to collect missing information by contacting authors using a standardized template to solicit the data required for inclusion of the study. In addition, the team employed several methods to impute missing information where possible and extensively tested the adequacy of these procedures. This allowed effect sizes for a large share of the included studies to be computed. However, the main empirical analysis was based on 2.259 of the 3.629 coded treatment effect estimates, for which it was possible to compute the SMD. While this sample is much larger than in most other systematic reviews, it remains difficult to assess the degree to which missing information may impact the empirical findings (i.e., whether reporting quality is correlated with effect size magnitudes).

One obvious limitation of the review is that it was not possible to conduct the detailed assessment of the risk of bias in included studies based on the detailed tool suggested by the Campbell Collaboration. The main reason was that most reports did not provide the information needed to objectively code the information required in the Campbell Collaboration tool. A careful assessment of research designs provided an alternative, which may nonetheless have missed important factors such as sample attrition.

Echoing other reviews in the social sciences (e.g., Tripney et al., 2013), the review found that the methods for calculating comparable effect sizes from studies using more complex multivariate econometric methods are underdeveloped and require further research. However, the review benefitted from the experience of the principal investigators and was carried out with frequent guidance from the Methods Coordinating Group of the Campbell Collaboration.

Finally, the search and selection of studies focused specifically on quantitative impact evaluations using a rigorous (quasi-) experimental design. While the team believes that this is a strength of the review, the method may have disregarded important findings from studies that were rather more qualitative in nature or did not attempt to provide causal effect estimates.

5.6 IMPLICATIONS FOR RESEARCH

The quality of the existing evidence for youth employment interventions is comparatively good, with an increasing share of experimental evaluations conducted in recent years. While this assessment of the study design could only provide a partial picture, the analysis showed a relatively high overall level of rigour of included studies: Only 9 per cent of studies were judged to have a low level of rigour, based on their research design and empirical methodology. However, clear indications for publication (or small-sample) bias were apparent, based on the sample of included studies.

A number of issues which placed limitations on this review could be mitigated with additional or improved primary research on youth employment interventions.

1) Existing research is spread unevenly across the globe. While the evidence gathered was global in nature, capturing 31 countries and all regions of the world, slightly more than half of the evidence derived from interventions in high-income countries. While it was possible to include a number of recent experimental studies from middle- and low-income countries – notably sub-Saharan Africa and Latin America and the Caribbean – there was a distinct lack of evidence from Asia, Central Europe, the Pacific, the Middle East and North Africa. Furthermore, the evaluations of youth employment

interventions in low- and middle-income countries were concentrated on rather small-scale, NGO-implemented interventions and there was a lack of evidence for larger, nationwide governmental programmes.

2) A notable observation regarding the quality of impact assessment reports is that too few studies provided evidence about heterogeneous treatment effects for different sub-groups of the interventions, such as female or low-income youth. Similarly, as significant differences in effect size magnitude by length of time since programme exit were observed, it is clear that more research is needed to (re-)assess the effectiveness of youth employment interventions in the long run.

3) More evidence and comparative analyses are needed to assess relative effectiveness across intervention components and between intervention types. The review team believes that the practitioners would greatly benefit from more evidence of interventions with multiple treatment arms which compare the effectiveness of combining different intervention design features.

4) To gain a better understanding of the employment effects on young people, it is important to further observe their transitions from the informal economy to the formal economy. The extent of informality among youth calls for further research into successful approaches to facilitate an effective transition into formal sector jobs and formalized businesses.

5) In order for systematic reviews and meta-analyses to be useful tools in economics, it would be welcome if the authors of primary studies reported the information required to standardize treatment effects across different outcome measures in a more detailed, complete, or standardized way. This relates, in particular, to the follow-up mean of the outcome variable in the control group, as well as pooled (or comparison group) standard deviations. Only 13 of the 113 reports in the initial sample provided all the information needed to compute standardized mean differences without having to contact the authors or, in a second step, impute the missing information. For another 13 reports (representing seven interventions), it was not possible to compute SMDs even after taking these steps and their findings therefore could not be included in the effect-size based quantitative meta-analysis.

6) Frequently, it would also be welcome if authors provided more detail on reporting their study design and empirical identification strategy as well as occurrence and potential causes of attrition. Based on the reported details, it was often difficult to judge the internal validity (or risk of bias) of studies due to a lack of information about potential biases, such as attrition, selection or mismeasurement.

7) Finally, the review originally set out to compare the cost-effectiveness of different intervention types. This was not possible as very few studies indicated the cost of implementation in published reports. Approaching the "how" is therefore not an easy task and much remains to be done to improve the research and reporting standards and advocate for more and better evidence about the impact of youth employment interventions.

6 References

6.1 REFERENCES TO INCLUDED REPORTS

Adoho, Franck, Shubha Chakravarty, Dala Korkoyah Jr., Mattias Lundberg, & Afia Tasneem (2014). *The impact of an adolescent girls employment program: The EPAG project in Liberia*. World Bank Policy Research Working Paper 6832.

Aedo, Cristián, & Sergio Nuñez (2004). *The impact of training policies in Latin America and the Caribbean: The case of Programa Joven*. Inter-American Development Bank, Research Department.

Aedo, Cristián, & Marcelo Pizarro Valdivia (2004). *Rentabilidad económica del programa de capacitación laboral de jóvenes Chile Joven*. Report.

Ahmed, Ali, Shubha Chakravarty, Mattias Lundberg, & Plamen Nikolov (2014). *The role of training programs for youth employment in Nepal: Interim impact evaluation report on the Employment Fund*.

Almeida, Rita K., Mahdi Barouni, Stefanie Brodmann, Rebekka Grun, & Patrick Premand (2012). *Entrepreneurship training and self-employment among university graduates: Evidence from a randomized trial in Tunisia*. The World Bank.

Alvares de Azevedo, Thomaz, Jeff Davis, & Munene Charles (2013). *Testing what works in youth employment: Evaluating Kenya's Ninaweza program*. Volume 1: A Summative Report and Volume 2: Appendices. Global Partnership for Youth Employment (GPYE).

Alzuá, María L., & Pablo Brassiolo (2006). *The impact of training policies in Argentina: An evaluation of Proyecto Joven*. Inter-American Development Bank, Office of Evaluation and Oversight (OVE).

Alzúa, María L., Paula Nahirñak, & Belissario Alvarez de Toledo (2007). *Evaluation of Entra 21 using quantitative and qualitative data*. Q-Squared Working Paper No. 41.

Alzúa, María L., Guillermo Cruces, & Carolina Lopez Erazo (2013). *Youth training programs beyond employment. Evidence from a randomized controlled trial.*

Arandarenko, Mihail, Aleksandra Nojkovic, & Marko Vladisavljevic (2014). *Impact evaluation of employment programmes targeting disadvantaged youth in Serbia*. Foundation for the Advancement of Economics.

Attanasio, Orazio, Adriana Kugler, & Costas Meghir (2011). Subsidizing vocational training for disadvantaged youth in Colombia: Evidence from a randomized trial. *American Economic Journal: Applied Economics*, *3*, 3, 188–220.

Bampasidou, Maria (2012). "Unbundling the degree effect in a job training program for disadvantaged youths." PhD, University of Florida.

Bampasidou, Maria, Carlos A. Flores, Alfonso Flores-Lagunes, & Daniel J. Parisian (2014). The role of degree attainment in the differential impact of job corps on adolescents and young adults. In *Factors affecting worker well-being: the impact of change in the labor market* (*Research in Labor Economics*, Volume 40, 113–156). Emerald Group Publishing Limited.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, & Munshi Sulaiman (2014). *Women's empowerment in action: Evidence from a randomized control trial in Africa*.

Barza, Rachael. 2011. "Essays in labor economics." PhD, Harvard University.

Blanco, German, Carlos A. Flores, & Alfonso Flores-Lagunes (2013a). Bounds on average and quantile treatment effects of Job Corps training on wages. *Journal of Human Resources*, *48*, 3, 659–701.

Blanco, German, Carlos A. Flores, & Alfonso Flores-Lagunes (2013b). The effects of Job Corps training on wages of adolescents and young adults. *American Economic Review*, *103*, 3, 418–422.

Blasco, Jaume, David Casado, Jordi Sanz, & Federico A. Todeschini (2014). Evaluación del impacto de los programas PCPI del Servicio de Empleo de Catalunya. *Revista de evaluación de programas y políticas públicas/Journal of public programs and policy evaluation (REPPP)*, *2*, 187–217.

Blattman, Christopher, Jeannie Annan, Eric Green, & Julian Jamison (2013). *Building women's economic and social empowerment through enterprise. An experimental assessment of the Women's Income Generating Support (WINGS) Program in Uganda*. Innovations for Poverty Action (IPA).

Blattman, Christopher, Eric P. Green, Julian Jamison, & Jeannie Annan (2014). *Employing and empowering marginalized women: A randomized trial of microenterprise assistance*.

Blattman, Christopher, Nathan Fiala, & Sebastian Martinez (2013). Generating skilled self-employment in developing countries: Experimental evidence from Uganda. *The Quarterly Journal of Economics*, forthcoming.

Bloom, Howard S., Larry L. Orr, Stephen H. Bell, George Cave, Fred Doolittle, Winston Lin, & Johannes M. Bos (1997). The benefits and costs of JTPA Title II-A programs: Key findings from the National Job Training Partnership Act study. *Journal of Human Resources*, *32*, 3, 549–576.

Blundell, Richard, Monica Costa Dias, Costas Meghir, & John Van Reenen (2004). Evaluating the employment impact of a mandatory job search program. *Journal of the European Economic Association*, *2*, 4, 569–606.

Brodaty, Thomas (2007). La politique active de l'emploi en faveur des jeunes: Les dispositifs ont-ils touche leur cible? (Are French youth employment active labour market policies well targeted? With English summary). *Annales d'Economie et de Statistique*, *85*, 3–40.

Broecke, Stijn (2013). Tackling graduate unemployment in North Africa through employment subsidies: A look at the SIVP programme in Tunisia. *IZA Journal of Labor Policy*, *2*, 9, 1–19.

Bruhn, Miriam, & Bilal Zia (2013). Stimulating managerial capital in emerging markets: The impact of business training for young entrepreneurs. *Journal of Development Effectiveness*, *5*, 2, 232–266.

Calero, Carla, Carlos Henrique Corseuil, Veronica Gonzales, Jochen Kluve, & Yuri Soares (2014). *Can arts-based interventions enhance labor market outcomes among youth? Evidence from a randomized trial in Rio de Janeiro*.

Caliendo, Marco, Steffen Künn, & Ricarda Schmidl (2011). *Fighting youth unemployment: The effects of active labor market policies*. Forschungsinstitut zur Zukunft der Arbeit/Institute for the Study of Labor (IZA).

Cansino Muñoz-Repiso, José Manuel, & Antonio Sánchez Braza (2011). Effectiveness of public training programs reducing the time needed to find a job/Eficacia de los programas públicos de formación en la reducción del tiempo necesario para encontrar un empleo. *Estudios de Economía Aplicada*, *29*, 1, 1–26.

Card, David, Pablo Ibarrarán, Ferdinando Regalia, David Rosas-Shady, & Yuri Soares (2011). The labor market impacts of youth training in the Dominican Republic. *Journal of Labor Economics*, *29*, 2, 267–300.

Carling, Kenneth, & Laura Larsson (2005). Does early intervention help the unemployed youth? *Labour Economics*, *12*, 3, 301–319.

Cave, George, Hans Bos, Fred Doolittle, & Cyril Toussaint (1993). *Jobstart: Final Report on a Program for School Dropouts*.

Centeno, Luis, Mario Centeno, & Alvaro A. Novo (2009). Evaluating job-search programs for old and young individuals: Heterogeneous impact on unemployment duration. *Labour Economics*, *16*, 1, 12–25.

Centeno, Mário, & Álvaro A. Novo (2006). Avaliação do impacto de um programa de apoio à procura de emprego. *Boletim Económico, Banco de Portugal* Verão.

Chang, Young Eun, Aletha C. Huston, Danielle A. Crosby, & Lisa A. Gennetian (2007). The effects of welfare and employment programs on children's participation in Head Start. *Economics of Education Review*, *26*, 1, 17–32.

Charpail, Christine, Tristan Klein, & Serge Zilberman (2005). Évaluation des politiques d'emploi: La deuxième génération des panels des bénéficiaires. *DARES Document d'Etudes*, *95*.

Chen, Xuan (2013). "Partial identification of average treatment effects in program evaluation: Theory and applications." PhD, Economics, University of Miami.

Cho, Yoon Y., Davie Kalomba, Ahmed Mushfiq Mobarak, & Victor Orozco (2013). *Gender differences in the effects of vocational training: Constraints on women and drop-out behavior*. World Bank.

Chong, Alberto, & Jose Galdo (2006). *Training quality and earnings: The effects of competition on the provision of public-sponsored training programs*. Inter-American Development Bank (mimeo).

Chong, Alberto, & Jose Galdo (2012). Does the quality of public-sponsored training programs matter? Evidence from bidding processes data. *Labour Economics*, *19*, 970–986.

Chong, Alberto, Jose Galdo, & Jaime Saavedra (2008). Informality and productivity in the labor market in Peru. *Journal of Economic Policy Reform*, *11*, 4, 229–245.

Corseuil, Carlos H., Miguel Foguel, Gustavo Gonzaga, & Eduardo Pontual Ribeiro (2014). *Youth turnover in Brazil: Job and worker flows and an evaluation of a youth-targeted training program*.

Costa Dias, Monica, Hidehiko Ichimura, & Gerard J. van den Berg (2013). Treatment evaluation with selective participation and ineligibles. *Journal of the American Statistical Association*, *108*, 502, 441–455.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, & Philippe Zamora (2013). Do labor market policies have displacement effects? Evidence from a

clustered randomized experiment. *The Quarterly Journal of Economics*, *128*, 2, 531–580.

Crépon, Bruno, Esther Duflo, Elise Huillery, William Parienté, & Juliette Seban (2014). *Les effets du dispositif d'accompagnement à la création d'entreprise CréaJeunes: Résultats d'une expérience contrôlée*. Rapport d'évaluation.

De Giorgi, Giacomo (2005). The New Deal for Young People five years on. *Fiscal Studies*, *26*, 3, 371–383.

Díaz, Juan J., & Miguel Jaramillo (2006). *An evaluation of the Peruvian "youth labor training program" – Projoven*. Working Paper: OVE/WP-10/06. Inter-American Development Bank.

Egebark, Johan, & Niklas Kaunitz (2014). *Do payroll tax cuts raise youth employment?* Research Institute of Industrial Economics (IFN).

Ehlert, Christoph, Jochen Kluve, & Sandra Schaffner (2012). Temporary work as an active labor market policy: Evaluating an innovative program for disadvantaged youths. *Economics Bulletin*, *32*, 2, 1765–1773.

Elías, Víctor, Fernanda Ruiz Núñez, Ricardo D. Cossa, & D. Bravo (2004). *An econometric cost-benefit analysis of Argentina's youth training program*. Inter-American Development Bank.

Espinoza Peña, Henry (2010). "Impact evaluation of a job-training programme for disadvantaged youths: The case of Projoven." PhD, Maastricht University.

Fein, David, Erik Beecroft, & John D. Blomquist (1994). *The Ohio Transitions to Independence demonstration: Final impacts for jobs and work choice*. Abt Associates.

Fiala, Nathan (2014). *Stimulating microenterprise growth: Results from a loans, grants and training experiment in Uganda*.

Flores-Lagunes, Alfonso, Arturo Gonzalez, & Todd Neumann (2010). Learning but not earning? The impact of Job Corps training for Hispanic youth. *Economic Inquiry*, *48*, 3, 651–667.

Franklin, Simon (2014). *Job search, transport costs and youth unemployment: Evidence from urban Ethiopia*.

Frölich, Markus, & Martin Huber (2014). *Direct and indirect treatment effects: Causal chains and mediation analysis with instrumental variables*. IZA Discussion Paper No. 8280.

Frumento, Paolo, Fabrizia Mealli, Barbara Pacini, & Donald B. Rubin (2012). Evaluating the effect of training on wages in the presence of noncompliance,

nonemployment, and missing outcome data. *Journal of the American Statistical Union*, *107*, 498, 450–466.

Galdo, Jose, Miguel Jaramillo, & Veronica Monatalva (2008). *Household wealth and heterogeneous impacts of a market-based training program: The case of PROJOVEN in Peru*.

General Accounting Office (1996). *Job Training Partnership Act: Long-term earnings and employment outcomes*. United Nations General Accounting Office.

Gong, Tao (2005). "Two essays on youths' labor supply." PhD, Economics and Finance, Middle Tennessee State University.

Griffith, James (2001). An approach to evaluating school-to-work initiatives: Post-secondary activities of high school graduates of work-based learning. *Journal of Vocational Education and Training*, *53*, 1, 37–60.

Groh, Matthew, Nandini Krishnan, David McKenzie, & Tara Vishwanath (2012). *Soft skills or hard cash? The impact of training and wage subsidy programs on female youth employment in Jordan*. World Bank, Policy Research Working Paper Series No. 6141.

Groh, Matthew, David McKenzie, Nour Shammout, & Tara Vishwanath (2014). *Testing the importance of search frictions, matching, and reservation prestige through randomized experiments in Jordan*. World Bank – Impact Evaluation series. Policy Research Working Paper No. WPS 7030.

Hall, Georgia S. (2000). "The impact of a school-to-work program ProTech: A study of post-high school outcomes." PhD, Boston College.

Hämäläinen, Kari, Ulla Hämäläinen, & Juha Tuomala (2014). *The labour market impacts of Youth Guarantee: Evidence from a Finnish reform*.

Heckman, James J., & Jeffrey Smith (1999). The pre-programme earnings dip and the determinants of participation in a social programme: Implications for simple programme evaluation strategies. *The Economic Journal*, *109*, 457, 313–348.

Heckman, James J., & Jeffrey Smith (2000). The sensitivity of experimental impact estimates (evidence from the national JTPA study). In David G. Blanchflower and Richard B. Freeman (Eds.), *Youth employment and joblessness in advanced countries* (pp. 331–356). National Bureau of Economic Research.

Heckman, James J., Hidehiko Ichimura, & Petra E. Todd (1997). Matching as an econometric evaluation estimator: Evidence from evaluating a job training programme. *Review of Economic Studies*, *64*, 4, 605–654.

Hicks, Joan Hamory, Michael Kremer, Isaac Mbiti, & Edward Miguel (2013). *Vocational education in Kenya: Evidence from a randomized evaluation among youth*.

Hollenbeck, Kevin, & W. J. Huang (2006). *Net impact and benefit-cost estimates of the workforce development system in Washington State*. W.E. Upjohn Institute for Employment Research.

Hollenbeck, Kevin, & Richard J. Willke (1991). *The employment and earnings impacts of the targeted Jobs Tax Credit*. W.E. Upjohn Institute for Employment Research, Staff Working Papers.

Ibarrarán, Pablo, Laura Ripani, Bibiana Taboada, Juan M. Villa, & Brígida García (2014). Life skills, employability and training for disadvantaged youth: Evidence from a randomized evaluation design. *IZA Journal of Labor & Development*, *3*, 10.

Ibarrarán, Pablo, & David Rosas (2007). *Impact evaluation of a labor training program in Panama*.

Jackson, Russell H., R. Malené Dixon, Ann McCoy, Carol Pistorino, Paul Zador, Cynthia Thomas, John Lopdell, Juanita Lucas-McLean, Frank Bennici, Andy Sum, Neeta Fogg, Ron D'Amico, Andrew Weigand, & Lee Bruno (2007). *Youth Opportunity Grant initiative: Impact and synthesis report*. US Department of Labor, Decision Information Resources.

Jaramillo, Miguel, & Sandro Parodi (2003). *Jóvenes emprendedores*. Lima, Instituto Apoyo.

Jaramillo, Miguel, & Sandro Parodi (2005). *Evaluación experimental del programa "Formación de Líderes Empresariales en el Departamento de Huancavelica*."

Jensen, Robert (2012). Do labor market opportunities affect young women's work and family decisions? Experimental evidence from India. *The Quarterly Journal of Economics*, *127*, 2, 753–792.

Knight, G. (2002). "Evaluation of the Australian wage subsidy Special Youth Employment and Training Program, SYETP." PhD thesis, University of Sydney.

Kornfeld, Robert, & Howard S. Bloom (1999). Measuring program impacts on earnings and employment: Do unemployment insurance wage reports from employers agree with surveys of individuals? *Journal of Labor Economics*, *17*, 1, 168–197.

Larsson, Laura (2003). Evaluation of Swedish youth labor market programs.

Journal of Human Resources, *38*, 4, 891–927.

Lee, David S. (2009). Training, wages, and sample selection: Estimating sharp bounds on treatment effects. *Review of Economic Studies*, *76*, 1071–1102.

Levinsohn, James, Neil Rankin, Gareth Roberts, & Volker Schoer. 2014. *Wage subsidies and youth employment in South Africa: Evidence from a randomised control trial.*

Maitra, Pushkar, & Subha Mani (2014). *Learning and earning: Evidence from a randomized evaluation in India*. GCC Working Paper Series, CC 14-05.

Maynard, Rebecca, Walter Nicholson, & Anu Rangarajan (1993). *Breaking the cycle of poverty: The effectiveness of mandatory services for welfare-dependent teenage parents*. Mathematica Policy Research.

McClanahan, Wendy S., Cynthia L. Sipe, & Thomas J. Smith (2004). *Enriching summer work: An evaluation of the Summer Career Exploration Program*. Public/Private Ventures.

Meager, Nigel, Peter Bates, & Marc Cowling (2003). An evaluation of business start-up support for young people. *National Institute Economic Review*, *186*, 1, 59–72.

Mensch, Barbara S., Monica J. Grant, Mary P. Sebastian, Paul C. Hewett, & Dale Huntington (2004). *The effect of a livelihoods intervention in an urban slum in India: Do vocational counseling and training alter the attitudes and behavior of adolescent girls?* Working Paper No. 194. The Population Council.

Millenky, Megan, Dan Bloom, Sara Muller-Ravett, & Joseph Broadus (2011). *Staying on course: Three-year results of the National Guard Youth ChalleNGe Evaluation*. MDRC.

Miller, Cynthia, Kristin E. Porter, Fannie M. Tseng, Yasuyo Abe, & Johannes M. Bos (2005). *The challenge of repeating success in a changing world: Final report on the Center for Employment Training Replication Sites*.

Neumark, David, & Donna Rothstein (2006). School-to-career programs and transitions to employment and higher education. *Economics of Education Review*, *25*, 4, 374–393.

Ñopo, Hugo, Miguel Robles, & Jaime Saavedra (2008). Occupational training to reduce gender segregation: The impacts of ProJoven. *Revista Economía*, *62*, 33–54.

Ñopo, Hugo, & Jaime Saavedra (2003). *Evaluación del impacto de mediano plazo de Projoven: Resultados de las mediciones realizadas a los seis, doce y*

dieciocho meses de culminado el programa. GRADE report.

Pessoa e Costa, Sofia, & Stéphane Robin (2009). *An illustration of the returns to training programmes: The evaluation of the "Qualifying Contract" in France*. Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Quint, Janet C., Johannes M. Bos, & Denise F. Polit (1997). *New chance: Final report on a comprehensive program for young mothers in poverty and their children*. Manpower Demonstration Research Corporation.

Richardson, James (1998). *Do wage subsidies enhance employability? Evidence from Australian youth*. Centre for Economic Performance Working Paper No. 387. London School of Economic and Political Science.

Rodríguez-Planas, Núria (2012). Longer-term impacts of mentoring, educational services, and learning incentives: Evidence from a randomized trial in the United States. *American Economic Journal: Applied Economics*, *4*, 4, 121–39.

Roger, Muriel, & Philippe Zamora (2011). Hiring young, unskilled workers on subsidized open-ended contracts: A good integration programme? *Oxford Review of Economic Policy*, *27*, 2, 380–396.

Saniter, Nils Peter (2014). "Educational interventions and labor market success." PhD, Freie Universität Berlin, Germany.

Santa María, Mauricio, Paula Acosta, Alexander Estacio, Natalia Millán, Mauricio Olivera, Mónica Parra Torrado, Carlos Alberto Castañeda, Yanira Marcela Oviedo, and Carlos Andrés Pérez (2009a). *Evaluación de impacto de tres programas y línea de base de un programa del servicio nacional de aprendizaje – SENA*. Informe Final. Fedesarrollo – Centro de Investigación Ecónomica y Social.

Santa María, Mauricio, Mauricio Olivera, Paula Acosta, Tatiana Vásquez, & Anwar Rodríguez (2009b). *Evaluación de impacto programa Jóvenes con Futuro*. Working Paper. Fedesarrollo – Centro de Investigación Ecónomica y Social.

Santiago Consultores Asociados (1999). *Evaluacion Ex-post Chile Joven Fase II*. Report.

Schneider, Hilmar, Arne Uhlendorff, & Klaus F. Zimmermann (2011). *Ökonometrie vs. Projektdesign: Lehren aus der Evaluation eines Modellprojekts zur Umsetzung des Workfare-Konzepts* (Econometrics vs. project design: Lessons from the evaluation of a Workfare project. With English summary). IZA Discussion Paper No. 5599.

Schochet, Peter, John Burghardt, & Sheena McConnell (2008). Does Job Corps work? Impact findings from the national Job Corps study. *American Economic Review*, *98*, 5, 1864–1886.

Schochet, Peter, Sheena McConnell, & John Burghardt (2003). *National Job Corps study: Findings using administrative earnings records data*. Final Report. Mathematica Policy Research.

SENCE (2008). *Evaluación de impacto del programa de Formación en Oficios para Jóvenes de Escasos Recursos, Ejecución 2005*. Informe Final.

SENCE: Mónica Riquelme Peña, Margarita Cifuentes, Bárbara Morales, & Álvaro Espinoza (2006). *Informe Final Preliminar Estudio: Evaluación de Impacto Programa Aprendices*.

Steiner, Roberto, Paula Acosta, Norberto Rojas, Natalia Millán, Carlos Castañeda, & José L. González (2010). *Evaluación de impacto del programa jóvenes rurales emprendedores del servicio nacional de aprendizaje – SENA*.

Universidad de Chile (2012). *Evaluación de Impacto del Programa de Subsidio al Empleo Joven*. Centro de Microdatos, Departamento de Economía, Universidad de Chile.

van Gameren, Edwin (2010). *Evaluación de impacto del programa de apoyo al empleo*. Centro de Estudios Económicos El Colegio de México.

Villalobos Barría, Carlos, & Stephan Klasen (2014). *The impact of SENAI's vocational training program on employment, wages, and mobility in Brazil: What lessons for Sub Saharan Africa?* International Research Initiative on Brazil and Africa (IRIBA).

Webb, Matthew, Arthur Sweetman, & Casey Warman (2014). *How targeted is targeted tax relief? Evidence from the Unemployment Insurance Youth Hires Program*. Queen's University, Department of Economics.

Wilkinson, David (2003). *New Deal for Young People: Evaluation of unemployment flows*. Policy Studies Institute, Research Discussion Paper 15.

Zhang, Junni L., Donald B. Rubin, & Fabrizia Mealli (2009). Likelihood-based analysis of causal effects of job-training programs using principal stratification. *Journal of the American Statistical Association*, *104*, 485, 166–176.

6.2 REFERENCES TO EXCLUDED REPORTS[44]

Examples of excluded reports and the reasons for excluding them are presented in the following table:

Reference of excluded report	Main reason for exclusion
Blanchflower, D. G., & Lynch, L. M. (1994). Training at work: A comparison of US and British youths. In *Training and the private sector: international comparisons* (pp. 233–260). University of Chicago Press.	Study design
Bonnal, L., Fougere, D., & Sérandon, A. (1997). Evaluating the impact of French employment policies on individual labour market histories. *The Review of Economic Studies*, *64*, 4, 683–713.	Study design
Dehejia, R. H., & Wahba, S. (1999). Causal effects in nonexperimental studies: Reevaluating the evaluation of training programs. *Journal of the American Statistical Association*, *94*, 448, 1053–1062.	Target group
De Mel, S., McKenzie, D., & Woodruff, C. (2014). Business training and female enterprise start-up, growth, and dynamics: Experimental evidence from Sri Lanka. *Journal of Development Economics*, *106*, 199–210.	Target group
Arcand, J. L., Dyer, P., Puerto Gonzalez, S., Gardiner, D., & Garbouj, M. (2013). MEDA Maroc's 100 Hours to Success impact evaluation: Baseline study.	Study ongoing
Pessoa e Costa, S., & Robin, S. (2007). The impact of training programmes on wages in France: An evaluation of the 'qualifying contract' using propensity scores (No. 2007-18). *Bureau d'Economie Théorique et Appliquée, UDS, Strasbourg.*	More recent study available
Costa Dias, M., Ichimura, H., & Van den Berg, G. J. (2008). The matching method for treatment evaluation with selective participation and ineligibles. *Available at SSRN 1136442*.	More recent study available
Dorsett, R. (2006). The New Deal for Young People: Effect on the labour market status of young men. *Labour Economics*, *13*, 3, 405–422.	Only relative effects
Turner, S. L., & Conkel, J. L. (2010). Evaluation of a career development skills intervention with adolescents living in an inner city. *Journal of Counseling & Development*, *88*, 4, 457–465.	Outcomes

[44] Further information about excluded reports is available from the authors on request.

6.3 ADDITIONAL REFERENCES

3ie (2013). *Impact evaluation glossary*. Available at: http://ridie.3ieimpact.org/downloads/Impact%20Evaluation%20Glossary.pdf [26 Feb. 2016].

Almeida, R., Orr, L., & Robalino, D. (2014). Wage subsidies in developing countries as a tool to build human capital: Design and implementation issues. *IZA Journal of Labor Policy*, *3*, 12.

Auer, P., Efendioglu, U., & Leschke, J. (2008). *Active labour market policies around the world: Coping with consequences of globalization*. Geneva, ILO.

Baird, S., Ferreira, F. H. G., Özler, B., & Woolcock, M. (2013). *Relative effectiveness of conditional and unconditional cash transfers for schooling outcomes in developing countries: A systematic review*. Campbell Systematic Reviews, 8.

Bell, D., & Blanchflower, D. (2009). *What should be done about rising unemployment in the UK?* IZA Discussion Paper No. 4040.

Betcherman, G., Dar, A., & Olivas, K. (2004). *Impacts of active labor market programs: New evidence from evaluations with particular attention to developing and transition countries*. Social Protection Discussion Paper Series, No. 0402. Washington, DC, World Bank.

Betcherman, G., Godfrey, M., Puerto, S., Rother, F., & Stavreska, A. (2007). *A review of interventions to support young workers: Findings of the Youth Employment Inventory*. Social Protection Discussion Paper Series, No. 715. Washington, DC, World Bank.

Biavaschi, C., Eichhorst, W., Giulietti, C., Kendzia, M. J., Muravyev, A., Pieters, J., Rodriguez Planas, N., Schmidl, R., & Zimmermann, K. F. (2012). *Youth unemployment and vocational training*. IZA Discussion Paper No. 6890.

Bloom, H. (2006). *The core analytics of randomized experiments for social research*. MDRC Working Papers on Research Methodology.

Boone, J., & van Ours, J.C. (2004). *Effective active labour market polices*. IZA Discussion Paper No. 1335.

Bonnal, L., Fougere, D., & Sérandon, A. (1997). Evaluating the impact of French employment policies on individual labour market histories. *Review of Economic Studies*, *64*, 4, 683–713.

Bördős, K., Csillag, M., & Scharle, Á. (2016). What works on wage subsidies for young people: A review of issues, theory, policies and evidence. Employment

Working Paper No. 199. Geneva, ILO. Available at http://www.ilo.org/wcmsp5/groups/public/---ed_emp/documents/publication/wcms_466538.pdf [05 Apr. 2016].

Borenstein, M., Cooper, H., Hedges, L. V., & Valentine, J. C. (2009). Effect sizes for continuous data. In H. Cooper, V. Hedges & J. C. Valentine (Eds.), *The handbook of research synthesis and meta-analysis* (pp. 221–235). New York, NY, Russell Sage Foundation.

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

Brewer, L. (2013). *Enhancing youth employability: What? Why? and How? Guide to core work skills*. Geneva, International Labour Office, Skills and Employability Department.

Brodaty, T., Crépon, B., & Fougère, D. (2000). *Using matching estimators to evaluate youth employment programs: Evidence from France, 1986–1988*. Econometric Society World Congress 2000 Contributed Paper No. 0576. New York, NY, Econometric Society.

The Campbell Collaboration. (2014). *Campbell Collaboration Systematic Reviews: Policies and Guidelines* Campbell Systematic Reviews 2014: Supplement 1. DOI: 10.4073/csrs.2014.1 Available at http://www.campbellcollaboration.org/artman2/uploads/1/C2_Policy_Document_V1.0_May_2014.pdf [26 Feb. 2016].

Card, D., Kluve, J., & Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, *120*, 548, F452–F477.

Card, D., Kluve, J., & Weber, A. (2015). *What works? A meta analysis of recent active labour market programme evaluations*. NBER Working Paper 21431.

Cho, Y., & Honorati, M. (2013). *Entrepreneurship programs in developing countries: A meta regression analysis*. World Bank Policy Research Working Paper No. 6402. Washington, DC, World Bank.

Cho, Y., Margolis, D. N., Newhouse, D., & Robalino, D. A. (2012). *Labor markets in low- and middle-income countries: Trends and implications for social protection and labor policies*. Human Development Social Protection Discussion Paper Series. Washington, DC, World Bank.

Cho, Y., Robalino, D., & Watson, S. 2014. *Supporting self-employment and small-scale entrepreneurship: Potential programs to improve livelihoods for vulnerable workers*. Washington, DC, World Bank.

Cunningham, W., McGinnis, L., Verdú, R. G., Tesliuc, C., & Verner, D. (2008). *Youth at risk in Latin America and the Caribbean*. Washington, DC, World Bank.

Cunningham, W., Sanchez-Puerta, M. L., & Wuermli, A. (2010). *Active labor market programs for youth: A framework to guide youth employment interventions*. Washington, DC, World Bank.

Dar, A., & Tzannatos, Z. (1999). *Active labor market programs: A review of the evidence from evaluations*. Social Protection Discussion Paper Series No. 9901. Washington, DC, World Bank.

Deeks, J. J., Altman, D. G., & Bradburn, M. J. (2001) Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis, in M. Egger, G. D. Smith and D. G. Altman (eds), *Systematic reviews in health care: Meta-analysis in context*, second edition, BMJ Publishing Group, London.

DerSimonian, R., & Laird, N. (1986). Meta-analysis in clinical trials. *Controlled Clinical Trials*, *7*, 3, 177–188

Dorsett, R. (2006). The new deal for young people: Effect on the labour market status of young men. *Labour Economics*, *13,* 3, 405–422.

Doucouliagos, C. H., & Stanley, T. D. (2009). Publication selection bias in minimum-wage research? A meta-regression analysis. *British Journal of Industrial Relations*, *47*, 2, 406–428.

Doucouliagos, C. H., Stanley, T. D., & Giles, M. (2012). Are estimates of the value of a statistical life exaggerated? *Journal of Health Economics*, *31*, 1, 197–206.

Duvendack, M., Hombrados, J. G., Palmer-Jones, R., & Waddington, H. (2012). Assessing "what works" in international development: Meta-analysis for sophisticated dummies. *Journal of Development Effectiveness*, *4*, 3, 456–471.

Duvendack, M., Palmer-Jones, R., Copestake, J. G., Hooper, L., Loke, Y., & Rao, N. (2011). *What is the evidence of the impact of microfinance on the well-being of poor people?* London, EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Egger, M., Davey Smith, G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, *315*, 629–634.

Eichhorst, W., & Rinne, U. 2015. *An assessment of the Youth Employment Inventory and implications for Germany's development policy*. IZA Research Report No. 67.

Fares, J., & Puerto, O. S. (2009). *Towards comprehensive training*. Social Protection Discussion Paper No. 0924. Washington, DC, World Bank.

Fields, G. S. (2011). Labor market analysis for developing countries. *Labour Economics*, *18*, S16–S22.

Franco, A., Malhotra, N., & Simonovits, G. (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science*, *345*, 6203, 1502–1505.

Galasso, E., Ravallion, M., & Salvia, A. (2001). *Assisting the transition from workfare to work: Argentina's Proempleo experiment*. Washington, DC, World Bank.

Gindling, T. H., & Newhouse, D. (2012). *Self-employment in the developing world*. World Bank Policy Research Working Paper No. 6201. Washington, DC, World Bank.

Grimm, M., & Paffhausen, A. L. (2015). Do interventions targeted at micro-entrepreneurs and small and medium-sized firms create jobs? A systematic review of the evidence for low and middle income countries. *Labour Economics*, *32*, 67–85.

Grosh, M., del Ninno, C., Tesliuc, E., & Ouerghi, A. (2008). *For protection and promotion: The design and implementation of effective safety nets*. Washington, DC, World Bank.

Heckman, J. J., LaLonde, R. J., & Smith, J. A. (1999). The economics and econometrics of active labour market programs, in O. Ashenfelter and D. Card (Eds.), *Handbook of Labour Economics*, *3*, Elsevier, Amsterdam.

Heckman, J., Lochner, L., & Cossa, R. 2002. *Learning-by-doing vs. on-the-job training: Using variation induced by the EITC to distinguish between models of skill formation*. Working Paper No. 9083. National Bureau of Economic Research.

Hedges, L., & Olkin, I. (1985). *Statistical models for meta-analysis*. New York: Academic Press.

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, *1*, 39–65.

Hedges, L. V., & Pigott, T. D. (2001). The power of statistical tests in meta-analysis. *Psychological Methods*, *6*, 203–217.

Higgins, J., & Green, S. (2011). *Cochrane handbook for systematic reviews of interventions*. Version 5.1.0 [updated March 2011]. Available at: http://handbook.cochrane.org/ [22 Feb. 2016].

Ibarrarán, P., & Rosas Shady, D. (2009). Evaluating the impact of job training programmes in Latin America: Evidence from IDB funded operations. *Journal of Development Effectiveness*, *1*, 2, 195–216. Available at: http://econpapers.repec.org/article/tafjdevef/v_3a1_3ay_3a2009_3ai_3a2_3ap_3a195-216.htm [30 Mar. 2016].

Imbens, G. W., & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*, 1, 5–86.

Independent Evaluation Group (IEG). (2013). *World Bank and IFC support for youth employment programs*. Washington, DC, World Bank.

ILO (2012). *Global Employment Trends for Youth 2012*. Geneva, International Labour Office.

ILO (2015a). *Global Employment Trends for Youth 2015: Scaling up investments in decent jobs for youth*. Geneva, International Labour Office.

ILO (2015b). *Small and medium-sized enterprises and decent and productive employment creation*. Report IV to the International Labour Conference, 104th Session, 2015. Geneva, International Labour Office.

J-PAL. (2013). *J-PAL Youth Initiative Review Paper*. Cambridge, MA, Abdul Latif Jameel Poverty Action Lab.

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

Kluve, J. (2006). *The effectiveness of European active labor market policy*. IZA Discussion Paper No. 2018.

Kluve, J. (2010). The effectiveness of European active labour market programmes. *Labour Economics*, *17*, 6, 904–918.

Kluve, J., Lehmann, H., & Schmidt, C. M. (1999). *Active labour market policies in Poland: Human capital enhancement, stigmatization or benefit churning*. IZA Discussion Paper No. 30.

Kluve, J., Puerto, S., Robalino, D., Rother, F., Weidenkaff, F., Stoeterau, J., Romero, J. M., & Witte, M. (2014). *Protocol: Interventions to improve labour market outcomes of youth: A systematic review of training, entrepreneurship promotion, employment services, mentoring, and subsidized employment interventions*. The Campbell Collaboration.

Kluve, J., & Schmidt, C. M. (2002). Can training and employment subsidies combat European unemployment? *Economic Policy*, *17*, 35, 409–448.

Knapp, G., & Hartung, J. (2003). Improved tests for a random effects meta-regression with a single covariate. *Statistics in Medicine*, *22*, 17, 2693–2710

Kritikos, A. (2014). *Entrepreneurs and their impact on jobs and economic growth*. IZA World of Labor, 8 pp.

Leroy, J. L., Gadsden, P., & Guijarro, M. (2012). The impact of daycare programmes on child health, nutrition and development in developing countries: A systematic review. *Journal of Development Effectiveness*, *4*, 3, 472–496.

Lieuw-Kie-Song, M., Philip, K., Tsukamoto, M., & Van Imschoot, M. (2010). *Toward the right to work: Innovations in Public Employment Programmes (IPEP)*. Geneva, International Labour Office.

Lieuw-Kie-Song, M., Puerto, S., & Tsukamoto, M. (Forthcoming). Boosting youth employment through public works. Employment Working Paper. Geneva, Employment Policy Department, International Labour Office.

Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Applied Social Research Methods Series, Vol. 49. SAGE Publications.

Majurin, E. (2014). *Start and Improve Your Business (SIYB): Implementation guide*. Geneva, International Labour Office.

Manpower Group (2013). *Talent shortage survey research results*. Milwaukee, WI, Manpower Group

Mishra, A., & Cameron, D. (2013). *Register of Impact Evaluation Published Studies (RIEPS) Protocol*. Washington, DC, The International Initiative for Impact Evaluation.

Nataraj, S., Perez-Arce, F., Kumar, K. B., & Srinivasan, S. V. (2013). The impact of labor market regulation on employment in low-income countries: A meta-analysis. *Journal of Economic Surveys*, *28*, 3, 551–572.

Neumark, D., & Grijalva, D. 2013. *The employment effects of state hiring credits during and after the Great Recession*. Working Paper No. 18928. National Bureau of Economic Research.

O'Leary, J. C. (1998). *Evaluating the effectiveness of active labor programs in Poland*. Upjohn Institute Technical Report No. 98-012.

Organisation for Economic Co-operation and Development (OECD). (2013). *Glossary of statistical terms, Active Labour Market Programmes*. Available at: http://stats.oecd.org/glossary/detail.asp?ID=28 [26 Feb. 2016].

Pasali, S. (2015). *Youth employment in Africa: What works and why?* Mimeo.

Piza, C., Cravo, T., Taylor, L., Gonzalez, L, Musse, I., Furtado, I., Sierra, A. C., & Abdelnour S. (2016). *The impact of business support services for small and medium enterprises on firm performance in low- and middle-income countries: A systematic review*. Campbell Systematic Reviews, 1.

Premand, P., Marguerie, A., Crépon, B., Bertrand, M. (2015). Projet Emploi Jeune et Développement des Compétences (PEJEDEC) evaluation d'impact du programme de Travaux à Haute Intensité de Main d'œuvre (THIMO): Résultats intermédiaires à court terme. Available at http://www.pejedec.org/content/download/policy_note_thimo_pejedec_pdf.pdf [30 Mar. 2016].

Robalino, D., Margolis, D., Rother, F., Newhouse, D., & Lundberg, M. (2013). *Youth employment: A human development agenda for the next decade*. Social Protection and Labor Discussion Paper, No. 1308. Washington, DC, World Bank Group.

Schoof, U. (2006). *Stimulating youth entrepreneurship: Barriers and incentives to enterprise start-ups by young people*. SEED Working Paper No. 76. Geneva, International Labour Office.

Sparreboom, T., & Staneva, A. (2014). *Is education the solution to decent work for youth in developing economies?: Identifying qualifications mismatch from 28 school-to-work transition surveys*. Geneva, International Labour Office.

Stanley, T. D. (2013). Does economics add up? An introduction to meta-regression analysis. *European Journal of Economics and Economic Policies: Intervention*, *10*, 2, 207–220

Stanley, T. D., Doucouliagos, C., & Steel, P. (2015). *Does ICT generate economic growth? A meta-regression analysis*. Economics Series 2015,_No. 9. Deakin University, Faculty of Business and Law, School of Accounting, Economics and Finance.

Tripney, J., Hombrados, J., Newman, M., Hovish, K., Brown, C., Steinka-Fry, K., & Wilkey, E. (2013). *Technical and vocational education and training (TVET) interventions to improve the employability and employment of young people in low- and middle-income countries: A systematic review*. Campbell Systematic Reviews, 9.

Valerio, A., Parton, B., & Robb, A. (2014). *Entrepreneurship education and training programs around the world: Dimensions for success*. Washington, DC, World Bank.

van Lieshout, S., Sievers, M., & Aliyev, M. (2012) *Start and Improve Your Business*

Global Tracer Study 2011. Geneva, International Labour Office.

Waddington, H., & Hombrados, J. G. (2012). *Tool to assess risk of bias and internal validity of social experiments and quasi-experiments*. 3ie (mimeo).

Waddington, H., White, H., Snilstveit, B., Garcia Hombrados, J., Vojtkova, M., Davies, P., Bhavsar, A., Eyers, J., Perez Koehlmoos, T., Petticrew, M., Valentine, J. C., & Tugwell, P (2012). How to do a good systematic review of effects in international development: A tool kit. *Journal of Development Effectiveness*, *4*, 3, 359–387.

Walther, A., & Pohl, A. (2005). *Thematic study on policy measures concerning disadvantaged youth*. Study commissioned by the European Commission, DG Employment and Social Affairs in the framework of the Community Action Programme to Combat Social Exclusion 2002–2006.

Wilson, T. (2013). *Youth unemployment: Review of training for young people with low qualifications*. BIS Research Paper No. 101, Department for Business, Innovation and Skills.

World Bank (2015). *The role of skills training for youth employment in Nepal: An impact evaluation of the employment fund*. Adolescent Girls Initiative (AGI); February 2015. Results Series. Washington, DC, World Bank Group.

Youth Employment Network (YEN), & International Youth Foundation (IYF) (2009). *Private sector demand for youth labour in Ghana and Senegal – Ghana and Senegal study findings*.

7 Information about this review

7.1 REVIEW AUTHORS

Lead review author:

The lead author is the person who develops and co-ordinates the review team, discusses and assigns roles for individual members of the review team, liaises with the editorial base and takes responsibility for the ongoing updates of the review.

Name:	**Susana Puerto**
Affiliation:	International Labour Organization
Address:	4 route des Morillons
City, state, province or county:	Geneva
Postal code:	1211
Country:	Switzerland
Phone:	+41-22-799-65-84
Email:	Puerto-gonzalez@ilo.org

Name:	**Jochen Kluve**
Affiliation:	Humboldt-Universität zu Berlin and RWI
Country:	Germany
Email:	jochen.kluve@hu-berlin.de

Co-author:

Name:	**David Robalino**
Affiliation:	The World Bank
Country:	USA
Email:	drobalino@worldbank.org

Co-author:

Name:	**Jose Manuel Romero**
Affiliation:	The World Bank
Country:	USA
Email:	jromero1@worldbank.org

Co-author:

Name:	**Friederike Rother**
Affiliation:	The World Bank
Country:	USA
Email:	frother@worldbank.org

Co-author:

Name:	**Jonathan Stoeterau**
Affiliation:	RWI
Country:	Germany
Email:	Jonathan.Stoeterau@rwi-essen.de

Co-author:

Name:	**Felix Weidenkaff**
Affiliation:	International Labour Organization
Country:	Switzerland
Email:	weidenkaff@ilo.org

Co-author:

Name:	**Marc Witte**
Affiliation:	University of Oxford
Country:	United Kingdom
Email:	marc.witte@sant.ox.ac.uk

7.2 SOURCES OF SUPPORT

The following institutions provided financial and other support for the current review:

- International Initiative for Impact Evaluation (3ie): Systematic Review Call 5 through funding from the Canadian International Development Agency (CIDA)
- The Youth Employment Network
- The International Labour Organization
- The World Bank
- Leibniz Association

7.3 ACKNOWLEDGEMENTS

The review authors are grateful to Viviana Perego, Selsah Pasali, Yamila Simonovsky, Misina Cato, Jonas Jessen, Karishma Tiwari, Cheng Qian, and Bienvenue Tien for excellent research assistance. We are grateful to Hugh Waddington for the technical support, advice and thoughtful comments and suggestions. Thanks also go to Azita Berar Awad, Gianni Rosas, Drew Gardiner, Jonas Bausch, Niall O'Higgins, Yamila Simonovsky, Louise Fox, Valli Corbanese, Janice Tripney, John Eyers, Martina Vojtkova, Michael Grimm, Anna Luisa Paffhausen, Kristin Hausotter, Clare Ignatowski, Markus Pilgrim and participants of the 2016 ILO/IZA conference, of the 2015 IZA workshop on employment institutions, of the 2015 World Bank/IZA conference on employment and development, of 3ie workshops and symposia in Washington DC, Belfast, Manila and Dhaka and to several anonymous referees at 3ie and the Campbell Collaboration for very helpful comments. The authors would like to acknowledge the support provided by Selsah Pasali whose background analysis of the sample of studies from Africa supported the qualitative analysis, as well as Viviana Perego for her support and analysis of the evidence gap map. Financial support from 3ie and the Leibniz Association is gratefully acknowledged.

7.4 DECLARATIONS OF INTEREST

Selected work in progress and publications

Kluve, J. "Temporary work as an active labor market policy: Evaluating an innovative program for disadvantaged youths" (with C. Ehlert and S. Schaffner).

Publications

Kluve, J. (2015). *What works? A meta analysis of recent active labor market program evaluations* (with D. Card and A. Weber). IZA Discussion Paper No. 9236.

Kluve, J., & Yuri Soares (2014). *Can arts-based interventions enhance labor market outcomes among youth? Evidence from a randomized trial in Rio de Janeiro* (with C. Calero, C. H. Corseuil, V. Gonzales and Y. Soares).

Kluve, J. (2010). Active labour market policy evaluations: A meta-analysis (with D. Card and A. Weber). *The Economic Journal*, *120*, F452–F477.

Kluve, J. (2010). The effectiveness of European active labor market programs. *Labour Economics*, *17*, 904–918.

Puerto, S. (2009). *Towards comprehensive training*. SPDP 924 (with J. Fares). Washington, DC, World Bank.

Puerto, S., & Rother, F. (2007). *Global inventory of interventions to support young workers: Synthesis report* (with G. Betcherman, M. Godfrey and A. Stavreska). World Bank, SPDP 0715.

Puerto, S. (2007). *Interventions to support young workers in Latin America and the Caribbean*. World Bank (mimeo).

Puerto, S. (2007). *Learning from international experiences*. YEI, World Bank (mimeo).

Puerto, S. (2007). *Labor market impacts for youth: A meta-analysis of the YEI*. World Bank (mimeo).

Puerto, S., & Rother, F. (2007). *Interventions to support young workers in OECD countries*. World Bank (mimeo).

Robalino, D., and Rother, F. (2013). *Youth employment: A human development agenda for the next decade* (with D. Margolis, D. Newhouse and M. Lundberg).

Robalino, D. (2012). *The right skills for the job: Rethinking training policies for workers* (with R. Almeida and J. Behrmann). World Bank.

Robalino, D. (2010). *From right to reality: Towards an integrated and equitable social protection system that works in Latin American and Caribbean labor markets* (with H. Ribe and I. Walker). Directions in Development. Washington, DC, World Bank.

Rother, F. (2007). *Interventions to support young workers in sub-Saharan Africa*. Washington, DC, World Bank (mimeo).

Rother, F. (2012). Training for the unemployed, low-income, and the low-skilled workers (with J. Kluve and M. L. Sanchez-Puerta). In *The right skills for the job: Rethinking effective training policies for workers*, by R. Almeida, J. Behrmann and D. Robalino. Washington, DC, World Bank.

The review's principal investigators have co-authored and authored reviews of ALMPs that have allowed for statistical rigour, contributing a series of meta-analyses to the labour economics literature and, specifically, providing applicable lessons on what works to improve labour market outcomes of youth. There has not been direct involvement with Campbell, Cochrane, EPPI-Centre, Collaboration for Environmental Evidence, Centre for Reviews and Dissemination, or the Joanna Briggs Institute.

8 Tables for the Appendix

8.1 CHARACTERISTICS OF INCLUDED PROGRAMMES – MAIN CATEGORY: SKILLS TRAINING [45] [46] [47]

[45] Note: three programmes appear in more than one of the following tables, due to the fact that they are composed of multiple interventions, each with a different main intervention type. These programmes are (1) Jordan New Opportunities for Women (Jordan NOW), in Jordan, which is composed of both skills training and subsidized employment interventions; (2) Economic Empowerment of Adolescent Girls (EPAG) in Liberia, which features both a skills training and an entrepreneurship promotion intervention; and (3) School-to-Work Opportunities Act (STWOA), in the United States, which provides interventions in skills training, employment services and subsidized employment.

[46] In the following tables, abbreviations will be used: ATE: average treatment effect; diff-in-diff: difference-in-differences; IPW: inverse probability weighting; IV: instrumental variables; RCT: randomized control trial.

[47] The signs for the overall treatment effects (+/-/0) in the following tables refer exclusively to the sign of average standardised mean differences (SMDs) as identified by the analysis. Outcomes marked with an asterisk are imputed based on average positive and statistically significant (PSS) estimates and average t-statistic (tstat), when it was not possible to compute SMDs given the information provided in the studies. When neither SMDs nor PSS were obtainable from the studies, outcomes are marked with the symbol °.

Argentina (regional) – Entra 21			
Main intervention: Two programmes, provided by NGOs SES and ADEC, with similar features. Skills training intervention providing technical and soft skills training in ICT-related activities through in-class learning and internships. Includes job placement component as each implementer is committed to inserting 40 per cent of their graduates into the labour market. Co-financed by the International Youth Foundation and the Multilateral Investment Fund of the Inter-American Development Bank. **Other features:** Targeted at disadvantaged unemployed youth with secondary education, SES in five regions of the country, ADEC in the city of Cordoba. The programme also provides the trainees with transportation expenses, books, training materials and clothes. SES participants engage in community service to partially repay the costs of their training. ADEC beneficiaries have a tutor and receive a monthly stipend.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Alzúa, Nahirñak & Alvarez de Toledo, 2007	matching; cross-sectional originally collected+survey data	empl.prob. +* (PSS 0.38, tstat 1.43)	earnings +* (PSS 1, tstat 5.66)
Alzúa, Cruces & Lopez Erazo, 2013	multivariate linear; RCT	empl.prob. + (SMD 0.29, PSS 0.56, tstat 1.54)	wage +* (PSS 1, tstat 2.46)
Argentina (national) – Proyecto Joven			
Main intervention: Skills training, 14–20 weeks' duration, for an average of 200 hours. Two phases: 1. technical knowledge – 6–12 weeks, beneficiaries are taught knowledge and technical skills for a particular occupation; 2. internships – 8 weeks. Implemented by the Government, NGOs and the private sector; co-financed by the Government and the Inter-American Development Bank. **Other features:** Targets young people from poor households with, at most, secondary education, no working experience and who are unemployed, underemployed or inactive. The programme provides transportation expenses, a subsidy for females with young children, medical check-ups, books, materials and work clothing. Training centres are hired through an international bidding process, and are responsible for recruiting firms willing to accept interns from Proyecto Joven. The firm does not have to pay any stipend or wage to the trainees.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Aedo & Nuñez, 2004	matching; originally collected cross-sectional data	empl.prob. + (SMD: 0.07, PSS:0.25, tstat 1.96)	earnings + (SMD 0.13, PSS 0.25, tstat 1.96)
Elías et al., 2004	matching+diff-in-diff; originally collected panel data	empl.prob. + (SMD 0.02, PSS 0, tstat 1) wage + (SMD 0.09, PSS 0, tstat 0.86)	earnings + (SMD 0.01, PSS 0, tstat 0.28)
Alzúa & Brassiolo, 2006	matching; originally collected cross-sectional data	empl.prob. +* (PSS 0.33, tstat 1.80)	earnings +* (PSS 0.33, tstat 0.94)

Brazil (local) – Entra 21			
Main intervention: Two main providers: 1. Brazil CEPRO – professional training project that offers ICT training. Includes coursework on life skills training and technical training in office skills. Following basic training, participants choose from among four specialized courses in computer installation and maintenance, telephone systems, computer graphics or website development (600 hours). Additional training is offered in either entrepreneurship or employability. Then job placement phase in local businesses. 2. Brazil IH – Training programme providing ICT training, life skills instruction and job placement assistance. Participants have the option of focusing on one of three areas: lodging, food and beverages or tourism. In addition to training, participants undertake internships with a local hospitality business. Course duration is 480 hours over a five-month period, internships last one month. **Other features:** Brazil CEPRO – targets disadvantaged youth in Cotia, Sao Paolo. Provides the option to tailor courses according to participants' interests. Job placement is offered, exploiting CEPRO's contacts and established relationships with local businesses. Brazil IH – private, non-profit organization, involves business, government and civil society sector leaders. Targets 480 disadvantaged youth in north-eastern Brazil. On completion of the programme, graduates are awarded a nationally recognized certification.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Alzúa, Nahirñak & Alvarez de Toledo, 2007	matching; originally collected+survey cross-sectional data	empl.prob. -* (PSS 0, tstat -4.27)	earnings -* (PSS 0, tstat -2.51)
Brazil (local) – Galpao			
Main intervention: Small-scale labour training programme consisting of a combination of vocational, academic and life skills training. Approximately six months' duration, five hours a day, five days a week. The treatment includes 300 hours of vocational training (construction-related, soldering, woodworking), 180 hours of training on academic and basic skills, including remedial courses in both Mathematics and Portuguese, and 120 hours of life skills training. **Other features:** Targets low-educated youth in Rio de Janeiro. Vocational component is tailormade to suit the beneficiaries. Delivered through a pedagogic method that utilizes arts and dance. The corresponding life skills training component is combined with a technical component teaching vocational skills.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Kluve et al., 2014	multivariate linear; RCT	empl.prob. + (SMD 0.15, PSS 0.13, tsat 0.17) earnings + (SMD 0.10, PSS 0, tstat 0.82)	hours worked – (SMD -0.03, PSS 0, tstat 0.06)

Brazil (national) – Lei do Aprendiz		
Main intervention: Targeted ALMP conducted by the Ministry of Labour. Payroll subsidies to firms that hire and train young workers under temporary contracts. Purely subsidized programme, although this is temporary during training. **Other features:** Training courses are provided by official professional qualification agencies or by training institutions certified by the Ministry of Labour. If an apprentice has not yet completed primary school (an eight-year schooling stage), they are required to enrol at school.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Corseuil et al., 2014	matching+IV; administrative panel data	empl.prob. + (0.03, PSS 0.5, tstat 0.24) wage +* (PSS 1, tstat 7.40) hours worked -* (PSS 0, tstat -5.80)
Brazil (national) – Senai Vocational Training		
Main intervention: Organized at the national and state levels as a private, non-profit organization, financed, managed and led by the industry. The SENAI developed as an institution providing training as a standalone operation, not linked to a particular job in a firm, requiring participants to find employment for themselves on graduation. **Other features:** Directed at youth. The SENAI is principally financed by all industrial companies through a tax of 1 per cent on all payrolls, which serves as the basis of their contribution to the social security system. SENAI aims to actively match the labour demands of industry.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Villalobos Barría & Klasen, 2014	IPW; cross-sectional survey data	empl.prob. + (SMD 0.16, PSS 1, tstat 1.96) empl.qual. + (SMD 0.07, PSS 0.78, tstat 1.09) wage + (SMD 0.04, PSS 0.67, tstat 0.18) hours worked – (SMD -0.02, PSS 0.22, tstat -4.46) earnings + (SMD 0.10, PSS 0.78, tstat 0.43)

Chile (national) – Apprentices Hiring Programme			
Main intervention: Subsidy to employers hiring interns to promote on-the-job training. Programme also includes technical skills training delivered in-classroom by hiring firms or contracted skills training providers. Duration varies between four months and two years. **Other features:** Targets disadvantaged, first time jobseekers. Monitoring of employers and sanctions for non-compliers. Subsidy amounts to 40 per cent of minimum wage + una tantum reduction in employer social security payments to contribute to interns' training costs.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
SENCE: Riquelme Peña et al., 2006	diff-in-diff; orginally collected pooled cross-sectional data	empl.prob. + (SMD 0.47, PSS 1, tstat 1.74) unempl.duration° wage°	unempl.prob. (SMD 0.37) earnings +* (PSS 1, tstat 2.17)
Chile (national) – Chile-Joven			
Main intervention: Skills training intervention with in-class learning and internship component. Skills training also features a business skills element. Demand driven, relies on private-sector implementers to identify job market opportunities. Publicly financed. **Other features:** Targets young people from poor households, not enrolled in education, with no working experience or unemployed, underemployed or inactive. Benefits include transport and meal subsidies, unless the internship is paid, plus free insurance against workplace accidents.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Santiago Consultores Asociados, 1999	diff-in-diff; originally collected panel data	empl.prob. + (SMD 0.43) unempl.duration° wage°	unempl.prob. + (SMD 0.55) earnings°
Aedo, Pizarro & Valdivia, 2004	matching+diff-in-diff; originally collected panel data	empl.prob. + (SMD 0.45, PSS 0.56, tstat 2.25)	wage 0* (PSS 0.22, tstat 1.35)
Chile (national) – Formación en Oficios para Jóvenes de Escasos Recursos			
Main intervention: Two pathways: Skills training for dependent employment and skills training for self-employment. Four phases: 1. classes and laboratories, ICT training (500 hours); 2. on-the-job training, internships (384 hours); 3. technical assistance, specific to business skills training (24 hours); 4. follow-up and monitoring (six months).			

Other features: Targets disadvantaged youth, preferably school dropouts. Many entreprises pay beneficiaries a modest salary during their internships. Strict supervision and monitoring of entreprises where internships take place to ensure the training content of the internship is appropriate.

Reference	Data collection and analysis	Overall treatment effect	
SENCE, 2008	matching+diff-in-diff; survey panel data	empl.prob. + (SMD 0.05, PSS 0.125, tstat 1.40) unempl.durat. 0* (PSS 0.15, tstat 0.66) wage 0* (PSS 0)	hours worked +* (PSS 0.3, tstat 1.96) earnings +* (PSS 0.80, tstat 2.27)

Colombia (national) – Formación Técnica y Tecnológica (FT&T)

Main intervention: Technical and life skills vocational training. Apprenticeship component. Designed and financed by the Government.

Other features: Different pathways for technicians and technologists. Leads to the attainment of a vocational qualification.

Reference	Data collection and analysis	Overall treatment effect	
Santa María et al. 2009a	diff-in-diff+matching; originally collected panel data	empl.prob. + (SMD 0.17, PSS 1, tstat 2.20) earnings 0 (SMD 0.00, PSS 0, tstat 0.82)	empl.qual. +* (PSS 1, tstat 3.55)

Colombia (local) – Jóvenes En Acción

Main intervention: Three months' in-classroom training + three months' on-the-job training. Training institutions may be either profit-making or not for profit.
Other features: Targets the lowest socio-economic stratum of the population living in urban areas. Offered in the seven largest cities of the country. Vocational skills provided by the courses are very diverse, ranging from administrative to manual occupations, as well as courses in skilled occupations, including IT specialists, data entry, surveyors and accountant assistants. Internships are unpaid.

Reference	Data collection and analysis	Overall treatment effect	
Attanasio, Kugler & Meghir, 2011	multivariate linear and nonlinear; RCT	empl.prob. +* (PSS, tstat) empl.qual. +* (PSS 1, tstat 3.15) wage + (SMD 0.12, PSS 0.5, tstat 2.31)	hours worked + (SMD 0.01, PSS 0.5, tstat 0.00) earnings – (SMD -0.02, PSS 0.33, tstat 0.91)

Colombia (regional) – Programa de capacitación Jóvenes con Futuro (JCF)		
Main intervention: Skills training programme with three components: 1. soft skills, ICT training and remedial education (280 hours); 2. technical skills training (480 hours); 3. on-the-job training (440 hours). Nine months' total duration. Publicly financed, implemented by NGO. **Other features:** Targeted at disadvantaged youth with clean criminal records, not in employment or education, with at least some secondary level education and who are not benefitting from other programmes. Selection process through admission test. Individual mentoring and psychological support to beneficiaries and their families. Transport subsidies provided.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Santa María et al., 2009a	regression-adjusted diff-in-diff+diff-in-diff+matching; originally collected panel data	empl.prob. + (SMD 0.27, PSS 0.5, tstat 2.37)
Dominican Republic (national) – Programa Juventud y Empleo		
Main intervention: Training and counselling services. Courses (maximum duration 350 hours) split into two parts: Basic skills training and technical/vocational training, followed by two-month internship in private sector firms. Co-financed by the Government and the Inter-American Development Bank. Private implementers. **Other features:** Targeted at low-income youth with less than secondary education who are not enrolled in schooling. Special emphasis on enrolling women. Partial reimbursement for transportation costs and meals, provision of insurance against workplace accidents. Provision of training services outsourced to private training institutions, which are selected through a competitive bidding process. Proposals from potential training providers have to include written commitments from at least one local firm to offer internships to trainees graduating from the provider's programme.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Card et al., 2011	multivariate linear; RCT	empl.prob. + (SMD 0.02, PSS 0, tstat 0.00) earnings + (SMD 0.10, PSS 0.14, tstat 1.36) hours worked – (SMD -0.06, PSS 0, tstat -0.73)
Ibarrarán et al., 2014	IV+multivariate linear; RCT	empl.prob. + (SMD 0.05, PSS 0.11, tstat 0.69) unempl.durat +* (PSS 0, tstat -2.01) particip.rate + (SMD 0.06, PSS 0, tstat 0.91) earnings 0* (PSS 0, tstat 1.21)

France (national) – Contrat de Qualification		
Main intervention: Fixed-term labour contract between a youth and a private employer. Specific feature is the emphasis on formal training. Training usually occurs in a private or public teaching institution (such as a high school) with which the firm has signed a training convention. The contract lasts between six and 24 months, and can be renewed only once. Training duration at least 25 per cent of the total duration of the contract. **Other features:** Targets low-educated youth. All unemployed youth are automatically eligible, except those holding a vocational/technological high school diploma and those holding a higher education diploma. The training received is endorsed by a diploma.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Charpail, Klein & Zilberman, 2005	mean comparison corrected for selection bias; survey panel data	unempl.duration ° empl.qual. ° wage ° empl.prob. + earnings +
Pessoa e Costa & Robin, 2009	PSM; panel survey data	(SMD 0.12, PSS 1, tstat 2.81) (SMD 0.08, PSS 1, tstat 2.37) wage + (SMD 0.12, PSS 1, tstat 3.37)
Germany (pilot) – Arbeit Sofort!		
Main intervention: Pilot workfare project, consisting of training and employment modules, as well as basic skills in German and Mathematics and technical understanding. Individual profiling leading to individual action plans with the objective of integration into the labour market. Provision of formal trainee contracts for a maximum of six months. Publicly financed and implemented. **Other features:** Targets young people in receipt of social assistance in Berlin's Charlottenburg-Wilmersdorf district. Excludes those with clear professional or educational prospects for the future and likely to be only temporarily in need of social assistance. Mandatory: non-participation results in loss of social assistance benefits. Participants receive monthly compensation slightly above the social assistance rates and reimbursement of rental costs. Payment can begin on the first day of the programme.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Schneider, Uhlendorff & Zimmermann, 2011	matching; administrative panel data	empl.prob. + (SMD 0.12, PSS 0, tstat 0.97)

Germany (national) – Further training (medium to long term) (FT)		
Main intervention: Classroom training, may vary between part- or full-time courses. Five to seven months' median duration. **Other features:** Predominantly focused on unemployed youth with vocational qualification, who require additional qualifications to succeed in the labour market.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Caliendo, Kunn & Schmidl, 2011	IPW; cross-sectional administrative data	empl.prob. 0* (PSS 0.45, tstat 0.09)
Germany (national) – Mandatory internships		
Main intervention: Mandatory internships during higher education. **Other features:** Targeted at highly educated first time jobseekers. Mandatory. Modest salary paid to interns.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Saniter, 2014	IV; cross-sectional survey data	empl.qual. + (SMD 0.01, PSS 0, tstat 0.23) wage + (SMD 0.10, PSS 0.5, tstat 2.32)
Germany (national) – Preparatory training (PT)		
Main intervention: Subsidized internship within a firm where predominantly basic practical skills and literacy are developed. **Other features:** Targeted at unemployed youth with secondary education. Some employers might use this as a probationary period before offering a full apprenticeship position within the firm.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Caliendo, Kunn & Schmidl, 2011	IPW; cross-sectional administrative data	empl.prob. -* (PSS 0, tstat -3.99)
Germany (national) – Short-term training (STT)		
Main intervention: Courses to improve auxiliary skills. Two months' duration. **Other features:** Targeted at unemployed youth with secondary education. Focuses on skills that are important in the application process (e.g., through computer classes or language courses).		
Reference	**Data collection and analysis**	**Overall treatment effect**
Caliendo, Kunn & Schmidl, 2011	IPW; cross-sectional administrative data	empl.prob. +* (PSS 0.61, tstat 1.70)

India (local) – Livelihoods Training for Adolescent Living		
Main intervention: Experimental intervention providing vocational counselling and training and assistance with opening savings accounts. **Other features:** Targets young girls in slum areas of Allahabad in Uttar Pradesh. The intervention is delivered to groups of approximately 20 women meeting once a week at the home of a peer educator. Participants receive reproductive health training sessions, vocational counselling, savings formation information and follow-up support from the peer educator. The project also makes arrangements for older girls to attend Government-run vocational courses.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Mensch et al., 2004	PSM+multivariate nonlinear; originally collected panel data	hours worked 0* (PSS 0, tstat -1.64)
India (local) – Satya/Pratham programme		
Main intervention: Training programme in stitching and tailoring. Six months' duration. Participants required to commit to up to two hours per day in a five-day week. Financed and implemented by two local NGOs. **Other features:** Targeted at young women with at least five or more grades of schooling residing in slum areas of New Delhi. Extensive advertising campaign. Participants pay a deposit, which is refunded with a bonus upon successful completion of the programme to incentivize regular attendance. Certificate issued upon completion.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Maitra & Mani, 2014	regression-adjusted diff-in-diff; RCT	empl.prob. + (SMD 0.30, PSS 1, tstat 2.32) earnings + (SMD 0.24, PSS 1, tstat 2.05) hours worked + (SMD 0.31, PSS 1, tstat 2.79)

Jordan (pilot) – Jordan New Opportunities for Women (Jordan NOW)			
Main intervention: Free intensive training in soft skills identified by employers as important for new graduates to acquire. 45 hours over a nine-day period with a maximum of 30 participants in each training group. Financed by the World Bank, implemented by local NGO. **Other features:** Targets young female graduates. Sessions based on active participation and cooperative learning rather than lectures. Training takes place in 17 sessions offered throughout six governorates to maximize access (identical training facilities and training content across sessions). Sessions held during daylight hours at locally known and trusted institutions to minimize the effect of social and cultural restrictions on mobility.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Groh et al., 2012	multivariate linear; RCT+administrative panel data	empl.prob. + (SMD 0.05, PSS 0.08, tstat 0.78) hours worked 0* (PSS 0, tstat 0.49)	participation rate + (SMD 0.07, PSS 0.3, tstat 1.12) earnings 0* (PSS 0, tstat 0.20)
Kenya (local) – Ninaweza Youth Empowerment Programme			
Main intervention: Employability programme providing ICT training, training in life skills, internships and job-placement support. Duration of training is eight weeks, of internships eight weeks; job placement lasts for six months. **Other features:** Targets unemployed young women with secondary education who live in the informal settlements around Nairobi. Implementers make sure that the technical training responds to market demands. Both monetary and non-monetary benefits provided.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Alvares de Azevedo, Davis & Charles, 2013	tabulation; RCT	empl.prob. + (SMD 0.15, PSS 0.10, tstat 1.06) empl.qual. + (SMD 0.25, PSS 0, tstat 1.50) business creation + (SMD 0.17, PSS 0, tstat -1.13)	hours worked + (SMD 0.12, PSS 0, tstat 0.71) earnings + (0.34, PSS 1, tstat 2.98)

Kenya (regional) – Technical and Vocational Voucher Programme (TVVP)			
Main intervention: Randomized evaluation of a youth vocational education intervention. Half of the applicants are randomly awarded a voucher for vocational training; of these, half are randomly awarded a voucher that can only be used in public (Government) institutions, the other half receives a voucher to be used in either private or public institutions. The voucher covers almost the full tuition costs for most private vocational education programmes and Government-run rural village polytechnics or technical training institutes. Most schools offer mandatory entrepreneurship programmes, others offer mandatory remedial subjects, and others encourage their trainees to sit in on some training in closely related fields. **Other features:** Targets out-of-school youth in western Kenya. The programme includes a cross-cutting information intervention, which exposes a randomly selected half of all treatment and control participants to information about the actual returns to vocational education. Young women are encouraged to enrol in more lucrative male-dominated trades.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Hicks et al., 2013	IV+multivariate linear; RCT+survey panel data	empl.prob. – (SMD -0.13, PSS 0, tstat -1.49) hours worked + (SMD 0.03, PSS 0, tstat 0.33) consumption + (SMD 0.15, PSS 0, tstat 0.45) profits – (SMD -0.05, PSS 0, tstat -1.00)	unempl.prob. + (SMD 0.03, PSS 0, tstat 0.34) earnings – (SMD -0.01, PSS 0, tstat -0.26) wage + (SMD 0.45, PSS 1, tstat 2.29)
Liberia (local) – Economic Empowerment of Adolescent Girls (EPAG)			
Main intervention: The programme combines classroom-based technical and life skills training, with a focus on skills with high market demand, augmented by follow-up support to enter wage employment or start a business. Financed by Government, World Bank, the Nike Foundation and the Governments of Australia, the United Kingdom, Norway, Denmark and Sweden; implemented by Government and NGOs. **Other features:** Targets girls with basic literacy and numeracy skills but who are not enrolled in school for several months prior to programme initiation, residing in one of nine target communities in and around Monrovia. Implemented by four NGOs selected by the Liberian Ministry of Gender and Development through a competitive bidding process. Participants are assigned to a "Job Skills" track or a "Business Development Services" track. Wherever possible, the participant's track preference is honoured. Service providers are responsible for developing training curricula and making arrangements for free childcare services. Performance bonuses are awarded to training providers that successfully place their graduates in jobs or micro-enterprises. Contests and competitions are also held among beneficiaries. Strong monitoring component.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Adoho et al., 2014	regression-adjusted diff-in-diff; RCT	empl.prob. + (SMD 0.29, PSS 0.67, tstat 3.22)	earnings +* (PSS 0.5, tstat 2.80)

Malawi (national) – Apprenticeship Training Programme and Entrepreneurial Support for Vulnerable Youth			
Main intervention: Apprenticeship programme. Financed by Government and the Global Fund, implemented by Government and the private sector. **Other features:** Targeted at vulnerable youth, mainly defined as orphans or school dropouts. A pool of potential trainers is identified in each district, then the master craftspeople are selected from this pool, based on their expertise and business performance in the neighbourhood. Trainees receive a small stipend to cover meals and accommodation.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Cho et al., 2013	IV; RCT	hours worked +* (PSS 0.08, tstat 0.16) business creat. – (SMD -0.28, PSS 0, tstat -1.63)	earnings 0* (PSS 0, tstat -0.98)
Mexico (national) – Becate			
Main intervention: Skills training and internship programme. Offering training for one to three months, five to eight hours a day for five days a week. Publicly financed, implemented by the private sector. **Other features:** Targeted at young unemployed who have just graduated from technical or vocational school. Payment, reimbursment of transport costs and insurance against work-related injuries. Strong monitoring component.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
van Gameren, 2010	regression-adjusted diff-in-diff; survey panel data	empl.prob. +* (PSS 0.83, tstat 2.25)	earnings +* (PSS 0.5, tstat 1.52)

Nepal (national) – Employment Fund			
Main intervention: The Employment Fund authorizes training programmes and employment services. Training courses cover technical skills in a wide range of trades. In addition, female students receive 40 hours of life-skills training, and a subset receive a short course in basic business skills. Broad range of training and employment providers from formal technical and vocational training institutions, public and private providers, to skilled artisans offering apprenticeships. **Other features:** Targets poor youth with less than ten years of formal education. Providers selected through a competitive bidding system. Upon verification, providers receive an outcome-based payment. Differential pricing mechanism that awards a higher incentive to service providers who agree to train and place more disadvantaged groups. The programmes may also include skills certification upon completion of training or job placement assistance.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Ahmed et al., 2014	diff-in-diff+matching; RCT	empl.prob. + (SMD 0.27, PSS 0.48, tstat 2.37) earnings + (SMD 0.44, PSS 0.75, tstat 3.42)	hours worked + (SMD 0.36, PSS 0.81, tstat 3.47)
Panama (national) – Procajoven			
Main intervention: Demand-driven training, job orientation and job placement activities. Two distinct modalities: 1. Insertion Modality (MI) – short-term programme, including classroom training in job-readiness skills and technical training followed by internship in a firm. Classroom training lasts 270 hours, internship 172 hours. 2. Transition Modality (MT) – provides job-readiness training and a longer internship (344 hours). Financed by the Government and the Inter-American Development Bank, implemented by Government and the private sector. **Other features:** Targets disadvantaged and unemployed youth. MI targets low-income youth. MT targets first time jobseekers with completed secondary education. In both cases, a direct transfer is made to participants for transportation and meals. Training institutions selected through competitive public bids.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Ibarrarán & Rosas, 2007	kernel weighted difference in means; originally collected cross-sectional data	empl.prob. + (SMD 0.18, PSS 0.11, tstat 1.01) empl.qual. + (SMD 0.09, PSS 0, tstat 0.50) wage 0* (PSS 0, tstat -1.50)	hours worked + (SMD 0.16, PSS 0.18, tstat 1.02) earnings + (SMD 0.20, PSS 0.06, tstat 0.64)

Peru (national) – Projoven			
Main intervention: Basic or semi-skilled training in blue-collar occupations. Two phases: 1. learning phase – training courses are provided by training centres for a three-month period; 2. internship phase at private firms for three months, during which the trainee receives a market wage paid by the internship firm. After these three months, the firm may or may not hire the trainee on a more permanent basis. Publicly funded. Public and private training institutions. **Other features:** Targets disadvantaged youth with low levels of formal education, with minimal or no labour market experience, and currently underemployed, unemployed or out of the labour force. Stipend for transportation and lunch plus supplement for women with children under six years old to cover childcare expenses. Selection of the training course providers relies on bidding processes targeting best training courses at the most competitive prices. Training centres must present formal agreements with private firms that guarantee internships remunerated by the firm at no less than the monthly minimum wage payment for each beneficiary.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Ñopo & Saavedra, 2003	matching+diff-in-diff; originally collected panel data	empl.prob. 0 (SMD -0.00) earnings°	hours worked °
Chong & Galdo, 2006	matching+diff-in-diff; originally collected panel data	earnings +* (PSS 0.43, tstat 1.83)	
Diaz & Jaramillo, 2006	matching+diff-in-diff; originally collected panel data	empl.prob. + (SMD 1.83, PSS 0.46, tstat 1.95) earnings +* (PSS 0.81, tstat 3.10)	hours worked +* (PSS 0.36, tstat 1.36)
Chong, Galdo & Saavedra, 2008	matching+diff-in-diff; originally collected panel data	empl.prob. 0* (PSS 0, tstat 0.45)	empl.qual. +* (PSS 1, tstat 6.05)
Galdo, Jaramillo & Montalva, 2008	matching+diff-in-diff; originally collected panel data	empl.prob. + (SMD 0.09, PSS 0.22, tstat 0.67)	earnings +* (PSS 0.44, tstat 2.38)
Ñopo, Robles & Saavedra, 2008	matching+diff-in-diff; originally collected panel data	empl.prob.0 (SMD -0.00) earnings + (SMD 0.49)	hours worked ° wage + (SMD 0.24)

Espinoza Peña, 2010	matching+diff-in-diff; RCT	empl.prob. – (SMD -0.01, PSS 0.06, tstat -0.13)	earnings 0 (SMD 0.00, PSS 0.11, tstat -0.81)
Chong & Galdo, 2012	matching+diff-in-diff; originally collected panel data + administrative data	empl.prob. +* (PSS 0.33, tstat 1.96) earnings +* (PSS 0.72, tstat 1.96)	empl.qual. +* (PSS 0.83, tstat 1.96)

Serbia (regional) – Youth Employment and Migration (YEM)

Main intervention: Three components:
1. institution-based training – competency-based training organized by a training provider. Between one and six months' duration;
2. on-the-job training (pre-employment qualification) – competency-based training organized in a partner enterprise. Between one and six months' duration. There is no obligation on the enterprise to retain trainees, unless the firm trains more than nine young persons at any given time;
3. programme for persons with disabilities – institution-based and/or on-the-job training followed by subsidized employment.

Other features: Targets youth with low educational attainment, long unemployment spells and those considered "hard-to-place" due to their personal and household characteristics (youth at risk of social exclusion). Relaxed entry criteria and the possibility of longer programme duration are envisaged for the most disadvantaged youth.

Reference	Data collection and analysis	Overall treatment effect	
Arandarenko, Nojkovic & Vladisavljevic, 2014	matching; cross-sectional survey data	empl.prob. + (SMD 0.01, PSS 0, tstat 0.14) wage + (SMD 0.37, PSS 0.5, tstat 0.96)	unempl.prob. – (SMD -0.07, PSS 0, tstat -0.27)

Spain (regional) – Programa de cualificacion profesional inicial (PCPI)		
Main intervention: Skills training programme with three modules: 1. vocational training; 2. general and life skills training; 3. voluntary remedial education leading to secondary school degree. Between 800 and 1,100 hours' duration plus 150–250 hours in job training centres. Financed by the Government and European Social Fund, both public and private implementers. **Other features:** Targeted at youth who finish secondary school but do not get a secondary school degree. Low class sizes to allow for more individualized support. Individual profiling and mentoring component. One privately implemented sub-programme is not subsidized and is therefore not free.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Blasco et al., 2014	matching; administrative cross-sectional data	empl.prob. 0* (PSS 0.08, tstat 0.19)
Spain (national) – Programa de Escuelas taller y Casas de Oficio		
Main intervention: Vocational training programme. Two steps: theoretical education and professional internship. Duration one or two years. Financed by Government and the European Social Fund and managed in a decentralized way. **Other features:** Targeted at the young unemployed. Offered free of charge, participation is voluntary. The study only focuses on the province of Seville.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Cansino Muñoz-Repiso & Sánchez Braza, 2011	matching; administrative cross-sectional data	unempl.duration -* (PSS 1, tstat -7.27)

Sweden (national) – Labour Market Training			
Main intervention: Courses of various length and content, both vocational and non-vocational. Courses last up to 12 months. Publicly financed and implemented. **Other features:** Targets individuals who are unemployed or at risk of unemployment and who have low levels of education and skills. However, individuals with a high school diploma may also form part of the target group. Potentially, all age groups are targeted, but especially youths. Mandatory: rejecting an offer to participate can lead to suspension of unemployment benefits. Participants should continue their job-seeking activities during the programme.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Larsson, 2003	matching; administrative cross-sectional data	empl.prob. –* (PSS 0, tstat -1.81)	earnings - (SMD -0.14, PSS 0, tstat -3.78)
Sweden (national) – Utvecklingsgarantin (UVG)			
Main intervention: Guarantee of early programme participation for the young unemployed. Local employment offices are responsible for the youth during the first 90 days of unemployment. If the individual is still openly unemployed after 90 days, they are sent to the municipal programme office which must assign them to some activity within ten days. The content of the activity varies among participants, and includes workplace practice, training or a combination of both. **Other features:** Local employment office sending the individual to the municipality pays a constant compensation per participant and working day to cover the cost of administration and the early intervention programme. Compensation to participants depends on what the individual received as openly unemployed, but participants without any previous compensation for unemployment still receive a moderate amount. If the individual rejects an offer to participate in a programme without any acceptable reason, they can lose their unemployment or social assistance benefits.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Carling & Larsson, 2005	diff-in-diff+tabulation; administrative repeated cross-sectional data	unempl.duration + (SMD 0.01, PSS 0.5, tstat 2.20)	
Uganda (national) – Empowerment and Livelihood for Adolescents (ELA)			
Main intervention: Vocational and life skills training. Comprises a series of courses on income-generating activities (preferably for self-employment), complemented by financial literacy courses. Publicly financed, implementers can be either public or private. **Other features:** Targets adolescent girls. The intervention is delivered from designated adolescent development clubs, fixed meeting places within communities. Clubs also host popular recreational activities. Club activities are led by a female mentor. Interventions are tailored.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Bandiera et al., 2014	multivariate linear; RCT	empl.prob. + (SMD 0.13, PSS 0.67, tstat 3.48)	consumption + (SMD 0.18, PSS 1, tstat 4.92)

United States (pilot) – CET replication sites			
Main intervention: Replication of Center for Employment Training (CET) in San José, California, in different settings in California. Provision of training in a work-like setting and of employment services. Employers are involved in the design and delivery of training. Instruction in basic skills integrated into the training. **Other features:** Targets out-of-school youth. Trainees are enrolled with little prescreening. Requires full-time commitment. Trainees are allowed to progress as they master competencies, without any fixed schedule.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Miller et al., 2005	multivariate linear; RCT	empl.prob. – (SMD -0.01, PSS 0.03, tstat -0.11) earnings + (SMD 0.05, PSS 0.03, tstat -0.37) wage + (SMD 0.07, PSS 0.08, tstat 0.49)	hours worked + (SMD 0.17, PSS 0.31, tstat 1.21) hh income 0* (PSS 0, tstat 0.06)
United States (regional) – High School Career and Technical Education Programmes (Washington)			
Main intervention: Training programme for academic and employment success. Can encompass counselling, tutoring, job training, mentoring or work experience. Other strategies include summer employment, study skills training or basic skills remedial education to obtain a GED. Publicly implemented and financed. **Other features:** Targets low-income youth with an educational deficiency in the Washington State. Individual profiling to develop individual plans. At least 30 per cent of the funding must be used to provide activities for out-of-school youth.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Hollenbeck & Huang, 2006	matching; administrative panel data	empl.prob. + (SMD 0.11, PSS 0.86, tstat 1.84) earnings + (SMD 0.13, PSS 1, tstat 2.44)	hours worked + (SMD 0.15, PSS, tstat 2.58) salary + (SMD 0.09, PSS 0.86, tstat 2.58)

United States (national) – Job Corps			
Main intervention: Remedial education and vocational programme. Services vary from intensive education instruction, equivalent to a high school academic year, to vocational training, and on-the-job training plus counselling, job placement and health services. Average of 440 hours' academic instruction (remedial education in reading, maths and writing skills plus GED programme of high school equivalency) and 700 hours of vocational training, over an average of 30 weeks. Mainly private contractors and around one-quarter operated by US Department of Agriculture and US Department of Interior. Publicly financed. **Other features:** Targets low-income youth who are legal residents of the United States. Most participants live at the local Job Corps centre while enrolled. The residential component of the programme includes formal social skills training, meals and dormitory-style life. With the help of counsellors, participants develop individualized, self-paced programmes with open-exit educational philosophy. Students receive a stipend after leaving Job Corps.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Schochet, McConnell & Burghardt, 2003	tabulation; RCT	empl.prob. 0 (SMD -0.00, PSS 0.31, tstat 0.15)	earnings +* (PSS 0.46, tstat 0.85)
Schochet, Burghardt & McConnell, 2008	IV+tabulation; RCT and administrative data	empl.prob. +* (PSS 0.67, tstat -1.11)	earnings +* (PSS 0.65, tstat 0.61)
Lee, 2009	tabulation+multivariate nonlinear; RCT	empl.prob. +* (PSS 1, tstat 3.74) earnings + (SMD 0.4, PSS 0.4, tstat 1.86)	hours worked + (SMD 0.02, PSS 0.5, tstat 0.80) wage +* (PSS 1, tstat 2.70)
Zhang, Rubin & Mealli, 2009	principle stratification approach; RCT	earnings °	
Flores-Lagunes, Gonzalez & Neumann, 2010	IV+multivariate linear; RCT	earnings + (SMD 0.10, PSS 1, tstat 1.96)	
Bampasidou, 2012	tabulation; RCT	empl.prob. + (SMD 0.09, PSS 0.33, tstat 1.96)	earnings + (SMD 0.12, PSS 0.8, tstat 1.96)
Frumento et al., 2012	principle stratification approach; RCT	empl.prob. + (SMD 0.01)	wage °
Blanco, Flores & Flores-Lagunes, 2013a	ATE bounds with nonparametric means; RCT	wage + (SMD 0.09, PSS 1, tstat 3.08)	
Blanco, Flores & Flores-Lagunes, 2013b	nonparametric bounds+quantile treatment effects; RCT	wage + (SMD 0.11, PSS 0.5, tstat 2.47)	

Chen, 2013	IV; RCT	empl.prob. + (SMD 0.09, PSS 1, tstat 1.96)	earnings + (SMD 0.10, PSS 1, tstat 1.96)
Bampasidou et al., 2014	nonparametric averages; RCT	empl.prob. + (SMD 0.08, PSS 0.67, tstat 2.15)	earnings + (SMD 0.11, PSS 1, tstat 3.39)
Frölich & Huber, 2014	IV; RCT	earnings + (SMD 0.09, PSS 1, tstat 1.98)	
United States (national) – JobStart			
Main intervention: Four components – education in basic academic skills, occupational skills training, training-related support services, job placement assistance. At least 200 hours of basic education and at least 500 hours of occupational training. Publicly financed, implementers can be public, private or NGOs. **Other features:** Targets economically disadvantaged school dropouts with poor reading skills. Instruction in basic academic skills is based on individualized curricula. Training-related support services include assistance with transportation and childcare, counselling, work-readiness and life skills training and needs-based or incentive payments tied to programme performance.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Cave et al., 1993	multivariate linear; RCT	empl.prob. + (SMD 0.01, PSS 0, tstat -0.16) earnings 0* (PSS 0, tstat -0.06)	hours worked 0* (PSS 0, tstat -0.37) wage 0* (PSS 0, tstat 0.34)

United States (national) – JTPA			
Main intervention: Supports job training for individuals facing barriers to employment and needing special training to obtain productive employment. It consolidates multiple federally funded training programmes, whereby states have the flexibility to design and implement a statewide approach to job training, based on the concept of one-stop career centres. Comprises three "service strategies": classroom training (occupational skills and basic education), a mix of on-the-job training and/or job-search assistance and other services. Services can be provided by high schools, community colleges, vocational schools and community-based organizations. **Other features:** Targets out-of-school youths with significant barriers to employment. Funds are distributed to over 600 local service delivery areas, where staff plan and implement programmes with guidance from a local private industry council.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
General Accounting Office, 1996	tabulation; RCT	empl.prob. + (SMD 0.03, PSS 0.02, tstat 1.96)	earnings 0 (SMD 0.00, PSS 0.35, tstat 1.23)
Bloom et al., 1997	multivariate linear; RCT	earnings +* (PSS 0.25, tstat 1.96)	
Heckman, Ichimura & Todd, 1997	tabulation; RCT	earnings +* (PSS 0.5, tstat 0.39)	
Heckman & Smith, 1999	tabulation; RCT	earnings +* (PSS 0.25, tstat 0.39)	
Kornfeld & Bloom, 1999	tabulation; RCT+administrative data	empl.prob. °	earnings °
Heckman & Smith, 2000	multivariate linear; RCT	earnings 0* (PSS 0, tstat -0.85)	

United States (local) – Ohio Transitions to Independence Demonstration			
Main intervention: Publicly implemented and financed. Two components targeting different groups: 1. JOBS – mandatory employment and training programme. Various education and training activities, ranging from post-secondary education to job skills training. 36 months' duration. 2. Work Choice – mandatory assessment of job readiness and vocational counselling plus voluntary JOBS activities. 12 months' duration. **Other features:** Targeted at Ohio's disadvantaged, welfare-recipient parents of all ages, but beneficiaries are mostly youths. JOBS targets adult recipients whose youngest child is aged six or over, Work Choice targets single parents whose youngest child is aged between one and six. Individual profiling and tailored "employability plan". JOBS programme uses an assessment-driven approach giving caseworkers substantial discretion in making placements. Heavily oriented towards education and training activities, intensive use of post-secondary education and paid public service work. For Work Choice, participants are offered childcare and extended Medicaid coverage if they should cease to be eligible for Aid to Dependent Children benefits due to earnings.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Fein, Beecroft & Blomquist, 1994	multivariate linear; RCT	empl.prob. + (SMD 0.04, PSS 0.73, tstat 1.96)	earnings 0* (PSS 0.05, tstat -0.98)
United States (local) – Quantum Opportunity Program (QOP)			
Main intervention: Randomized programme offering mentoring by full-time case workers, academic tutoring and life skills activities during the four years of high school. 750 hours per year. Implemented by Government and NGO, financed by Government and the Ford Foundation. **Other features:** Targeted at low-performing high school students in 11 schools from seven sites. Financial incentives provided. Case-managers develop a highly personal, long-lasting connection with each youth. Provision of supportive services.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Rodríguez-Planas, 2012	multivariate nonlinear; RCT+administrative panel data	empl.prob. – (SMD -0.05, PSS 0.05, tstat -0.60) earnings + (SMD 0.07, PSS 0, tstat 0.60)	hours worked – (SMD -0.04, PSS 0, tstat -0.41) wage 0 (SMD -0.00, PSS 0, tstat -0.37)

United States (national) – School-to-Work Opportunities Act (STWOA)		
Main intervention: Federal Act providing funding to support increased career preparation activities in the country's public schools. Six months' duration. Three components: 1. school-based initiatives – cooperative education, providing a practical way to apply knowledge learned in academic and technical studies and to gain real world-of-work experiences coupled with Tech Prep, preparing students for technical careers and other post-secondary educational experiences in technical and four-year colleges; 2. work-based activities – job shadowing, internships, apprenticeships, school-based enterprises; 3. connecting activities – development of partnerships with employers and post-secondary institutions. **Other features:** Targets first-time jobseekers with secondary education. Monitoring component. Emphasis on career links to academic curriculum, career-awareness activities, classroom instruction and curriculum integrating high academic standards with the knowledge and skills for post-secondary education or for a skill certificate.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Hall, 2000	tabulation; cross-section mail survey	empl.prob. +* (PSS 0.5, tstat 1.96) wage + (SMD 0.86, PSS 0, tstat 0.21)
Griffith, 2001	multivariate linear; originally collected + administrative cross-sectional data	hours worked ° wage ° earnings°
United States (local) – Summer Career Exploration Programme		
Main intervention: Private-sector summer youth programme. Provides job-related counselling, basic skills training and career exploration to help youth learn more about the world of work, their career interests and the importance of doing well academically and going on to college. Involves summer work in the private sector, pre-employment training, preparation for a job interview and learning about the college admission process and the offer provided by various colleges. Participants work for 25 hours a week for six weeks during the summer. Created by the William Penn Foundation and supported by a collaborative of foundations, corporations and trusts. **Other features:** Targets teenagers in the Philadelphia area. Emphasis on placing teens in jobs based on career interest, and use of college student monitors as role models. Participants also receive support from work-site supervisors, who provide the teens with specific work-related instruction and soft skills advice. Participants receive compensation equivalent to minimum wage.		
Reference	**Data collection and analysis**	**Overall treatment effect**
McClanahan, Sipe & Smith, 2004	tabulation; originally collected cross-sectional data	empl.prob. – (SMD -0.20, PSS 0) salary ° earnings °

United States (regional) – Workforce Investment Act (WIA)			
Main intervention: Aims to prepare youth for academic and employment success. Staff members work with each young person to develop a plan that may encompass counselling, tutoring, job training, mentoring or work experience. Other strategies include summer employment, study skills training or basic skills instruction in preparation for obtaining a GED. Publicly financed and implemented. **Other features:** Targets low-income youth with an educational deficiency. Voluntary. Youth are assessed to determine academic and skill level and support service needs.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Hollenbeck & Huang, 2006	regression+matching; administrative cross-sectional data	empl.prob. + (SMD 0.09, PSS 0.67, tstat 2.27) earnings + (SMD 0.02, PSS 0.33, tstat 0)	hours worked + (SMD 0.04, PSS 0.33, tstat 0) wage + (SMD 0.03, PSS 0.33, tstat 2.58)

8.2 CHARACTERISTICS OF INCLUDED PROGRAMMES – MAIN CATEGORY: ENTREPRENEURSHIP PROMOTION

Bosnia and Herzegovina (national) – Partner Microcredit Foundation Experiment			
Main intervention: Business and financial literacy programme. Business training (three days, nine hours total) through local NGO. **Other features:** Targets young loan clients of a non-profit lending organization. Clients paid approximately US$35 and offered free transportation to training location. Total cost of providing the course, including transportation and cash compensation, about US$245 per participant.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Bruhn & Zia, 2013	multivariate linear; RCT	investment + (SMD 0.06, PSS 0.25, tstat 0.69) business survival + (SMD 0.03, PSS 0, tstat -0.46)	profits 0* (PSS 0, tstat 0.64) no.employees 0* (PSS 0, tstat -0.09)
Colombia (national) – Jóvenes Rurales Emprendedores			
Main intervention: Business training programme. Two to six months' duration. Publicly financed and implemented. **Other features:** Targets low-income unemployed youth in rural areas to promote activities in the agricultural sector, the agro-industrial sector as well as services and industries.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Steiner et al., 2010	diff-in-diff+matching; panel survey data collected by authors	empl.prob. + (SMD 0.29, PSS 1, tstat 2.82) empl.qual. – (SMD -0.03, PSS 0, tstat -0.21)	hours worked – (SMD -0.04, PSS 0, tstat -0.35) wage + (SMD 0.21, PSS 0, tstat 1.89)

France (national) – CréaJeunes			
Main intervention: Four to six weeks of training, coaching, business plan presentation and 18 months of post-business creation support and access to finance. Three to four days per week almost full time. Publicly financed, implemented by NGOs and volunteers. **Other features:** Targets young unemployed people from disadvantaged urban areas. Training also includes life skills and personal development training sessions.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Crépon et al., 2014	tabulation; RCT	empl.prob. – (SMD -0.03, PSS 0, tstat -0.49) empl.qual. + (SMD 0.04, PSS 0, tstat 0.66) consumption 0* (PSS 0, tstat -0.24) profits 0* (PSS 0, tstat 0.06) sales 0* (PSS 0, tstat -1.51)	hours worked 0* (PSS 0, tstat -0.44) earnings 0* (PSS 0, tstat 0.40) wage 0* (PSS 0, tstat -0.77) business creation 0 (SMD 0.04, PSS 0, tstat 0.66) business survival 0 (SMD 0, PSS 0, tstat 0.10)
Liberia (local) – Economic Empowerment of Adolescent Girls (EPAG)			
Main intervention: Combination of six months of classroom-based training followed by six months of follow-up support for young women to enter wage employment or start a business. Two tracks: 1. Job Skills (JS) track, also includes short module on self-employment; 2. Business Development Services (BDS) track. Financed by Government, World Bank, the Nike Foundation and the Governments of Australia, the United Kingdom, Norway, Denmark and Sweden; implemented by Government and NGOs. **Other features:** Targets young women. Performance bonuses to training providers that successfully place their graduates in jobs or micro-enterprises. Programme designed around the girls' needs. Frequent and unannounced monitoring visits to ensure that service providers maintain high-quality learning environment. Training participants sign "Trainee Commitment Forms" at the start of the training. Small stipends and a completion bonus contingent upon attendance, free childcare at every training site, assistance to open a savings account at a local bank. Delivered in small groups, each with a coach or mentor, to foster support networks and boost attendance.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Adoho et al., 2014	regression adjusted diff-in-diff; RCT	empl.prob. + (SMD 0.29, PSS 0.67, tstat 3.22)	earnings 0* (PSS 0, tstat 2.80)

Peru (regional) – Calificación de Jóvenes Creadores de Microempresas		
Main intervention: Business training, business advisory and finance programme with business plan competition. Publicly financed and implemented. Two phases: 1. pre-phase of business advisory (two hours) and business training (24 hours) to develop a business plan; 2. post-phase of business advisory (12 hours), business training (36 hours), internship (170 hours) and access to finance. **Other features:** Targeted at low-income young people owning an informal business for less than a year or interested in starting a business. Only winners of business plan competition eligible for participation in second phase.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Jaramillo & Parodi, 2003	matching; originally collected+administrative cross-sectional data	jobs created +* (PSS 1, tstat 2.20)
Peru (local) – Formación de Líderes Empresariales		
Main intervention: Business training (100 hours), business advisory services (12 hours), participation in business fairs (80 hours) and access to credit intervention. Publicly financed and implemented. **Other features:** Targets young microenterprise owners in rural areas. Winners of business plan competition eligible for access to credit.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Jaramillo & Parodi, 2005	tabulation; RCT	business creation + (SMD 0.20)
Peru (regional) – Formación Empresarial De La Juventud (Project JUMP)		
Main intervention: Business training (four weeks), business advisory and finance assistance. Four business advisory visits over three months during post-business creation phase. Publicly financed and implemented. **Other features:** Targeted at poor young people. Personalized coaching to develop business plan. Winners of business plan competition eligible for funding.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Jaramillo & Parodi, 2003	matching; originally collected+ administrative cross-sectional data	business creation +* (PSS 1, tstat 2.10)

Tunisia (national) – Turning Theses into Enterprises			
Main intervention: Business training (20 days' full time), including behavioural skills training, and personalized coaching to develop a business plan as well as access to finance for winners of business plan competition. Entrepreneurship courses organized by public employment office, supervision from university professors in business plan development. **Other features:** Targets university students in applied undergraduate curriculum. Opportunity for students to graduate with a business plan instead of following the standard curriculum. External private sector coaches. Winners of business plan competition eligible to receive seed capital up to US$10,000.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Almeida et al., 2012	IV+multivariate linear; RCT	empl.prob. + (SMD 0.09, PSS 0.4, tstat 1.28) hours worked 0* (PSS 0, tstat 0.68) earnings 0* (PSS 0, tstat 0.01)	unempl.prob. – (SMD -0.02, PSS 0, tstat -0.27) empl.qual. – (SMD -0.01, PSS 0, tstat -0.12)
Uganda (regional) – Start and Improve Your Business (SIYB) programme			
Main intervention: Randomized experiment. Beneficiaries receive one of the following: 1. a cash grant; 2. a loan; 3. business skills training and a loan; 4. business skills training and a cash grant. Business skills training lasted for 32 hours in total. Cash grants delivered through free savings accounts. Loans (50 per cent collateral required) provided by local microfinance organization and guaranteed by the ILO. **Other features:** Targets microenterprise owners in semi-urban locations. Cash grants US$200. Loans US$180–220.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Fiala, 2014	multivariate linear; RCT	consumption – (SMD -0.01, PSS 0, tstat -0.30) no.employees – (SMD -0.04, PSS 0.06, tstat -0.03) capital + (SMD 0.03, PSS 0, tstat 0.02)	profits + (SMD 0.01, PSS 0.08, tstat 0.08) sales – (SMD -0.02, PSS 0, tstat -0.06)

Uganda (regional) – Women's Income Generation Support (WINGS)			
Main intervention: Business skills training (three days, 24 hours in total), cash grants (US$150) and follow-up support to start non-farming businesses. Business plan developed after training provision. Approved business plans are eligible for start-up grant. Financed and implemented by NGO. **Other features:** Targets young women in war-affected region. Three follow-up visits by trained community workers to monitor and support the business activities. Additional option of group training to form business support networks and spousal inclusion in training (four days, 32 hours in total).			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Blattman et al., 2013	multivariate linear; RCT	hours worked +* (PSS 1, tstat 1.96) consumption +* (PSS 1, tstat 1.96)	earnings +* (PSS 1, tstat 1.96)
Blattman et al., 2014	multivariate linear; RCT	empl.prob. + (SMD 0.86, PSS 1, tstat 6.93) earnings +* (PSS 1, tstat 3.85) business survival + (SMD 0.84, PSS 1, tstat 4.91)	hours worked +* (PSS 0.5, tstat 2.05) business creation + (SMD 0.86, PSS 1, tstat 1.96)
Uganda (regional) – Youth Opportunities Programme (YOP)			
Main intervention: Government grant programme to help beneficiaries become self-employed artisans through one-time unsupervised cash transfers. Beneficiaries invited to form groups and submit grant proposals for non-agricultural vocational training and enterprise start-up. **Other features:** Targets poor and unemployed young adults. Grants around US$382. No formal Government monitoring after cash transfer made.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Blattman, Fiala & Martinez, 2014	multivariate linear; RCT	unempl.prob. + (SMD 0.07, PSS 0.5, tstat 1.59) earnings + (SMD 0.17, PSS 0.83, tstat 2.86) no. employees + (SMD 0.21, PSS 1, tstat 3.29)	hours worked + (SMD 0.19, PSS 1, tstat 4.06) consumption + (SMD 0.13, PSS 0.67, tstat 2.17) capital + (SMD 0.22, PSS 0.83, tstat 3.46)

United Kingdom – The Prince's Trust		
Main intervention: Business start-up assistance for young people providing start-up loan or grant and business mentors (one to two hours' contact per month). Publicly financed. **Other features:** Size of start-up loans and grants approximately £1,500–£3,500.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Meager, Bates & Cowling, 2003	multivariate linear and nonlinear; cross-sectional administrative data	empl.prob. +* (PSS 0.5, tstat 4.22) wage 0* (PSS 0, tstat 0.19) earnings -* (PSS 0, tstat -3.64)

8.3 CHARACTERISTICS OF INCLUDED PROGRAMMES – MAIN CATEOGRY: EMPLOYMENT SERVICES

Ethiopia (local) – Franklin Subsidized Transport Experiment			
Main intervention: Distributes money to cover transport costs to travel to the centre of the city to receive information about new jobs. Duration eight to 11 weeks. Designed and implemented by individual researcher, financed by the World Bank and the International Growth Centre. **Other features:** Targets young unemployed educated jobseekers living on the outskirts of Addis Ababa. Subsidy covers two trips to the city centre per week. Designed to be a low-cost programme, which could be easily scalable.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Franklin, 2014	multivariate linear; RCT	empl.prob. + (SMD 0.14, PSS 0.25, tstat 1.26) wage 0* (PSS 0, tstat 0.58)	hours worked +* (PSS 1, tstat 2.19)
Finland (national) – Finnish Vocational Labour Market Training (LMT)			
Main intervention: Counselling, intensified monitoring and individualized job-search plans that guarantee activation measures for the unemployed. Individualized job-search plan has to be signed within three months from registration. Job-search training. Different pathways for skilled and unskilled youth. Skilled receive job coaching, work practice and job placements with employment subsidies. Programmes for unskilled include career planning and information on different educational possibilities, mainly through work practice in specific youth workshops offered by municipalities. Publicly financed and implemented. **Other features:** Targets unemployed youth. The job-search plan has to explicitly name the activation measures agreed upon, which are mutually binding: Implementer is obliged to offer the activation measures included in the plan within three months from signing the contract, and non-compliance on the part of the jobseeker can be sanctioned. Work practice entails non-salaried employment with compensation at the minimum unemployment allowance.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Hämäläinen, Hämäläinen & Tuomala, 2014	regression-adjusted diff-in-diff; administrative panel data	unempl.prob. 0 (SMD 0.00, PSS 0, tstat 0.44) unempl.durat. - (SMD -0.07, PSS 0, tstat -1.50)	earnings – (SMD -0.05, PSS 0, tstat -1.27)

France (regional) – Counselling and Job Placement for Young Graduate Job Seekers		
Main intervention: Job placement assistance programme. Publicly financed, private providers (both for-profit and NGOs). Two phases: 1. agency counsels jobseeker and helps them to find a durable job with a length of at least six months; 2. agency supports beneficiary in their job. **Other features:** Targeted at young, educated jobseekers who have been unemployed for at least six months. Run in ten experimental regions. Strong mentoring component. The private provider is paid in three stages, with each payment conditional on the fulfilment of a corresponding objective: The last payment is conditional on the individual finding a job with a contract of at least six months' duration and staying employed for at least six months.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Crépon et al., 2013	multivariate linear+IV; RCT	empl.prob. + (SMD 0.05, PSS 0.33, tstat 1.17) wage – (SMD -0.01, PSS 0, tstat -0.22) earnings 0 (SMD -0.00, PSS 0.11, tstat 0.01)
Germany (national) – Job search assistance (JS)		
Main intervention: Job-search monitoring and assessment of the career opportunities of individuals. Publicly financed and implemented. **Other features:** Targets the young unemployed with secondary education. Strong monitoring component.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Caliendo, Künn & Schmidl, 2011	matching; administrative cross-sectional data	empl.prob. +* (PSS 0.61, tstat 1.53)

Germany (national) – Mandatory visits to job information centres			
Main intervention: Mandatory visits to job information centres while still attending school. Job information centres are public establishments providing detailed and comprehensive information on occupations, vocational training and apprenticeships, higher education, job tasks, earnings prospects, local labour market conditions and, if required, counselling to facilitate individuals' labour market-related choices. **Other features:** Targets students in secondary school. The centres are designed to combine visitors' autonomous retrieval of information with assistance from professional job counsellors. Mandatory: eligible students have to visit a job information centre no later than two years before they leave secondary education.			
Reference	**Data collection and analysis**	**Overall Treatment effect**	
Saniter, 2014	diff-in-diff; originally collected panel data+administrative panel data	empl.prob. 0 (SMD -0.00, PSS 0, tstat -0.19) unempl.durat. 0* (PSS 0, tstat 0.65) wage 0* (PSS 0.17, tstat -0.02)	unempl.prob. + (SMD 0.17, PSS 0, tstat 1.50) earnings 0* (PSS 0, tstat 0.34)
India (regional) – BPO recruiting services			
Main intervention: Recruiting services to help participants obtain jobs in the business process outsourcing industry, providing information on employment opportunities in the sector, interview skills lessons, mock interviews and assessment of English language skills. Three annual sessions, each lasting between four and six hours. Designed and implemented by individual researcher, financed through private and public funds. **Other features:** Targets women in randomly selected rural villages in the states of Haryana, Punjab, Rajasthan and Uttar Pradesh. No restrictions but the jobs are mainly for those with a secondary school degree, some English language ability and experience with computers: this means that the beneficiaries are mainly young women. Intensive advertisement campaign. The jobs are competitive, so no guarantee of employment.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Jensen, 2012	diff-in-diff+IV; RCT	empl.prob. + (SMD 0.01, PSS 0.5, tstat 0.27)	consumption 0* (PSS 0, tstat -0.66)

Jordan (local) – Jordan New Opportunities for Women 2.0 (Jordan NOW 2.0)		
Main intervention. Two experimental job-search assistance programmes. Implemented by private sector training services firm, financed by TFESSD trust fund, the Knowledge for Change trust fund and the Research Support Budget of the World Bank. Three phases: 1. testing participants on technical, soft skills and personal traits; 2. individual profiling; 3. matching candidates with firms to set up an interview. **Other features:** Targeted at unemployed youth with tertiary education in Amman. Experiment 1 initially restricted to females, then opened to males as well; experiment 2 excludes those already employed or those not seeking a job. Intensive advertising campaign.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Groh et al., 2014	multivariate linear; RCT+administrative cross-sectional data	empl.prob. + (SMD 0.03, PSS 0, tstat 0.42) salary – (SMD -0.03, PSS 0, tstat -0.22)
Portugal (national) – Programa Inserjovem		
Main intervention: Job-search support initiatives, involving vocational guidance, counselling, monitoring, intensive job-search assistance and small basic skills training. If deemed necessary, beneficiaries can participate in a number of vocational or non-vocational training courses. Publicly financed and implemented. **Other features:** Targeted at first-time jobseekers, who must be enrolled in the programme prior to completion of six months of unemployment. Participation is mandatory: non-enrolment leads to loss of unemployment benefits. Elaboration of a Personal Employment Plan, whereby the beneficiary is expected to meet on a regular basis with the placement officer and to actively search for a job. Unjustified rejection of job offers leads to cancellation of registration.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Centeno & Novo, 2006	diff-in-diff+matching; administrative panel data	unempl.duration 0* (PSS 0, tstat 0.91)
Centeno, Centeno & Novo, 2009	diff-in-diff+matching; administrative panel data	unempl.duration 0* (PSS 0, tstat -0.69)

United States (national) – School-to-Work Opportunities Act (STWOA)		
Main intervention: Federal Act providing funding to support increased career preparation activities in the country's public schools. Six months' duration. Three components: 1. school-based initiatives – Cooperative Education, providing a practical way to apply knowledge learned in academic and techical studies and to gain real world-of-work experiences and Tech Prep, preparing students for technical careers and other post-secondary educational experiences in technical and four-year colleges; 2. work-based activities – job shadowing, internships, apprenticeships, school-based enterprises; 3. connecting activities – development of partnerships with employers and post-secondary institutions. **Other features:** Targets first-time jobseekers with secondary education. Monitoring component. Emphasis on career links to academic curriculum, career-awareness activities, classroom instruction and curriculum integrating high academic standards with the knowledge and skills for post-secondary education or for a skill certificate.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Gong, 2005	matching; survey+administrative panel data	wage + (SMD 0.01, PSS 0.14, tstat 1.09)
Neumark & Rothstein, 2006	multivariate linear; cross-sectional survey data	empl.prob. 0* (PSS 0.17, tstat 0.51)

8.4 CHARACTERISTICS OF INCLUDED PROGRAMMES – MAIN CATEGORY SUBSIDIZED EMPLOYMENT

Australia (national) – Special Youth Employment and Training Programme (SYETP)		
Main intervention: Publicly financed. Two versions: 1. Standard SYETP – flat rate subsidy of A$75 a week paid to employers for 17 weeks, equivalent in value to half the average teenage wage. 2. Extended SYETP: subsidy of A$100 for 17 weeks, and A$75 for a further 17 weeks. **Other features:** Standard SYETP targets youth claiming unemployment benefits and not studying full-time for at least four of preceding 12 months. Extended SYETP targets youth who have been unemployed for at least eight of the preceding 12 months. To obtain the subsidy, employers register their vacancies with the Commonwealth Employment Service (CES), and have to accept workers referred by the CES. Employers agree on a training plan for the individual worker with the CES (which could, in practice, be even simple orientation services).		
Reference	**Data collection and analysis**	**Overall treatment effect**
Richardson, 1998 Knight, 2002	multivariate nonlinear; cross-sectional survey data matching+multivariate nonlinear; cross-sectional survey data	empl.prob. + (SMD, 0.53) empl.prob. 0* (PSS 0, tstat 0.73)
Canada (national) – Youth Hires		
Main intervention: For eligible workers, any employment insurance premiums paid in 1999 and 2000 in excess of the firm's 1998 premiums are refunded to the employer. Publicly financed and implemented. **Other features:** Targeted at youth. Firms receive no rebate for wage increases to any individual worker whose annual earnings exceed the maximum insurable limit. Extensive awareness-raising campaign. Premium rebate presented as being automatic and without administrative burden.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Webb, Sweetman & Warman, 2014	regression-adjusted diff-in-diff; survey panel data	empl.prob. + (SMD 0.06, PSS 0.17, tstat 1.39) participation rate + (SMD 0.08, PSS 0.67, tstat 1.80) empl.qual. – (SMD -0.10, PSS 0, tstat -1.53) wage – (SMD -0.05, PSS 0, tstat -0.86) unempl.prob. 0 (SMD 0.00, PSS 0, tstat 0.18) hours worked – (SMD -0.01, PSS 0, tstat -0.46) earnings – (SMD -0.04, PSS 0, tstat -0.82)

Chile (national) – Subsidio al Empleo Joven		
Main intervention: 30 per cent subsidy of the monthly wage: 10 per cent reduction in social security contributions for employers plus 20 per cent wage subsidy directly paid to worker. Both employees and self-employed eligible. Four years' duration. Publicly financed and implemented. **Other features:** Targeted at disadvantaged youth in formal jobs. Incentives to complete secondary education. Extension possible if beneficiary enrols in higher education and for pregnant young women.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Universidad de Chile, 2012	RDD; administrative cross-sectional data	empl.prob. +* (PSS 0.94, tstat 1.86) empl.qual. +* (PSS 0.83, tstat 1.64) participation rate +* (PSS 1, tstat 1.96)
France (national) – Contrat Jeune en Entreprise		
Main intervention: Wage subsidy for employers hiring an eligible young worker on an open-ended contract. The subsidy amount depends on the wage paid, and is proportional to the part-time ratio for part-time workers. Employers receive the subsidies for two years and then half the monthly allowance in the third year. Publicly financed and implemented. **Other features:** Targets disadvantaged young people who dropped out of school before passing the secondary school examination that would qualify them for entry to university. All employers are entitled to claim the subsidy. No dismissal, except for professional misconduct, is allowed during the first three years of the contract.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Roger & Zamora, 2011	regression-adjusted diff-in-diff; survey panel data	empl.prob. + (SMD 0.04, PSS 0, tstat 0.42)
France (national) – Stages d'Initiation à la Vie Professionnelle (SIVP)		
Main intervention: Subsidized employment programme, tripartite contract between the State, the worker and the employer. Only private employers eligible. Flat wage rate depending on the age of the beneficiary plus top-up by employer of 17–27 per cent of the minimum wage (employer's social security contributions on this amount are waived). 36 months' duration, non-renewable. **Other features:** Targeted at first-time jobseekers or youth claiming unemployment benefits for over 12 months. Monitoring provided by an individual tutor within the firm and by an external monitoring organization (25 hours per month).		
Reference	**Data collection and analysis**	**Overall treatment effect**
Brodaty, 2007	matching; panel survey data	empl.prob. +* (PSS 0.76, tstat 3.16)

France (national) – Stages de formation		
Main intervention: Internship placement. Six to nine months' duration. Flat stipend paid by the Government. **Other features:** Targeted at first-time jobseekers with, at most, secondary education who dropped out of school before passing the secondary school examination that would qualify them for entry to university. Any employer eligible. Also features 200–500 hours of training financed by the Government.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Brodaty, 2007	matching; panel survey data	empl.prob. +* (PSS 0.74, tstat 2.93)
France (national) – Travaux d'Utilité Collective (TUC)		
Main intervention: Social Development and Community Public Works project. Three to 24 months' duration. Salary of 1250F paid by Government plus top-up of up to 500F paid by employer (employer's social security contributions on this amount are waived). **Other features:** Targeted at unemployed youth. Those participants who are relatively elder, with respect to their fellow beneficiaries, must have been registered as unemployed for at least one year. Possibility of complementary soft skills/employability skills training.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Brodaty, 2007	matching; panel survey data	empl.prob. 0* (PSS 0.11, tstat 1.18)
Germany (national) – Job creation schemes (JCS)		
Main intervention: Working opportunity in areas of public interest (infrastructure, social work). Low level of remuneration subsidized by employment agency. Maximum duration 12 months, but may be extended if likely to lead to regular employment. **Other features:** Targeted at unemployed youth with secondary education who have very little previous labour market experience and potentially low labour market attachment. Placement subordinate to placement in training or regular employment: Parallel qualification measures should be implemented, but could be omitted if not deemed necessary.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Caliendo, Künn & Schmidl, 2011	matching; administrative cross-sectional data	empl.prob. -* (PSS 0, tstat 2.23)

Germany (national) – JUMP wage subsidies (JWS)		
Main intervention: Wage subsidy to regular employment with minimum 15 hours per day at the maximum amount of 60 per cent (40 per cent) of the full wage, for a maximum duration of one (two) years. Employers have to guarantee a period of post-subsidy employment equivalent to half the subsidized period. **Other features:** Targeted at unemployed youth with secondary education. No minimum duration in unemployment necessary.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Caliendo, Künn & Schmidl, 2011	matching; administrative cross-sectional data	empl.prob. +* (PSS 0.64, tstat 2.34)
Germany (national) – SGB III wage subsidies (WS)		
Main intervention: Wage subsidy to regular employment at the maximum amount of 50 per cent of the full wage, for a maximum period of one year. Employers have to guarantee a period of post-subsidy employment of the same duration as the subsidized period, for a maximum of 12 months. **Other features:** Targeted at unemployed youth with secondary education. No minimum duration in unemployment necessary.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Caliendo, Künn & Schmidl, 2011	matching; administrative cross-sectional data	empl.prob. +* (PSS 0.57, tstat 1.38)
Jordan (pilot) – Jordan New Opportunities for Women (Jordan NOW)		
Main intervention: Experimental scheme providing a job voucher that beneficiaries can take to a firm while searching for jobs. The job voucher pays the employer an amount equal to the mandatory minimum monthly wage of 150JD per month for a maximum of six months within an 11-month period. Financed by the World Bank, implemented by local NGO. **Other features:** Targets young female graduates. To be eligible to use the voucher, a firm has to provide proof of registration, have a bank account to receive payment and provide an offer letter with the graduate's name and specification of work duties. After the start of employment, both the firm and graduate are required to confirm the employment with the programme administrator each month, with periodic monitoring and random visits. Extensive advertising campaign.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Groh et al., 2012	multivariate linear; RCT+administrative panel data	empl.prob. + (SMD 0.05, PSS 0.08, tstat 0.78) hours worked 0* (PSS 0, tstat 0.49) participation rate + (SMD 0.07, PSS 0.33, tstat 1.12) earnings 0* (PSS 0, tstat 0.20)

South Africa (regional) – Youth Wage Subsidies for South Africa			
Main intervention: RCT randomly allocating a voucher to participants allowing firms that employ them to be compensated for a portion of the wages. Individuals need to be employed full-time in a formal non-government business. The subsidy amount is capped at half the wage or R833 per month and can be claimed for a maximum of six months or up to R5,000. Financed by 3ie through the Global Development Network, the South African National Treasury and Department of Labour, the European Union through the Programme to Support Pro-Poor Policy Development and the World Bank through the Strategic Impact Evaluation Fund. Implemented by Government and individual researchers. **Other features:** Targets youth in the Johannesburg metropolitan area in Gauteng province, the eThekwini (greater Durban) metropolitan area of KwaZulu-Natal province and the urban area of Polokwane and surrounding rural areas of the Limpopo province. Firms have to be officially registered for tax and be paying unemployment insurance. Subsidies are transferable between companies; an individual takes the unclaimed subsidy with them should they leave a firm.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Levinsohn et al., 2014	multivariate linear+IV; RCT	empl.prob. + (SMD 0.21, PSS 1, tstat 3.08) earnings + (SMD 0.01, PSS 0, tstat 0.25)	participation rate + (SMD 0.02, PSS 0, tstat 0.19)
Sweden (national) – Swedish employer-paid payroll tax			
Main intervention: Payroll tax reduction of 15.2 per cent. Six out of seven mandatory fees are halved. **Other features:** Targets young workers. No conditionality. Any employer eligible.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Egebark & Kaunitz, 2014	regression-adjusted diff-in-diff; panel survey data	empl.prob. + (SMD 0.01, PSS 0.5, tstat 2.73)	wage + (SMD 0.04, PSS 0.5, tstat 2.11)

Sweden (national) – Youth Practice		
Main intervention: Subsidized employment programme, placing participants in both the private and the public sector. Six months' duration with a possible extension to 12 months. Wage subsidy as employers pay a very small fraction of the participation allowance. Participation is preceded by at least four months' active job search as openly unemployed, and beneficiaries should allocate between four and eight hours a week to jobseeking activities at the local employment office. Publicly financed, both public and private implementers. **Other features:** Targeted at the young unemployed with a high school diploma. If the participant is entitled to unemployment benefits, they receive an allowance equal to the benefit. Jobs include a mixture of work and training, leading to human capital accumulation. Participants should be a supplementary resource for the employer and not displace regular employment.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Larsson, 2003	matching; administrative cross-sectional data	empl.prob. – (SMD -0.04, PSS 0, tstat -0.82) earnings -* (PSS 0, tstat -2.16)
Costa Dias, Ichimura & van den Berg, 2013	IV+matching; administrative panel data	empl.prob. 0* (PSS 0, tstat -0.98)
Tunisia (national) – Stage d'Initiation à la Vie Professionnelle (SIVP)		
Main intervention: Employment subsidy ranging between 100 and 250 TND depending on the level and the subject of the degree studied and the year of study completed. One year duration. About two-thirds of the graduate's wage is paid by the employer; the Government subsidizes the other third. Employer exempt from social security contributions (for an additional year under certain conditions). **Other features:** Targets university graduates who are looking for their first job and have been unemployed for at least three months (less for certain degrees facing particular recruitment difficulties). If employers break the contract, they must reimburse the subsidies and pay the social security contributions. There are restrictions governing the proportion of subsidized workers that can be hired.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Broecke, 2013	matching; originally collected cross-sectional data	empl.prob. + (SMD 0.28, PSS 0.87, tstat 2.97) empl.qual. + (SMD 0.28, PSS 0.4, tstat 1.20) earnings 0 (SMD -0.00, PSS 0, tstat -0.04)

Turkey (national) – 2008 Employment Package			
Main intervention: Tax cuts and subsidies applied to payroll tax burden on employers for beneficiaries: 100 per cent subsidized in first year, 80 per cent in second and so on through the fifth year of employment. Implemented and financed by the Government. **Other features:** Targeted at young males. Only individuals who were not registered for social security for the six months before the law's passage qualify. Newly hired people must fill new positions.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Barza, 2011	diff-in-diff; repeated cross-sectional survey data	empl.prob. – (SMD -0.05, PSS 0.5, tstat 0.02)	unempl.prob. + (SMD 0.21, PSS 1, tstat 1.96)
United States (national) – School-to-Work Opportunities Act (STWOA)			
Main intervention: Federal Act providing funding to support increased career preparation activities in the country's public schools. Six months' duration. Three components: 1. school-based initiatives – cooperative education, providing a practical way to apply knowledge learned in academic and technical studies and to gain real world-of-work experiences plus Tech Prep, preparing students for technical careers and other post-secondary educational experiences in technical and four-year colleges; 2. work-based activities – job-shadowing, internships, apprenticeships, school-based enterprises; 3. connecting activities – development of partnerships with employers and post-secondary institutions. **Other features:** Targets first-time jobseekers with secondary education. Monitoring component. Emphasis on career links to academic curriculum, career-awareness activities, classroom instruction and curriculum integrating high academic standards with the knowledge and skills for post-secondary education or for a skill certificate.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Gong, 2005	matching; survey+administrative panel data	wage + (SMD 0.01, PSS 0.14, tstat 1.09)	
Neumark & Rothstein, 2006	multivariate linear; cross-sectional survey data	empl.prob. 0* (PSS 0.17, tstat 0.51)	

United States (national) – Targeted Jobs Tax Credit (TJTC)			
Main intervention: Wage subsidy in the form of a tax credit to employers of recently hired eligible workers. Any employer is eligible and is provided with a credit of 40 per cent of wages paid in the first 12 months of employment, up to a maximum of $6000. A minimum of 90 days or 120 hours of employment is required to claim a credit. Publicly financed. **Other features:** Targeted at economically disadvantaged youth, economically disadvantaged Vietnam young veterans, economically disadvantaged ex-offenders, handicapped persons receiving or having completed vocational rehabilitation, general assistance recipients and social security insurance recipients. Can be obtained as a voucher issued to the beneficiary applying for any job or an employment service certification requested by the employer if the new employee is considered eligible but does not have a voucher.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Hollenbeck & Willke, 1991	multivariate linear; administrative panel data	empl.prob. -* (PSS 0.08, tstat -2.08)	wage 0* (PSS 0.42, tstat -0.32)

8.5 CHARACTERISTICS OF INCLUDED PROGRAMMES – UNSPECIFIED MAIN CATEGORY

Germany (pilot) – ALMP for disadvantaged youth in Germany		
Main intervention: Three components: 1. individual coaching; 2. classroom training; 3. temporary work. An employment agency is involved for job placement. Financed by Government; implementers can be public or private.		
Other features: Targets low-skilled young unemployed. Participants receive salary. Individual profiling and skills assessment take place at the temporary work agency. Content of training is allowed to differ between individuals.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Ehlert, Kluve & Schaffner, 2012	OLS; cross-sectional administrative data	empl.prob. + (SMD 0.27, PSS 0.37, tstat 1.56)
United Kingdom (national) – New Deal for Young People		
Main intervention: Combination of job-search assistance, training, wage subsidies and job experience. Financed by Government, implementers can be public or private. Multistage programme: 1. Gateway: intensive job search; 2. if option 1 unsuccessful – option to choose between subsidized employment, full-time education or training course, work in the voluntary sector or government job; 3. Follow-through: similar to Gateway, but of shorter duration. **Other features:** Targeted at young individuals who have been claiming unemployment insurance for six months. Mandatory. Sanction for non-compliers is temporary withdrawal of welfare benefits. Each individual is assigned a personal adviser, with whom they meet at least biweekly.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Wilkinson, 2003	regression-adjusted diff-in-diff; pooled cross-sectional administrative data	empl.prob. +* (PSS 0.67, tstat 1.96) unempl.prob. -* (PSS 1, tstat -5.89)
Blundell et al., 2004	diff-in-diff+PSM; originally collected panel data+administrative data	empl.prob. + (SMD 0.15, PSS 0.11, tstat 1.19)
De Giorgi, 2005	RDD; originally collected panel data+cross-sectional survey data	empl.prob. +* (PSS 1, tstat 3.22)

United States (regional) – National Guard Youth Challenge Programme		
Main intervention: Financed by Government, implementers can be public or private. Three phases: 1. Pre ChalleNGe – orientation and assessment period, two weeks; 2. Residential Phase – eight core components to promote positive youth development: Leadership/Followership, Responsible Citizenship, Service to Community, Life-Coping Skills, Physical Fitness, Health and Hygiene, Job Skills, Academic Excellence. Most programmes help participants prepare for the GED exam, but a few of them offer a high school diploma, 20 weeks; 3. Post-residential Phase, featuring a mentoring programme, one year. **Other features:** Targets disadvantaged and low-educated youths and school dropouts. Large military component; participants live in military bases and are enrolled as "cadets". Strong mentorship component, but mentors are self-selected.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Millenky et al., 2011	OLS; RCT	empl.prob. + (SMD 0.09, PSS 0.33, tstat 1.42) wage + (SMD 0.12, PSS 1, tstat 2.10) earnings +* (PSS 0.5, tstat 2.34)
United States (regional) – New Chance		
Main intervention: Array of services including instruction in basic academic skills, career exposure and employability development classes, occupational skills training, work experience, job-placement assistance, health and family planning classes and services, life skills classes. 12–18 months' duration. Publicly financed, implemented by Government and NGOs. **Other features:** Voluntary for young mothers receiving welfare; however, daily attendance at all classes is expected five days a week. Non-monetary benefits are provided in the form of childcare and free meals.		
Reference	**Data collection and analysis**	**Overall treatment effect**
Chang et al., 2007	OLS; RCT	empl.prob. – (SMD -0.04, PSS 0) hh income °
Quint, Bos & Polit,1997	OLS; RCT	empl.prob. – (SMD -0.01, PSS 0, tstat -0.34) unempl.prob. – (SMD -0.01, PSS 0, tstat -0.12) particip.rate – (SMD -0.06, PSS 0, tstat -1.20) empl.qual. – (SMD -0.02, PSS 0, tstat) hours worked 0* (PSS 0, tstat -1.52) earnings 0* (PSS 0, tstat -0.85) hh income 0* (PSS 0, tstat -0.29) wage 0* (PSS 0, tstat 0.07)

United States (pilot) – Teenage Parent Demonstration			
Main intervention: Wide array of employment-oriented services, including attending regular high school, enrolling in alternative education programmes, participating in job training, working. Financed by Government; implementers can be public or private. **Other features:** Mandatory programme for teenage mothers receiving welfare. Non-compliance is sanctioned by welfare benefits loss. Each participant has a case manager who develops their tailored service plan. Strong monitoring component. Incentives provided in the form of a modest daily stipend and childcare and transport assistance.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Maynard, Nicholson & Rangarajan, 1993	OLS; RCT	empl.prob. + (SMD 0.12, PSS 0.34, tstat 1.93) hh income °	earnings + (SMD 0.09, PSS 0.26, tstat 1.61)
United States (national) – Youth Opportunity Grant Initiative			
Main intervention: Several programmes in the different localities opening a centre. Total of 15 youth-development activities available (job-readiness training and life-skills training being the most common). Sports and recreation, short-term unsubsidized employment, internships, community service and mathematics and reading remediation also offered. Publicly financed. **Other features:** Targets disadvantaged youth residing in high-poverty communities in urban, rural and Native American reservation communities. Level of participation per participant varies widely across grantees and enrolees. Strong monitoring component. Focus on educational achievement, literacy/numeracy and remedial classes. Grantees encouraged to establish partnerships with public, private and non-profit organizations to leverage resources.			
Reference	**Data collection and analysis**	**Overall treatment effect**	
Jackson et al., 2007	diff-in-diff+matching; originally collected+survey+administrative panel data	empl.prob. 0* (PSS 0.12, tstat -0.48) particip.rate ° wage +* (PSS 0.2, tstat 1.10)	unempl.prob. 0* (PSS 0, tstat -1.96) hours worked 0* (PSS 0, tstat -0.69)

8.6 SEARCH

8.6.1 List of included sources

Source	Type of source	Search date
7th IZA/World Bank Conference: Employment and Development, November 2012 http://www.iza.org/conference_files/worldb2012/	Conference website	4 February 2014
8th IZA/World Bank Conference on Employment and Development, August 2013 http://www.iza.org/conference_files/worldb2013/	Conference website	4 February 2014
IZA conferences and meetings http://www.iza.org/en/webcontent/events/izaconferences_html	Conference website	4 February 2014
2012 Global Youth Economic Opportunities Conference, September 2012 http://www.youtheconomicopportunities.org/	Conference website	4 February 2014
2013 Global Youth Economic Opportunities Conference, September 2013 http://www.youtheconomicopportunities.org/	Conference website	4 February 2014
Ideas4Work Conference: Youth Employability and Entrepreneurship in Africa, January 2013 http://www.iyfnet.org/library/ideas4work-conference-youth-employability-entrepreneurship-africa	Conference website	4 February 2014
Arab Youth & Entrepreneurship: Holistic Approaches to Nurturing Local Ecosystems, February 2013 http://aye2013.org/	Conference website	4 February 2014
Youth Entrepreneurship Knowledge sharing Symposium for East Africa, September 2013 http://www.ilo.org/addisababa/events-and-meetings/WCMS_221001/lang--en/index.htm	Conference website	2 February 2014

Source	Type of source	Search date
Regional Youth Employment Consultation in Latin America, International Development Research Centre (IDRC) and Instituto de Pesquisa Econômica Aplicada (IPEA), 5–6 December 2013	Conference website	2 February 2014
World Bank Regional Impact Evaluation Workshop: Quality education, skills and productivity among youth in Africa, Dakar, Senegal, September 30 – October 4, 2013	Conference website	1 March 2014
Doha Evidence Symposium: Increasing Youth Productivity in the Middle East and North Africa, 6–8 March 2014	Conference website	12 March 2014
Electronic Theses Online Service (EThOS) http://ethos.bl.uk	Dissertations and theses databases	4 February 2014
Networked Digital Library of Theses and Dissertations – SCIRUS ETD Search http://www.ndltd.org/serviceproviders/scirus-etd-search	Dissertations and theses databases	15 December 2013
ProQuest Dissertations & Theses Database www.proquest.co.uk/en-UK/catalogs/databases/detail/pqdt.shtml	Dissertations and theses databases	16 December 2013
ASSIA (Applied Social Sciences Index and Abstracts)	General database	15 December 2013
RePEc (Research Papers in Economics)/IDEAS Economics and Finance Research: http://ideas.repec.org/	General database	15 December 2013
Sociological Abstracts	General database	15 December 2013
SSRN (Social Science Research Network) http://www.ssrn.com/	General database	15 December 2013
Social Science Citation Index (SSCI) and Arts and Humanities Citation Index (AHCI) of Web of Science	General database	15 December 2013

Source	Type of source	Search date
http://thomsonreuters.com/web-of-knowledge/		
ABI/INFORM Global	General database	15 December 2013
EconLit	General database	15 December 2013
ERIC (Education Resources Information Centre)	General database	15 December 2013
IBSS (International Bibliography of the Social Sciences)	General database	15 December 2013
JSTOR http://www.jstor.org/	General database	15 December 2013
PAIS International	General database	15 December 2013
OpenGrey http://www.opengrey.eu/	Grey literature databases	4 February 2014
Directory of Open Access Repositories (OpenDOAR) http://www.opendoar.org/	Grey literature databases	4 February 2014
African Development Bank Evaluation Reports http://www.afdb.org/en/documents/evaluation-reports/	Institutional website	4 February 2014
Asian Development Bank (ADB) Evaluation Resources http://www.adb.org/site/evaluation/resources	Institutional website	3 February 2014
Bureau for Research and Economic Analysis of Development (BREAD) http://www.ibread.org/	Institutional website	3 February 2014
Campbell Collaboration http://www.campbellcollaboration.org	Institutional website	5 February 2014
Centre for Economic Policy Research (CEPR) http://www.cepr.org	Institutional website	3 February 2014
Center for Economic Studies (CESifo) http://www.cesifo-group.de/ifoHome.html	Institutional website	3 February 2014

Source	Type of source	Search date
ESRC (Economic and Social Research Council) http://www.esrc.ac.uk/	Institutional website	3 February 2014
Global Development Network (GDN) http://www.gdnet.org/index.html	Institutional website	5 February 2014
Institute for Development Policy and Management (IDPM) at the University of Manchester	Institutional website	4 February 2014
Institute for Fiscal Studies – Centre for the Evaluation of Development Policy http://www.ifs.org.uk/edepo/index.php	Institutional website	5 February 2014
Institute for the Study of Labour (IZA) http://www.iza.org	Institutional website	4 February 2014
Institute of Development Studies (IDS) http://www.ids.ac.uk/	Institutional website	5 February 2014
Inter-American Development Bank Office of Evaluation and Oversight http://www.iadb.org/en/office-of-evaluation-and-oversight	Institutional website	2 February 2014
Millennium Challenge Corporation (MCC) http://www.mcc.gov/pages/results/evaluations	Institutional website	2 February 2014
National Bureau of Economic Research (NBER) http://www.nber.org	Institutional website	15 December 2013
Overseas Development Institute (ODI) http://www.odi.org.uk/	Institutional website	2 February 2014
Poverty and Economic Policy Research Network (PEP): Project List https://www.pep-net.org/	Institutional website	2 February 2014
UNDP International Policy Centre for Inclusive Growth (IPC-IG) http://www.ipc-undp.org/	Institutional website	3 February 2014
United States Department of Labour, Employment and Training Administration, Research Publication Database	Institutional website	4 February 2014

Source	Type of source	Search date
http://wdr.doleta.gov/research/		
University of California Center for Effective Global Action (CEGA): Research Projects http://cega.berkeley.edu/	Institutional website	4 February 2014
USAID Development Experience Clearinghouse https://dec.usaid.gov/	Institutional website	5 February 2014
World Bank Independent Evaluation Group (IEG) http://ieg.worldbankgroup.org	Institutional website	4 February 2014
World Bank Labor Markets http://www.worldbank.org/labormarkets	Institutional website	4 February 2014
Google Scholar	Website/ gateway	15 December 2013
Bibliographic information of the 2013 World Development Report and the 2012 Independent Evaluation Group report *Youth Employment Programs: An Evaluation of World Bank and IFC Support*	Other reports and reviews	4 February 2014
Bibliographic information of other reviews and meta-analyses: ▪ Betcherman, Gordon, Godrey, Martin, Puerto, Susana, Rother, Friederike, and Stavreska, Antoneta (2007). *A Review of Interventions to Support Young Workers: Findings of the Youth Employment Inventory*. SP Discussion Paper No. 0715. World Bank, Washington, DC. ▪ Card, David E., Kluve, Jochen and Weber, Andrea Michaela (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, *120*, 548, F452–F477, 201. ▪ Cho, Yoonyoung and Honorati, Maddalena (2013). *Entrepreneurship programs in developing countries: A meta regression analysis*. World Bank Policy Research Working Paper No. 6402. ▪ Kluve, Jochen (2010). The effectiveness of European active labour market programmes. *Labour Economics*, *17*, 6, 904–918. ▪ Interventions for Employment Creation in Micro, Small and Medium-sized Enterprises in Low- and	Other reports and reviews	4 February 2014

Source	Type of source	Search date
Middle-income Countries (title registration) ▪ Post-Basic Technical and Vocational Education and Training (TVET) Interventions to Improve Employability and Employment of TVET Graduates in Low- and Middle-Income Countries: A Systematic Review ▪ J-PAL Youth Initiative Review Paper		
3ie Database of Systematic Reviews http://www.3ieimpact.org/evidence/systematic-reviews/	Specialized database	3 February 2014
3ie Registry for International Development Impact Evaluations (RIDIE) http://ridie.3ieimpact.org/	Specialized database	3 February 2014
3ie Register of Impact Evaluation Published Studies (RIEPS) http://www.3ieimpact.org/evidence/impact-evaluations/	Specialized database	3 February 2014
Abdul Latif Jameel Poverty Action Lab (J-PAL) Evaluation and Publication Database http://www.povertyactionlab.org/	Specialized database	3 February 2014
British Library for Development Studies (BLDS) http://blds.ids.ac.uk/	Specialized database	15 December 2013
ELDIS http://www.eldis.org/	Specialized database	3 February 2014
Innovations for Poverty Action (IPA) Database http://www.poverty-action.org/project-evaluations/search	Specialized database	3 February 2014
JOLIS library catalogue – International Finance Corporation, International Monetary Fund, and World Bank http://jolis.worldbankimflib.org/e-nljolis.htm	Specialized database	4 February 2014
Labordoc (ILO) http://labordoc.ilo.org/	Specialized database	16 December 2013
Research for Development http://r4d.dfid.gov.uk/	Specialized database	2 February 2014

Source	Type of source	Search date
World Bank Poverty Impact Evaluations Database http://www1.worldbank.org/prem/poverty/ie/evaluationdb.htm	Specialized database	3 February 2014
Youth Employment Inventory (YEI) http://www.youth-employment-inventory.org/	Specialized database	3 February 2014

From November 2014 to January 2015, the review team contacted experts and authors of included studies, screened reference lists of included studies and conducted citation tracking in order to identify additional studies.

8.6.2 Screening questionnaire

1. **Has the study been published in 1990 or later?**
   - No → Exclude (EndNote: into *_1. Exclude on date)
   - Yes → next question
   - Unsure →next question
2. **Does the target group consist only or mainly of young people (aged 15–35 years)?**
   - No → Exclude (EndNote: into *_2. Exclude on target group)
   - Yes → Include
   - Unsure → next question
3. **Has the research been conducted in any of the following experimental or quasi-experimental designs?**
   a. Experimental:
      i. RCT
      ii. Natural experiment
   b. Quasi-experimental
      i. Difference-in-difference or triple difference
      ii. Regression discontinuity
      iii. Instrumental variables
      iv. Propensity score matching
      v. Panel analysis
      vi. Pipeline/stepped-wedge analysis
      vii. Time-series designs
      viii. Non-equivalent control group design (cohort designs, post-test designs, …)
   - No → Exclude (EndNote: into *_3. Exclude on study design)
   - Yes → next question
   - Unsure → next question
4. **Does the intervention include any of the following components?**
   a. Training and skills development:
      i. Technical skills training
      ii. Business skills training
      iii. Literacy or numeracy skills training
      iv. Behavioural, life skills or soft skills training
   b. Entrepreneurship promotion:
      i. Business advisory/mentoring
      ii. Business skills
      iii. Access to markets and value chains
      iv. Credit or access to credit
      v. Grants (monetary or in-kind)
      vi. Microfranchising
   c. Employment services:
      i. Job placement/intermediation services
      ii. Job counselling/job-search assistance/mentoring
      iii. Financial assistance for job search
   d. Subsidized employment:
      i. Linking beneficiaries to subsidized employment in private enterprises
      ii. Public work in infrastructure development projects
      iii. Social development and community works and services projects (e.g., children's care, sick and elderly care, security, health)
   - No → Exclude (EndNote: into *_4. Exclude on intervention)
   - Yes → next question
   - Unsure → next question
5. **Does the paper measure impact on any of the following labour market outcomes?**
   a. Employment outcomes
      i. Employment (empirical probability models)
      ii. Unemployment (empirical probability models)
      iii. Participation rate

iv. Hours worked
v. Unemployment duration
vi. Quality of employment

b. Earnings outcomes
i. Earnings/income
ii. Household income
iii. Consumption
iv. Salary and/or wage

c. Business performance
i. Profits
ii. Sales
iii. Number of employees and jobs created
iv. Capital and investment
v. Business creation
vi. Business survival

- No → Exclude (EndNote: into *_5. Exclude on outcomes)
- Yes → next question
- Unsure → next question

6. **Is the study of reasonable quality? (Full report review)**
   a. We can obtain all of the following details about study methodology:[48]
      i. The intervention (including setting, beneficiary population, benefits)
      ii. Sample characteristics (age, sample size)
      iii. Study type and analytical model: statistical tests on the studied associations, with the coefficients and significance levels reported
      iv. The methodology used to control for confounding factors and selection bias
   - No → Exclude (EndNote: into *_6. Exclude on study quality)
   - Yes → next question
   - Unsure → next question
7. **Is the study among the following categories of publication status?**
   a. Peer-reviewed journal
   b. Working paper
   c. Mimeo
   d. Book
   e. Policy/position paper
   f. Evaluation/technical report
   g. Dissertation/thesis
   - No → Exclude (EndNote: into *_7. Exclude on publication status) (Exclude if editorial, commentary, process evaluations, single-participant studies/anecdotal)
   - Yes
   - Unsure

[48] Based on Table 1 of Leroy, Gadsden & Guijarro (2012).

8.6.3 Search terms for electronic databases

The search terms for electronic databases include the most frequent and relevant exposure, outcome and subject terms which were identified during the scoping search through a frequency test of 107 keywords in a group of 32 preselected potentially relevant studies from the Youth Employment Inventory (available at: www.youth-employment-inventory.org/) based on a first draft of inclusion and exclusion criteria.[49]

1. Exposure terms

Retraining, training, skill, skills, entrepreneur*, program*, intervention, measur*, scheme, project, activation, subsidy, subsidies, subsidized, subsidised, upgrade, assistance, internship, intern, interns, business, counseling, counselling, mentor*, advisory, coaching, placement, insertion

2. Outcome terms

Unemployment, unemployed, employed, employment, participation, labor, labour, earning*, job*, wage*, income*, salar*, profit, revenue, work

3. Subject terms

Youth*, young, adolescent*, schoolleaver*, school leaver*, high school graduate*

4. Impact terms[50]

Labordoc ILO thesaurus terms: ES: evaluación, FR: évaluation
OpenDOAR: (impact AND (evaluat* OR assess* OR analy* OR estimat*))
Google Scholar: (Impact OR effect OR evaluation OR random)
Combine: 1 AND 2 AND 3

[49] However, we found that the number of results increases in a disproportional way. For example, the advanced search string for ABI/INFORM Global yielded 2,906 results without the term "student*", but 4,419 results with the term "student*". Therefore, we decided to exclude the term "student*" from all advanced search strings using Boolean operators since it seemed to capture too many irrelevant, purely education-related results.

[50] To ensure inclusion of papers which do not specifically report their research design in the title or abstract, the search excluded methodology terms. However, impact filters were useful for sources such as OpenDOAR, which displayed only a limited number of results in the scoping search. The selection of impact terms was based on the 3ie Register of Impact Evaluation Published Studies (RIEPS) Protocol (Mishra & Cameron, 2013).

8.6.4 Search strings for electronic databases: Examples

1. RePEc (Research Papers in Economics)/IDEAS Economics and Finance Research

Search strategy: Advanced search using Boolean operators, combination of exposure, outcome, subject terms. Searched in abstract, searched for synonyms, exact words, from 1990 until 15 December 2013:

(Retraining | retrain | retrained | training | trained | skill | skills | entrepreneur | entrepreneurs | entrepreneurial | entrepreneurship | intervention | interventions | subsidy | subsidies | subsidized | subsidised | internship | internships | intern | interns | counseling | counselling | mentor | mentoring) + (Unemployment | unemployed | employed | employment | labor | labour | earning | earnings | job | jobs | wage | wages | income | incomes | salary | salaries | profit | profits | revenue | revenues | work) + (Youth | youths | young | adolescent | adolescents | schoolleaver | schoolleavers)

2. EconLit

Search strategy: Advanced search using Boolean operators, combination of exposure outcome and subject terms. Searched in abstract, from 1990 until 15 December 2013.

((AB,TI(retraining OR training OR skill OR skills OR entrepreneur* OR program* OR intervention OR measur* OR scheme OR project OR activation OR subsidy OR subsidies OR subsidized OR subsidised OR upgrade OR assistance OR internship OR business OR counseling OR counselling OR mentor* OR advisory OR coaching OR placement OR insertion) OR (I21 OR I28 OR I24 OR I22 OR J24)) AND (TI,AB(youth* OR young* OR schoolleaver* OR "school leaver*" OR "high school graduat*") OR J13) AND ((J23 OR J4 OR J2 OR J64 OR J31 OR J38 OR E24 OR F16 OR J68 OR J45) OR TI(unemployment OR unemployed OR employed OR employment OR participation OR ((labour OR labor) adj3 (market* OR trend* OR mobility OR demand OR conditions OR force OR migration OR unskilled OR opportunit* OR supply OR casual)) OR earning* OR job* OR wage* OR income* OR salar* OR profit OR revenue OR work))) AND pd(19900101-20131231)

3. ERIC (Education Resources Information Centre)

Search strategy: Advanced search using Boolean operators, combination of exposure outcome and subject terms, from 2000 until 18 December 2013. Topic restrictions: adolescents OR youth employment OR young adults OR youth programs OR disadvantaged youth OR youth OR high school graduates OR youth problems.

(TI,AB(retraining OR training OR skill OR skills OR entrepreneur* OR program* OR intervention OR measur* OR scheme OR project OR activation OR subsidy OR subsidies OR subsidized OR subsidised OR upgrade OR assistance OR internship OR business OR counseling OR counselling OR mentor* OR advisory OR coaching OR placement OR insertion)) AND (TI(unemployment OR unemployed OR employed OR employment OR participation OR ((labour OR labor) adj3 (market* OR trend* OR mobility OR demand OR conditions OR force OR migration OR unskilled OR opportunit* OR supply OR casual)) OR earning* OR job* OR wage* OR income* OR salar* OR profit OR revenue OR work) OR (SU.EXACT("Public Sector") OR SU.EXACT("Self Employment") OR SU.EXACT("Job Search Methods") OR SU.EXACT("Occupational Mobility") OR SU.EXACT("Job Development") OR SU.EXACT("Labor Economics") OR SU.EXACT("Compensation (Remuneration)") OR SU.EXACT("Salary Wage Differentials") OR SU.EXACT("Labor Education") OR SU.EXACT("Employment Patterns") OR SU.EXACT("Labor Market") OR SU.EXACT("Employment Programs") OR SU.EXACT("Workers Compensation") OR SU.EXACT("Labor Supply") OR SU.EXACT("Unemployment") OR SU.EXACT("Employment") OR SU.EXACT("Employment Opportunities") OR SU.EXACT("Labor Demands") OR SU.EXACT("Labor Turnover"))) AND (TI,AB(youth* OR young OR schoolleaver* OR "school leaver*" OR "high school graduat*") OR (SU.EXACT("Youth") OR SU.EXACT("Young Adults") OR SU.EXACT("Adolescents"))) AND (subt.exact("adolescents" OR "youth employment" OR "young adults" OR "youth programs" OR "disadvantaged youth" OR "youth" OR "high school graduates" OR "youth problems") AND pd(20000101-20131218))

8.7 CODE DESCRIPTION

This section contains a description of the variables that were coded at the study level. Each variable name is followed by a description. Opposite each variable name/description there is a description of how the variable should be coded.

Variable group: Identifiers	
es_id: Effect size Estimate ID	*numeric, running numbers, headed by study_ID and es. (Example: Fifth estimate in third study of 15th programme in YEI = 015s03es05)*
study_id: Impact evaluation study ID	*numeric, running numbers, headed by progr_ID and s. (Example: Third study of 15th programme in YEI = 015s03)*
progr_id: Programme ID	*numeric, based on YEI Programme ID*
date_publ: Year of publication of study	*YYYY*
status_publ: Publication status of study	*1 = Peer-reviewed journal,* *2 = Working paper,* *3 = Mimeo,* *4 = Book,* *5 = Policy/position paper,* *6 = Evaluation/technical report,* *7 = Dissertation/thesis,*

Variable group: Description of data used and empirical methods	
data_src_coll: Data source: Collected original data for study	*0 = No,* *1 = Yes,*
data_src_surv: Data source: Survey data	*0 = No,* *1 = Yes,*
data_src_admi: Data source: Administrative data	*0 = No,* *1 = Yes,*
ie_design: Impact evaluation research design	*1 = RCT,* *2 = Natural experiment,* *3 = Pipeline,* *4 = Only panel,* *5 = Only cross-section,*

ie_method: Statistical methodology	*1 = Instrumental variable (IV),* *2 = Matching (PSM, etc.),* *3 = Regression-adjusted DiD,* *4 = Simple DiD,* *5 = DiD and matching combined,* *6 = Regression discontinuity design (RDD),* *7 = Multivariate linear (OLS, ANOVA, others less likely),* *8 = Multivariate non-linear regression model (Probit, Logit, Tobit),* *9 = Tabulation (simple differences in mean),* *10 = Other (specify in comments),*
ie_datatype: Structure of data underlying the treatment effect estimate	*1 = Cross-sectional data,* *2 = Pooled (repeated) cross-sectional data,* *3 = Panel data,*
ie_uoa_err: Treatment effect estimate is probably subject to unit of analysis error	*0 = No,* *1 = Yes,*
ie_itt: Intention-to-treat estimation specifically mentioned	*0 = No (only if specified that estimator does not measure ITT),* *1 = Yes,*

Variable group: Intervention category	
int_cat_skil: Intervention category: Skills training	*0 = No,* *1 = Yes,*
int_cat_entr: Intervention category: Entrepreneurship promotion	*0 = No,* *1 = Yes,*
int_cat_serv: Intervention category: Employment services	*0 = No,* *1 = Yes,*
int_cat_subs: Intervention category: Subsidized employment (wage subsidies, public works and employment guarantee schemes, public services)	*0 = No,* *1 = Yes,*
int_cat_main: Main category of intervention	*1 = Skills training,* *2 = Entrepreneurship promotion,* *3 = Employment services,* *4 = Subsidized employment,* *5 = Unspecified,*
int_cat_subc: Evaluation estimates effect for a sub-component of a more comprehensive programme	*0 = No,* *1 = Yes,*

Variable group: Programme characteristics: Skills training	
skil_type_tech: Type of skills training: Technical skills	*0 = No,* *1 = Yes,*
skil_type_busi: Type of skills training: Business skills	*0 = No,* *1 = Yes,*

skil_type_lite: Type of skills training: Literacy and/or numeracy	*0 = No,* *1 = Yes,*
skil_type_soft: Type of skills training: Behavioural, life skills, soft skills	*0 = No,* *1 = Yes,*
skil_deli_dist: Skills training delivered: Distant learning (e.g., books, online training)	*0 = No,* *1 = Yes,*
skil_deli_clas: Skills training delivered: In classroom	*0 = No,* *1 = Yes,*
skil_deli_work: Skills training delivered: At the workplace (e.g., internships, on-the-job training schemes, non-apprenticeship schemes)	*0 = No,* *1 = Yes,*
skil_deli_appr: Skills training delivered: Apprenticeship schemes (in shops with master craftsmen/women)	*0 = No,* *1 = Yes,*
skil_durat: Duration of skills training programme: Total number of hours per individual (averages)	*numeric,*
skil_prov_publ: Provider of the skills training: Public training institution	*0 = No,* *1 = Yes,*
skil_prov_priv: Provider of the skills training: Private training institution	*0 = No,* *1 = Yes,*
skil_prov_ngo: Provider of the skills training: NGO, Foundation, CBO, CSO	*0 = No,* *1 = Yes,*
skil_paym_lump: Payment system to training providers: Lump-sum budget	*0 = No,* *1 = Yes,*
skil_paym_serv: Payment system to training providers: Payment for services delivered	*0 = No,* *1 = Yes,*
skil_paym_resu: Payment system to training providers: Payment by outcomes	*0 = No,* *1 = Yes,*
skill_sele_nati: Selection of skills: Identified by national government	*0 = No,* *1 = Yes,*
skill_sele_loca: Selection of skills: Identified by regional/local government	*0 = No,* *1 = Yes,*
skill_sele_civi: Selection of skills: Identified by civil society	*0 = No,* *1 = Yes,*
skill_sele_priv: Selection of skills: Identified by private sector	*0 = No,* *1 = Yes,*
skill_sele_dono: Selection of skills: Identified by donors/development agencies	*0 = No,* *1 = Yes,*

Variable group: Programme characteristics: Entrepreneurship promotion	
entr_typ_advi: Type of intervention: Business advisory/mentoring	*0 = No,* *1 = Yes,*
entr_typ_skil: Type of intervention: Business skills	*0 = No,* *1 = Yes,*
entr_typ_acce: Type of intervention: Access to markets and value chains	*0 = No,* *1 = Yes,*
entr_typ_cred: Type of intervention: Credit or access to credit	*0 = No,* *1 = Yes,*
entr_typ_gran: Type of intervention: Grants (monetary or in-kind)	*0 = No,* *1 = Yes,*
entr_typ_fran: Type of intervention: Microfranchising	*0 = No,* *1 = Yes,*
entr_prov_publ: Provider of the entrepreneurship services: Public institution	*0 = No,* *1 = Yes,*
entr_prov_priv: Provider of entrepreneurship services: Private institution	*0 = No,* *1 = Yes,*
entr_prov_ngo: Provider of the entrepreneurship services: NGO, foundation, CBO, CSO	*0 = No,* *1 = Yes,*
entr_paym_lump: Payment system to service providers: Lump-sum budget	*0 = No,* *1 = Yes,*
entr_paym_serv: Payment system to service providers: Payment for services delivered	*0 = No,* *1 = Yes,*
entr_paym_resu: Payment system to service providers: Payment by outcomes	*0 = No,* *1 = Yes,*
entr_sele_comp: Selection process: Business plan/idea competition	*0 = No,* *1 = Yes,*
entr_sele_surv: Selection process: Survey, interview or test	*0 = No,* *1 = Yes,*
entr_sele_none: Selection process: Any youth within target population is eligible	*0 = No,* *1 = Yes,*

Variable group: Programme characteristics: Employment services	
serv_type_coun: Type of employment service: Job counselling/job-search assistance/mentoring	*0 = No,* *1 = Yes,*
serv_type_plac: Type of employment service: Job placement	*0 = No,* *1 = Yes,*
serv_type_fina: Type of employment service: Financial assistance for job search	*0 = No,* *1 = Yes,*
serv_prov_publ: Provider of the employment services: Public institution	*0 = No,* *1 = Yes,*
serv_prov_priv: Provider of employment services: Private institution	*0 = No,* *1 = Yes,*

serv_prov_ngo: Provider of the employment services: NGO, foundation, CBO, CSO	*0 = No,* *1 = Yes,*
serv_paym_lump: Payment system to service providers: Lump-sum budget	*0 = No,* *1 = Yes,*
serv_paym_serv: Payment system to service providers: Payment for services delivered	*0 = No,* *1 = Yes,*
serv_paym_resu: Payment system to service providers: Payment by outcomes	*0 = No,* *1 = Yes,*

Variable group: Programme characteristics: Subsidized employment	
subs_typ_secu: Type of subsidy: Reduction in employer social security contributions	*0 = No,* *1 = Yes,*
subs_typ_wage: Type of subsidy: Reduction in employer labour/wage costs	*0 = No,* *1 = Yes,*
subs_typ_paym: Type of subsidy: Direct payment to the individual (e.g., voucher)	*0 = No,* *1 = Yes,*
subs_set_abs: Subsidy setting: The absolute level of the subsidy is fixed by the Government	*0 = No,* *1 = Yes,*
subs_set_rel: Subsidy setting: The absolute level of the subsidy is variable	*0 = No,* *1 = Yes,*
subs_durat: Maximum duration of the subsidy in months per individual	*numeric,*
subs_empl_any: Eligible employers: Any employer is eligible	*0 = No,* *1 = Yes,*
subs_empl_form: Eligible employers: Only employers who offer formal contracts/only formal employers	*0 = No,* *1 = Yes,*
subs_empl_sect: Eligible employers: Only employers in certain sectors	*0 = No,* *1 = Yes,*
subs_empl_numb: Eligible employers: Only employers with certain number of employees	*0 = No,* *1 = Yes,*
subs_cond_none: Conditionality for eligibility: None	*0 = No,* *1 = Yes,*
subs_cond_empl: Conditionality for eligibility: The labour contract must be at least for half-time employment	*0 = No,* *1 = Yes,*
subs_cond_cont: Conditionality for eligibility: Employers need to offer a contract after the subsidy expires	*0 = No,* *1 = Yes,*
pubw_type_infr: Public work type: Infrastructure development projects (e.g., public works in rural and urban areas – construction, and maintenance of public works)	*0 = No,* *1 = Yes,*

pubw_type_serv: Public work type: Social development and community works and services projects (e.g., children's care, sick and elderly care, security, health)	*0 = No,* *1 = Yes,*
pubw_durat: Duration of the works in months per individual (average)	*numeric,*
pubw_sele_nati: Works/services selected by regional/national government	*0 = No,* *1 = Yes,*
pubw_sele_loca: Works/services selected by local/regional government	*0 = No,* *1 = Yes,*
pubw_sele_civi: Works/services selected by civil society (e.g., NGOs, communities, youth organizations)	*0 = No,* *1 = Yes,*
pubw_sele_priv: Works/services selected by private sector	*0 = No,* *1 = Yes,*
pubw_sele_dono: Works/services selected by donors	*0 = No,* *1 = Yes,*
pubw_wage_min: Wage setting: Programme wage in relation to the minimum wage (as stated in paper)	*1 = Lower,* *2 = Equal,* *3 = Greater,* *4 = There is no minimum wage policy,*
pubw_wage_ave: Wage setting: Programme wage in relation to the market wage for unskilled labour (as stated in paper)	*1 = Lower,* *2 = Equal,* *3 = Greater,* *4 = There is no market wage policy,*
pubw_targ_self: Targeting of participants: Self-selection targeting (e.g., through wage setting)	*0 = No,* *1 = Yes,*
pubw_targ_geo: Targeting of participants: Geographic targeting	*0 = No,* *1 = Yes,*
pubw_exec_info: Works/services execution by: Informal contractors	*0 = No,* *1 = Yes,*
pubw_exec_priv: Works/services execution by: Formal private contractors	*0 = No,* *1 = Yes,*
pubw_exec_publ: Works/services execution by: Formal public contractors	*0 = No,* *1 = Yes,*

Variable group: Programme characteristics: General	
targ_age_you: Target group of intervention: Youth (15–35 years)	*0 = No,* *1 = Yes,*
targ_age_ayou: Target group of intervention: All, but mainly youth	*0 = No,* *1 = Yes,*
targ_age_sta: Target group of intervention: Start age bracket	*numeric,*

targ_age_end: Target group of intervention: End age bracket	*numeric,*
targ_gend: Target group of intervention: Gender	*1 = male,* *2 = female,* *3 = both,*
targ_educ_prim: Target group of intervention: Education = low education (primary or lower)	*0 = No,* *1 = Yes,*
targ_educ_seco: Target group of intervention: Education = secondary education (or equiv.)	*0 = No,* *1 = Yes,*
targ_educ_high: Target group of intervention: Education = higher education (above secondary)	*0 = No,* *1 = Yes,*
targ_loc: Target group of intervention: Location	*1 = urban,* *2 = rural,* *3 = both,*
targ_unemp: Target group of intervention: Only unemployed at intervention start	*0 = No,* *1 = Yes,*
targ_emp: Target group of intervention: Already employed/entrepreneur at intervention start	*0 = No,* *1 = Yes,*
targ_first: Target group of intervention: Only first-time jobseekers	*0 = No,* *1 = Yes,*
targ_welf: Target group of intervention: Welfare recipient at Intervention start	*0 = No,* *1 = Yes,*
targ_lowi: Target group: Low income/disadvantaged/at risk/vulnerable youth	*0 = No,* *1 = Yes,*
targ_disab: Target group of intervention: Disability at intervention start	*0 = No,* *1 = Yes,*
targ_mand: Target group of intervention: Participation mandatory	*1 = No,* *2 = Yes,* *3 = Voluntary/Self-select/Apply,*
prog_bfit: Welfare benefits scheme provided during programme participation	*0 = No,* *1 = Yes,*
prog_ince_part: Incentives provided to programme participants	*1 = Non-monetary benefits (e.g., childcare, catering, transport),* *2 = Monetary benefits (e.g., stipend, transport allowance),* *3 = Both non-monetary and monetary,* *4 = Salary,* *5 = None,*
prog_moni: Monitoring of participants or compliance of beneficiaries	*0 = No,* *1 = Yes,*
prog_sanc: Sanctions for non-participation or non-compliance (e.g., linking programme participation to benefit reception)	*0 = No,* *1 = Yes,*
prog_start: Starting date of programme	*YYYYmMM,*

prog_end: Ending date of programme	*YYYYmMM,* *a = ongoing,*
prog_reg: Region of country where programme is implemented	*Region, naming according to YEI database/WDI. See sheet YEI Progr_ID Names*
prog_coun: Country where programme is implemented	*Country name according to YEI database/ WDI. See sheet YEI Progr_ID Names*
prog_scale: Scale of programme	*1 = national,* *2 = regional,* *3 = local,* *4 = pilot,*
prog_dura: Average duration a single cohort stays in programme, in months	*numeric,*
prog_gend: Programme design includes gender considerations	*0 = No,* *1 = Yes,*
prog_desi_gov: Design of programme: Government	*0 = No,* *1 = Yes,*
prog_desi_ngo: Design of programme: NGO/non-profit	*0 = No,* *1 = Yes,*
prog_desi_mult: Design of programme: Multilateral	*0 = No,* *1 = Yes,*
prog_desi_donor: Design of programme: Donor-organized NGO	*0 = No,* *1 = Yes,*
prog_desi_priv: Design of programme: Private sector	*0 = No,* *1 = Yes,*
prog_desi_other: Design of programme: Other	*0 = No,* *1 = Yes,*
prog_awar: Programme includes awareness raising about the programme to eligible participants	*0 = No,* *1 = Yes,*
prog_impl_gov: Implementer of programme: Government	*0 = No,* *1 = Yes,*
prog_impl_ngo: Implementer of programme: NGO/non-profit	*0 = No,* *1 = Yes,*
prog_impl_mult: Implementer of programme: Multilateral	*0 = No,* *1 = Yes,*
prog_impl_donor: Implementer of programme: Donor-organized NGO	*0 = No,* *1 = Yes,*
prog_impl_priv: Implementer of programme: Private sector	*0 = No,* *1 = Yes,*
prog_impl_other: Implementer of programme: Other	*0 = No,* *1 = Yes,*
prog_finan_gov: Financing of programme: Government	*0 = No,* *1 = Yes,*
prog_finan_ngo: Financing of programme: NGO/non-profit	*0 = No,* *1 = Yes,*

prog_finan_mult: Financing of programme: Multilateral	*0 = No,* *1 = Yes,*
prog_finan_donor: Financing of programme: Donor country	*0 = No,* *1 = Yes,*
prog_finan_indv: Financing of programme: Individual donors (foundations, companies, etc.)	*0 = No,* *1 = Yes,*
prog_finan_benef: Financing of programme: Beneficiaries	*0 = No,* *1 = Yes,*
prog_finan_empl: Financing of programme: Employer of beneficiaries	*0 = No,* *1 = Yes,*
prog_finan_other: Financing of programme: Other	*0 = No,* *1 = Yes,*

Variable group: Study level variables	
st_name: Study name	*String: Author_Year_Title_Publication*
st_outc: Study reports outcome not able to code	*0 = No,* *1 = Yes,*
st_outc_desc: Description of outcomes not coded	*String: i.e., contract, days, tenure, education*
st_subg: Study reports sub-group analysis not able to distinguish through existing variables	*0 = No,* *1 = Yes,*
st_subg_desc: Description of sub-group analysis not coded	*i.e., single mothers, etc.*
st_reles: Study also estimates effects of intervention relative to alternative treatment (not coded)	*0 = No,* *1 = Yes,*
st_reles_desc: Description of relative effects estimated	*String: Treatment vs. alternative treatment*
st_costs: Study cites costs of the programme	*0 = No,* *1 = Yes,*
st_cba: Study includes cost-benefit analysis	*0 = No,* *1 = Yes,*
st_procod: Study includes programme elements that cannot be reflected in coding sheet	*0 = No,* *1 = Yes,*
st_procod_desc: Description of programme elements that cannot be coded	*String*
st_eprob: Author mentions empirical identification problems or shortcomings of the method	*0 = No,* *1 = Yes,*
st_eprob_desc: Description of empirical identification problems as stated by the author	*String*
st_impl_prob: Study mentions programme implementation problems	*0 = No,* *1 = Yes,*
st_geneq: Study mentions or estimates general equilibrium effects of the programme	*0 = No,* *1 = Yes,*

st_geneq_desc: Description of general equilibrium effects	*String: i.e., deadweight loss, substitution effects,*
st_comm: Other comments	*String: Note everything that was difficult for coding this study. Note any particularities with regard to intervention design.*

We also coded effect size specific information at the effect size/outcome level. A single study may analyse more than one outcome or group. For this reason there may be multiple effect size observations for a single study. The effect size variables are listed below.

Variable group: Effect size level information	
es_outc_cat: Outcome category	*1 = Employment outcome,* *2 = Earnings outcome,* *3 = Business performance outcome,*
es_outc: Outcome for which effect size is measured	*If out_cat = 1:* *1 = Employment probability, 2 = Unemployment probability, 3 = Participation rate, 4 = Hours worked, 5 = Unemployment duration, 6 = Quality of employment (e.g., contract, fixed, benefits),* *If out_cat = 2:* *7 = Earnings/income, 8 = Household income, 9 = Consumption, 10 = Salary/wage,* *If out_cat = 3:* *11 = Profits, 12 = Sales, 13 = No. of employees/jobs created, 14 = Capital & investment, 15 = Business creation, 16 = Business survival,*
es_outc_occu: Occupation category for which outcome is measured	*1 = Dependent employment,* *2 = Self-employment,* *3 = Both,*
es_outc_stat: Status of occupation for which outcome is measured	*1 = Formal,* *2 = Informal,* *3 = Both,*
es_outc_cond: Effect size measures effect conditional on some other primary outcome	*0 = No,* *1 = Yes,*
es_outc_cdesc: Description of condition	*String: Describe the outcome that has to be satisfied to enter the sample population (e.g., employed),*
es_subg: Indication that estimation sample is different from targeted intervention population (author's statement or obvious deviation)	*0 = No,* *1 = Yes,*

es_subs: Treatment effect is estimated for a sub-sample of the entire study population (only if estimates for total (broader) sample are also reported) (e.g., sub-sample analysis of females)	*0 = No,* *1 = Yes,*
es_age_s: Group for which effect is estimated: Start age	*numeric,*
es_age_e: Group for which effect is estimated: End age	*numeric,*
es_gend: Group for which effect is estimated: Gender	*1 = male,* *2 = female,* *3 = both,*
es_educ_prim: Group for which effect is estimated: Education = low education (primary or lower)	*0 = No,* *1 = Yes,*
es_educ_seco: Group for which effect is estimated: Education = secondary education (or equiv.)	*0 = No,* *1 = Yes,*
es_educ_high: Group for which effect is estimated: Education = higher education (above secondary)	*0 = No,* *1 = Yes,*
es_welf: Group for which effect is estimated: Only welfare recipients	*0 = No,* *1 = Yes,*
es_lowi: Group for which effect is estimated: Low income/disadvantaged/at risk/vulnerable	*0 = No,* *1 = Yes,*
es_loc: Group for which effect is estimated: Location	*1 = urban,* *2 = rural,* *3 = both,*
es_page: Page number where this effect size was found	*numeric,*
es_type: Type of effect size measure	*1 = Dichotomous/binary,* *2 = Continous,* *3 = Correlational,*
es_mmeth: Method of measurement of effect size	*1 = Pretest/posttest comparison,* *2 = Posttest comparison,* *3 = Follow-up comparison,*
es_type_dich: If es_type = 1	*1 = No. of events (treatment/comparison),* *2 = Event rates (treatment/comparison),* *3 = 2x2 contingency table (both events, treatment,comparison, both-non events),* *4 = Odds ratio,* *5 = Log odds ratio,* *6 = Risk ratio,* *7 = Risk difference,*

es_type_cont: If es_type = 2	*1 = Means (treatment/comparison),* *2 = Raw difference in means,* *3 = (Covariate) adjusted difference in means (= unstandardized regression coefficient),* *4 = Standardized mean difference (= standardized regression coefficient),* *5 = Log difference in means,* *6 = Log standardized mean difference,* *7 = t-value, f-value, p-value (from a paired t-test),* *8 = Frequency table (2 groups sample sizes),*
es_sign_type: Type of significance test	*1 = t-value,* *2 = p-value,* *3 = F-value (df = 1),* *4 = Chi-square (df = 1),* *5 = Standard error of coefficient estimate (from regression or matching),* *6 = Variance,* *7 = Confidence intervals (lower, upper),* *8 = Sample sizes (treatment/comparison),* *9 = Standard deviations (treatment/comparison),* *10 = Common standard deviation,* *11 = Total sample size,* *12 = Standard deviation of difference,*
es_raw: Estimated treatment effect as reported in study	*numeric,*
es_sign_yes: Estimated treatment effect is significant at the 5 per cent level	*0 = No,* *1 = Yes,*
es_sign_valu: Value of significance test	*numeric,*
es_ci_up: Upper value of confidence intervals if reported	*numeric,*
es_ci_low: Lower value of confidence intervals if reported	*numeric,*
treat_outc: Outcome of treatment group at endline if reported (e.g., mean) (use values after matching/covariate adjustment)	*numeric,*
treat_n: Number of observations in treatment group for which effect is estimated	*numeric,*
treat_sd: Standard Deviation in treatment group (post-intervention)	*numeric,*
contr_outc: Outcome of comparison group at endline if reported (e.g., mean) (use values after matching/covariate adjustment)	*numeric,*
contr_n: Number of observations in comparison group for which effect is estimated	*numeric,*

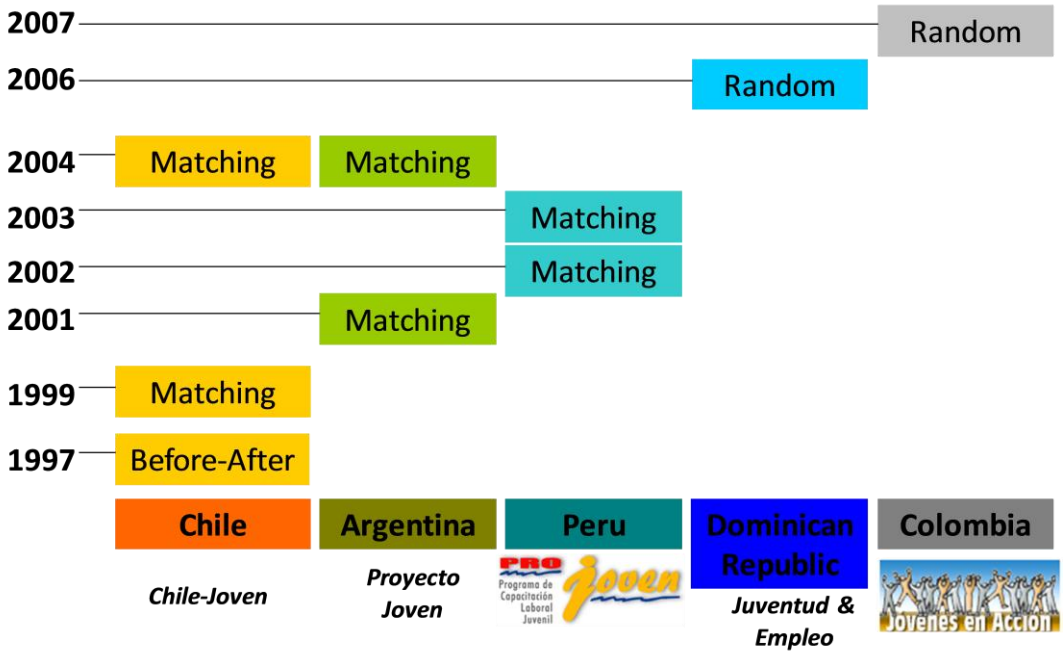
contr_sd: Standard deviation in comparison group (post-intervention)	*numeric,*
total_mean: Mean in the total sample	*numeric,*
total_n: Number of observations total	*numeric,*
total_sd: Standard deviation in total sample (post-intervention)	*numeric,*
pooled_sd: Pooled standard deviation (post-intervention)	*numeric,*
es_direct: Effect size direction	*1 = positive (higher values equal more positive outcomes)* *2 = negative (higher values equal more negative outcomes)*
es_term: Duration between individual exiting the intervention and data measurement	*1 = Short term (less than 12 months),* *2 = Medium term (12–24 months),* *3 = Long term (more than 24 months)*
es_date: Date at which effect size is measured	*YYYYmMM*
General coding	*a = Not applicable; 888 = Unsure TBD; 999 = Not specified*

In addition to the variables above, information was collected about the following programme-related variables, which were considered relevant for the analysis. To minimize the number of missing values for these variables, information was extracted from the study as well as from sources outside the study (which is the core unit of analysis), including project reports and project websites.

Variable group: Additional	
sup_incentsrv: Incentives to service providers (payments conditional on outcomes of programme participants)	*0 = No,* *1 = Yes,*
sup_incentpart: Incentives to participants (for programme participation and/or performance)	*0 = No,* *1 = Yes,*
sup_monit: Monitoring mechanisms	*0 = No,* *1 = Yes,*
sup_profile: Participant profiling for services provided	*0 = No,* *1 = Yes,*

9 Figures for the Appendix

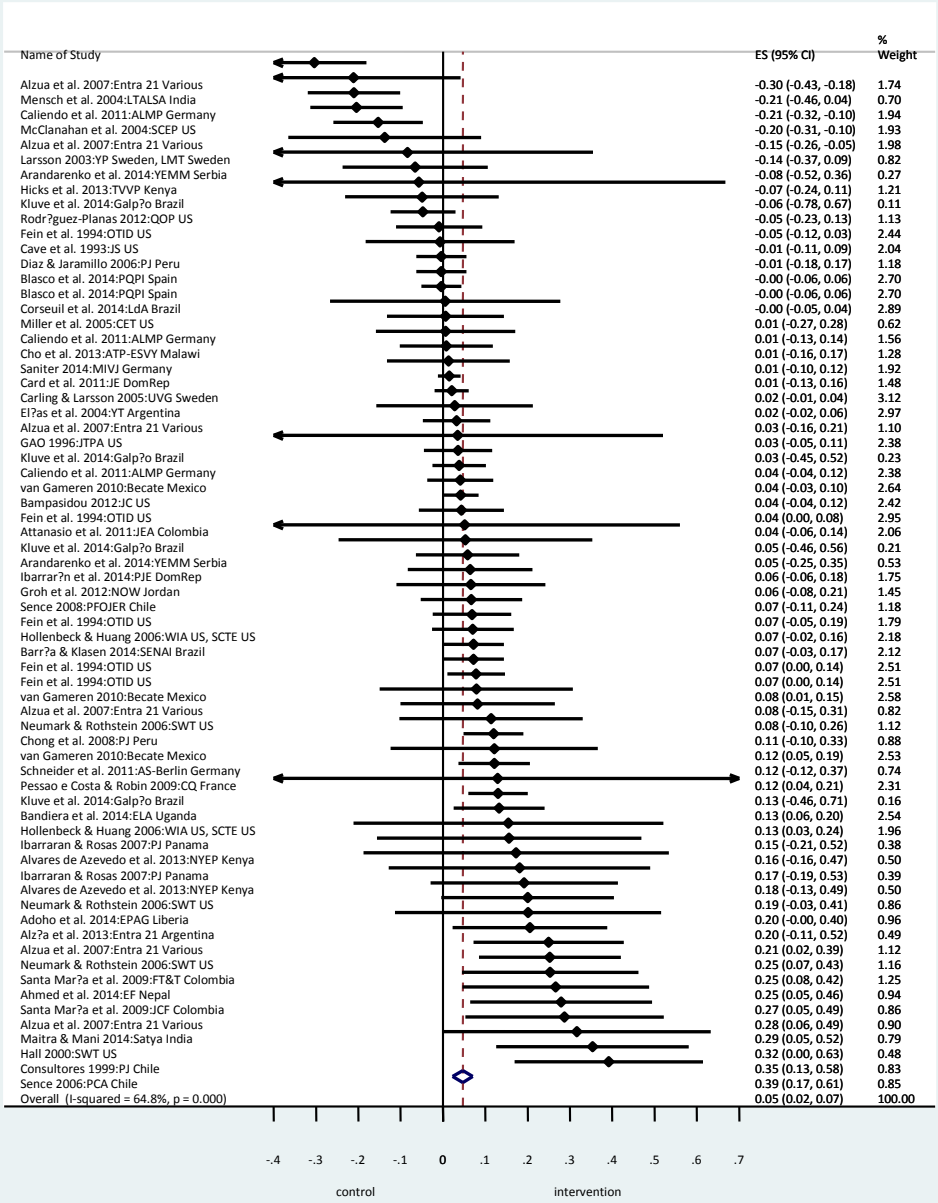
Figure 48: Learning gains on methodologies: An illustration from Latin America and the Caribbean



9.1 DISAGGREGATED FOREST PLOTS

9.1.1 Employment outcomes

Figure 49: Disaggregated forest plot for employment outcomes. Main category of intervention: Skills training



No. of SMDs/Studies: Total: 1330/105

Note: Imputation: full, SMDs limit = .75, SMD_SE limit = .75

Note: The PECO programme in Spain (Cansino Muñoz-Repiso & Sánchez Braza, 2011) (0.82 SMD; 95% CI = 0.45, 1.18) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 50: Disaggregated forest plot for employment outcomes. Main category of intervention: Entrepreneurship promotion

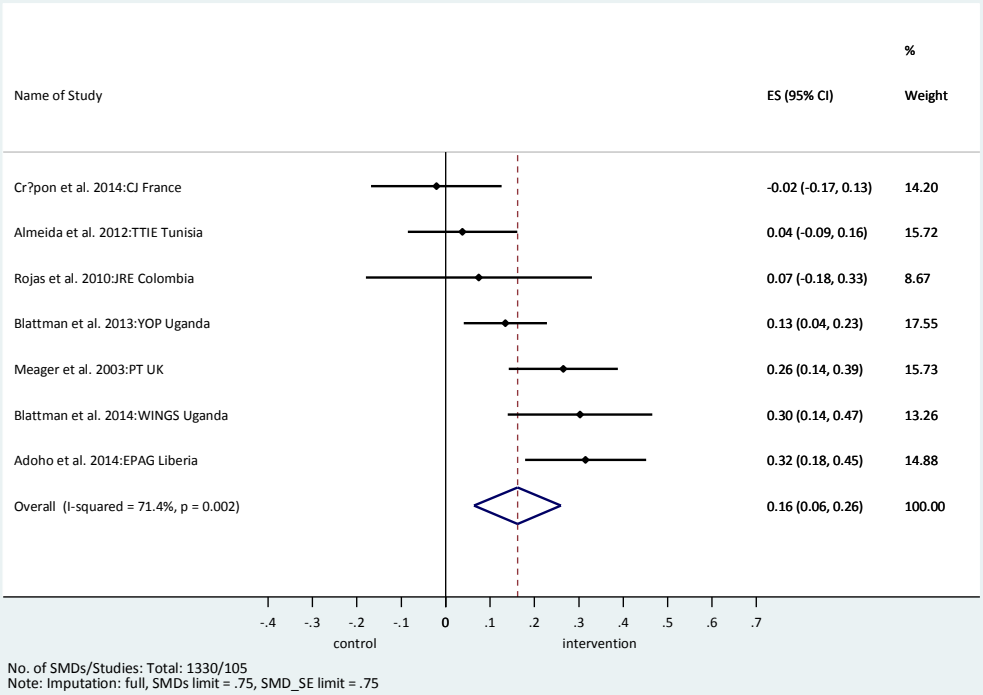


Figure 51: Disaggregated forest plot for employment outcomes. Main category of intervention: Employment services

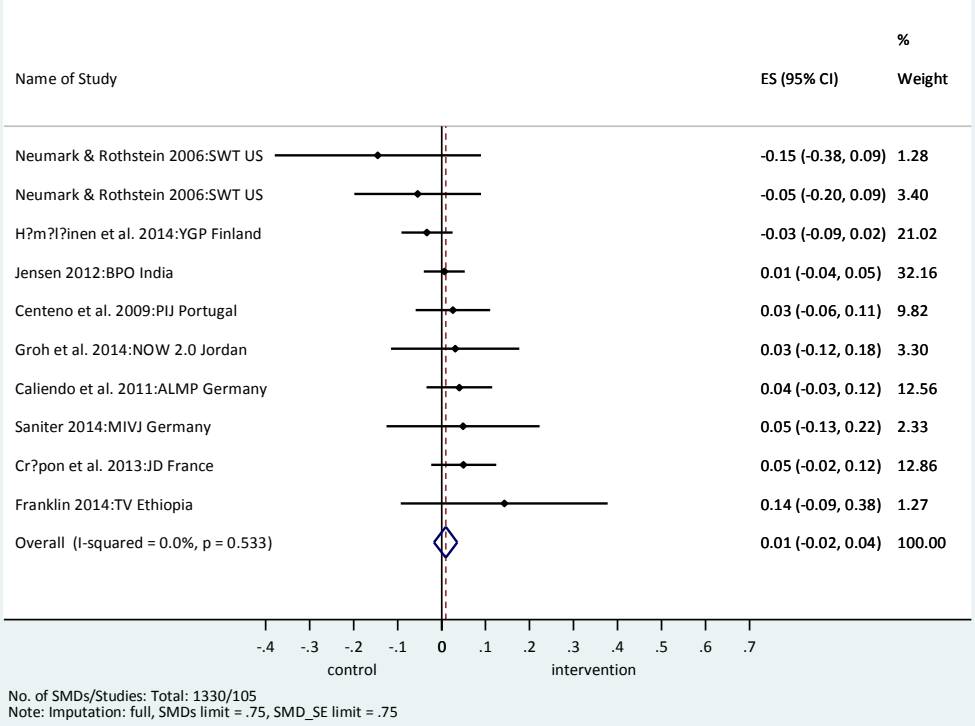


Figure 52: Disaggregated forest plot for employment outcomes. Main category of intervention: Subsidizied employment

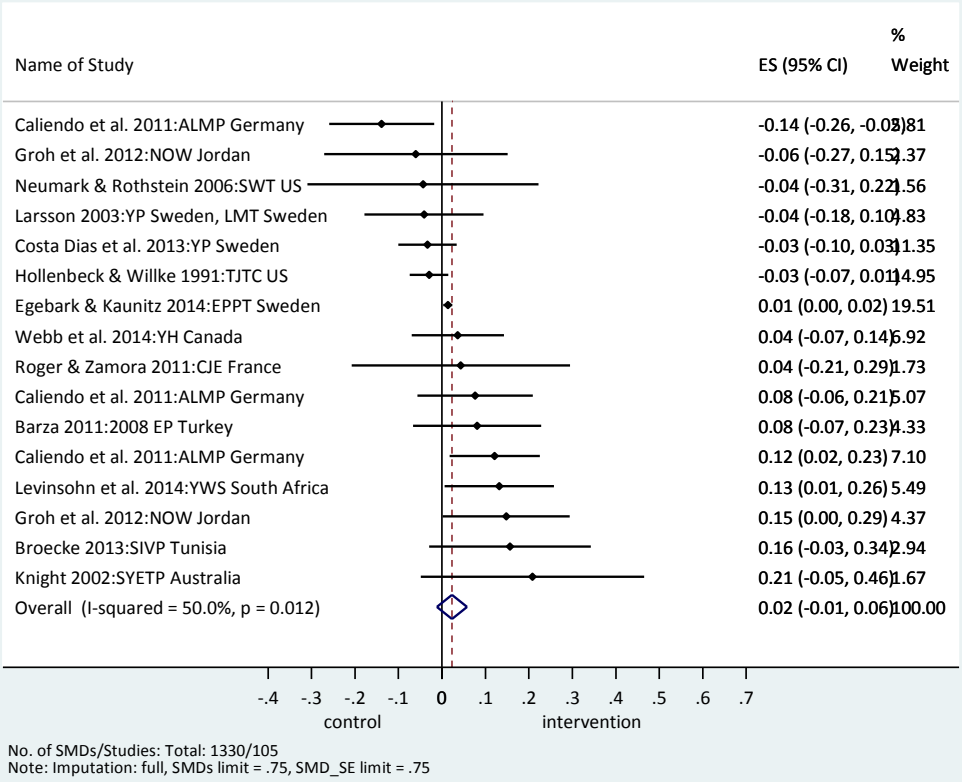
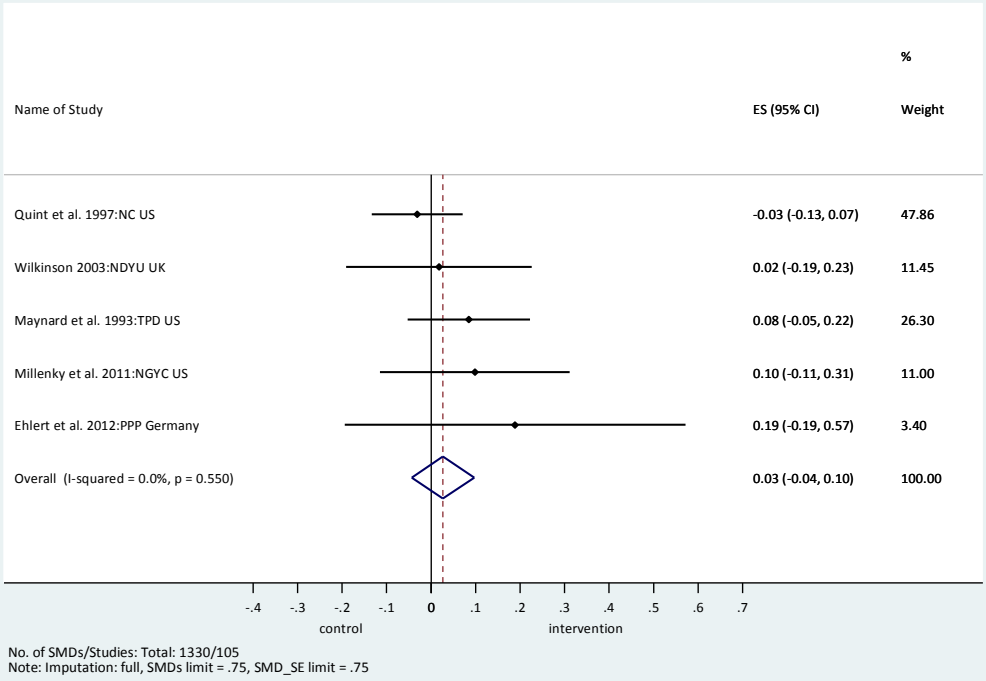
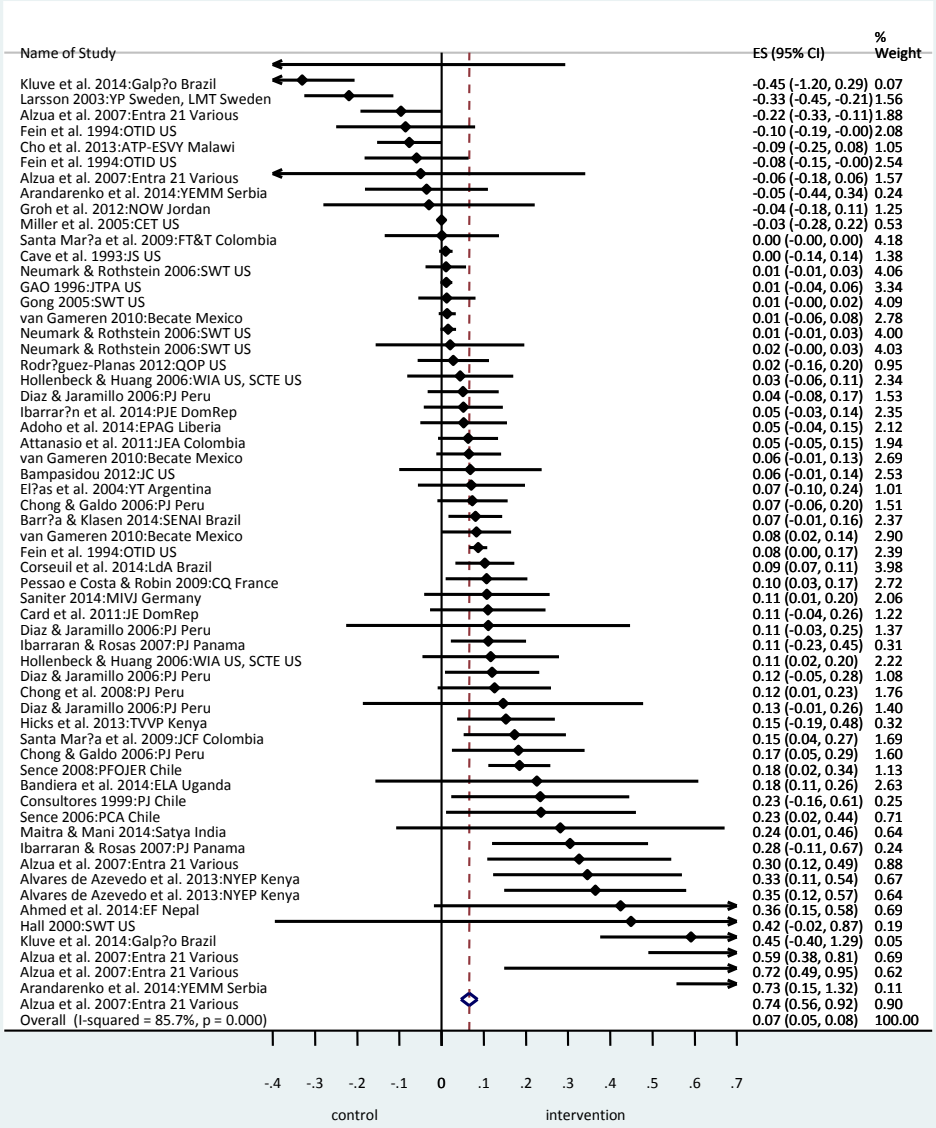


Figure 53: Disaggregated forest plot for employment outcomes. Main category of intervention: Unspecified



9.1.2 Earnings and income outcomes

Figure 54: Disaggregated forest plot for earnings and income outcomes. Main category of intervention: Skills training



No. of SMDs/Studies: Total: 638/89
Note: Imputation: full, SMDs limit = .75, SMD_SE limit = .75
Note: The Galpão programme in Brazil (Kluve et al., 2014) (0.82 SMD; 95% CI = (-0.03, 1.66) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

Figure 55: Disaggregated forest plot for earnings and income outcomes. Main category of intervention: Entrepreneurship promotion

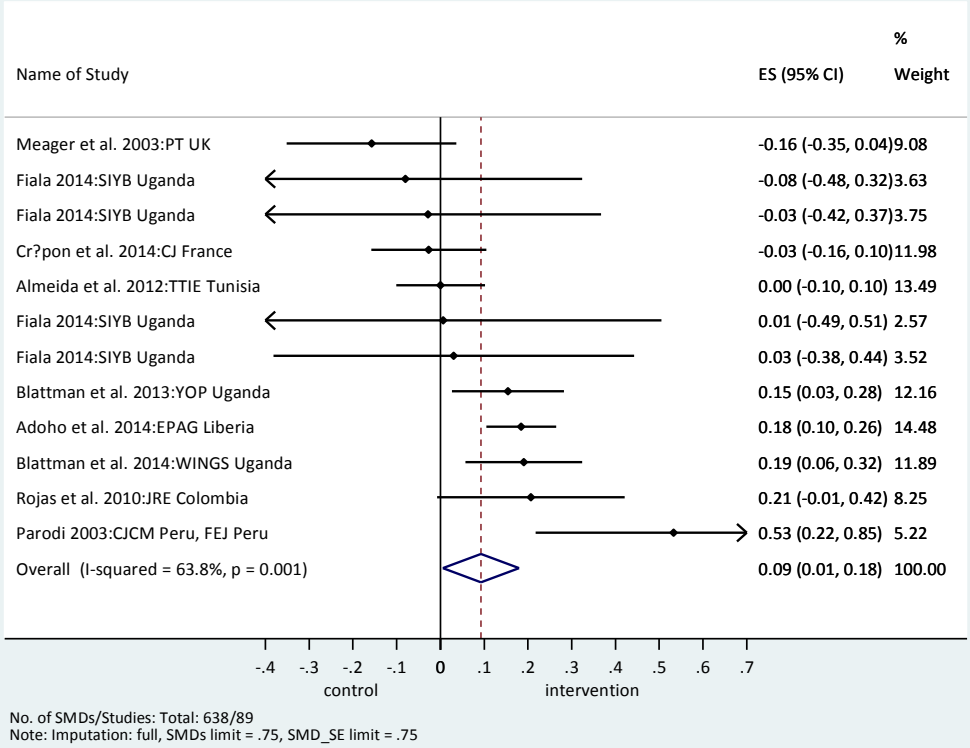


Figure 56: Disaggregated forest plot for earnings and income outcomes. Main category of intervention: Employment services

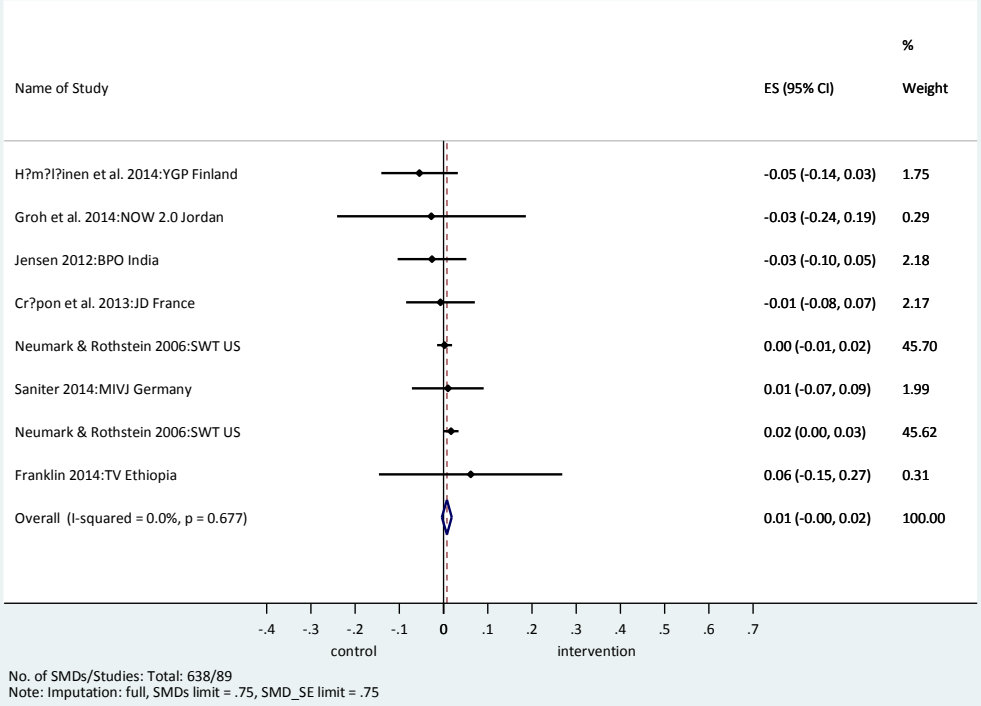


Figure 57: Disaggregated forest plot for earnings and income outcomes. Main category of intervention: Subsidized employment

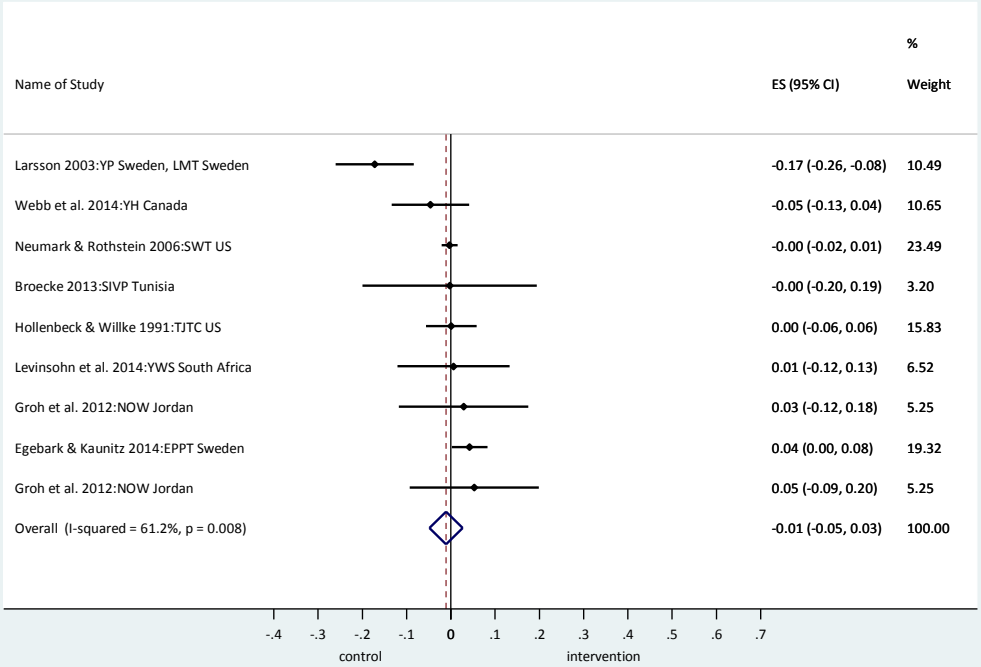
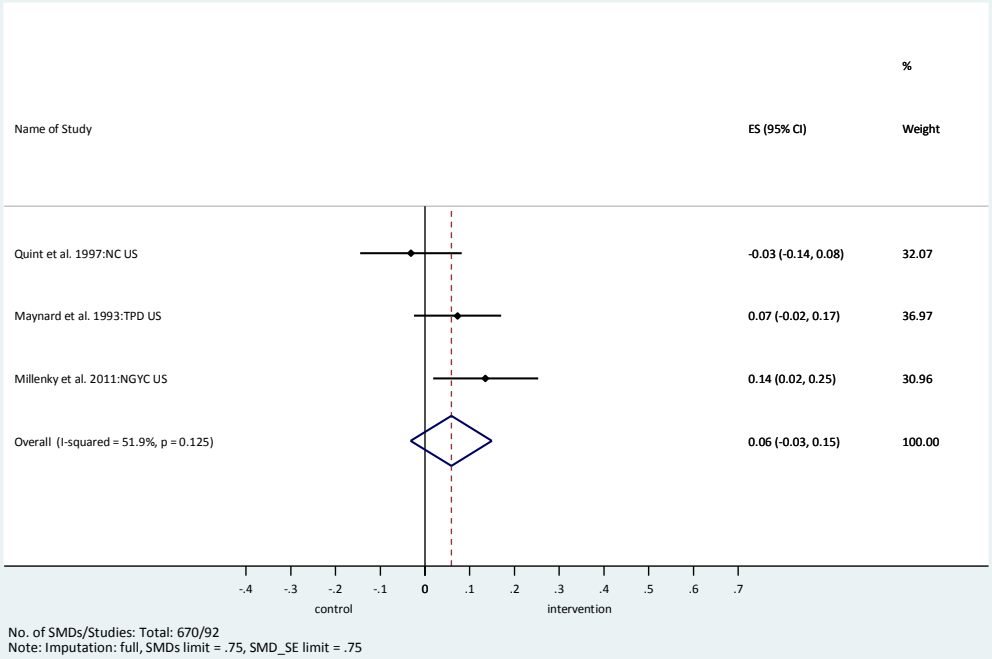
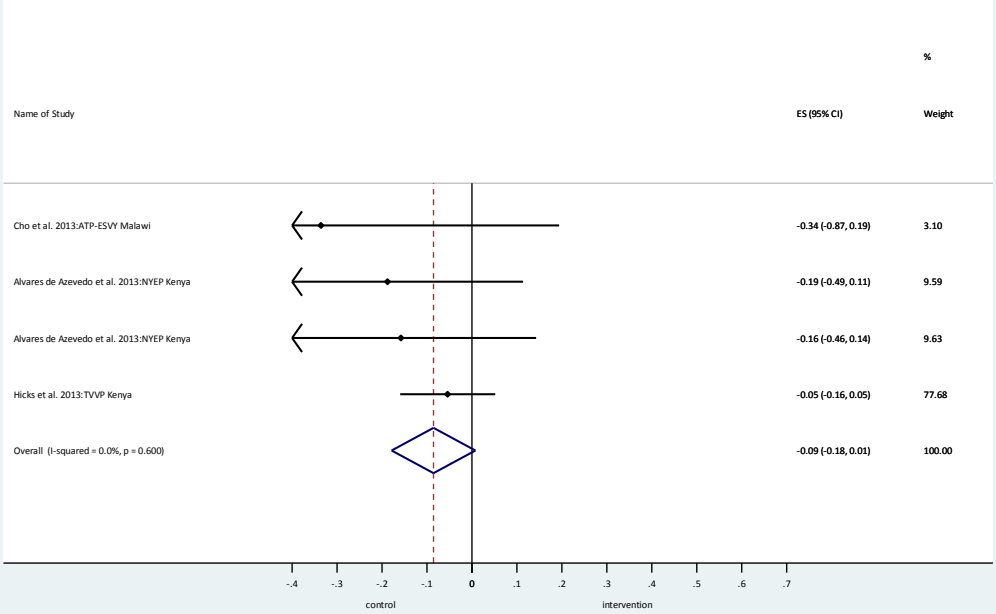


Figure 58: Disaggregated forest plot for earnings and income outcomes. Main category of intervention: Unspecified



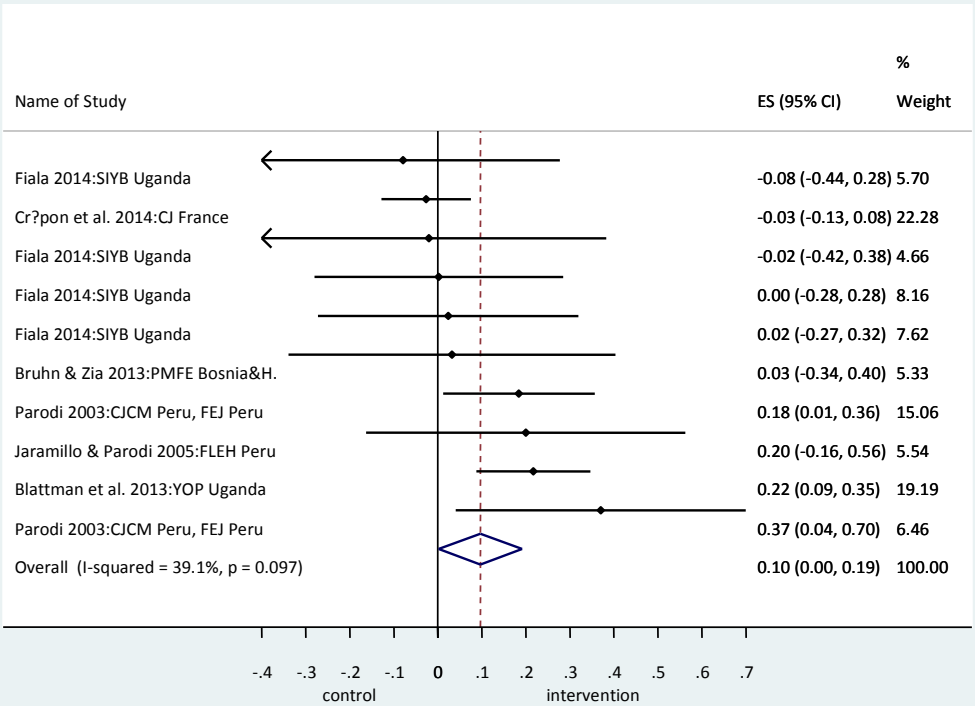
9.1.3 Business outcomes

Figure 59: Disaggregated forest plot for business outcomes. Main category of intervention: Skills training



No. of SMDs/Studies: Total: 169/14
Note: Imputation: full, SMDs limit = .75, SMD_SE limit = .75

Figure 60: Disaggregated forest plot for business outcomes. Main category of intervention: Entrepreneurship promotion



No. of SMDs/Studies: Total: 169/14
Note: Imputation: full, SMDs limit = .75, SMD_SE limit = .75
Note: The WINGS programme in Uganda (Blattman et al., 2014) (0.82 SMD; 95% CI = 0.13, 1.50) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the forest plot above.

The disaggregated forest plots for business outcomes of main categories employment services, subsidized employment and unspecified are not displayed due to lack of observations.

Figure 61: Funnel plot of employment outcomes, aggregated at study level

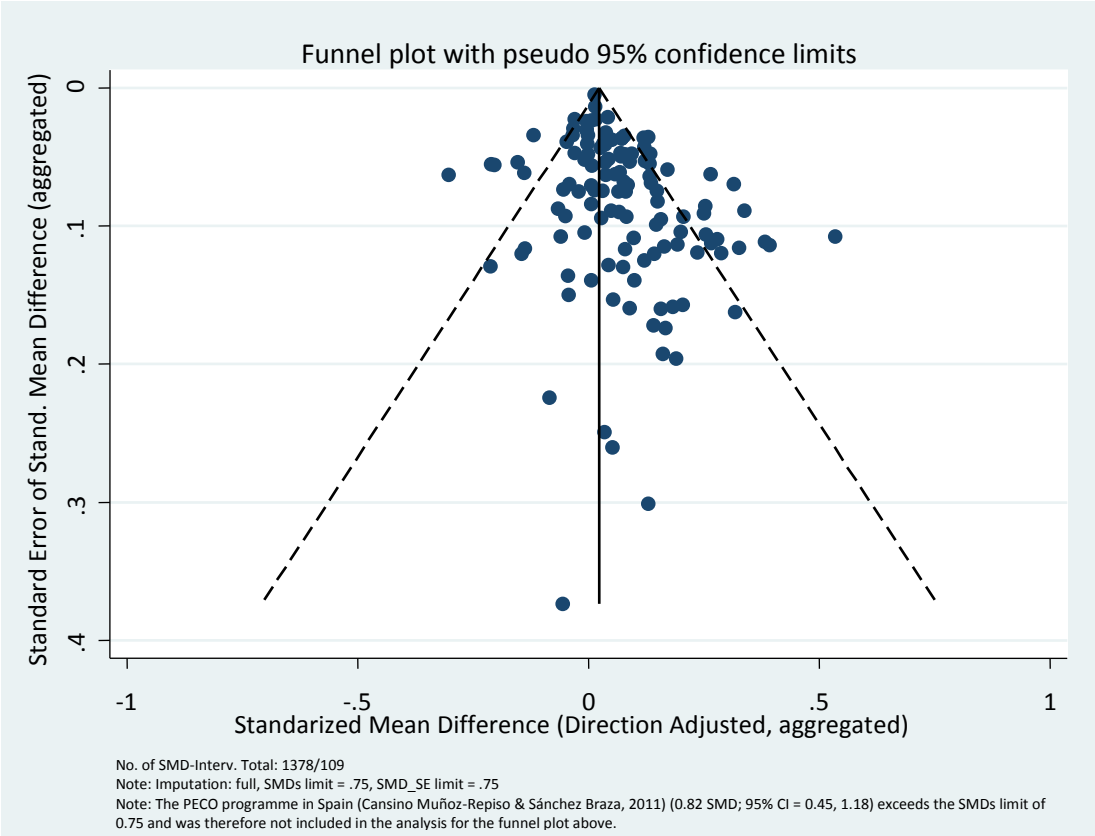


Figure 62: Funnel plot of earnings and income outcomes, aggregated at study level

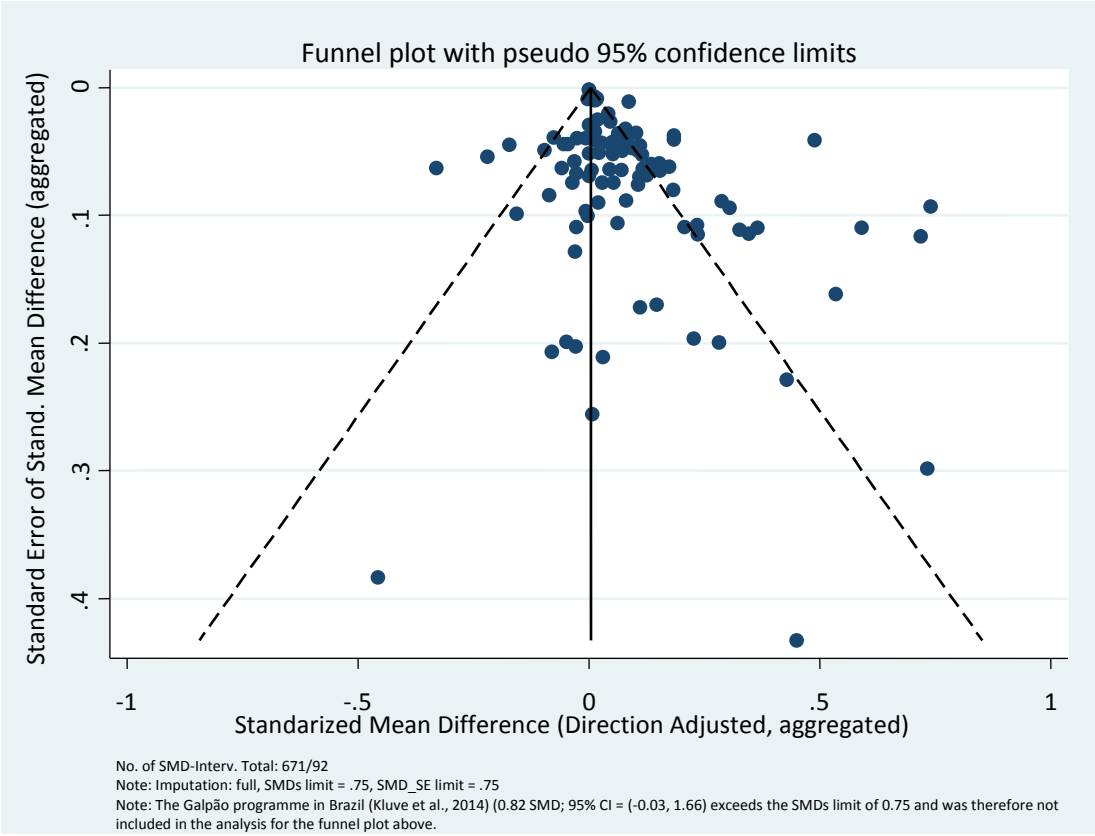


Figure 63: Funnel plot of business performance outcomes, aggregated at study level

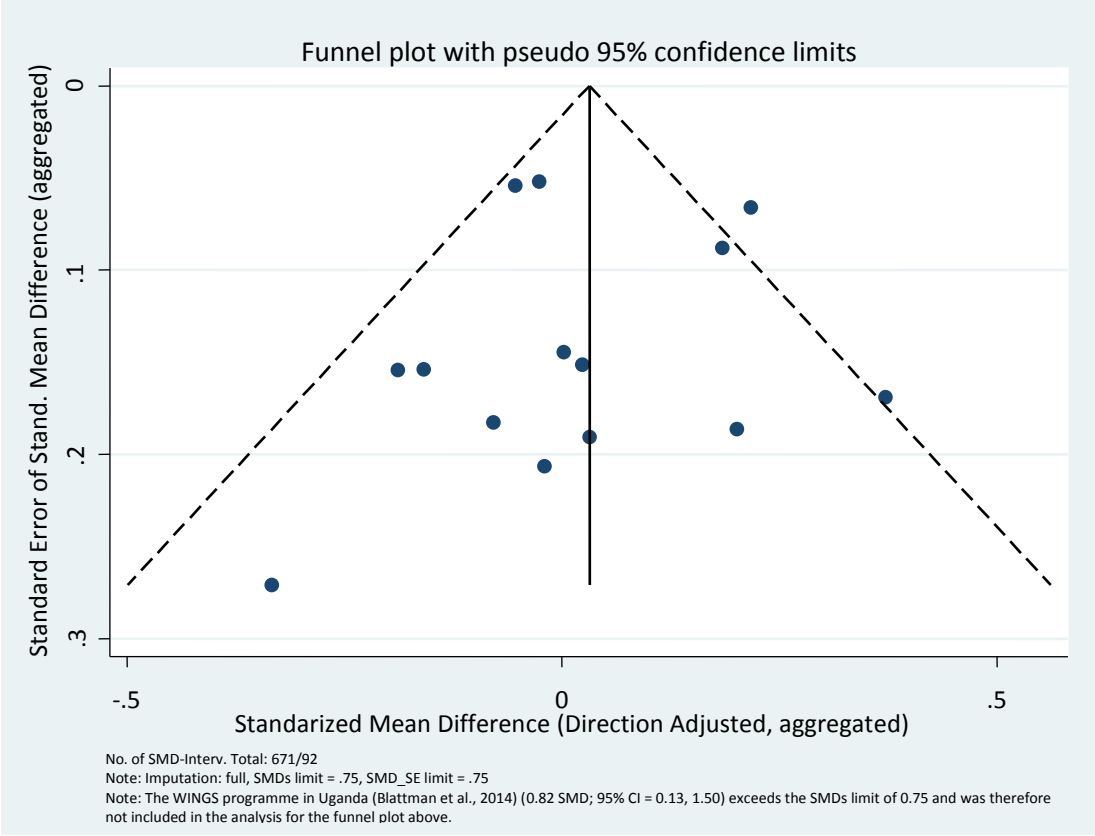


Figure 64: Funnel plot of employment outcomes, disaggregated (on effect size estimate level)

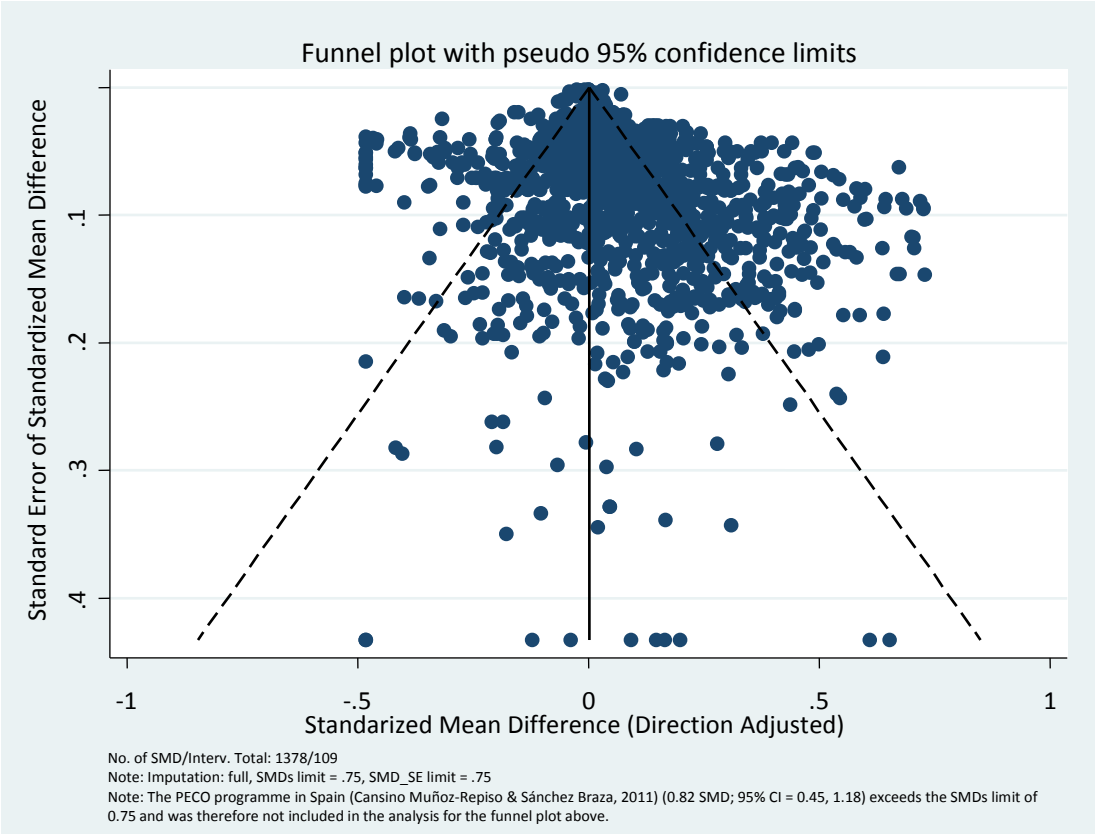


Figure 65: Funnel plot of earnings and income outcomes, disaggregated (on effect size estimate level)

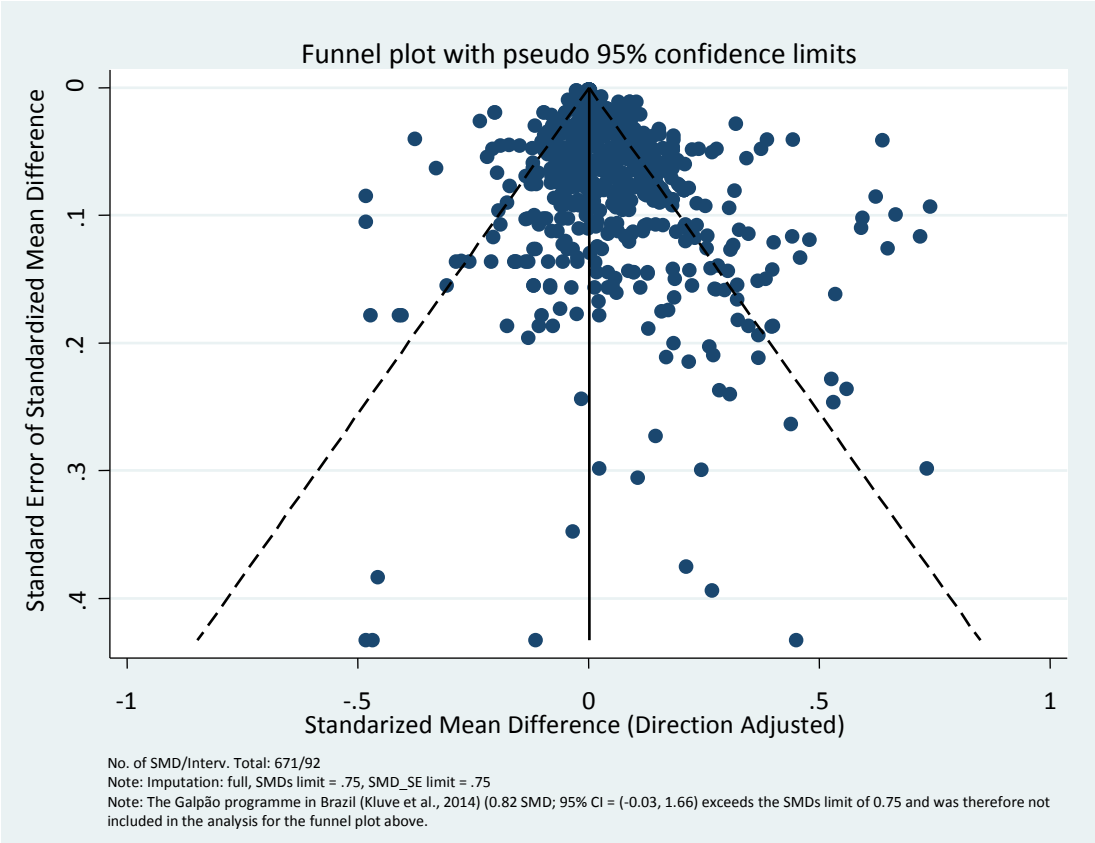
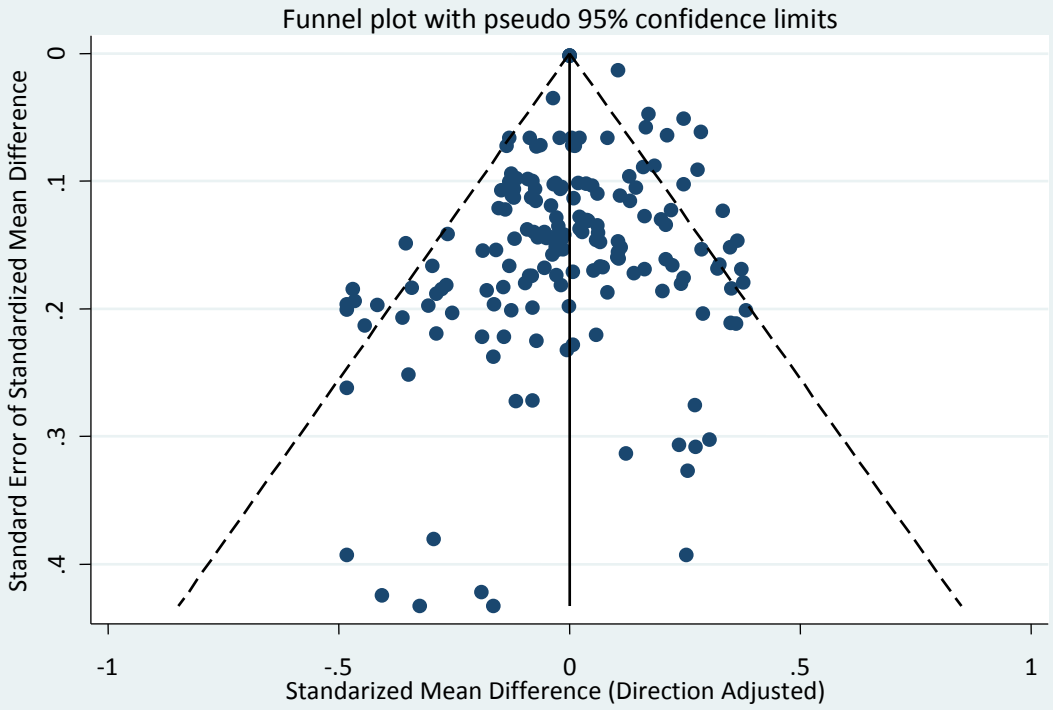


Figure 66: Funnel plot of business performance outcomes, disaggregated (on effect size estimate level)



No. of SMD/Interv. Total: 671/92
Note: Imputation: full, SMDs limit = .75, SMD_SE limit = .75
Note: The WINGS programme in Uganda (Blattman et al., 2014) (0.82 SMD; 95% CI = 0.13, 1.50) exceeds the SMDs limit of 0.75 and was therefore not included in the analysis for the funnel plot above.