



Protocol: Interventions to Improve Labour Market Outcomes of Youth: A Systematic Review of Active Labour Market Programmes

Jochen Kluge, Susana Puerto, Jonathan Stoeterau, Felix Weidenkaff, Marc Witte, David Robalino, Friederike Rother and José Manuel Romero

Submitted to the Coordinating Group of:

<input type="checkbox"/>	Crime and Justice
<input checked="" type="checkbox"/>	Education
<input type="checkbox"/>	Disability
<input checked="" type="checkbox"/>	International Development
<input type="checkbox"/>	Nutrition
<input type="checkbox"/>	Social Welfare
<input type="checkbox"/>	Other:

Plans to co-register:

<input checked="" type="checkbox"/>	No		
<input type="checkbox"/>	Yes	<input type="checkbox"/> Cochrane	<input type="checkbox"/> Other
<input type="checkbox"/>	Maybe		

Date Submitted: 18 December 2013

Date Revision Submitted: 14 October 2014

Approval Date:

Publication Date: 03 November 2014

Note: Campbell Collaboration systematic review Protocol Template version date: 24 February 2013

Table of Contents

1	BACKGROUND	4
1.1	THE IMPORTANCE OF YOUTH EMPLOYMENT	4
1.2	THE INTERVENTION: ACTIVE LABOUR MARKET PROGRAMMES FOR YOUTH	6
1.3	HOW THE ALMPs ARE SUPPOSED TO WORK	7
1.3.1	<i>Training and skills development</i>	8
1.3.2	<i>Entrepreneurship promotion</i>	12
1.3.3	<i>Employment services</i>	16
1.3.4	<i>Subsidized employment</i>	19
1.4	WHY THE REVIEW IS NEEDED	24
2	OBJECTIVES OF THE REVIEW	30
3	REVIEW METHODOLOGY	32
3.1	CRITERIA FOR INCLUSION AND EXCLUSION OF STUDIES IN THE REVIEW	32
3.1.1	<i>Population</i>	32
3.1.2	<i>Intervention</i>	32
3.1.3	<i>Comparison</i>	33
3.1.4	<i>Outcomes</i>	34
3.1.5	<i>Study designs</i>	34
3.2	SEARCH STRATEGY FOR IDENTIFICATION OF RELEVANT STUDIES	39
3.2.1	<i>Method to find and select relevant studies</i>	39
3.2.1.1	Managing and documenting the search and selection process	40
3.2.2	<i>Data sources and search strategies</i>	40
3.2.2.1	Scoping search	40
3.2.2.2	Primary search	42
3.2.2.2.1	General databases	42
3.2.2.2.2	Specialized databases	43
3.2.2.3	Search terms	43
3.2.2.4	Complementary search	44
3.2.2.4.1	Hand-searching of relevant websites/gateways	44
3.2.2.4.2	Reference lists and citation tracking	47
3.2.2.4.3	Institutions, networks and experts	47
3.2.2.5	Addressing potential biases	49
3.2.2.6	Study selection	49
3.3	DESCRIPTION OF METHODS USED IN THE COMPONENT STUDIES	51
3.4	CRITERIA FOR DETERMINATION OF INDEPENDENT FINDINGS	52
3.5	DETAILS OF STUDY CODING CATEGORIES	55
3.6	STATISTICAL PROCEDURES AND CONVENTIONS	57
3.6.1	<i>Standardizing effect sizes</i>	57
3.6.2	<i>Unit of analysis error</i>	59
3.6.3	<i>Treatment of missing information</i>	59
3.6.4	<i>Data synthesis</i>	60
3.6.4.1	Descriptive analysis of studies and interventions	60
3.6.4.2	Meta-analysis	61
3.6.4.3	Assessment of heterogeneity	61
3.6.4.4	Sensitivity analysis and detection of biases	62
3.7	TREATMENT OF QUALITATIVE RESEARCH	63
4	ACKNOWLEDGEMENTS	64
5	REFERENCES	65

6	APPENDIX.....	71
6.1	SCREENING QUESTIONNAIRE	71
6.2	LIST OF SOURCES FOR SCOPING SEARCH.....	74
6.3	SEARCH TERMS FOR ELECTRONIC DATABASES	78
6.4	SEARCH STRING FOR ELECTRONIC DATABASES: EXAMPLE: ERIC	79
6.5	KEYWORDS FOR SEARCH IN ENGLISH, SPANISH, FRENCH, GERMAN AND PORTUGUESE.....	80
6.6	CODE DESCRIPTION	81
7	SOURCES OF SUPPORT.....	97
8	DECLARATIONS OF INTEREST	98
9	REVIEW AUTHORS	100
10	ROLES AND RESPONSIBILITIES	102
11	PRELIMINARY TIMEFRAME	106
12	PLANS FOR UPDATING THE REVIEW.....	107
13	AUTHORS' RESPONSIBILITIES.....	108
14	PUBLICATION IN THE CAMPBELL LIBRARY.....	109

FIGURES

FIGURE 1: GLOBAL YOUTH UNEMPLOYMENT AND UNEMPLOYMENT RATE, 2000–2015P	4
FIGURE 2: SEARCH AND SELECTION PROCESS	39

TABLES

TABLE 1. TRAINING AND SKILLS DEVELOPMENT INTERVENTIONS: RESULTS CHAIN.....	10
TABLE 2. ENTREPRENEURSHIP INTERVENTIONS: RESULTS CHAIN.....	13
TABLE 3. EMPLOYMENT SERVICES INTERVENTIONS: RESULTS CHAIN	17
TABLE 4: SUBSIDIZED EMPLOYMENT INTERVENTIONS: RESULTS CHAIN.....	22
TABLE 5. EXISTING REVIEWS.....	27
TABLE 6. ONGOING REVIEWS	ERROR! BOOKMARK NOT DEFINED.
TABLE 7. INTERVENTIONS OF YOUTH EMPLOYMENT PROGRAMMES.....	33

BOXES

BOX 1. TRAINING AND SKILLS DEVELOPMENT: JUVENTUD Y EMPLEO IN THE DOMINICAN REPUBLIC	9
BOX 2. ENTREPRENEURSHIP PROMOTION: START AND IMPROVE YOUR BUSINESS (SIYB).....	15
BOX 3. EMPLOYMENT SERVICES: COUNSELLING AND JOB PLACEMENT FOR YOUNG GRADUATE JOBSEEKERS IN FRANCE	19
BOX 4. SUBSIDIZED EMPLOYMENT: NEW DEAL FOR YOUNG PEOPLE (NDYP) IN THE UNITED KINGDOM.....	21

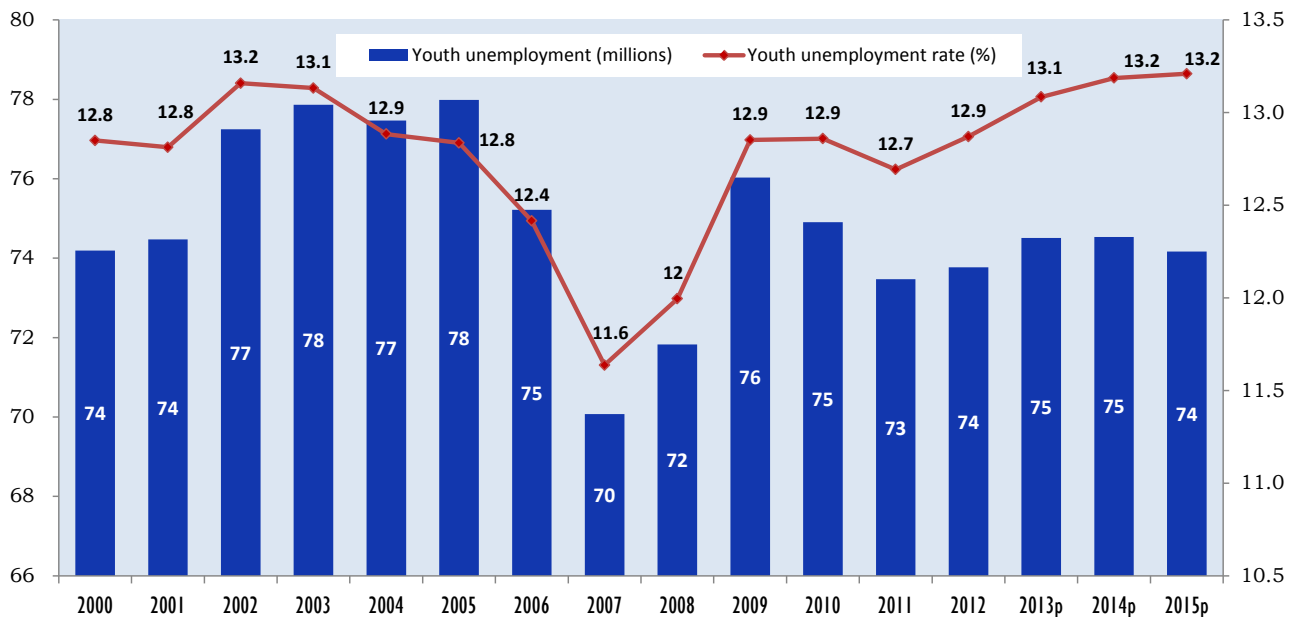
1 BACKGROUND

1.1 The importance of youth employment

There are about 1.2 billion youth worldwide, aged 15 to 24, and nearly 75 million of them are looking for a job (ILO, 2014). Such a sizable youth cohort presents an opportunity for growth but can also become a source of instability if youth unemployment is not addressed by effective interventions.

Young men and women have been in the spotlight ever since the hefty impact on youth employment caused by the economic crisis became apparent. The youth unemployment rate saw its largest annual increase between 2008 and 2009: rising from 12 to 12.9 per cent, with youth in industrialized countries being particularly affected. During the same period (2008–2009), youth unemployment rates in developed economies and the European Union (EU) and in Central and South-Eastern Europe (non EU) and Commonwealth Independent States (CIS) increased by 4.1 and 3.2 percentage points, respectively. These are the largest annual increases in youth unemployment rates ever recorded in any region.

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



Note: p = projection. Source: ILO (2014).

Stimulus packages, consultations, and private and public investments targeting youth came into vogue. Despite the response to the crisis, the global youth unemployment rate remains stubbornly high and today it is projected to stand at 13.2 per cent (ILO, 2014), as shown in

Figure 1. Does this signify insufficient investment or reflect the fact that it takes time for youth employment programmes to have a positive impact?

Global youth unemployment is estimated to stand at 74.5 million in 2014, an increase of 4.4 million since 2007. Higher unemployment and lower labour force participation among youth contributed to a decrease in the global youth employment-to-population ratio to 41.1 per cent in 2014, compared with 44 per cent in 2007. Part of this decrease is due to rising enrolment in education but the figures also reflect increasing levels of discouragement among youth (ILO, 2014).

The ratio of youth-to-adult unemployment rates illustrates the extent to which the economic crisis impacted regional and country-wide youth labour markets. Unemployment among youth is higher than among adults. Without exposure to a working environment, youth lack the job and soft skills that develop over time with experience. These inherent deficits “naturally” translate into youth unemployment rates that are twice – and sometimes three times – as high as the rates for adults. While the global ratio of youth-to-adult unemployment rates has hardly changed in recent years, its regional dynamics have resulted in some areas experiencing youth unemployment which is four times greater than the adult rate, exposing the vulnerability of young jobseekers to economic shocks, particularly in developing countries. According to the ILO, between 2009 and 2010, at the height of the economic crisis, the youth-to-adult ratio of unemployment rate in South-East Asia and the Pacific increased from 4.6 to 6. At the country level, ILO estimates for 2014 place the youth-to-adult ratios of unemployment rate in Egypt, Indonesia, and Sri Lanka at 5.8, 7.1, and 8, respectively (ILO, 2014).

The effects of the crisis on youth went beyond unemployment. The economic downturn also hit employed youth hard. The number of non-standard jobs, including temporary employment and part-time work, increased significantly, as many youth were drawn into lower quality work, temporarily postponing their career aspirations to secure an income, however meagre. In Europe, an increasing number of employed youth have unwillingly accepted non-standard jobs. Youth part-time employment as a share of total youth employment in Europe was 25 per cent in 2011. A further 40.5 per cent of employed youth in the region were engaged under temporary contracts (ILO, 2013).

In most regions, the youth unemployment rate is following an upward trend (ILO, 2014), which gives a long-term perspective to the challenge of addressing the issue. With a rather sluggish recovery and poor economic growth, job prospects for youth are not expected to improve substantially in the medium term. This economic landscape and the potential negative consequences of unemployment on human capital, health, happiness, crime, and socio-political stability (Bell and Blanchflower, 2009) call for further research and comparative analyses into the most effective measures to improve the labour market outcomes of youth. This requires (i)

learning from past experiences (both successes and failures), and (ii) the provision of a quantifiable estimate of the magnitude of impacts arising from youth employment programmes.

1.2 The intervention: Active labour market programmes for youth

Labour market interventions to be examined by this systematic review fall into the category of Active Labour Market Programmes (ALMP), 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 programs for youth when in transition from school to work, labour market programs to provide or promote employment for unemployed and other persons (excluding young and disabled persons) and special programs 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. labour demand and supply matching and market efficiency through, for instance, job-search assistance, labour market information, and pre-screening of programme applicants. ALMPs can also enhance labour supply through 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).

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 the problems.

For the purposes of this review, ALMPs are clustered into the following typology of interventions:

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

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 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 the businesses that programme participants start or already own.

Exposure to ALMPs is expected to create a spillover effect among non-programme 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 may have a negative impact on the performance of non-participants. For example, there is evidence that wage subsidy programmes may 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), therefore decreasing the overall employment impact of the programme. To address this issue, increased attention must be given to programme design features that can help to compensate for negative spillovers, 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 drawn to determine whether one event in the sequence actually yields the expected changes in labour market outcomes. Once the theories of change are clear, the systematic review will seek to examine whether the evidence supports the expected transition channels and causality.

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, separating the results chains for each category seeks to provide further transparency in the assumptions and support the interpretation of results revealing potential causal mechanisms.

In the interest of a well-defined intervention description, we omit those activities and outputs that are not strictly linked to labour market effects. Similarly, we focus narrowly on individual-level labour market outcomes, leaving aside other potential side-effects, such as increased psycho-social 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 study will 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) 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 the chances of young people in the labour market (Biavaschi et al., 2012). Skill 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 et al., 2007; Fares & Puerto, 2009). The objective of skill training programmes is to develop employment-relevant skills of jobseekers. Broadly speaking, employability skills refer to a set of job-specific technical skills, but also include non-technical, soft (or behavioural) 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 (Cunningham, Sanchez-Puerta & Wuerml, 2010; Youth Employment Network & IYF, 2009).

In this analysis, we classify training programmes according to the skill-set which they target (Table 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 skill 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).
- Second, business skills training, often taught as part of programmes that aim to increase entrepreneurial activities among youth. Such entrepreneurial training programmes cover a large variety of factors that are believed to determine business success (ranging from financial skills to problem-solving skills).
- 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”).
- 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 very 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 comprehensive skills training programme is provided in Box 1.

Box 1. 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 a comprehensive active labour market programme (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 that focused 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, Ibarraán, Regalia, Rosas and Soares, 2011; Ibarraán, Ripani, Taboada, Villa and García, 2012.

A number of conditions determine whether skills training programmes are successful in bringing *additional* youth into work – most notably, consistency 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.

Table 1. Training and skills development interventions: Results chain

Inputs	Activities		Outputs	Outcomes
<ol style="list-style-type: none"> 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	<ol style="list-style-type: none"> 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	<ol style="list-style-type: none"> 1. Increased employability 2. Increased probability of employment 3. Increase in number of hours worked 4. Reduced time to find job/shorter unemployment duration 5. Increased ability to retain job/longer job duration 6. Better quality of employment (contract type, number of hours worked) 7. Increased earnings or consumption
	Business skills training	As above	<ol style="list-style-type: none"> 1. Improved management skills 2. Improved understanding of business mechanisms 3. Improved financial literacy 	<ol style="list-style-type: none"> 1. As above 2. Increased business performance (efficiency, profits, investments, output of entrepreneurs)¹
	Literacy or numeracy skills training	As above	<ol style="list-style-type: none"> 1. Improved reading and writing skills 2. Improved mathematical skills 	

¹ Additional entrepreneurship-related outcomes are listed in section 1.3.2 below.

Inputs	Activities		Outputs	Outcomes
	Behavioural, life skills or soft skills training	As above	<ol style="list-style-type: none"> 1. Improved psychosocial characteristics 2. Improved decision-making skills 3. Improved communication and teamwork skills 4. Increased self-management, self-esteem 5. Improved physical and mental health 	As above

Assumptions

		<ol style="list-style-type: none"> 1. Target group participates in training (there is awareness about the programme's existence) 2. Contracted training institutions/employers conduct training 	<ol style="list-style-type: none"> 1. Participants complete/attend the training 2. Training addresses participants' constraints (e.g. existing skill shortages) 3. Participants learn in training/training increases skill level/training is well matched to interests and abilities of participants 4. Training induces expected behavioural change 	<ol style="list-style-type: none"> 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 increased 5. Employers value certified training 6. Participants gain recognized and valued qualifications 7. Adequate economic, social, institutional and administrative conditions are in place
--	--	---	--	--

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.

1.3.2 Entrepreneurship promotion

Entrepreneurship promotion programmes aim to lower the barriers and costs for young unemployed and underemployed 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 lack of personal savings and resources to 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. We disaggregated such programmes into three types:

- First, those providing or facilitating access to credit (including microfinance programmes).
- Second, those providing start-up grants.
- Third, those fostering micro-franchising mechanisms.

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

In addition to access to finance, there are programmes that offer business advisory services and mentoring for soon-to-be or already self-employed youth. Finally, some interventions try to reduce the barriers to business creation by assisting prospective entrepreneurs to enter markets or existing value chains. The above-described interventions and their underlying theory of change are shown in Table 2. We included training related ALMPs, which deliver the skills relevant for starting or maintaining a business (such as management skills, leadership skill and financial literacy), described in section 1.3.1 above.

Table 2. Entrepreneurship interventions: Results chain

Inputs	Activities		Outputs	Outcomes
<ol style="list-style-type: none"> 1. Budget 2. Staff 3. Local counterparts 4. Trainers 5. Partnerships 6. Facilities 7. Equipment 8. Supplies 9. Technical expertise 10. Curricula 	Business advisory services/mentoring (e.g. assessment of business plans)	Advice delivered	<ol style="list-style-type: none"> 1. Improved management skills (e.g. accounting practice, stocks management, investments) 2. Improved understanding of business mechanisms 3. Improved financial literacy/behaviour 4. Increased knowledge of business laws and regulations 5. Improved knowledge of business possibilities 6. Reduced risk/uncertainty in starting a business 	<ol style="list-style-type: none"> 1. Increased business investment, performance and competitiveness (e.g. efficiency, output, profits, sales, number of employees, investments, business survival) 2. Additional businesses started 3. Increased earnings or consumption among young entrepreneurs
	Access to markets and value chains	<ol style="list-style-type: none"> 1. Support business networks 2. Provide technology necessary for value chain inclusion 	<ol style="list-style-type: none"> 1. Increased knowledge of markets and networks 2. Increased access to business networks and supply chains 	As above
	Credit or access to credit	<ol style="list-style-type: none"> 1. Provide credit to young entrepreneurs 2. Match entrepreneurs with credit agencies 	<ol style="list-style-type: none"> 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	<ol style="list-style-type: none"> 1. Match participants with franchisors 2. Intermediate between franchisors and potential franchisees 3. Distribute information about franchising 4. Assist in setting up franchise 	Increased incentives (lower barriers) to start own business/franchise	As above

Inputs	Activities		Outputs	Outcomes
		business 5. Support exiting franchisees		
Assumptions		<ol style="list-style-type: none"> 1. Correct group is targeted (e.g. participants are credit constrained) 2. Target group participates in programme (there is awareness about the programme's existence) 	<ol style="list-style-type: none"> 1. Participants learn from advisory service (sufficient skill level) 2. Advice prompted expected behavioural change 3. Credit or grant is used for enterprise 4. Credit agency/franchisor does not exploit entrepreneur 	<ol style="list-style-type: none"> 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

Many entrepreneurship programmes take a comprehensive approach; for example, combining access to credit with business skills training or post-programme consultation (i.e. mentoring and coaching). Most microfranchise programmes follow this pattern – delivering finance, business networks and mentoring services.

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.

Box 2. Entrepreneurship promotion: Start and Improve Your Business (SIYB)

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 economies and economies in transition. The SIYB programme is a system of interrelated training packages and supporting materials for small-scale entrepreneurs. The programme is designed by the International Labour Organization (ILO) and implemented with support from certified trainers in partner institutions in more than 100 countries. Initially developed in the 1980s, it has now been translated into more than 40 languages and had reached more than 4.5 million potential and existing entrepreneurs by 2010. 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 Start Your Business (SB) 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 report 54 per cent greater profits.

Sources: based on information available at: www.ilo.org/siyb [12 Oct. 2014]; van Lieshout, Sievers & Aliyev, 2012; Fiala, 2014.

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.² Whether an entrepreneur will finally hire additional workers may also depend on the macroeconomic and labour market environment. Box 2 describes the programme Start and Improve Your Business (SIYB), a widely used and adapted entrepreneurship training package designed by the ILO and tailored for youth.

1.3.3 Employment services

Employment service programmes are generally based on the (matching and) intermediation approach to active labour market policy. Intervention types within employment services are shown in Table 3. *Job-placement programmes* acknowledge the existence of information asymmetries 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 (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 and networks to find job openings); and, on the other hand, they provide information to potential employers about available and 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*, include job-search training, educational or career guidance, counselling and monitoring programmes. Such programmes primarily address disadvantaged or demotivated youth who are disconnected from the labour market. Their primary aim is to improve the intensity, motivation and effectiveness of job-searches by participants. 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 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 being the main providers of employment services, some others have moved into subcontracting, opening an important role for private employment agencies to address mismatches and information failures in the labour market. Box 3 illustrates a subcontracting model applied by a French public employment agency to facilitate counselling and job-placement for educated youth.

² It is important to note that the theory of change analysis does not details potential general equilibrium effects, such as substitution.

Table 3. Employment services interventions: Results chain

Inputs	Activities		Outputs	Outcomes
<ol style="list-style-type: none"> 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	<ol style="list-style-type: none"> 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 	<ol style="list-style-type: none"> 1. Improved matching of jobseekers and employers 2. Increased intensity (motivation) and efficiency of job-search 	<ol style="list-style-type: none"> 1. Increased labour-market participation 2. Increased probability of employment 3. Increase in hours worked 4. Reduced time to find job/shorter unemployment duration 5. Increased ability to keep a job/longer job duration 6. Better quality of employment (contract type, hours worked) 7. Increased earnings or consumption
	Job counselling/job-search assistance/mentoring	<ol style="list-style-type: none"> 1. Provide career and personal development advice 2. Provide job-search advice or training 3. Ensure active and efficient job-search 	<ol style="list-style-type: none"> 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, child-care)	Greater ability to find and accept jobs (e.g. enhanced mobility)	As above
Assumptions		<ol style="list-style-type: none"> 1. Target group (unemployed and employers) takes up the service 	<ol style="list-style-type: none"> 1. Correct target group identified (participants are constrained by lack 	<ol style="list-style-type: none"> 1. Existing labour demand for employment service beneficiaries

Inputs	Activities		Outputs	Outcomes
		<ul style="list-style-type: none"> offer (there is awareness about the programme's existence) 2. Participants complete/attend the service 3. Participants comply with conditionalities and service requirements 4. Service matches the needs and abilities of participants 	<ul style="list-style-type: none"> 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 	<ul style="list-style-type: none"> 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

Box 3. 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 followed 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 compete mainly with other educated workers and in weak labour markets.

Sources: based on information available at: www.youth-employment-inventory.org [12 Oct.2014]; Crépon, Duflo, Gurgand, Rathelot and Zamora, 2013.

There is indication that involvement in employment services (and in ALMPs in general) creates a stigmatizing effect on participants (Boone & van Ours, 2004, and 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 non-compliance with 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 4), both of which are designed to increase the job and training opportunities available to young

unemployed. 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 effects.

Wage subsidies come in numerous forms and can be provided through various mechanisms, ranging from direct transfers (vouchers) to reductions in social security contributions and tax credits.

First, 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.

Second, 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 better 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 facing 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 US with rather mixed results concerning the effectiveness of wage subsidies as tools for fostering job creation. The UK programme, New Deal for Young People, demonstrated the subsidy's ability to lower unemployment among youth relative to other available interventions, including subsidized education and training. The programme is explained in Box 4. While evidence in developing countries is more limited, it is also mixed, showing negative impacts of wage subsidies in Jordan (Groh, Krishnan, McKenzie & Vishwanath, 2012) but positive and sustained impacts in South Africa (Levinsohn et al., 2014).

Box 4. Subsidized employment: 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 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 enter 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 a 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 [12 Oct. 2014].

The second type of labour market intervention analysed in this category is *labour-intensive public employment programmes*. Public employment programmes are the umbrella for public works programmes and employment guarantee schemes. Despite the strong association of these programmes with infrastructure and construction works, there is a growing trend towards works in the social sector, environmental services and multi-sectoral, community driven programmes (Lieuw-Kie-Song, Philip, Tsukamoto & Van Imschoot, 2010).

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.

Table 4: Subsidized employment interventions: Results chain

Inputs	Activities		Outputs	Outcomes
<ol style="list-style-type: none"> 1. Budget 2. Staff 3. Local counterparts 4. Trainers 5. Partnerships 6. Facilities 7. Equipment 8. Supplies 9. Technical expertise 10. Curricula 	Linking beneficiaries to subsidized employment in private enterprises		<ol style="list-style-type: none"> 1. More positive attitudes towards employment/increased incentives to apply or to work 2. Development of a work ethic and work habits 3. Increased technical skills/work experience 4. Incentives to continue education 5. Participant regain labour market contact/integrate jobseekers into networks 	<ol style="list-style-type: none"> 1. Increased probability of (formal, well-regulated) employment beyond the programme duration 2. Reduced time to find job/shorter unemployment duration 3. Increased ability to retain a job/longer job-duration 4. Better quality of employment (contract type, hours worked, salary) 5. Increased earnings or consumption
	Public work in infrastructure development projects	Matching unemployed to financed public works in infrastructure development projects	<ol style="list-style-type: none"> 1. As above 2. Improved sense of social participation and opportunities 	As above
	Social development and community works and services projects (e.g. children's care, sick and elderly care, security, health)	Matching unemployed to financed public work in community and services projects	<ol style="list-style-type: none"> 1. As above 2. Improved sense of social participation and opportunities 3. Improved social skills 	As above
Assumptions		Target group (unemployed and employers) participates in programme (and it is	<ol style="list-style-type: none"> 1. Participants are motivated to work and sufficiently qualified (adequately matched) 2. Participants learn on the job (i.e. 	<ol style="list-style-type: none"> 1. Existing labour demand for the skills acquired by beneficiaries of the interventions 2. Correct barriers and constraints for

Inputs	Activities	Outputs	Outcomes
		<p>experience increases in skill levels)</p> <p>3. Induces (positive/expected) behavioural changes/no adverse behavioural changes</p> <p>4. Subsidies are not exploited by firms</p> <p>5. Correct target group: participants would not have obtained the job without the subsidy³</p>	<p>youth on the labour market are addressed</p> <p>3. Work experience adequately signals higher employability/skills</p> <p>4. Skills/work experience match labour market demands</p> <p>5. No stigmatizing effects⁴</p> <p>6. Danger of programme/welfare dependency among participants taken into account⁵</p>

³ Targeting is critical to avoid misuse of subsidized employment programmes.

⁴ Programmes may create a stigmatizing effect on participants, particularly the most highly educated. In France, workfare participants carried a stigma that hindered their transition to better and more durable jobs (Bonnal, Fougere & Sérandon, 1997; and Brodaty, Crepon & Fougere, 2000). Adequate programme marketing and publicizing is important to address stigma effects.

⁵ Programs 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 aiming 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.

While public employment programmes have been often recommended as a measure in times of crises (e.g. seasonal shocks or economic recession)⁶ 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 et al., 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)⁷ 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 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 such 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 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 of youth employment programmes is more common among

⁶ 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 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.

⁷ 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.

developed countries and particularly scarce in Asia, Africa and the Middle East and North Africa, Asia and sub-Saharan Africa. While contextual variables, such as income level and development, seem to play a role in shaping the probability of positive outcomes from youth ALMPs (Betcherman, Godfrey, Puerto, Rother & Stavreska, 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 urban contexts, informal rather than formal settings, and in post-conflict and fragile-state environments.

- *The question of programme focus:* Most evaluations exist in the area of training and skills development, while evidence on other types of youth employment interventions, such as subsidized employment, employment services and entrepreneurship, is relatively scarce. There is a significant knowledge gap on 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 relative 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.
- *The range of beneficiaries:* More evidence is needed to understand 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 will take stock of the available evidence and examine changes in labour market outcomes prompted by labour market interventions for youth.

Assessing the impact of active labour market programmes 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 economies in transition, 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 changed 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).⁸

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)⁹, an online global repository of information on labour market programmes for youth, offers records of 116 impact evaluation studies of youth employment interventions worldwide. While rigour varies between studies, there is a recognized transition towards randomized experiments and stylized methods of evaluating impact.

This systematic review will examine experimental and quasi-experimental evaluations of ALMPs that target youth. The goal is to look at the available evidence in order to fill the knowledge gap on the impact and effectiveness of these interventions in a systematic, rigorous manner. Section 3 provides further information on the methodology proposed 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 5 presents the available evidence on both completed and ongoing reviews, identifying key differences to this review and summarizing its added value. To the best of our knowledge, this is the first systematic review of the impact of employment interventions on youth labour market outcomes that collates global evidence from youth ALMPs, examines the most relevant outcomes along the causal chain and identifies study effect sizes through a rigorous meta-analysis.

While reviews by Gonzalez et al., Grimm and Paffhausen and Cho and Honorati do not focus on youth, there may be a number of youth entrepreneurship interventions and studies that will overlap with the systematic review presented in this protocol.

⁸ In a review of evaluation methods used in ALMPs, Heckman, La Londe and Smith (1999) identified a number of methodological lessons – from the recognition of a 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 relate to these lessons by providing a framework that relies on credible comparison groups and minimizes selection bias.

⁹ Available at: www.youth-employment-inventory.org [12 Oct. 2014].

Table 5. Existing and on-going 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 presented in this protocol will build on the work of Betcherman et al. with 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 work 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 timely, recent impact evaluations to identify and discuss knowledge gaps. The review does not rely on a statistical meta-analysis or study effect sizes.
IEG (2013)	In 2012, the Independent Evaluation Group of the World Bank 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

	<p>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, the review described in this title registration will cover countries from all levels of development and will disregard any training programme that is delivered in a formal education setting.</p>
<p>Reviews looking at ALMPs in general, not focused on youth</p>	<p>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, Kluve and Weber (2010). The sample of programmes for youth is limited, as are the findings. Only Card et al. (2010) offer a relatively rigorous search and quantitative analysis of impact based on study significance, but none of these studies analysed effect size magnitudes. Other studies with similar limitations include studies that looked at programmes implemented in OECD countries only, e.g. Heckman et al. (1999), Kluve and Schmidt (2002), and Kluve (2006 and 2010).</p>
<p>Reviews of entrepreneurship interventions, not focused on youth</p>	<p>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 and snowball search.</p>
	<p>Grimm and Paffhausen (2013) propose to 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 lead to such impact. The review relies on a rigorous search strategy and provides effect sizes in selected outcomes, namely number of employees and/or growth rate.</p>
<p>Reviews looking at labour market regulations, not focusing on youth</p>	<p>A recent review by Nataraj, Perez-Arce, Kumar and Srinivasan (2013) looks at the impact of regulations (such as minimum wages, regulation 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 impact and effectiveness of interventions serving young people.</p>

On-going review

Gonzalez, Piza, Taylor, Cravo and Abdelnour (2013) will explore the impact of business support services, both direct and indirect, for small and medium-sized enterprises on firm performance in low- and middle-income countries. The review promises a rigorous search methodology as well as insights into the impact of such services in Africa, which will be of particular interest to this review. There is no mention of effect sizes assessment.

2 OBJECTIVES OF THE REVIEW

Addressing the youth employment challenge ranks high on the development agenda with a view to improving 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, based on rigorous evidence. The current political and economic context makes even more imperative the need to provide evidence on what does, what does not work, why, and how to improve labour market outcomes of young men and women.

Policymakers are regularly looking for ideas on how to support youth in the labour market. Their requests come with an increased requirement for evidence. 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 knowledge, technical and financial assistance are often made to the World Bank from client countries. Non-government organizations, donors and employment practitioners in general are also looking intensively at success factors to support youth.

In this context, this systematic review seeks to support policymakers and practitioners with evidence-based recommendations on what works to effectively support youth in the labour market. The review team plans to achieve this by summarizing and integrating empirical research on the impact of interventions aiming to improve labour market outcomes of youth. The systematic review aims to investigate the impact of labour market interventions on labour market outcomes of young people as well as examining whether the evidence supports the underlying assumptions about what ALMPs for youth are designed to achieve.

The following research questions will frame the analysis on what are effective measures, which will ultimately help decision-makers in the allocation of their resources and 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 will look at skills training, entrepreneurship promotion, employment services, mentoring and subsidized employment interventions.
2. Which of these interventions are the most effective?

The review will also look at a more contextual research question; namely, how applicable is the evidence to countries in Africa? This question was put forward by the review's main donor, Canadian CIDA, and will frame the review's advice on youth employment interventions that can effectively counteract joblessness in the African context. By synthesizing the evidence on the

relative effectiveness of different labour market interventions for youth, this systematic review will contribute to closing the knowledge gap in this field with a real impact on the nearly 75 million young men and women that are now actively looking for a job.

3 REVIEW METHODOLOGY

3.1 Criteria for inclusion and exclusion of studies in the review

The review's criteria for relevance include the following conditions:

- (i) the studies must be of the youth employment interventions listed in the research question or combinations of these interventions
- (ii) the studies must reflect completed experimental and quasi-experimental evaluations which measure impacts on eligible labour market outcomes (section 3.1.4).

The review team will incorporate findings from qualitative evaluations and assessments linked to the studies that comply with the criteria for inclusion. Relevance decisions, to be made by Principal Investigators (PIs), will be based on a reading of abstracts and, in cases when this is insufficient, on the reading of full reports. In order to ensure the reliability of relevance decisions, the review team will assess a sample of three studies. This exercise will help to ensure uniformity in decision making.

The selection of studies will be based on the inclusion and exclusion criteria outlined by the screening questionnaire presented in the Appendix. The review's eligibility criteria are detailed in the following subsections.

3.1.1 Population

The ALMPs considered in the study (i) target the unemployed or those with low levels of skills or limited work experience or who are generally disadvantaged in the labour market and (ii) aim to promote employment and/or earnings/wage growth among the target population, rather than simply providing income support (Heckman et al., 1999). In terms of age groups, the study focuses on programmes that are designed for – or target primarily – young men and women aged between 15 and 35.

3.1.2 Intervention

Eligible programmes considered for analysis, as defined in Table 6, include training and skills development, entrepreneurship promotion, employment services and subsidized employment.

Table 6. Interventions of youth employment programmes

Interventions	Description
Training and skills development	Comprises programmes outside the formal education system (and therefore does not consider Technical Vocational and 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	Aims to provide entrepreneurial skills as well as physical, financial and social capital for self-employment.
Employment services	Delivers job counselling, job-search assistance, and/or mentoring services, which are often complemented by job placements and technical or financial assistance.
Subsidized employment	Looks mainly at programmes providing wage subsidies or interventions aiming to reduce the labour cost for employers as well as labour-intensive programmes or public works which provide short-term employment to youth in infrastructure or social development and community projects.

3.1.3 Comparison

Eligible comparison groups include those which receive no intervention, another youth employment intervention or are due to receive the intervention in a pipeline design. In general, studies will focus on the binary case, where two treatment states exist: exposure to treatment and non-exposure (control).

Since the extension of the binary model to multi-valued treatments is straightforward, corresponding methods have been developed. The case of a randomized control trial (RCT) with several “arms” (or treatment states) is particularly simple. Hence, we will include studies with multi-valued treatments. We will code relative effects (e.g. effects that are measured relative to an alternative treatment) for those studies that also include a non-exposure comparison group. We will code relative effects separately for studies which include a non-exposure (control) state since the interpretation of results that compare two actual treatments is different, in particular in relation to other studies’ findings. The comparison group must either be constructed using an experimental design or using one of the quasi-experimental methods discussed in section 3.1.5.

3.1.4 Outcomes

The study identified the following primary outcomes of interest, which will be the focus of the analysis and provide the framework for potential policy and programme recommendations:

- **Employment outcomes**, including employment, unemployment, participation rates, hours worked, unemployment duration and quality of employment.¹⁰
- **Earnings outcomes**, including reported earnings and income, household income, consumption and salary and/or wage.
- **Business performance**, including profits, sales, number of employees and jobs created, capital and investment, business creation and business survival.¹¹

The systematic review will include studies that measure change in at least one of these outcomes among intervention participants and relative to non-intervention participants based on a counterfactual analysis. When available, the analysis will include effects on population subgroups outlined by age cohorts, gender, education level, income level and location, among other dimensions (see coding tool for a full set of combinations in section 6.6). No restriction is placed on the duration of follow-up data.

The review also includes outcomes which are measured conditional on some other outcome. Commonly, the review covers studies estimating the effect of treatment on wages by either (i) setting all wages of unemployed individuals to zero or (ii) excluding all unemployed individuals from the treatment and control groups. The latter would be the treatment effect on wages conditional on being employed. The review treats such conditional treatment effects as a subgroup analysis. Generally, unconditional treatment effects are considered to be the more policy-relevant estimate and are the basis for the review's analysis.

The review will not include studies that focus only on intermediary outcomes without measuring impact on the primary outcomes listed above. The review will consider net employment or general equilibrium effects as outcomes of interest if a study also measures the above described primary outcomes and meets the study inclusion criteria.

3.1.5 Study designs

The key conceptual point is that both exposure to treatment (factual) and non-exposure (control

¹⁰ The review will consider outcomes from empirical probability models whenever measured and available.

¹¹ The reviewers acknowledge that there are a number of ongoing systematic reviews looking at business performance outcomes. Provided that such reviews do not have a focus on youth, this review will consider them as primary outcomes of interest.

or counterfactual) states are defined for the individual unit, along with realizations of the outcome variable for each state. The causal model, known as the Potential Outcome Model, is as follows:

- Treatment indicator: D (=1 if participating; =0 if not participating). In general, $D = \{0,1,\dots,T\}$, most applications focus on the binary case, $D = \{0,1\}$, i.e. an exposure state and a non-exposure state.
- Outcome measure: Y .
- Observable characteristics (socio-demographic characteristics) that may influence treatment assignment and the outcome are captured in the Vector \mathbf{X} . Assumption: Covariates \mathbf{X} are predetermined, i.e. they are not influenced by participation.
- Individuals (units of observation): $i = 1,\dots,N$.

The potential outcome model in the binary case is therefore:

$$Y = Y_0 \quad \text{if } D = 0,$$

$$Y = Y_1 \quad \text{if } D = 1,$$

This illustrates the missing data problem that impact evaluation research faces (in the binary case):

	$D = 1$ (participants)	$D = 0$ (non-participants)
Y_1	Observable	Unobservable (Counterfactual)
Y_0	Unobservable (Counterfactual)	Observable

The model has several important implications. First, the causal effect at the individual level $\Delta = Y_1 - Y_0$ cannot be observed. Hence, we need an adequate parameter to summarize individual causal effects, and empirically impact evaluations will focus on the estimation of average causal effects.

Various types of averaging are possible in this regard, and the most commonly used evaluation parameters are the Average Treatment Effect (ATE) and, especially, the Average Treatment Effect on the Treated (ATET):

$$\text{ATE: } E(\Delta) = E(Y_1 - Y_0) = E(Y_1) - E(Y_0)$$

$$\text{ATET: } E(\Delta | D = 1) = E(Y_1 - Y_0 | D = 1) = E(Y_1 | D = 1) - E(Y_0 | D = 1)$$

where $E(\cdot)$ denotes the expectations operator, and the second term in the last equation is the counterfactual (= the average outcome of the participants in the case of non-participation). This counterfactual $E(Y_0 | D = 1)$ is, in econometric terms, not identified. A parameter is identified if it can be estimated with ever-increasing precision as the sample size increases. This, however, is the case for the first term $E(Y_1 | D = 1)$, which is identified from observable data and could be estimated with perfect precision if N were infinitely large.

The evaluation problem is therefore: what assumption allows the unobservable counterfactual average to be replaced by an alternative, identified population average? Such an identifying assumption, if correct, allows for construction of a counterfactual situation and identifies the population parameter. Note that, due to the unobservable nature of the counterfactual, every impact evaluation requires one or more identifying assumptions. These must be justified and made plausible by the researcher, since they cannot be statistically tested (at least not fully), hence cannot be right or wrong a priori nor proven right or wrong a posteriori.

This – simple and well-known – delineation of the causal model underlying impact evaluations is explained here, because it facilitates the description of the various methods that the component studies of this systematic review use. These methods are competently and comprehensively reviewed in several articles and books, such that a detailed explanation is beyond the scope of this protocol and, in fact, not required to explain the review's procedure. Important resources in this regard include, for instance, articles focusing predominantly on the econometrics of impact evaluation (e.g. Heckman et al., 1999, Imbens & Wooldridge, 2009) and several (hand-) books providing guidance for practitioners (e.g. Gertler et al., 2011, Khandker et al., 2010). It is worth noting that many of the empirical methods for causal analysis now commonly used in evaluation research have been developed explicitly for the purpose of evaluating training and other employment programmes (Imbens & Wooldridge, 2009).

The key virtue of formulating the problem of causal inference using Rubin's causal model is that it makes apparent the central role of the relationship between treatment assignment and potential outcomes. Moreover, this allows the set of available methods to be correspondingly classified into three groups, following Imbens and Wooldridge (2009):

1. Causal inference based on RCTs
2. Methods based on unconfoundedness (selection on observables)
3. Selection on unobservables.

We will consider these three categories and the methods they comprise in turn.

1. *Randomized experiments*: The most straightforward case for analysis occurs when assignment to treatment is randomized (in a controlled way by an experimenter) and, thus, independent of covariates \mathbf{X} as well as the potential outcomes Y . In such classical RCTs it is 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 US), with some increasing trend over the last decade, although still not on a very large scale. In recent years, RCTs have increasingly been used in development economics – analysing very diverse questions – and we expect to identify a fair number of experiments addressing the type of intervention that is of interest to this systematic review. Still, it is likely that the majority of relevant impact evaluations will not be based on RCTs.
2. *Methods for causal inference under unconfoundedness*: More common is the case in which researchers analyse data from non-experimental (also called “observational”) studies. Non-experimental data generally create 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 are fairly well understood (Imbens & Wooldridge, 2009). All these labels refer to some variant of the assumption that adjusting treatment and control groups for differences in observed covariates \mathbf{X} (i.e. pretreatment variables) removes all biases in comparisons between treated and control units (Imbens & Wooldridge, 2009). This case is of great practical relevance, with many impact evaluation studies relying on some form of this assumption: specifically, this category comprises classical **regression methods**, e.g. adjusting for covariates in a **linear regression**. Another method that is based on the unconfoundedness assumption and has been applied increasingly often is **statistical matching**, generating balanced samples in \mathbf{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 \mathbf{X} .
3. *Selection on unobservables*: Without unconfoundedness, there is no general approach to estimating treatment effects (Imbens & Wooldridge, 2009). Various methods have been proposed for special cases (see Imbens and Wooldridge, 2009) and three of them are important for our systematic review. One method is the **instrumental variables** 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 design (RD)** 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 control 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 control 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 control 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 control 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 control groups.

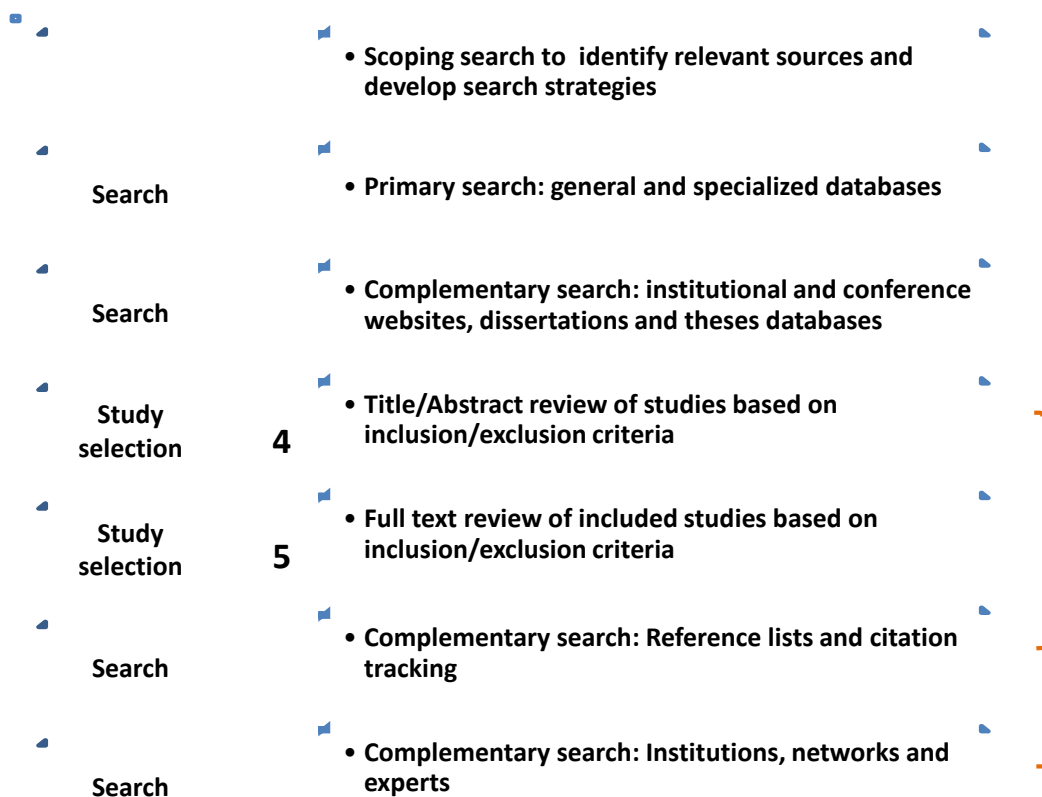
These three categories comprise the impact evaluation methods that are relevant for this systematic review. We note, however, that caution will be needed when synthesizing effect sizes from these different studies, due to potential problems relating to comparability of different treatment effect estimators (see also section 3.4) and partial versus bivariate effect sizes (Keef & Roberts, 2004; Becker and Wu, 2007). As noted below, we will therefore collect data on, and assess sensitivity of findings to, the treatment effect estimator and use of adjusted analysis.

3.2 Search strategy for identification of relevant studies

3.2.1 Method to find and select relevant studies

The search for relevant literature is based on a variety of sources in order to ensure that published and unpublished studies (“grey literature”) relevant to the review question are included in the search process. The search process includes (i) a primary search – searching of general and specialized databases – and (ii) a complementary search – hand-searching of relevant websites, literature snowballing and contacting experts. Through this comprehensive search process, the review seeks to arrive at a comprehensive and unbiased set of studies. To this end, country restrictions will not be applied to the search and selection process. The search will include search terms in English, Spanish, French, German and Portuguese, but there will not be any language restrictions in the selection process.

Figure 2: Search and selection process



Source: Authors, based on Hammerstrøm, Wade and Jørgensen, 2010.

In accordance with the review's focus on contemporary research, the selection process will focus on studies published since 1990. An appropriate filter to identify such studies will be applied to the search where possible. The screening and study classification process will be based on the inclusion and exclusion criteria (see section 3.1). The study search and selection process will follow the procedure defined in Figure 2.

3.2.1.1 Managing and documenting the search and selection process

The search and selection process will be conducted in parallel by at least three researchers. The review team uses EndNote to manage and document the process. The software allows decision tracking for each identified citation throughout the search. Bibliographic information of studies from electronic databases will be imported into EndNote as well as databases with compatible formats. The review team will explore using Zotero as an intermediary software when information cannot be easily exported into EndNote and, ultimately, will enter the details manually into EndNote if needed. Manual entry will also be executed on studies that cannot be excluded on title or abstract and are not already recorded. EndNote will support duplicates checks across different databases.

To enable transparency and reproducibility the review team will keep records of the search process. The search log includes the database, the database interface, the type of database, the customized search strategy, the language of search terms, the search string, the number of records obtained, the date of search and initials of the researcher. The overall search and screening process, including counts of studies included and excluded in the screening process, will be documented in a flow-diagram in the final report.

3.2.2 Data sources and search strategies

3.2.2.1 Scoping search

The search aims to arrive at a comprehensive and unbiased set of relevant studies. To this end, the review team systematically tested and screened 85 potentially relevant sources to identify pertinent sources and develop customized search strategies. The sources include general databases, specialized databases, institutional websites, conference websites, dissertations and theses databases and grey literature databases that are potentially relevant for the systematic review. A list of sources identified and used for the scoping search is presented in the Appendix.

For each source, the review team tested and documented several search strategies in the review's search log and identified one or more preferred search strategies that yielded a comprehensive and precise set of potentially relevant results for each source. Depending on the possibility of carrying out advanced searches, the review team applied the following search strategies to strike a balance between sensitivity and precision of results:

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 1990 onwards. Where available, the review team added relevant topics or the database's thesaurus terms for exposure, outcome and subject terms. The search terms for electronic databases are presented in the Appendix. Where applicable, search terms were truncated. The list of thesaurus terms was identified by screening the frequencies of thesaurus terms of 32 potentially relevant studies.¹² The standard search using Boolean operators included six test searches in title and abstract which were used to refine the search strategy for each database:

1. (Exposure terms OR thesaurus terms)
2. (Exposure terms OR thesaurus terms) AND (Outcome terms OR thesaurus terms)
3. (Exposure terms OR thesaurus terms) AND (Outcome terms OR thesaurus terms) AND (Subject terms OR thesaurus terms)
4. (Exposure terms OR thesaurus terms) AND (Outcome terms OR thesaurus terms) AND (Subject terms OR thesaurus terms) AND publication date: 1990–current
5. (Exposure terms OR thesaurus terms) AND (Outcome terms OR thesaurus terms) AND (Subject terms OR thesaurus terms) AND publication date: 2000–current
6. (Exposure terms OR thesaurus terms) AND (Outcome terms OR thesaurus terms) AND (Subject terms OR thesaurus terms) AND (Methodology terms OR thesaurus terms).

For websites or databases with basic search functions, the review team adjusted the search terms due to limited functionality of search functions. The preferred search strategies were based on keyword searches and/or topic/theme searches. The review team analysed the frequencies of 107 relevant exposure, outcome and subject terms in the title and abstract of 32 potentially relevant studies. A list of the most frequent exposure, outcome, and subject terms is shown below and was validated in two databases (Social Science Research Network (SSRN) and Innovations for Poverty Action (IPA) publications).

1. Exposure terms: training job
2. Outcome terms: employment labour
3. Subject terms: youth young adolescent student.

The team tested 20 keyword combinations of these search terms, each consisting of one exposure or outcome term combined with one subject term. Each combination was used to conduct a search, separately, for each of these sources. In addition, the review team identified

¹² The PIs of the systematic review selected 32 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.

relevant themes and topics to complement the keyword search. For databases/websites which do not allow the combination of keywords, separate keyword searches were conducted for the nine terms.

To determine the relevance of sources, the review team screened the titles of the first 50 results of each preferred search strategy. Depending on the number of potentially relevant studies included among the 50 results, the source was classified as being of low, medium or high relevance. The relevance of sources was considered low if, within the first 50 results, there was not a quantitative study related to a youth employment intervention.

This iterative scoping search was instrumental in identifying and developing customized search strings that yielded relevant results. However, the preferred search strategies for 85 databases and websites yielded more than 85,000 results. Based on the review of preferred search strategies during the scoping search, databases and websites will not be included in the final search strategy if the review team does not have access to the source (e.g. SocIndex), the source is considered of low relevance (e.g. African Economic Outlook), or the source is covered by another source (e.g. ILO working papers are included in Labordoc). The following sections present the final search strategy.

3.2.2.2 Primary search

The primary search of databases includes a wide range of general databases and those databases specialize in literature relevant to development economics and labour market issues. If additional relevant databases are identified through our search process, we will consider searching these as well. Based on the review during scoping search, relevant electronic bibliographic databases to be searched include those listed below.

3.2.2.2.1 General databases

- ABI/INFORM Global Cambridge Scientific Abstracts (CSA)¹³
- ASSIA (Applied Social Sciences Index and Abstracts) (CSA)
- EconLit (CSA)
- ERIC (Education Resources Information Centre) (CSA)
- IBSS (International Bibliography of the Social Sciences)
- JSTOR <http://www.jstor.org/>
- PAIS International (CSA)
- RePEc (Research Papers in Economics)/IDEAS Economics and Finance Research: <http://ideas.repec.org/>
- Social Science Citation Index (SSCI) of Web of Science
- Sociological Abstracts (CSA)

¹³ ABI stands for Abstracted Business Information.

- SSRN (Social Science Research Network) <http://www.ssrn.com/>

3.2.2.2 Specialized databases

- 3ie Database of systematic reviews <http://www.3ieimpact.org/evidence/systematic-reviews/>
- 3ie Registry for International Development Impact Evaluations (RIDIE) <http://ridie.3ieimpact.org/>
- 3ie Register of Impact Evaluation Published Studies (RIEPS) <http://www.3ieimpact.org/evidence/impact-evaluations/>
- Abdul Latif Jameel Poverty Action Lab (J-PAL) Evaluation and Publication Database <http://www.povertyactionlab.org/>
- British Library for Development Studies (BLDS) <http://blds.ids.ac.uk/>
- ELDIS <http://www.eldis.org/>
- Innovations for Poverty Action (IPA) Database <http://www.poverty-action.org/project-evaluations/search>
- JOLIS library catalogue – International Finance Corporation, International Monetary Fund and World Bank <http://jolis.worldbankimflib.org/e-nljolis.htm>
- Labordoc (ILO) <http://labordoc.ilo.org/>
- Research for Development <http://r4d.dfid.gov.uk/>
- World Bank Poverty Impact Evaluations Database <http://www1.worldbank.org/prem/poverty/ie/evaluationdb.htm>
- Youth Employment Inventory (YEI) <http://www.youth-employment-inventory.org/>

3.2.2.3 Search terms

The search terms reflect the inclusion criteria defined in section 3.1 and try to strike a balance between sensitivity (e.g. finding all articles in a topic area) and specificity (e.g. finding only relevant articles). For electronic databases with advanced search functions, we classify search terms according to three categories which will be combined by the Boolean operator “AND” to identify potentially relevant studies in each database. To ensure inclusion of papers which do not specifically report their research design in the title or abstract, the search excludes methodology terms. However, impact filters¹⁴ may be useful for sources such as OpenDOAR, which display only a limited number of results. Within each set, the search terms will be combined by the Boolean operator “OR”. These categories are

1. (ST-1) “exposure” terms – representing the interventions included in the review
2. (ST-2) “outcome” terms – including all outcomes for which we measure effect sizes
3. (ST-3) “subject” terms – reflecting the review’s focus on youth employment programmes.

The preferred search is 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,

¹⁴ The selection of impact terms is based on the 3ie Register of Impact Evaluation Published Studies (RIEPS) Protocol (Mishra and Cameron, 2013).

SSRN, IDEAS and Social Science Citation Index, will be searched for studies published since 1990 in order to include potentially relevant studies between 1990 and 2000. The search terms for electronic databases and an example for ERIC are presented in the Appendix. The search strategy will be modified according to the specifications of each database. Also, when possible, synonyms and wildcards will be applied as appropriate. To account for terminology differences across disciplines, database thesauri will be consulted to ensure that all appropriate synonyms have been included. The use of synonyms will also account for British or American English spelling. Where available, we will also rely 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 due to limited functionality of search functions. During the scoping search, 20 keyword combinations of search terms, each consisting of one exposure or outcome term combined with one subject term, had been tested. Due to the high number of duplicates between the search strings, the review team will only apply the most relevant search strings which yield a comprehensive and precise set of results. The search strings (employ* OR labour OR labor) AND (youth) appeared to yield the most relevant results. In addition, the review team identified relevant themes and topics to complement the keyword search. For databases/websites which do not allow the combination of keywords, separate keyword searches will be conducted for “employ*”, “labour”, “labor” or “youth”.

In general, the search process will be iterative and flexible to enable it to adapt over the search period. For example, we will include additional relevant subject headings identified in included studies throughout the search process. Similarly, we will browse the subject/thesaurus term indexes to work out which additional search terms might identify papers of interest. This will also be documented in the final report. The final review will report details of the modifications for each database, along with the results.

3.2.2.4 Complementary search

The primary search will be supplemented by a complementary search to include further studies for inclusion. The complementary search includes screening and hand-searching of relevant websites/gateways and conference websites, literature snowballing as well as contacting experts and relevant institutions.

3.2.2.4.1 Hand-searching of relevant websites/gateways

The search will be undertaken for relevant institutional websites and recent conferences as well as dissertation and theses databases. We believe that all relevant journals are covered by the primary search. We will therefore not hand-search websites of relevant individual journals unless they are not indexed. To identify studies that have not yet been indexed by the indexing tools, the review team will manually scan the table of contents in the most current issues of

journals in which a large number of included studies have been found.

Relevant institutional websites:

The hand-searching strategy will be customized for each website. Search terms will be used for websites that include a search facility. Otherwise, relevant sections (for example, “documents” or “publications”) will be searched.

- African Development Bank Evaluation Reports
<http://www.afdb.org/en/documents/evaluation-reports/>
- Asian Development Bank (ADB) Evaluation Resources
<http://www.adb.org/site/evaluation/resources>
- Bureau for Research and Economic Analysis of Development (BREAD)
<http://ipl.econ.duke.edu/bread>
- Campbell Collaboration Library: systematic reviews
<http://www.campbellcollaboration.org>
- Center for Economic Studies (CESifo) <http://www.cesifo-group.de/ifoHome.html>
- Centre for Economic Policy Research (CEPR) <http://www.cepr.org>
- ESRC (Economic and Social Research Council) <http://www.esrc.ac.uk/publications/>
- Global Development Network (GDN) <http://www.gdnet.org/index.html>
- Institute for Development Policy and Management (IDPM) at the University of Manchester <http://www.sed.manchester.ac.uk/idpm/>
- Institute for Fiscal Studies – Centre for the Evaluation of Development Policy
<http://www.ifs.org.uk/edepo/index.php>
- Institute for the Study of Labor (IZA) <http://www.iza.org>
- Institute of Development Studies (IDS) <http://www.ids.ac.uk/>
- Inter-American Development Bank Office of Evaluation and Oversight
<http://www.iadb.org/en/office-of-evaluation-and-oversight>
- Millennium Challenge Corporation (MCC)
<http://www.mcc.gov/pages/results/evaluations>
- National Bureau of Economic Research (NBER) <http://www.nber.org>
- Overseas Development Institute (ODI) <http://www.odi.org.uk/>
- Poverty and Economic Policy Research Network (PEP): Project List <http://www.pep-net.org/publications/pep-projects/>
- UNDP International Policy Centre for Inclusive Growth (IPC-IG) <http://www.ipc-undp.org/>
- United States Department of Labour, Employment and Training Administration, Research Publication Database <http://wdr.doleta.gov/research/>
- University of California Center for Effective Global Action (CEGA): Research Projects
<http://cega.berkeley.edu/research/>
- USAID Development Experience Clearinghouse <https://dec.usaid.gov/>
- World Bank Independent Evaluation Group (IEG) <http://ieg.worldbankgroup.org>
- World Bank Labor Markets <http://www.worldbank.org/labormarkets>

If the review team identifies further relevant institutional websites or databases, for example through the 3ie Register of Impact Evaluation Published Studies (RIEPS) Protocol (Mishra and Cameron, 2013), the review team will include these websites in the complementary search. The overlaps in specialized databases and relevant institutional websites as indicated in the Center for Global Development (2013) list of impact evaluation databases will be addressed by starting the search with the most comprehensive specialized databases, such as the 3ie Impact Evaluation Database and searching other specialized databases particularly for studies that have not yet been included. We expect that most studies included in NBER, CEPR, IZA and CESifo will be covered by IDEAS or SSRN. However, we will cross-check these websites for studies not covered by IDEAS or SSRN.

Conference websites:

Conference websites will be searched for possibly relevant studies. The systematic review team believes that the majority of possibly relevant ongoing and unpublished studies will be presented in one of the following conferences that were held in 2012 and 2013. Therefore, the review team will not search for conference abstracts or proceedings through specialist database sources. However, the list of conferences can be extended if the review team identifies other conferences that are deemed relevant for the research question throughout the search process.

- 7th IZA/World Bank Conference: Employment and Development, November 2012, http://www.iza.org/conference_files/worldb2012/
- 8th IZA/World Bank Conference on Employment and Development, August 2013, http://www.iza.org/conference_files/worldb2013/viewProgram?conf_id=2392
- IZA conferences and meetings in 2012 and 2013, http://www.iza.org/en/webcontent/events/izaconferences_html
- 2012 Global Youth Economic Opportunities Conference, September 2012, <http://www.youtheconomicopportunities.org/>
- 2013 Global Youth Economic Opportunities Conference, September 2013, <http://www.youtheconomicopportunities.org/>
- Ideas4Work Conference: Youth Employability and Entrepreneurship in Africa, January 2013, <http://iyflive.net/events/ideas-4-work>
- Arab Youth & Entrepreneurship: Holistic Approaches to Nurturing Local Ecosystems, February 2013, <http://aye2013.org/>
- Youth Entrepreneurship Knowledge sharing Symposium for East Africa, September 2013, http://www.ilo.org/addisababa/events-and-meetings/WCMS_221001/lang--en/index.htm
- Regional Youth Employment Consultation in Latin America , International Development Research Centre (IDRC) and Instituto de Pesquisa Econômica Aplicada (IPEA), December 2013
- Regional Impact Evaluation Workshop: Quality education, skills and productivity among youth in Africa, World Bank, September/October 2013
- Evidence Symposium: Increasing Youth Productivity in the Middle East and North Africa, ILO, March 2014,

http://www.ilo.org/empent/Eventsandmeetings/WCMS_234442/lang--en/index.htm

Dissertations and theses databases:

To include possibly relevant dissertations and theses that are not indexed in bibliographic databases, the review team will search national and international dissertation and theses databases. Search terms will be used for databases that include a search facility.

- Networked Digital Library of Theses and Dissertations (NDLTD) – SCIRUS ETD Search <http://www.scirus.com/srsapp/advanced/>
- ProQuest Dissertations and Theses Databases www.proquest.co.uk/en-UK/catalogs/databases/detail/pqdt.shtml
- Electronic Theses Online Service (EThOS) <http://ethos.bl.uk>

3.2.2.4.2 Reference lists and citation tracking

As outlined in Figure 2, the review team will conduct citation tracking. For this purpose it will rely on Google Scholar and IDEAS. Screening all studies which cited included studies will help to identify further studies for inclusion. The reference lists of included studies and relevant existing reviews and meta-analyses will be searched for studies. In addition, the review team will search the bibliographic information of the 2013 World Development Report (World Bank, 2014), the IEG report (IEG, 2013) and the J-PAL Youth Initiative review Paper (J-PAL, 2013). The bibliographic information contained within the reference lists will be hand-searched for potentially relevant studies. To our knowledge there is only one rigorous review of youth employment interventions that seeks to answer a similar question. The 2007 World Bank review by Betcherman et al. (2007) was based on the evidence gathered by the YEI, an online global repository of studies and information on labour market programmes for youth. Other reviews and meta-analyses such as Card et al. (2010), Cho and Honorati (2013) and Kluve (2010) will be screened for studies if the review team considers them relevant. The 3ie Database of systematic reviews and the Campbell Collaboration Library will be searched for relevant systematic reviews that can be screened for studies.

3.2.2.4.3 Institutions, networks and experts

The review team will contact authors of previous reviews or included studies, experts and focal points in relevant institutions, asking if they know of any studies (ongoing, published or unpublished) that might be relevant in addition to the studies that have been included after review of full reports.

Institutions and networks:

- Inter-Agency Network on Youth Development (IANYD) and UN Youth Flash Newsletter

- System Wide Action Plan on Youth (SWAP): Sub-working group on employment and entrepreneurship (including CEPAL, FAO, IFAD, ILO, ITU, UN-HABITAT, UNCDF, UNESCO, UNDP, UNIDO, UNEP, UNRWA, UNV, UN Women, UNWTO, WIPO and YEN)
- Global Partnership for Youth Employment (GPYE)
- Youth Employment Network (YEN) focal points in relevant multilateral institutions and their respective regional offices (such as IADB, FOMIN, OECD and World Bank)
- Focal points of the YEN Lead Country Network, a group of governments from developing and emerging countries that have voluntarily committed themselves to prioritizing youth employment on their national policy agendas
- YEN focal points in relevant national development agencies/ministries and their respective regional offices (such as GIZ, CIDA, SIDA, Federal Ministry for Economic Cooperation and Development, Germany)
- Key research institutions such as J-PAL, IZA, NBER

Advisory group of the systematic review:

- Kamilla Gurmede, former Director, J-PAL Africa
- Munshi Sulaiman, Postdoctoral Research Associate at Economic Growth Center, Yale University and former Research Coordinator for BRAC International Programme, Uganda
- Nathan Fiala, Assistant Professor, University of Connecticut, USA
- Luka Jonathan Mangset, Department of Enterprise Development and Promotion, Federal Ministry of Youth Development, Nigeria
- Roberto Suarez Santos, Deputy Secretary-General, International Organisation of Employers (IOE)
- Laura Brewer, Skills and Employability Specialist, International Labour Organization (ILO)
- Haroon Bhorat, Director of the Development Policy Research Unit, Professor of Economics at the University of Cape Town, South Africa
- Christian DaSilva, Canadian International Development Agency (CIDA)
- Helen Osborne, Director, Strategy and Performance, Youth Business International (YBI)
- Gerhard Ressel, Federal Ministry for Economic Cooperation and Development, Germany

Experts and focal points:

A list of experts and focal points is provided below:

- Gordon Betcherman, University of Ottawa
- David McKenzie, World Bank
- Mattias Lundberg, World Bank
- Furio Rosati, Understanding Children's Work (UCW) programme
- Robert Sauder, Canadian International Development Agency (CIDA)
- Emilie Milroy, Canadian International Development Agency (CIDA)
- Kristin Hausotter, Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ)
- David Rosas, Inter-American Development Bank
- Pablo Ibarrarán, Inter-American Development Bank

- Peter Wichmand, International Labour Organization (ILO)
- Mariel Sabra, Multilateral Investment Fund (FOMIN)
- Sandy Jambawai, African Development Bank
- Paul Gertler, University of California, Berkeley
- José Luis Polanco, Programs Coordination Unit, Ministry of Labour, Dominican Republic
- Carolina Robino, Programme Officer, International Development Research Centre (IDRC)
- Iris Boustros, Senior Programme Specialist, International Development Research Centre (IDRC)

3.2.2.5 Addressing potential biases

The review team will perform a comprehensive search to address potential biases, such as publication, time lag and language bias. The translation service offered by Google (<http://translate.google.com/>) will be used to translate information into English in cases where the title and/or abstract are not in English, Spanish, French, German or Portuguese. Keywords will be searched in internet search engines using Google Scholar to identify potentially relevant studies. During the scoping search, the keyword “employment” was used to test 85 sources on the availability of results in Spanish, French, German and Portuguese. For databases that yielded results, keyword searches in Spanish, French, German and Portuguese will be included in the search strategy. A list of keywords for search in English, Spanish, French, German and Portuguese is presented in the Appendix. While the review considers studies published after 1990, it is believed that most of the relevant studies that comply with the review’s inclusion criteria were published after 2000.

The review team expects to identify ongoing and unpublished studies in the grey literature 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 will be undertaken for the following grey literature databases:

- OpenGrey <http://www.opengrey.eu/>
- Directory of Open Access Repositories (OpenDOAR) <http://www.andoar.org/>

3.2.2.6 Study selection

The study selection will be piloted by four researchers who will screen the title, abstract and full report of a sample of reports independently and compare their results. Based on the review’s inclusion and exclusion criteria, discrepancies will be resolved by further review of the respective titles and abstracts. This process will be repeated until a high level of consistency in application of the selection criteria is achieved.

The remainder of the screening will then be carried out by individual reviewers (i.e. single

screening). However, in order to minimize bias, included and excluded results will be cross-checked by another researcher in EndNote. Discrepancies will be resolved by the two reviewers. For sources that can be imported to EndNote, the review team will screen title/abstract and full report of studies in EndNote after removing duplicates. In general, results will be cross-checked by another researcher. However, titles will be single-screened with quality checks on a random sample of studies.

For sources that do not allow us to export bibliographic information, we will screen title/abstract online and manually import potentially relevant results into EndNote if they are not already recorded. To this end, the review team will first screen studies that can be imported into EndNote and databases which are deemed particularly relevant. Another researcher will double-screen potentially relevant results from websites and all results of those websites which are deemed particularly relevant, such as the 3ie Impact Evaluation database, the World Bank Poverty Impact Evaluation database and the Youth Employment Inventory. The translation service offered by Google (<http://translate.google.com/>) will be used to translate information into English in cases where the title and/or abstract are not in English, Spanish, French, German or Portuguese. Within each step of the selection procedure (see Figure 2), researchers will work through exclusion criteria hierarchically, as defined in section 3.1 and the screening questionnaire provided in the Appendix.

- If the exclusion criteria are met, the study will be dropped from further screening but a record of this decision will be kept in the reference management software EndNote.
- If the study is “potentially relevant” according to the inclusion criteria, the study proceeds to the next round of screening.
- In case the title and/or abstract of a study do not provide enough information, or when the research is in doubt, the respective selection criteria will be classified as “unsure” and the study will proceed to the next round of screening.

If more than one study of the same programme meets the inclusion criteria or an updated version of the study is available, one of the principal investigators will determine which of them will be included (see section 3.4). In the event that full-text versions do not allow for a final decision, disputes will be resolved by one of the principal investigators. The involvement of research institutions such as the World Bank and the International Labour Organization in this systematic review ensures that full-text access to all publications found relevant in the first two steps of the search process should be possible. In the event that the full text of studies which have been deemed relevant during the selection process is not available or does not provide sufficient information to include or exclude the study, we will contact the authors to ask for English-language versions or further information about the impact evaluation. Screening against the selection criteria will then proceed as normal.

3.3 Description of methods used in the component studies

The review will collect component studies that conduct impact evaluations of youth interventions. The objective of an impact evaluation is to estimate quantitatively the causal effect of the intervention on the outcome it intends to influence. Modern evaluation research has come to utilize a counterfactual concept of causality, which in several steps of methodological development over the past decades has taken on the shape in which it is used today (Holland, 1986). This causal model defines the causal effect of a treatment as the difference between the factual outcome (“Of the 100 training participants x per cent found a job”) and the counterfactual case (“What percentage of the same 100 training participants would have found a job without the programme?”). Clearly, the counterfactual is a hypothetical construct and can never be observed in data, since no individual or group can be both exposed and not exposed to the intervention simultaneously. Holland (1986) refers to this as the fundamental problem of causal inference. In order to evaluate the effect of the treatment, we therefore always need to compare distinct units receiving the different levels of the treatment. Such a comparison can involve different physical units or the same physical unit at different times (Imbens and Wooldridge, 2009).

As indicated in section 3.1.5, the most commonly used evaluation parameter is the ATE. However, it should be noted that some evaluation studies do not estimate ATE but intention-to-treat parameters (ITT). The ITT case does not look at actual exposure to the treatment but generates treatment and control groups according to the initial assignment (sometimes a practicability issue, sometimes a design decision). The logic of the counterfactual causal model, however, is not affected by whether ITT or ATE parameters are estimated. Given the increased weight of ITT for policy recommendations (Bloom, 2006), this review will incorporate studies with ATE, ATET and ITT parameters. In the absence of ITTs, the review team will explore ways of computing them.

Examples of eligible studies:

Pablo Ibararán, Laura Ripani, Bibiana Taboada, Juan Villa, and Brigida García (2012): Life Skills, Employability and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design	Programme under study: Juventud y Empleo in the Dominican Republic. Random assignment (lottery) was applied on a group of potential participants identified by selected private training institutions. Individuals had previously applied to the programme and met the eligibility criteria. The study provides ITT estimators.
Oriana Bandiera, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim	Programme under study: Empowerment and Livelihood for Adolescents (ELA) programme. Uganda.

<p>Gulesci, Imran Rasul and Munshi Sulaiman (2012): Empowering Adolescent Girls: Evidence from a Randomized Control Trial in Uganda</p>	<p>Randomization as well as programme activities were conducted at village level, offering some leverage to reduce and control for the exposure of non-participants to the programme. All adolescent girls within the respective treatment villages (100 villages) were eligible to participate in the programme. The study estimates ITT and ATE impacts of the programme.</p>
<p>Patrick Premand, Stefanie Brodmann, Rita Almeida, Rebekka Grun, Mahdi Barouni (2012): Entrepreneurship training and self-employment among university graduates: Evidence from a randomized trial in Tunisia</p>	<p>The study runs an experiment comparing outcomes of a business plan programme against the standard curriculum for undergraduates in the final year of licence appliquée. Tunisia.</p> <p>Taking advantage of oversubscription, half of the applicants were randomly assigned to the entrepreneurship track and the other half were assigned to continue with the standard curriculum. Randomized assignment was conducted at the project level, stratified by gender and by the subject that students were reading. The study presents ITT and treatment-on-the-treated (TOT) estimates.</p>
<p>Orazio Attanasio, Adriana Kugler and Costas Meghir (2011): Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial</p>	<p>Programme under study: Jóvenes en Acción. Colombia.</p> <p>Programme assignment is randomized based on oversubscription captured by selected training institutions in the seven largest cities throughout the country. The study presents ITT parameters across different courses and training centres.</p>

3.4 Criteria for determination of independent findings

For the purpose of clarity, we set forth the following terminology for the remainder of the protocol:

A youth employment programme (e.g. “Jóvenes En Acción” in Colombia) may consist of several different interventions (e.g. training or job-search assistance). These interventions may be evaluated by more than one study (e.g. evaluation), which may each publish multiple reports (e.g. working papers, journal publications). We treat two reports as part of the same study if they are based on the same data and hence cannot be treated as independent, even if they are written by different authors. Each outcome category (e.g. employment outcome) may be measured by different outcome constructs (e.g. employment probability, unemployment probability). Each report may present different treatment effect estimates for the same outcome category— for example for different subgroups, different statistical methods, or outcome

constructs. These treatment effect estimates are converted into effect sizes (e.g. standardized mean differences (SMDs)) for analysis.

Hence, an intervention population (all participants) may be different from the study population (all in dataset), which may itself be different from the sample population for a specific treatment effect estimate on a specific outcome.

Furthermore, we use the following terminology to refer to different treatment effect estimates often reported in primary studies of this review: relative effect sizes are treatment effects measured against an alternative treatment (in contrast to those measured against the non-exposure state); conditional effect sizes are effects measured conditional on some other outcome; and, partial effect sizes are computed from multivariate models, controlling for covariates which are believed to affect the outcome.

In a meta-analysis, the unit of analysis is the study. To maintain the independence assumption, it is important that one effect size per outcome-construct and study (and for each subgroup, if reported) is included in the analysis (Borenstein, Cooper, Hedges & Valentine, 2009). This implies that different estimates within each study have to be combined into one effect size per subgroup. There are various ways in which individual effect size estimates may be correlated. For example, studies may analyse different subgroups of the treatment population, outcomes at different times but on the same units, multiple treatments or variations of the above.

We will pursue the following approaches:

Estimated treatment effects may be regarded as independent from each other when the underlying data are derived from different sample populations. This is the case if a study reports treatment effects for, say, different subgroups but not an overall estimate. We also regard this to be the case if we have multiple treatment effects from studies (which use different data samples) for the same programme. In such cases, we create a (within-study) summary effect size by estimating a random effects (RE) model prior to the pooled analysis across studies. The specification of the random effects model is discussed in section 3.6.4.2 of this protocol (based on Borenstein et al., 2009, equations 12.2–12.8).

In contrast, treatment effects cannot be regarded as independent of each other when multiple effects are estimated for the same sample of participants (that is, some participants contribute information to more than one effect estimate). This may arise if, for example, a study reports treatment effect estimates on:

- multiple outcome constructs for the same type of outcome (e.g. hourly wages and annual income as part of earning outcomes)

- multiple time-points for the same individual (e.g. repeated observations for several follow-up periods)
- multiple treatment groups and the same control group.

In such cases, we will create a synthetic effect size to obtain one (composite) effect size per study and outcome construct. To arrive at an unbiased effect size, we compute a sample-weighted average across dependent effect sizes per study using the formulas provided in chapter 24 of Borenstein et al. (2009). In order to account for differences in sample sizes, effect sizes will be averaged by using an inverse variance weighting of the individual effect size. Hence, effect sizes based on larger studies are given more weight in the combined effect size. To compute its variance (equation 24.5), we assume a correlation coefficient r equal to 1 (between effect sizes being combined).¹⁵

In the event that one study uses different methods to estimate treatment effects for the same outcome (and subgroup), the estimate with the lowest risk of bias will be used as the main record.¹⁶ This will also be the case, if several reports exist, using the same set of data but, for example, different estimation methods. If a study reports several types of treatment effect estimates (such as ATE, ATET, LATE or ITT, see section 3.3), we will code all of them and separate them in the final analysis. As the ITT estimate is usually regarded as the more policy-relevant parameter (Bloom, 2006), we will generally base our results on the later. If possible, we will compute ITT from ATE estimates based on information reported on compliance rates using the formula provided in Bloom (2006).

In some cases it may not be immediately apparent which estimate has the lowest risk of bias (RoB, see section 3.6)) (e.g. nearest neighborhood vs. kernel matching). If a report included in the review used several estimation methods within the same study which cannot be ranked by our RoB assessment, we will choose between different statistical methodologies following a procedure similar to that outlined in Tripney et al. (2013, section 2.5, p. 29ff), who draw from available practical guidance on this topic. In the event that the first and second coder cannot come to an agreement about the estimate with the lowest RoB, the decision will be taken by one of the PIs.

Finally, in cases where a single report presents effect sizes for more than one study (e.g. different programmes/datasets), the respective effect sizes will be coded as if they had come from separate reports.

¹⁵ Using this assumption about the covariance between correlated groups, we follow Baird, Ferreira, Özler and Woolcock (2013), arguing that it yields the most conservative composite effect size estimate. The assumption implies that the existence of multiple estimates for one outcome provides no improvement in precision, but only alters the effect size.

¹⁶ This will also be the case if one study publishes several reports with different estimation methods.

If it is unclear whether multiple reports provide independent findings, we will contact the authors for clarification. The decision will be recorded as part of the study selection process.

Once effect sizes have been combined at the study level, each (composite) effect size can be considered independent and analysed using meta-analysis methods as discussed in the following section.

3.5 Details of study coding categories

For our analysis, eligible studies will be coded on variables related to study methods, the nature of the intervention and its implementation, the characteristics of the subject samples, the outcome variables and statistical findings, and contextual features. The following describes some of the study level variables that will be coded for each of these types of characteristics.

- **Study methods:** impact evaluation research design (RCT, natural experiment, etc.), statistical methodology (instrumental variables, difference-in-difference, etc.), risk of bias, among others.
- **Nature of intervention and its implementation:** main category of intervention (skills training, entrepreneurship services, etc.), average duration single cohort stays in programme, scale of programme, among others.
- **Characteristics of the subject samples of analysis:** age, gender, status of their occupation (informal, formal, etc.) among others.
- **Contextual features:** setting, year and type of publication, etc.

In order to assess the risk of bias of primary studies, we are adapting the framework proposed in Duvendack et al. (2011, 2013) and utilised recently in Vaessen et al. (2014). The approach combines an assessment of both research design and the method of statistical analysis. By placing RCTs at one end of the spectrum and cross-section designs at the other, the tool aims to reflect the potential capacity of these statistical methods to control for the potential biases of quasi-experimental designs.¹⁷

We will assess risk of bias according to the design-based approach.

We will use these variables in sensitivity and moderator analyses. Variables that could be used as potential effect size moderators. These include: age groups, gender, level of education,

¹⁷ As the authors indicate, this framework should not be taken to endorse a universal 'hierarchy of methods' (White 2009); but rather to provide an objective and efficient framework to assess the potential risk of bias of randomized and quasi-randomized studies.

occupation categories and geographic location. In addition, a potential study level moderator that will be coded is the set of methodological characteristics used (see study methods above).

As mentioned in section 3.1.4, outcomes are grouped into three main categories: employment, earnings and business performance outcomes. The study outcomes for employment that will be extracted from the selected studies include employment probability, hours worked and unemployment duration, among others. The study outcomes for earnings are: income, consumption and salary/wage, among others. Those for business performance are: profits, sales, capital and investment, among others.

Specifically, the coding sheet is structured as follows. Variables related to the study and programme identification, date and type of publication are in the section labelled *Identifiers*. The next section presents the data used and study methods. The intervention characteristics are given in the section *Intervention category*. The following sections of study level information are related to the programme characteristics. As elaborated in section 1.3, we restrict our analysis to four types of programmes: skills training, entrepreneurship promotion, employment services and subsidized employment. The final section for study level information contains variables that provide a general description of the overall programme. It encompasses general information on the dates of the programme's start and end, the country in which the programme is being/has been implemented, its scale, the financing and implementing agencies and the involvement of private sector and civil society in the design and implementation of the programme.

Listed after the variables of study level information, we provide a section detailing “effect size level” variables, which contains information on the set of outcome variables mentioned above as well as summary statistics of the target group (such as age, gender, educational attainment and income status). In addition, it includes all reported effect size statistics that will be coded (number and mean of controls and treatments, regression coefficients, statistics for test of significance and the date of the effect size).

A comprehensive list of information (i.e. variables) that will be extracted (both study level and “effect size level” variables) from the studies is included in a summary of the coding tool in the Appendix.

The studies are being double-coded manually into a spreadsheet by a team of six coders. Two drafts of the coding tool were tested by coders based on 33 studies of impact evaluations of youth employment programmes. The coding tool has been updated as necessary based on additional feedback from the coding process as well as in response to comments and suggestions received from 3ie as well as during the Campbell Colloquium and 3ie's conference in Manila. Coding rules and definitions have been specified in a coding manual.

3.6 Statistical procedures and conventions

3.6.1 Standardizing effect sizes

Since the measurement of effects differs widely across studies, outcome measures need to be standardized for comparability. According to Cho and Honorati (2013, p. 15) “the effect size provides a measure of ‘how well did it work’ without confounding the effectiveness of a program with the sample size used in a study” and hence provides “a more reliable, comparable measure of program success”. The quantitative meta-analysis will be based on all studies for which the relevant information to compute comparable effect sizes can be gathered.

We will focus on calculating standardized mean differences (Hedges g)¹⁸ for both continuous outcome variables (e.g. income) and dichotomous outcome variables (e.g. employment probability). The standardized mean differences (SMD) represent the difference between the means in treated and control groups (the treatment effect) as measured relative to the variability of the respective outcome observed in that study. Although risk ratios or odds ratios are sometimes regarded as more appropriate for synthesizing dichotomous outcomes, testing the coding sheet on a prior set of studies showed that most studies do not report sufficient information. Hence, using odds ratios in the analysis would be likely to lead to a substantial loss of information (cf. Tripney et al., 2013).

From an initial survey of studies likely to be included in the review, it became obvious that most studies report either matching- or regression-based estimates of the treatment effect (even for RCT-based designs).¹⁹ Hence, we expect that SMDs in most cases will be computed using the formulas given in Borenstein et al. (2009) and Keef and Roberts (2004; cited in Waddington et al., 2012, p. 372f):

For studies using parallel group or matching-based strategies g and its standard error SE_g are computed as $g = \frac{\bar{Y}_t - \bar{Y}_c}{S_p} * \left[1 - \frac{3}{4 * (n_t + n_c - 2) - 1} \right]$ $SE_g = \sqrt{\left[\frac{n_t + n_c}{n_c * n_t} + \frac{g^2}{2 * (n_c + n_t)} \right]}$.

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 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 control group (Hedges’

¹⁸ Hedges g equals Cohen’s d but applies an adjustment for small sample sizes (Borenstein et al., 2009).

¹⁹ This is in line with experiences documented by previous systematic reviews in related fields, such as Baird et al. (2013) or Tripney et al. (2013).

approach). If neither S_p or S_t and S_c is available, we will try to derive an adequate estimate for the standard deviation based on assumptions (e.g. equivalence of the standard deviation in both groups). Similarly, we will attempt to impute missing information on sample. Such procedures will be documented in the final report.

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

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

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.²⁰

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 is rarely reported, we follow Cohen's approach and compute $\hat{\sigma}$ from the standard deviation of **the dependent variable across all observations** (S_t) (cf. Lipsey and Wilson, 2001): $\hat{\sigma} = \sqrt{\frac{S_t * (n_t - 1) - (\beta^2 * (n_c * n_t) / (n_c + n_t))}{n_t - 1}}$

In the event that the necessary information for calculating SE_g is not available,

we will approximate it by $SE_g = \frac{g}{t}$

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

For some studies, we will have to transform reported statistics (often t , F , p or z -values) prior to calculating effect sizes. We will transform this data following the procedures suggested in Lipsey and Wilson (2001).

For studies that report sufficient data on post-intervention treatment and control group outcomes (e.g. means, sample sizes and information from which to compute the pooled standard deviation) for the (matched) samples of participants and non-participants are available, effect sizes will be computed using the Lipsey/Wilson Effect Size Calculator.²¹ Since the Lipsey/Wilson Effect Size Calculator does not correct for potential biases in small-sample studies, we will convert these into Hedges' g using the adjustment factor suggested in Lipsey and Wilson (2001).

All effect sizes will be computed such that positive effect sizes represent positive outcomes (i.e.

²⁰ For studies with large n , $c(v)$ is considered equal to 1.

²¹ Effect Size Calculator available at:

http://www.campbellcollaboration.org/resources/effect_size_input.php [13 Oct. 2014].

the intervention increases wages, reduces unemployment duration). All effect sizes will be coded along with original values of effect sizes and significance tests. The latter should enable us to replicate and verify the calculation of SMDs as well as to check for any systematic relationship between reported statistical measures and effect sizes.

3.6.2 Unit of analysis error

We will also correct the standard errors for a possible unit of analysis error. 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. use cluster robust standard errors, variance components analysis).²² 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.

If we suspect that studies may be biased by a unit of analysis error, we will adjust standard errors (SE_g) by the formulae suggested in Higgins and Green (2011, p. 16.12ff):

$$SE_{corrected} = SE_{uncorrected} \times \sqrt{1 + (m - 1) \times ICC}$$

where m is the number of observations per cluster and ICC is the intra-cluster correlation coefficient, which is an estimate of the relative variability within and between clusters (Waddington et al., 2012, p. 370). This formula corrects for the fact that, due to dependence across observations, the effective sample size is smaller than the total sample size.

As the information necessary to apply the correction formula is often not available, we may have to resort to other methods to account for the unit of analysis problems (as described in Higgins and Green, 2011, p. 16.14). We will also conduct sensitivity analyses to investigate the robustness of the conclusions. The final report will identify any cluster-randomized trials in the review and state how the potential unit of analysis error has been dealt with.

3.6.3 Treatment of missing information

A first screening of potential studies for the meta-analysis has shown that many do not report sufficient data to compute effect sizes and their standard errors. Most often, we find that the post-intervention mean and/or standard deviation of the outcome variable is not reported. In such cases, we will either try to impute the data from available information or we anticipate having to apply certain assumptions to calculate effect sizes. In this event, we plan to allow assumptions similar to those applied in a previous Campbell-registered systematic review by

²² For example, when the intervention is delivered at a cluster level (e.g. village or household) but the analysis of impact is carried out at the individual level (Waddington et al., 2012).

Baird et al. (2013). The exact assumptions needed to compute effect sizes will be documented in the final report.

Where a study's documentation of effect sizes does not even allow these computations to be performed, we will contact the authors to request further information. If this approach is not successful, the study will be excluded from the meta-analysis on the basis of insufficient data. For these studies, we will present the available information and discuss their findings narratively. This process and the potential effect it may have on the review's outcome will be documented and discussed in the final report.

If we find that the necessary information to calculate the standardized mean difference (g) cannot be obtained for a large set of studies, we will try to compute response ratios (for continuous and binary outcomes) where possible. Response ratios are computed as the difference in the outcome between the treatment and control group relative to initial outcome in the control group (cf. Lipsey and Wilson, 2001). Where possible, we will approximate the standard error of the response ratio using the t statistics/ p -value of the regression coefficient or of the results of the t test for equality of means between groups after matching. We will compare results with those where computation of SMDs was possible as part of the sensitivity analysis. If applied, the respective methods will be detailed in the final report.

Some reports may not detail all intervention-specific variables that we aim to include as moderator variables into our meta-analytic model. A subsequent step in the research process – subject to further discussion – is therefore to check in which cases it is possible to supplement the missing programme information. First, we will gather information from working paper versions of the respective studies, since these tend to be more detailed than published articles. Second, we will try to collect the information from other literature on the respective programme and during the same time period. This may comprise other impact evaluations, process evaluations or programme documentation.

All these steps will be documented throughout the research process and will later be used to check whether our results are sensitive to the inclusion of data from external sources.

3.6.4 Data synthesis

Once the coding process has been concluded, the extracted data will be imported into Stata to conduct the statistical analysis.

3.6.4.1 Descriptive analysis of studies and interventions

The analysis will begin with a descriptive examination of all project studies that are included in the sample based on the study selection process (see section 3.1). This section will aim to present a conclusive picture of the existing literature on youth employment interventions. The

descriptive analysis will report the distribution of our sample by programme characteristics (e.g. intervention type, regional allocation, timing and population) as well as by study characteristics (e.g. publication date, methodology, outcome measure). It will be based mainly on descriptive summary tables.

3.6.4.2 Meta-analysis

The analysis will subsequently examine to what degree the collected data allow for a quantitative meta-analysis using a random effects statistical model. We assume that data will at least be sufficient to conduct a pooled meta-analysis within each of the three outcome categories (e.g. employment, earning and business).

As a first step of the meta-analysis, we will investigate the computed effect sizes and present summary tables and graphical overviews (such as forest plots) alongside with the 95 per cent confidence intervals. This will be done using the “metan” command in Stata. To account for differences in sample sizes for individual studies, we will synthesize effect sizes across studies using inverse-variance weighting of the individual effect sizes. This is recommended in settings which present significant contextual heterogeneity in terms of study population, intervention and implementation. If the number of observations allows, we will use the technique developed by Hedges, Tipton & Johnson (2010) to estimate robust standard errors in the meta-analysis model in order to account for statistical dependencies (in addition to the methods discussed in section 3.4).

The data synthesis will separate studies with different kinds of counterfactuals (i.e. “treatment vs. no treatment” studies are analysed separately from “treatment vs. alternative treatment” studies). If we find enough studies that assess the relative effectiveness of different programmes vis-à-vis each other (i.e. training alone vs. training and job-search assistance combined), we will consider estimating a network meta-analysis model as described in Grant and Calderbank-Batista (2013). If the number of such effect sizes is not sufficient, we will descriptively present major findings on the relative effectiveness of youth employment interventions, along with information on the respective methods.

3.6.4.3 Assessment of heterogeneity

This research project analyses a broad range of heterogeneous interventions, outcome measures and study designs. Due to the strong heterogeneity of methodological designs and interventions, the legitimacy of estimating a pooled model within outcome categories is difficult to foresee prior to our study selection process. We will investigate heterogeneity within outcome categories by (i) types of outcomes and (ii) by intervention category (if possible). To test for overly strong heterogeneity across studies we employ I^2 statistics, Q -statistics, as well as the τ^2 statistic. These

statistics test whether the percentage of variability in effect estimates may be due to heterogeneity rather than by chance. A significant Q ($P < .05$) and an I^2 value of at least 50 per cent will be considered as indicators of heterogeneity.

Following guidance from the Campbell Collaboration (2014, p.9), we will present findings for effect sizes separately by study design (randomized vs. quasi-experiments). Where we find no empirical evidence for differences in effects, we will pool estimates. In cases where we detect substantial heterogeneity within outcome categories, we will examine potential sources by stratifying the meta-analysis further along different dimensions. To avoid “comparing the incomparable” we expect to stratify data synthesis at least by intervention categories as defined in section 3.5. If sample sizes allow, we will stratify the analysis further by subgroups (e.g. by gender or age). Also, we aim to analyse the change in effect size by time elapsed after programme completion.²³

In the event that we have enough data points by outcome category, we will attempt to analyse the factors explaining heterogeneity by moderator analysis using a meta-regression model. According to Borenstein et al. (2009) a minimum of ten studies of sufficient quality are required to test the effect of individual moderator variables. The analysis will be carried out using a random effects meta-regression model, as described in Higgins and Green (2011). For this, we expect to employ the “metareg” command in Stata.

As discussed in section 3.5, we anticipate coding a broad set of variables related to intervention and study design as well as contextual factors to be used as potential moderator variables. This will permit us to analyse which characteristics are associated with larger and smaller treatment effects. The final set of moderator variables will depend on the available data. The exact specification of the random effects meta-regression model will be documented and discussed in the final report.

If a moderator analysis is not possible, we will discuss and explore the factors which may be driving the heterogeneity of results narratively.

3.6.4.4 Sensitivity analysis and detection of biases

We will conduct a range of sensitivity checks to test the robustness of our results. Sensitivity analysis will be carried out by restricting the meta-analysis to a subset of all studies included in the original meta-analysis. Most importantly, the sensitivity analysis will test the impact of our review decisions on the report (such as the chosen methods for standardizing effect sizes, the relevance of missing data or the decisions regarding the independency of the findings). Further,

²³ As part of this, we will construct ranges of follow-up duration (e.g. short term 1–3 months, long term 9–12 months, etc.) and aim to synthesize effect sizes at these different follow-up periods.

we will test for the influence of methodological designs, econometric specifications and data sources. We will also test the potential of exclusion of low-quality studies from the synthesis and the exclusion of statistical outliers. We will also explore heterogeneity in studies which use bivariate and multivariate analysis (i.e. bivariate versus partial effect sizes).

Finally, we will assess and discuss the danger of publication bias in our sample of included studies. First, by testing the effect of publication status in our meta-regression model; that is, we will include the standard error of the effect size as a covariate in meta-regression (Egger's test, Egger, 1997) in order to test statistically for publication bias. Second, by plotting the effect size against standard errors (funnel plots). We will do this analysis using the "metafunnel" and "metacum" commands in Stata.

3.7 Treatment of qualitative research

The review will keep records of additional research papers, documents, and other information on a particular youth employment intervention including qualitative research. Supplementary information will support the review team in (i) capturing programme details and contextual variables that may not be recorded in the impact evaluation studies, and accordingly (ii) acquiring a better understanding of the underlying theory of change or the causality behind intervention types and their combined delivery.

4 ACKNOWLEDGEMENTS

We would like to thank Hugh Waddington, John Eysers, Martina Vojtkova, Janice Tripney, Michael Grimm, Anna Luisa Paffhausen, Kristin Hausotter, and participants at the Campbell Collaboration Group Colloquium in Belfast as well as at the ADB/3ie Making Evaluation Matter Conference for their support and comments on the review's methodology. We would like to thank Viviana Perego, Selahattin Selsah Pasali, Karishma Tiwari and Cheng Qian for their support and contribution to the search and coding process. Thanks as well to Sergio Graziosi and Jeff Brunton at the EPPI-Centre for their support in testing EPPI reviewer. We are grateful for feedback on the study methodology, which was provided during the workshop organized by the Evaluation unit (EVAL) of the International Labour Organization (ILO) in Geneva in 2013. The financial support of the Canadian International Development Agency (CIDA), the World Bank, the Youth Employment Network and the ILO is greatly appreciated.

5 REFERENCES

- 3ie. (2013). Impact Evaluation Glossary. Available at: <http://ridie.3ieimpact.org/downloads/Impact%20Evaluation%20Glossary.pdf> [13 Oct. 2014].
- 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.
- Becker, B.J., Wu, M.-J. (2007). The synthesis of regression slopes in meta-analysis. *Statistical Science*, 22, 3, 414–429. DOI:10.1214/07-STS 243
- 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.
- Blundell, R., Costa Dias, M., Meghir, C. & Van Reenen, J. (2004). Evaluating the employment impact of a mandatory job search assistance program. *Journal of the European Economic Association*, 2, 4, 569–606.
- Bonnal, L., Fougere, D., and Sérandon, A. (1997). Evaluating the impact of French employment policies on individual labour market histories. *The review of Economic Studies*, 64, 4, 683–713.
- Boone, J. & van Ours, J.C. (2004). *Effective Active Labour Market Policies*. IZA Discussion Paper No. 1335.
- 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.

Brodaty, T., Crepon, B. & Fougere, 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 Systematic Reviews: Policies and Guidelines Campbell Systematic Reviews 2014: Supplement 1*. DOI: 10.4073/csrs.2014.1 Available at http://www.campbellcollaboration.org/resources/research/the_production.php [14 Oct. 2014].

Card, D., Kluve, J. & Weber, A. (2010). Active labour market policy evaluations: A meta-analysis. *The Economic Journal*, 120, 548, F452–F477.

Card, D., Ibararán, P., Regalia, F., Rosas, D. & Soares, Y. (2011). The labor market impacts of youth training in the Dominican Republic. *Journal of Labor Economics*, 29, 2, 267–300.

Center for Global Development (2013). Closing the evaluation gap: List of impact evaluation databases (updated Nov. 2013). Available at <http://international.cgdev.org/page/list-impact-evaluation-databases> [13 Oct. 2014].

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.

Crépon, B., Duflo, E., Gurgand, M., Rathelot, R. & Zamora, P. (2013). Do labor market policies have displacement effects? Evidence from a clustered randomized experiment. *The Quarterly Journal of Economics*, 128, 2, 531–580.

Cunningham, W., Sanchez-Puerta, M. L. & Wuerml, A. (2010). *Active labor market programs for youth: A framework to guide youth employment interventions*. World Bank Employment Policy Primer No. 16. 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.

Dorsett, R. (2006) The new deal for young people: Effect on the labour market status of young men. *Labour Economics*, 13, 405–422.

Egger, M., Davey Smith, G., Schneider, M. and Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *British Medical Journal*, 315, 629–634.

Fares, J. & Puerto, O. S. (2009). *Towards comprehensive training*. Social Protection Discussion Paper No. 0924. Washington, DC, World Bank.

Fiala, N. (2014). *Stimulating microenterprise growth: Results from a loans, grants and training experiment in Uganda* (25 Apr., 2014). Available at: <http://www.nathanfiala.com/Stimulating%20Microenterprise%20Growth.pdf>. [25 Apr. 2014]

- Galasso, E., Ravallion, M. & Salvia, A. (2001). *Assisting the transition from workfare to work: Argentina's proempleo experiment*. Washington, DC, World Bank.
- Gertler, P. J., Martinez, S., Premand, P., Rawlings, L. B. & Vermeersch, C. M. (2011). *Impact evaluation in practice*. 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.
- Gonzalez, L., Piza, C., Taylor, L., Cravo, T.A. & Abdelnour, S. (2013). *The impacts of business support services for small and medium enterprises on firm performance in low-and-middle-income countries: A systematic review*. The Campbell Collaboration. Available at: <http://campbellcollaboration.org/lib/project/284/> [14 Oct. 2014].
- Grant, E. S. & Calderbank-Batista, T. (2013). Network meta-analysis for complex social interventions: Problems and potential. *Journal of the Society for Social Work and Social Research*, 4, 406–420.
- Grimm, M. & Paffhausen, A. L. (2013). *Interventions for employment creation in micro, small and medium-sized enterprises in low- and middle-income countries*. The Campbell Collaboration. Available at: <http://campbellcollaboration.org/lib/project/285/> [14 Oct. 2014].
- Groh, M., Krishnan N., McKenzie, D. & Vishwanath, T. (2012). *Soft skills or hard cash? The impact of training and wage subsidy programs on female youth employment in Jordan*. World Bank, Washington, DC. Available at: <http://elibrary.worldbank.org/doi/book/10.1596/1813-9450-6141>.
- 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.
- Hammerstrøm, K., Wade, A. & Jørgensen, A.-M. K. (2010). *Searching for studies: A guide to information retrieval for Campbell Systematic Reviews*. Campbell Systematic Reviews 2010, Supplement 1. The Campbell Collaboration. Available at: <http://www.campbellcollaboration.org/lib/download/969/> [14 Oct. 2014].
- Heckman, J. J., LaLonde, R. J. & Smith, J. A. (1999). The economics and econometrics of active labour market programs. In O. Ashenfelter & D. Card (Eds.), *Handbook of Labor Economics 3*, Amsterdam, Elsevier.
- Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods*, 1(1), 39-65.
- Hempel, K. & Fiala, N. (2011). *Measuring success of youth livelihood interventions. A practical guide to monitoring and evaluation*. Washington, DC, Global Partnership for Youth Employment.
- 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/> [13 Oct. 2014].

Holland, P. W. (1986). Statistics and causal inference (with discussion). *Journal of the American Statistical Association*, *81*, 396, 945–970.

Ibarrarán, P., Ripani, L., Taboada, B., Villa, J. & García, B. (2012) *Life skills, employability and training for disadvantaged youth: Evidence from a randomized evaluation design*. Washington, DC, Inter-American Development Bank.

IEG. (2013). *World Bank and IFC support for youth employment programs*. Washington, DC, World Bank.

Imbens, G. W. & Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, *47*, 1, 5–86.

International Labour Organization. (2012). *The youth employment crisis: A call for action*. Resolution and conclusions of the 101st Session of the International Labour Conference. Geneva, International Labour Office.

International Labour Organization. (2013). *Global Employment Trends for Youth 2013: A generation at risk*. Geneva, International Labour Office.

International Labour Organization. (2014). *Global Employment Trends 2014: The risk of a jobless recovery*. 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.

Khandker, S. R., Koolwal, G. B. & Samad, H. A. (2010). *Handbook on impact evaluation: Quantitative methods and practices*. Washington, DC, World Bank.

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. & Schmidt, C. M. (2002). Can training and employment subsidies combat European unemployment? *Economic Policy*, *17*, 35, 409–448.

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.
- Levinsohn, J., Rankin, N., Roberts, G. & Schöer, V. (2014). *Wage subsidies and youth employment in South Africa: Evidence from a randomised control trial*. Stellenbosch Economic Working Papers: 02/14.
- 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.
- Lipsey, M. W. & Wilson, D. B. (2001). *Practical meta-analysis*. Applied Social Research Methods Series, Volume 49. SAGE Publications.
- 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*.
- O’Leary, J.C. (1998). *Evaluating The Effectiveness of Active Labor Programs in Poland*. Upjohn Institute Technical Report No.98-012; June 1998.
- OECD. (2013). Glossary of statistical terms, Active Labour Market Programmes. Available at: <http://stats.oecd.org/glossary/detail.asp?ID=28> [13 Oct. 2014].
- Sánchez-Meca, J., Marín-Martínez, F. & Chacón-MoscOSO, S. (2003). Effect-size indices for dichotomized outcomes in meta-analysis. *Psychological methods*, 8, 4, 448.
- 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.
- 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.
- Vaessen, J., Duvendack, M., Palmer-Jones, R., Leeuw, F.L., Bonilla, S., Rivas, A., Lukach, R., Bastiaensen, J., Holvoet, N., Hombrados, J. & Waddington, H. (2014). The effect of microcredit on women’s control over household spending: a systematic review and meta-analysis. *Campbell Systematic Reviews*, 10. Available at: <http://www.campbellcollaboration.org/lib/project/178/>
- van Lieshout, S., Sievers, M. & Aliyev, M. (2012). *Start and improve your business: Global tracer study 2011: ILO’s business management training programme*. Geneva, International Labour Office.
- Van Reenen, J.(2003). *Active labour market policies and the British New Deal for the young*

unemployed in context. NBER Working Paper Series Working Paper No. 9576.

Waddington, H., White, H., Snilstveit, B., Hombrados, J. G., Vojtkova, M., Davies, P., Bhavsar, A., Evers, J., Koehlmoos, T. P., Mark Peticrew, 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. (2012). *World Development Report 2013: Jobs*. Washington, DC, World Bank.

Youth Employment Network and International Youth Foundation. (2009). *Private sector demand for youth in Ghana and Senegal. Ghana and Senegal Study Findings*. Available at: http://www.ilo.org/public/english/employment/yen/downloads/psi/psi_study.pdf [13 Oct. 2014].

6 APPENDIX

6.1 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. Access to markets and value chains
 - iii. Credit or access to credit
 - iv. Grants (monetary or in-kind)
 - v. 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:²⁴
 - 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. Dissertations/theses
 - No → Exclude (EndNote: into *_7. Exclude on publication status) (Exclude if editorial, commentary, process evaluations, single-participant studies/anecdotal)

²⁴ Based on table 1 of Leroy et al. 2012.

- Yes → next question
- Unsure → next question

6.2 List of sources for scoping search

Search string ID	Source	Type of source
SSID_1	ASSIA (Applied Social Sciences Index and Abstracts) (CSA)	General database
SSID_2	ABI/INFORM Global (CSA)	General database
SSID_3	British Education Index (CSA)	General database
SSID_4	Australian Education Index (CSA)	General database
SSID_5	EconLit (CSA)	General database
SSID_6	ERIC (Education Resources Information Centre) (CSA)	General database
SSID_7	IBSS (International Bibliography of the Social Sciences) (CSA)	General database
SSID_8	JSTOR http://www.jstor.org/	General database
SSID_9	PAIS International (CSA)	General database
SSID_10	RePEc (Research Papers in Economics) / IDEAS Economics and Finance Research: http://ideas.repec.org/	General database
SSID_11	Sociological Abstracts (CSA)	General database
SSID_12	SocIndex (Social Science Citation Indexes) (CSA)	General database
SSID_13	SSRN (Social Science Research Network) http://www.ssrn.com/	General database
SSID_14	Social Science Citation Index (SSCI) and Arts and Humanities Citation Index (AHCI) of Web of Science http://thomsonreuters.com/web-of-knowledge/	General database
SSID_15	Africa-Wide (http://www.ebscohost.com/academic/africa-wide-information)	General database
SSID_16	Africa-Portal (http://www.africaportal.org/)	General database
SSID_17	3ie Database of Systematic Reviews http://www.3ieimpact.org/evidence/systematic-reviews/	Specialized database
SSID_18	3ie Registry for International Development Impact Evaluations (RIDIE) http://ridie.3ieimpact.org/	Specialized database
SSID_19	3ie Register of Impact Evaluation Published Studies (RIEPS) http://www.3ieimpact.org/evidence/impact-evaluations/	Specialized database
SSID_20	Abdul Latif Jameel Poverty Action Lab (JPAL) Evaluation and Publication Database http://www.povertyactionlab.org/	Specialized database
SSID_21	British Library for Development Studies (BLDS) http://blds.ids.ac.uk/	Specialized database
SSID_22	ELDIS http://www.eldis.org/	Specialized database
SSID_23	Innovations for Poverty Action (IPA) Database http://www.poverty-action.org/project-evaluations/search	Specialized database
SSID_24	JOLIS library catalogue – International Finance Corporation, International Monetary Fund, and World Bank http://jolis.worldbankimflib.org/e-nljolis.htm	Specialized database
SSID_25	Labordoc (ILO) http://labordoc.ilo.org/	Specialized database
SSID_26	Research for Development http://r4d.dfid.gov.uk/	Specialized database
SSID_27	SciDev Net (Science and Development Network) www.scidev.net/en/	Specialized database
SSID_28	World Bank Development Impact Evaluation Initiative (DIME): Database of ongoing Impact Evaluations http://www.worldbank.org/dime	Specialized database
SSID_29	World Bank Poverty Impact Evaluations Database http://www1.worldbank.org/prem/poverty/ie/evaluationdb.htm	Specialized database
SSID_30	Youth Employment Inventory (YEI) http://www.youth-employment-	Specialized database

	inventory.org/	
SSID_31	African Development Bank Evaluation Reports http://www.afdb.org/en/documents/evaluation-reports/	Institutional website
SSID_32	African Economic Outlook http://www.africaneconomicoutlook.org/en/	Institutional website
SSID_33	Asian Development Bank (ADB) Evaluation Resources http://www.adb.org/site/evaluation/resources	Institutional website
SSID_34	Asian Development Bank (ADB) Ongoing Evaluations http://www.adb.org/site/evaluation/ongoing-evaluations	Institutional website
SSID_35	Bureau for Research and Economic Analysis of Development (BREAD) http://ipl.econ.duke.edu/bread	Institutional website
SSID_36	Campbell Collaboration http://www.campbellcollaboration.org	Institutional website
SSID_37	Centre for Economic Policy Research (CEPR) http://www.cepr.org	Institutional website
SSID_38	Center for Economic Studies (CESifo) http://www.cesifo-group.de/ifoHome.html	Institutional website
SSID_39	Center for Global Development http://international.cgdev.org/	Institutional website
SSID_40	DFID: Evaluation Reports https://www.gov.uk/government/organisations/department-for-international-development/series/evaluation-reports	Institutional website
SSID_41	ESRC (Economic and Social Research Council) http://www.esrc.ac.uk/publications/	Institutional website
SSID_42	European Commission DG for Employment, Social Affairs and Inclusion http://ec.europa.eu/social/	Institutional website
SSID_43	European Foundation for the Improvement of Living and Working Conditions http://www.eurofound.europa.eu	Institutional website
SSID_44	European Training Foundation http://www.etf.europa.eu/	Institutional website
SSID_45	Global Development Network (GDN) http://www.gdnet.org/index.html	Institutional website
SSID_46	Institute for Development Policy and Management (IDPM) at the University of Manchester http://www.sed.manchester.ac.uk/idpm/	Institutional website
SSID_47	Institute for Fiscal Studies – Centre for the Evaluation of Development Policy http://www.ifs.org.uk/edepo/index.php	Institutional website
SSID_48	Institute for the Study of Labour (IZA) http://www.iza.org	Institutional website
SSID_49	Institute of Development Studies (IDS) http://www.ids.ac.uk/	Institutional website
SSID_51	Inter-American Development Bank Office of Evaluation and Oversight http://www.iadb.org/en/office-of-evaluation-and-oversight	Institutional website
SSID_52	Islamic Development Bank http://www.isdb.org	Institutional website
SSID_50	International Labour Organization http://www.ilo.org	Institutional website
SSID_53	Millennium Challenge Corporation (MCC) http://www.mcc.gov/pages/results/evaluations	Institutional website
SSID_54	National Bureau of Economic Research (NBER) http://www.nber.org	Institutional website
SSID_55	OECD Development Assistance Committee: Evaluation Resource Center (DEReC) http://www.oecd.org/derec/	Institutional website
SSID_56	Overseas Development Institute (ODI) http://www.odi.org.uk/	Institutional website
SSID_57	Poverty and Economic Policy Research Network (PEP): Project List http://www.pep-net.org/publications/pep-projects/	Institutional website
SSID_58	UNDP International Policy Centre for Inclusive Growth (IPC-IG) http://www.ipc-undp.org/	Institutional website
SSID_59	United States Department of Labour, Employment and Training Administration, Research Publication Database http://wdr.doleta.gov/research/	Institutional website

SSID_60	University of California Center for Effective Global Action (CEGA): Research Projects http://cega.berkeley.edu/research/	Institutional website
SSID_61	UNU WIDER http://www.wider.unu.edu/	Institutional website
SSID_62	USAID Development Experience Clearinghouse https://dec.usaid.gov/	Institutional website
SSID_63	World Bank Independent Evaluation Group (IEG) http://ieg.worldbankgroup.org	Institutional website
SSID_64	World Bank Labor Markets http://www.worldbank.org/labormarkets	Institutional website
SSID_65	7th IZA/World Bank Conference: Employment and Development, November 2012, http://www.iza.org/conference_files/worldb2012/	Conference website
SSID_66	8th IZA/World Bank Conference on Employment and Development, August 2013, http://www.iza.org/conference_files/worldb2013/call_for_papers	Conference website
SSID_67	IZA conferences and meetings http://www.iza.org/en/webcontent/events/izaconferences_html	Conference website
SSID_68	2012 Global Youth Economic Opportunities Conference, September 2012, http://www.youtheconomicopportunities.org/	Conference website
SSID_69	2013 Global Youth Economic Opportunities Conference, September 2013, http://www.youtheconomicopportunities.org/	Conference website
SSID_70	Ideas4Work Conference: Youth Employability and Entrepreneurship in Africa, January 2013, http://iyflive.net/events/ideas-4-work	Conference website
SSID_71	OECD International Economic Forum on Africa 2012: Promoting Youth Employment, October 2012, http://www.oecd.org/site/devfap12/	Conference website
SSID_72	African Development Bank – OECD West Africa Regional Conference on Youth Employment, 10 November 2012, http://www.oecd.org/site/devyewa/Agenda%20Dakar%20Conference.pdf	Conference website
SSID_73	Arab Youth & Entrepreneurship: Holistic Approaches to Nurturing Local Ecosystems, February 2013, http://aye2013.org/	Conference website
SSID_74	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
SSID_75	Electronic Theses Online Service (EThOS) http://ethos.bl.uk	Dissertations and theses databases
SSID_76	Cybertesis http://www.unesco.org/new/en/communication-and-information/portals-and-platforms/goap/key-organizations/latin-america-and-the-caribbean/cybertesis/	Dissertations and theses databases
SSID_77	DART-Europe E-theses Portal http://www.dart-europe.eu/basic-search.php	Dissertations and theses databases
SSID_78	Index to Theses in Great Britain and Ireland: www.theses.com/	Dissertations and theses databases
SSID_79	Networked Digital Library of Theses and Dissertations – SCIRUS ETD Search http://www.ndltd.org/serviceproviders/scirus-etd-search	Dissertations and theses databases
SSID_80	ProQuest Dissertations & Theses Database www.proquest.co.uk/en-UK/catalogs/databases/detail/pgdt.shtml	Dissertations and theses databases
SSID_81	OpenGrey http://www.opengrey.eu/	Grey literature databases
SSID_82	Directory of Open Access Repositories (OpenDOAR) http://www.opendoar.org/	Grey literature databases
SSID_83	Grey Literature Network Service http://www.greynet.org/	Grey literature databases
SSID_84	Regional Youth Employment Consultation in Latin America , International Development Research Centre (IDRC) and Instituto de Pesquisa	Conference website

6.3 Search terms for electronic databases²⁵

1. Exposure terms:

Retraining, training, skill, skills, entrepreneur*, program*, intervention, measur*, scheme, project, activation, subsidy, subsidies, subsidized, subsidised, upgrade, assistance, internship, intern, interns, business, counselling, 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*, schoollleaver*, school leaver*, high school graduate*

4. Impact terms:

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

²⁵ The search terms for electronic databases include the keywords (employment labor labour youth young adolescent student) which were identified through a frequency test of keywords in a group of 32 pre-selected papers during the scoping search. 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 seems to capture too many irrelevant, purely education-related results.

6.4 Search string for electronic databases: Example: ERIC

(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 counselling 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(20000101-20140101))

6.5 Keywords for search in English, Spanish, French, German and Portuguese

The keyword searches in Spanish, French, German and Portuguese combined (“employment” OR “labour”) AND “youth”:

Language	(1) Outcome terms	(2) Subject terms
English	employ* OR labour OR labor	youth
Spanish	empleo OR trabajo	joven*
French	emploi OR travail	jeune*
German	arbeit OR beschäftigung	jugend*
Portuguese	emprego OR trabalho	juventude

Combine (1) and (2)

6.6 Code description

This section contains a description of the variables that will be coded at the study level. Each variable name is followed by a description. Below each variable name/description there is a description of how the variable should be coded.

1. Variable group: Identifiers

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

numeric, format: 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 = Dissertations/Theses*

2. 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 = Panel,*
- 5 = Cross-Section,*

ie_itt: Intention-to-Treat Estimation specifically mentioned
0 = No (only if specified that estimator does not measure ITT),
1 = Yes

3. Variable group: Intervention category

int_cat_skil: Intervention Category:
 Skills Training

0 = No, 1 = Yes,

int_cat_entr: Intervention Category: Entrepreneurship Services

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 = Comprehensive,*

4. 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: Skill training delivered: Distance learning (e.g. books, online training)

0 = No, 1 = Yes,

skil_deli_clas: Skill training delivered: In classroom

0 = No, 1 = Yes,

skil_deli_work: Skill training delivered: At the workplace (e.g. internships, on-the-job training schemes, non-apprenticeship schemes)

0 = No, 1 = Yes,

skil_deli_appr: Skill training delivered: Apprenticeship schemes (in shops with master craftsmen/women)

0 = No, 1 = Yes,

skil_durat: Duration of skill training programme: Total number of hours per individual (averages)

numeric,

skil_prov_publ: Provider of the Skill Training: Public training institution

0 = No, 1 = Yes,

skil_prov_priv: Provider of the Skill Training: Private training institution

0 = No, 1 = Yes,

skil_prov_ngo: Provider of the Skill Training: NGO, Foundation, CBO, CSO

0 = No, 1 = Yes,

skil_paym_lump: Payment system to training providers: Lump-sum budget

0 = No, 1 = Yes,

skill_paym_serv: Payment system to training providers: Payment for services delivered
0 = No, 1 = Yes,

skill_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_regi: 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_donor: Selection of skills: Identified by donors/development agencies
0 = No, 1 = Yes,

5. Variable group: Programme characteristics: Entrepreneurship promotion

entr_typ_advi: Type of intervention: Business advisory/mentoring
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,

6. 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 the 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,

7. 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_num: 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 national government
0 = No, 1 = Yes,

pubw_sele_loca: Works/services selected by local 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_donor: 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,

8. Variable group: Programme description: General characteristics

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. child care, catering, transport)

2 = Monetary benefits (e.g. stipend, transport allowance)

3 = Salary

4 = 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 receipt of benefits)

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,

888 = unsure,

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_aware: 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,

We also code effect size specific information. These are at the effect size/outcome level. There may be more than one outcome or group that a single study analyses. For this reason there may be multiple effect size observations for a single study. The effect size variables are listed below.

9. Variable group: Effect size level information

es_outc_cat: Outcome category

1= Employment outcome,

2= Earning 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 and 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,
4 = Unpaid work,

es_outc_stat: Status of occupation for which outcome is measured

1 = Formal,
2 = Informal,
3 = Both,

es_subg: Indication that estimation sample is different from targeted programme population (authors statement or obvious deviation)

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 = Continuous,

3 = Correlational,

es_adj: Use of unadjusted versus (covariate) adjusted estimation method

1=Unadjusted

2=Adjusted

es_mmeth: Method of measurement of effect size

1 = Pretest comparison,

2 = Post-test comparison,

3 = Follow-up comparison,

es_meth_gr: Relationship between treatment/control group

1 = Unmatched groups (e.g. RCTs, cohort studies),

2 = One group/matched groups (crossover trials, pre-post designs),

es_type_dich: If es_type = 1

1 = No. of events (treatment/control),

2 = Event rates (treatment/control),

3 = 2x2 contingency table (both events, treatment, control, 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/control),

2 = Raw difference in means (= unstandardized regression coefficient),

3 = Standardized mean difference (= standardized regression coefficient, beta),

- 4 = Log raw difference in means,
- 5 = Log standardized mean difference,
- 6 = t-value, f-value, p-value (from a paired t-test),
- 7 = 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 regression,
- 6 = Variance,
- 7 = Confidence intervals (lower, upper),
- 8 = Sample sizes (treatment/control),
- 9 = Standard deviations (treatment/control),
- 10 = Common standard deviation,
- 11 = Total sample size,
- 12 = Standard deviation of difference,

es_raw: Raw effect size as reported in study for single effect size value (e.g. mean differences, unstandardized beta coefficient)

numeric,

es_sign_valu: Value of significance test

numeric,

es_ci_up: Upper value of confidence intervals if reported

numeric, 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

numeric,

contr_outc: Outcome of control group at endline if reported (e.g. mean) (use values after matching/covariate adjustment)

numeric,

contr_n: Number of observations in control group for which effect is estimated
numeric,

contr_sd: Standard deviation in treatment group
numeric,

total_mean: Mean in the total sample
numeric,

total_n: Number of observations total
numeric,

total_sd: Pooled standard deviation
numeric,

es_size_valu: Effect size as computed from
http://www.campbellcollaboration.org/resources/effect_size_input.php, report two decimals
with sign in front: plus if difference favours treatment group, minus if favours control
numeric,

es_sd_valu: Standard deviation of effect size as computed from
http://www.campbellcollaboration.org/resources/effect_size_input.php
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 the moment when the individual exits the intervention and data
measurement
1 = Short term (0–12 months),
2 = Medium term (12–24 months),
3 = Long term (24+ months),

es_date: Date at which effect size is measured
YYYYmMM

7 SOURCES OF SUPPORT

- International Initiative for Impact Evaluation (3ie)'s Systematic Review Call 5 through funding from the Canadian International Development Agency (CIDA)
- The Youth Employment Network
- The International Labour Organization
- The World Bank
- RWI, Berlin

8 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]

Rother, F. (forthcoming) "Youth Employment: A Human Development Agenda for the Next Decade" with Mattias Lundberg, David Margolis, David Newhouse, and David Robalino (forthcoming).

Publications

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*. World Bank, SPDP 0715 [with Betcherman, Godfrey and Stavreska].

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).

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 Jochen Kluve and Maria Laura Sanchez-Puerta). In *The right skills for the job: Rethinking effective training policies for workers*, by Almeida, Behrmann and Robalino. Washington, DC, World Bank.

Robalino, D. with Almeida, R. & Behrmann, J. (2012). *The right skills for the job: Rethinking training policies for workers*. 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 Helena

Ribe and Ian Walker). *Directions in Development*. Washington, DC, World Bank.

The review's PIs have co-authored and authored reviews of active labour market programmes that have allowed for statistical rigour, contributing a series of meta-analyses to the labour economics literature and more precisely 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.

9 REVIEW AUTHORS

Lead review author

Name: Jochen Kluve

Title: Prof. Dr.

Affiliation: Humboldt-Universität zu Berlin

Address: Spandauer Str. 1

City, State, Province or County: Berlin

Postal Code: 10178

Country: Germany

Phone: +49-30-2021-598-13

Email: jochen.kluve@hu-berlin.de

Co-author:

Name: Susana Puerto

Affiliation: International Labour Organization

Country: Switzerland

Email: puerto-gonzalez@ilo.org

Co-author:

Name: Jonathan Stoeterau

Affiliation: Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI)
and Humboldt-Universität zu Berlin

Country: Germany

Email: jonathanstoeterau@gmail.com

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: marcjosefwitte@gmail.com

Co-author:

Name: David Robalino

Affiliation: World Bank

Country: USA

Email: drobalino@worldbank.org

Co-author:

Name: Jose Manuel Romero

Affiliation: World Bank

Country: USA

Email: jromero1@worldbank.org

Co-author:

Name: Friederike Rother

Affiliation: World Bank

Country: USA

Email: frother@worldbank.org

10 ROLES AND RESPONSIBILITIES

Content

Jochen Kluge (Lead PI) will lead the different analytical stages of the review, bringing in his substantive teaching and applied expertise on labour economics and findings from his research on and analyses of active labour market programmes. Prof. Kluge will be in charge of content-related discussions and coordinate the review's data analysis and interpretation stages. He will, in particular, lead the meta-analytical work, providing substance on the underlying theory of change and the effect sizes on selected labour market outcomes. Prof. Kluge's exposure to impact evaluation design and implementation ensure adequate understanding of concepts and easy abstraction of results from evaluative evidence. His country work experience and research in Africa are great assets to the study's analysis on how applicable results are in the African context. A sample of his work includes the following two meta-analyses of ALMPs:

- <http://www.nber.org/papers/w16173>
- <http://www.tinbergen.nl/cost/london/kluge.pdf>

Susana Puerto (co-PI) will provide extensive sector/thematic support to the review, examining the review parameters and the underlying logical model. She will also support the analysis of effects from the sample of studies, the interpretation of results and the meta-analysis work. She brings in her applied expertise in labour economics, youth employment and evaluation as well as eight years of experience analysing youth employment programmes and providing advice to policymakers and practitioners on what works to improve labour market outcomes of youth. Over the last three years, Susana's work has focused on youth employment in sub-Saharan Africa, leading and supporting numerous impact evaluation initiatives in the region and promoting evidence-base policymaking through capacity building for African policymakers and practitioners implementing youth employment initiatives. She now manages the Youth Employment Network, an inter-agency programme of the ILO, the World Bank and the UN. Her experience with meta-studies is reflected in the following research papers:

- <http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/SP-Discussion-papers/Labor-Market-DP/0924.pdf>
- <http://www.youth-employment-inventory.org/downloads/7.pdf>

David Robalino (co-PI) will support the team on the design and methodological part of the meta-analysis. He will be an integral part of the team for the analysis and the definition of policy implications of the reviewed programmes. As a Lead Economist and Leader of the Labour and Youth Team in the Human Development Anchor of the World Bank, and Co-Director of the

Employment and Development programme at IZA – the Institute for the Study of the Labour – he can facilitate different platforms. He is also a co-founder of the Jobs-Knowledge Platform. Since joining the Bank, David has been working on issues related to youth, labour markets, and social insurance. He has worked in several countries in Latin America, the Middle East and North Africa, sub-Saharan Africa and Asia. David has published widely on issues related to labour markets, youth, and social insurance, and other related topics. Prior to joining the Bank, David was a researcher at the RAND Corporation where he was involved in research on health, population and labour, climate change and the development of quantitative methods for policy analysis under conditions of uncertainty. David also served in the Presidential Committee for Social Security Reform in Ecuador. A sample of David's publications can be found at:

- http://www.iza.org/en/webcontent/personnel/photos/index_html?key=4274

Friederike Rother (co-PI) will support the interpretation of results as well as content-related work and discussions on interventions and labour market outcomes. She brings in her applied knowledge and expertise on labour economics and youth employment programmes' design and implementation. Through her experience at the World Bank's Social Protection and Labour Markets anchor, Friederike provides technical support on issues related to labour markets, job creation, economic growth, and youth employment. Prior to joining the World Bank, Friederike served as a policy maker at the German Employers' Association. She also worked for the German Development Corporation GIZ. Her research work, particularly linked to young people, includes:

- <http://siteresources.worldbank.org/SOCIALPROTECTION/Resources/SP-Discussion-papers/Labor-Market-DP/0715.pdf>
- <http://www.youth-employment-inventory.org/downloads/6.pdf>
- <http://www.youth-employment-inventory.org/downloads/4.pdf>

Systematic review methods

The four PIs will lead the problem formulation process, opening up the question and clarifying scope, concepts/terms, parameters and variables of interest. The team will subsequently outline the review's protocol and set adequate criteria for inclusion and exclusion of studies. Jochen Kluge will provide his expertise on reviews and preparatory work for meta-analyses of active labour market programmes. In addition, it is worth noting that both Jochen Kluge and Susana Puerto have been invited to review systematic review applications in previous 3ie calls, which provides further familiarity with the standards of the grant organization and the structure and features of the review methods.

Susana Puerto and Friederike Rother are responsible for protocol writing as well as data

collection and evaluation ensuring recording of methodological details. Both team members participated and co-authored the assembly of the Youth Employment Inventory (YEI, www.youth-employment-inventory.org) between 2005 and 2007, which constructed a rigorous review methodology and coding protocol to ensure relevance and efficiency in the data collection and evaluation process (<http://www.youth-employment-inventory.org/downloads/1.pdf>). In 2007, the YEI compiled 289 studies of labour market programmes for youth. Today, the YEI has reached over 700 studies from various designs and level of rigour.

Statistical analysis

Jochen Kluge will take the lead of the statistical work, designing the econometrical model for the meta-analysis and its underlying logical model. If data allows, he will compare meta-analytic models for programme effect size. He will be supported by Susana Puerto and Jonathan Stoeterau in the statistical analysis and corresponding data cleaning and crunching. Jochen, Susana and David Robalino will add their expertise in the design and implementation of impact evaluations of labour market programmes and in the design of meta-analyses of labour market interventions. Jonathan Stoeterau, on the other hand, will provide statistical support with the econometrical models and relevant tests for the meta-analysis and the determination of programme effect size.

Information retrieval

Friederike Rother and Susana Puerto will be responsible for the data collection and evaluation stages of the review and, in relation to information retrieval, they will supervise the work of the research assistants in the literature search and information gathering from the studies. They will conduct team exercises to ensure all research assistants are working with the same criteria. As indicated above, Friederike and Susana have direct experience with information retrieval after building the YEI dataset in 2005–2007. Today, they continue cleaning and maintaining the inventory, demonstrating strong co-ownership and leadership in the collection of information and its evaluation.

José Manuel Romero, Jonathan Stoeterau, Felix Weidenkaff and Marc Witte will conduct a literature search of impact evaluation studies not yet incorporated into the YEI and examine in detail the studies of programmes already recorded in the YEI, looking closely at the variables of interest for this review. They will also look at past reviews and meta-studies of youth labour market interventions. Jonathan has previously worked as an external evaluation consultant for the German Agency for International Cooperation (GIZ), where he analysed individual-level employment effects of GIZ-supported educational reforms in Ethiopia by applying impact evaluation methods to a survey of university graduates. Specializing in applied econometrics through his graduate studies, he can build on strong methodological skills to assess the empirical foundation of impact evaluation studies. As an assistant on several research projects,

he gained proficiency in managing large datasets and implementing econometric approaches in Stata. Felix Weidenkaff will provide his expertise in literature review and information retrieval directly related to the improvement and expansion of the YEI. He developed a spin-off exercise of the YEI in Kenya and Egypt aimed at sizable country-level inventories of youth employment interventions that will allow further research on trends and investment in youth. He has contributed to analytical papers and literature reviews of women and youth entrepreneurship interventions. Marc Witte is a doctoral candidate in economics at the University of Oxford. His research is on labour markets in developing countries, with a special focus on technical training and youth employment in the informal sector. Having previously worked for the GIZ in Rwanda, the Rheinisch-Westfälisches Institut für Wirtschaftsforschung (RWI) and the Ecologic Institute (Berlin), he possesses not only the necessary econometrical skills to contribute to the review, but also relevant policy-shaping and on-field experiences. José Manuel Romero is a doctoral candidate in economics at The George Washington University. His areas of concentration are applied microeconometrics and development economics. He has co-authored impact evaluation studies related to microfinance, household risk coping mechanisms and food security in Malawi. In addition, he has over seven years' professional experience in economic and data analysis that includes experience working for the World Bank, International Monetary Fund and multiple research centers. The entire review team brings knowledge of five different languages, including English, French, Spanish, Portuguese and German, ensuring skills necessary to retrieve and interpret key information from the sample of studies.

11 PRELIMINARY TIMEFRAME

	2013	2014												2015			
	Dec	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Jan	Feb	Mar	Apr
Preparation of protocol	■																
Review of protocol (IDCG)	■	■	■	■	■	■											
Revised protocol							■	■	■	■	■						
Scoping search	■	■	■														
Study search	■	■	■	■	■	■											
Assessment of study relevance		■	■	■	■	■	■	■	■	■							
Extraction of data			■	■	■	■	■	■	■	■	■						
Synthesis / statistical analysis												■	■				
Preparation of draft report													■	■			
Review of draft report (IDCG)															■		
Revision of draft report																■	
Dissemination of final report																	■

12 PLANS FOR UPDATING THE REVIEW

The review team has set a number plans to continue research efforts around the assessment of what works in youth employment.

First, the ILO and the World Bank have set an area of collaboration around the Youth Employment Inventory (YEI) which also involves other partner institutions, namely, the German Federal Ministry for Economic Cooperation and Development, the Inter-American Development Bank and its sister agency the Multilateral Investment Fund. The work on the YEI directly impacts the maintenance and update of the review's database as well as the dissemination of findings.

Second, the ILO and the Multilateral Investment Fund are teaming up to carry out events to discuss evidence from impact evaluations of youth employment programmes, where the systematic review will feature. The first event will take place during the first quarter of 2015.

Third, RWI has entered into bilateral partnerships with the ILO, the World Bank and the Inter-American Development Bank for the development of impact evaluations that will support the expansion of the review's component studies.

13 AUTHORS' RESPONSIBILITIES

By completing this form, you accept responsibility for preparing, maintaining and updating the review in accordance with Campbell Collaboration policy. The Campbell Collaboration will provide as much support as possible to assist with the preparation of the review.

A draft review must be submitted to the relevant Coordinating Group within two years of protocol publication. If drafts are not submitted before the agreed deadlines, or if we are unable to contact you for an extended period, the relevant Coordinating Group has the right to de-register the title or transfer the title to alternative authors. The Coordinating Group also has the right to de-register or transfer the title if it does not meet the standards of the Coordinating Group and/or the Campbell Collaboration.

You accept responsibility for maintaining the review in light of new evidence, comments and criticisms, and other developments, and updating the review at least once every five years, or, if requested, transferring responsibility for maintaining the review to others as agreed with the Coordinating Group.

14 PUBLICATION IN THE CAMPBELL LIBRARY

The support of the Campbell Collaboration and the relevant Coordinating Group in preparing your review is conditional upon your agreement to publish the protocol, finished review and subsequent updates in the Campbell Library. Concurrent publication in other journals is encouraged. However, a Campbell systematic review should be published either before, or at the same time as, its publication in other journals. Authors should not publish Campbell reviews in journals before they are ready for publication in the Campbell Library. Authors should remember to include a statement mentioning the published Campbell review in any non-Campbell publications of the review.

I understand the commitment required to undertake a Campbell review, and agree to publish in the Campbell Library. Signed on behalf of the authors:

Form completed by: Susana Puerto

Date: 14 October 2014