**Do Youth Employment Programs Improve Labor Market Outcomes?**

**A Systematic Review[1]**

Jochen Kluve,[2] Susana Puerto,† David Robalino,‡ Jose Manuel Romero,‡ Friederike Rother,‡

Jonathan Stöterau,[2] Felix Weidenkaff,† Marc Witte*

This draft: June 2016

Preliminary, not for dissemination and citation

[2] Humboldt University Berlin and RWI; †International Labour Organization; ‡The World Bank; *University of Oxford. Corresponding authors: jochen.kluve@hu-berlin.de, puerto-gonzalez@ilo.org, drobalino@worldbank.org, frother@worldbank.org.

**Abstract**

This study reviews the evidence on the impact of youth employment programs on labor market outcomes. The analysis looks at the effectiveness of various interventions and the factors that influence program performance including country context, targeted beneficiaries, program design and implementation, and type of evaluation. We identify 113 counterfactual impact evaluations covering a wide range of methodologies, interventions, and countries. Using meta-analysis methods, we synthesize the evidence based on 2,259 effect sizes (Standardized Mean Differences, or SMD) and the statistical significance of 3,105 treatment effect estimates (Positive and Statistically Significant, or PSS). Overall, we find that just more than one-third of evaluation results from youth employment programs implemented worldwide show a significant positive impact on labor market outcomes – either employment rates or earnings. In general, programs have been more successful in middle- and low-income countries; this may be because these programs' investments are especially helpful for the most vulnerable population groups – low-skilled, low-income – that they target. We also conjecture that the more-recent programs might have benefited from innovations in design and implementation. Moreover, in middle and low income countries, skills training and entrepreneurship programs seem to have had a higher impact. This does not imply, however, that those programs should be strictly preferred to others; much depends on the needs of beneficiaries and program design. In high-income countries, the role of intervention type is less decisive – much depends on context and how services are chosen and delivered, a result that holds across country types. We find strong evidence that programs that integrate multiple interventions are more likely to succeed because they are better able to respond to the different needs of beneficiaries. We also find evidence about the importance of profiling and follow-up systems in determining program performance, and some evidence about the importance of incentive systems for services providers.

Keywords: Youth Employment, Active Labor Market Policy, Impact evaluations, Systematic review, Meta-Analysis

JEL codes: J21, J48, E24

# 1. Introduction

The economic crisis that began in 2007 reversed the gradual declining trend in global youth unemployment rates observed between 2002 and 2007, changing the course of the trend to increasing youth unemployment between 2007 and 2010. This turn of events quickly led to discouragement and significant withdrawal from the labor force among young workers and jobseekers. It is estimated that nearly 6.4 million youth worldwide moved into inactivity as a response to the crisis, while many others continue working yet living in poverty (ILO, 2012).

Today, the global youth unemployment rate has settled at 13.0 per cent, with nearly 73.3 million youth actively looking for jobs. While the unemployment rate among youth is expected to remain relatively constant in the near future, it is still well above its pre-crisis level of 11.7 per cent. Youth remain overrepresented among the unemployed and shaken by the changing patterns in the labor market. Two-fifths (42.6 per cent) of the global youth labor force were reported unemployed or working poor in 2013 (ILO, 2015). In addition, many youths are idle. In 2014, 12.4 percent of youth in the European Union (EU-28) were not in employment, education or training (NEET) (ILO 2015a).

As a result, many countries have implemented active labor market programs (ALMPs) aiming to connect youth to wage- or self-employment. These include general employment services (counseling, job search assistance, and intermediation), training, wage subsidies, and entrepreneurship programs. Unfortunately, although this has been changing over the last decade, the majority of programs have been implemented without proper evaluations, and our knowledge about their effectiveness has been limited.

While some previous studies have synthesized the evidence based on the effectiveness of ALMPs (e.g., Card et al. 2010 and 2015), very few reviews have focused specifically on programs and outcomes for youth. The most relevant review of labor market interventions for youth to date, Betcherman et al. (2007), has served as the basis for technical assistance and policy advice worldwide. Since then, a vast amount of research has been published, using experimental or quasi-experimental methods to determine the impact of new and innovative employment programs. While some recent reviews have covered this new evidence, these have not synthesized the existing empirical evidence using empirical methods such as meta-analysis (JPAL 2013), or they have focused on (potentially selective) subsets of the available evidence (IEG 2012). Other studies have put their emphasis on specific types of intervention or outcomes (Tripney et al. 2013, Grimm and Paffhausen 2015). There also remain knowledge gaps in trying to understand how issues related to design and implementation affect the effectiveness of different programs.

This Systematic Review looks at the available rigorous evaluations to understand the relative effectiveness of youth-targeted interventions and some of the key factors that influence their performance. We conducted a comprehensive search of available impact evaluations that yielded 113 studies considered of adequate quality and methodological rigor to be included in our meta-analysis. The systematic search and selection process allows us to address potential issues stemming from publication bias and selective reporting. The protocol, search, selection process, and quantitative results are presented comprehensively in Kluve et al. (2014, 2016). This paper builds on that Systematic Review, presents the search and data in a concise way, and presents new and additional narrative and quantitative evidence using the review's data base.

Specifically, we employ random-effects meta-analysis methods to synthesize and compare effect sizes reported in the primary studies. We further estimate multivariate meta-regression models to understand the determinants of success (effect sizes and the likelihood of a statistically significant effect), including information about the specific intervention design, country characteristics, and implementation features.

The paper is organized as follows: Following Section 1, Introduction, we propose a theory of change in Section 2 that will guide our analysis, linking program inputs to labor market employment outcomes of youth. In Section 3, we lay out our inclusion criteria and the search strategy. The characteristics of the studies and interventions included in the review are discussed in Section 4. Section 5 presents our data and methods. In Section 6, we present the empirical results of the meta-analysis, from both unconditional and multivariate analysis. Section 7 presents our conclusions.

## 2. Theory of change

The active labor market programs examined in this paper affect labor market outcomes for youth in multiple and interrelated ways. The interventions are clustered into four typologies: (i) training and skills development; (ii) entrepreneurship promotion; (iii) employment services; and (iv) subsidized employment interventions. The theory of change hypothesizes that participation in these ALMPs will ultimately improve the employment and earnings outcomes for youth. The idea behind this hypothesis is that youth can face different constraints that affect access to wage- or self-employment, constraints that can be addressed through targeted interventions. For instance, they might not have the necessary skills and/or work experience, they might not have information about job opportunities and/or knowledge about the jobs search process, and they are less likely to be able to access capital to start a business (Robalino et. al. 2013). This review is designed to illuminate which type of intervention (or combination of interventions) is most successful in addressing these constraints. Table 1 summarizes these.

Training and skills development, the first type of intervention, comprises programs outside the formal education system that offer skills training to young people to improve their employability and facilitate access to jobs.[3] The objective of skills training programs 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 behavioral) skills, such as self-management, teamwork, and communication. In this analysis, training and skills development programs are classified according to the skill-set they target, i.e., technical skills, business skills, literacy or numeracy skills, and non-technical skills, such as behavioral, life, or soft skills.

Entrepreneurship promotion interventions aim to provide advisory services, as well as facilitate access to finance and markets. The analysis, therefore, disaggregates programs into those providing or facilitating access to credit, providing start-up grants, or fostering micro-franchising mechanisms. In addition, the analysis includes programs that offer business advisory services and mentoring, business skills training, and access to markets and value chains.

Employment service programs are generally based on the (matching and) intermediation approach to active labor market policy. Employment services deliver job counseling, job-search assistance, and/or mentoring services for (re)activation purposes, which are often complemented by job placements and technical or financial assistance. The basic idea for providing employment services to youth is that young workers have difficulty signaling their skills and credentials and/or lack the networks or knowledge to search effectively for vacancies and connect with employers. As a result, these programs often focus on improving job-seeking skills and the efficiency of the matching process.

Subsidized employment interventions comprise wage-subsidy and (labor-intensive) public work programs. Wage subsidies provide incentives to employers to hire often first-time job-seekers for a given period of time by reducing labor costs. The programs can take many forms, from grants to employers or trainees to reductions in social security contributions. The main objective of these programs is to allow young workers to build work experience and acquire job-relevant skills (Almeida 2012). Public work programs and services, on the other hand, offer directly subsidized, temporary jobs to workers. These programs are mainly considered when there is not enough labor demand in the private sector (Grosh 2008) and often serve as a temporary safety net in times of crisis.

Table 1 provides a simplified results chain, aiming to map out the causal process from intervention delivery to labor market effects.

---

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

## Table 1. Types of interventions and constraints

| Type of constraint faced by youth | General Type of intervention used to address | Rationale | Services under this type of intervention | Risks | Illustrative examples |
|---|---|---|---|---|---|
| Information gap (lack of adequate information about job opportunities & lack of information about skills of young applicants by employers), limited access to networks, obstacles to applying for jobs (eg. high transport costs) | Employment and Intermediation Services | Creating mechanisms that make information exchange between(for) employers and workers less costly. | Information Systems/ Counseling, based on accurate labor market information Mentoring Training, Job search assistance, Support services | Displacement of employment (no new jobs created). | Programa Inserjovem (Portugal), Jordan New Opportunities For Women (NOW) |
| Limited Access to Credit; Lack of financial capital, Limited Social Networks, Limited know-how in setting up a business, bookkeeping, and similar skills, Value chain exclusion or disconnect | Youth Entrepreneur Promotion Programs | Directly supplying young entrepreneurs with access to the specific inputs needed for a business to succeed. | Microfinance, business skills training, assesments by experienced professionals, Facilitating access to value chains, mentorships that teach management and other know-how (marketing, business registration) | Moreal-hazard, low-potential porjects, and expensive assessments of business profitability. | Women's Income Generation Supports (WINGS, Uganda), The Prince's Trust (UK) |
| Inadequate supply of skills – technical, cognitive, and non-cognitive, Low Skill Level, No or little work experience, Skills Mismatch (youth are not trained for the jobs requested by employers), Missing "soft" non-cognitive skills, lack of basic skills (numeracy/literacy) | Skills training for young people | Training workers with the technical, vocational, non-cognitive skills that makes them desirable to firms | Different types of training: technical and vocational skills, business skills, literacy and/or numeracy, behavioral and non-cognitive skills that are implemented both in classrooms or on the job (OJT) | Governance of program. | Job Corps (US), Chile Joven |
| Little or no work experience of youths Minimum Wages and mandatory benefits (e.g. social security contribution) = Workers whose productivity is not high enough to outweigh the costs of employment of youth with little or no work experience | Subsidized Employment | Lowering hiring and labor costs of employing workers to allow them to gain work experiences which makes them more productive and propels them into their career path. | Direct payments to employers, tax deductions to employers, direct payments to workers, public works. | Deadweight loss (creation of more jobs than required). | JUMP wage subsidies (JWS, Germany), Youth Hires (Canada) |
| Specific household and individual conditions includding needs to take care of children, eldernly, or disabled family members, and unability to pay for transportation to access job opportunties. | Allowances or inkind support to cover specific services needs | Targetting specific constraints that inhibit labor market or entrepreneiral activities permits productive participation in productivite activities. | Subsidies for transportation to work, child care services. | Lack of effect due to other overarching constraints not being met. | Teenage Parent Demonstration (US) |

## 3. Search, selection, and coding of primary studies

### 3.1 Inclusion criteria

This Systematic Review focuses on studies that investigate the impact of interventions on labor market outcomes of young people and meet the following inclusion criteria:

*Population and context.* The review is global in coverage and considers interventions from all countries, regardless of their levels of development. Studies must have investigated ALMPs that are designed for – or primarily target – young women and men aged 15 to 35, in consideration of national definitions of youth. The ALMPs examined 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 labor 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).

*Intervention.* Eligible studies must have evaluated an ALMP that provided at least one of the four categories of intervention (also shown in Table 1) – training and skills development, entrepreneurship promotion, employment services, and/or subsidized employment.

The Review makes an important distinction between programs and interventions. A youth employment program is a single entity that may consist of one or several interventions. It is possible to find a comprehensive program that offers, for instance, both skills training and employment services. Program examples consisting of several interventions include the Job Corps program in the United States, Economic Empowerment for Adolescent Girls program in Liberia, Projoven program in Peru, and Employment Fund in Nepal.

Interventions are, therefore, components/tracks of programs, or in some contexts also identified as sub-programs within an overall larger program. They are defined based on characteristics such as the category of intervention or the population targeted. For example, if a program has a training track and an employment services track, and participants take one or the other, these are considered to be two interventions within the same program. Note that, by this definition of track, we are assuming that each intervention within a program has separate groups of participants that do not overlap. On the other hand, if a program has a single track that includes two different services, whereby the same participants take both training and employment services, this would be considered a single intervention. The identification of components within programs allows for the analysis of interactions across interventions (Section 5.3).

An additional consideration is made to identify primary intervention types under multi-pronged designs. The review defines "main category of intervention" as the largest and predominant

intervention type within a program. If several intervention types are equally distributed across the target population (i.e., an individual is exposed to more than one intervention type with the same level of intensity), the main category of intervention is classified as comprehensive.

*Comparison.* The Systematic Review included studies that measure change in at least one outcome of interest among intervention participants and relative to non-intervention participants based on a counterfactual analysis (comparing treated and control groups). Eligible comparison groups (counterfactual) include those that receive no intervention or are due to receive the intervention in a pipeline or waitlist study. Note that the comparison group of some studies might be exposed to interventions other than the evaluated intervention.

*Outcome.* Eligible studies must have reported at least one selected outcome variable measuring the primary outcomes of interest shown in Table 1, namely: employment outcomes, earnings outcomes, or business performance outcomes. The Review also captured outcomes that are measured conditional on other outcomes and excluded studies that focus only on intermediary outcomes without measuring impacts on the above-mentioned primary outcomes.

*Study design and methods.* The Review focuses on completed experimental and quasi-experimental evaluations, and considers the following research design categories and impact evaluation methods to quantitatively estimate the causal effect of the intervention on the outcome it intends to influence – (i) randomized experiments, (ii) methods for causal inference under un-confoundedness (classical regression methods, statistical matching, propensity score matching), and (iii) selection on unobservables (instrumental variables, regression discontinuity design, difference-in-differences).

The form of publication of eligible studies includes peer-reviewed journal, working paper, mimeo, book, policy or position paper, evaluation or technical report, and dissertation or thesis. Eligible studies can be published in any language as long as they meet all other eligibility criteria. The date of publication or reporting of the study must have been between 1990 and 2014.

### 3.2 Search methods

The search for relevant literature was based on a variety of sources to ensure that published and unpublished studies ("grey" or "gray" literature) relevant to the research question are included in the search process. The search process included (i) a primary search – searching of a wide range of general and specialized databases, and (ii) a complementary search – hand-searching of relevant websites; searching of dissertations, theses, and grey literature databases; literature snowballing; and contacting authors and experts. The search included search terms in English, Spanish, French, German, and Portuguese. For each source, the review team tested and documented several strategies and identified one or more preferred search strategies that yielded a comprehensive and precise set of

potentially relevant results. The full list of sources and specific searches conducted can be found in Kluve et al. (2014), the protocol for this Systematic Review.

### 3.3 Data extraction

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

At effect-size level, the coding tool captured sub-group analysis of employment, earnings, and business performance outcomes and estimated treatment effects by age cohorts, gender, education level, income level, and location, among other dimensions. For some sub-groups, such as those defined by education and income levels, we relied on the descriptions and classifications applied in the original studies. Types of outcomes were further disaggregated by occupation category (dependent vs. self-employment), status of occupation (formal vs. informal), and conditionality on other outcomes. To describe the data and empirical methods, the coding tool includes information about the research design, statistical methodology, type of significance test, type and method of measurement, date of data measurement, and data source. The coding tool also captured the form and year of publication.

For each category of intervention (i.e., skills training, entrepreneurship promotion, employment services, and subsidized employment), the coding tool extracts information about type of intervention, targeting and delivery mechanism, payment system and provider, duration of specific interventions, selection of participants, and conditionality of eligibility[4]. General program characteristics record the target group by age, gender, education level, income level, location, and employment status, as well as the type of organizations involved in designing, financing, and implementing the program. The coding tool keeps record of region, country, scale, and average duration of the program. Any awareness raising efforts and gender considerations integrated in program design and implementation were also captured. A set of select variable definitions, which includes some of these design features, is included in Table 1 of the Annex 2.

Three key program design features were also coded: the implementation of participant profiling, mechanisms to keep beneficiaries engaged in the program, and incentives to service providers. Profiling is the identification of individual factors that represent a risk in the labor market and assigning appropriate services based on this. Collecting the information to assess the main constraints (risks) in the labor market include caseworker mentor discretion, detailed screening,

---

[4] Program design and other information features coded were based on what the original design of the program was, whether or not the program was implemented as originally designed. This includes items such as program duration.

eligibility rules, or statistical profiling among others. A program is considered to profile if it uses information to assign participants specific services among an array of services offered or to determine the intensity of services (e.g. program duration). A key element of this is that the program collects information and proactively assigns services. Engagement mechanisms are features of the program to incentivize and increase the probability that beneficiaries will complete the program or perform well in the program. These are usually applied by either providing monetary (or in-kind) incentives for participants attending the programs or by individualized attention and guidance. Most frequently incentives include payments or living stipends that are contingent on attendance. This can also include incentives for participation such as taking away welfare or unemployment benefits. Individualized monitoring can be direct one-on-one supervision of program participants often through scheduled visits with a case manager or mentor. Incentives to service providers refers to instances where payment to service providers on the field from the respective funders are linked to the outcomes of the beneficiaries, either in terms of program completion, performance within the program (e.g. test scores) or labor market outcomes, such as gaining productive employment after program exit. While these are design features that have been observed in successful programs, such as some of the Jovenes program in Latin America and the Caribbean, empirical evidence on the whether these features systematically improve youth employment performance, no our knowledge, does not exist. On the other hand, this important information was often missing in the impact evaluation studies. To minimize the number of missing values in these design feature variables, additional information was researched from sources different to the study. Such sources included project reports and projects' websites.

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

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

Information extracted from included studies was discussed with a second reviewer and coding decisions involving assumptions were documented by the researcher. When information at effect size level could not be obtained by the researcher from the primary studies, authors of included studies were contacted to provide missing information and clarify discrepancies.[5] Additional details about inclusion criteria, search strategy, and data extraction are delineated in the protocol of the Systematic Review (Kluve et al. 2014).

**3.4 Search results**

The primary and complementary search resulted in 32,117 records based on search in more than 70 sources, including: 12 specialized databases; 11 general databases; 35 websites, such as institutional and conference websites; 5 dissertations, theses, and grey literature databases; and 9 other reviews and meta-analyses.

After removing duplicates using the reference management software EndNote, individual reviewers screened 28,375 records on title and abstract by applying the inclusion criteria of the screening questionnaire. A total of 1,141 records were identified for full-text screening.

To minimize bias, included and excluded results were cross-checked by another researcher and discrepancies were resolved by the two. This systematic screening process led to the identification of 86 studies that are considered to be of adequate content and methodological rigor to inform the Systematic Review. After extracting data from the preliminary set of 86 included studies, the review team screened 6,782 additional records that were identified through reference lists and citation tracking of included studies, hand-searching of key journals in which a large number of included studies were found, and contacting authors and experts. This search process led to the selection of 27 additional studies.

Overall, this comprehensive search and selection process resulted in 113 studies considered eligible for inclusion in this review. The search and screening process is illustrated in Figure 1.

**Figure 1. Results of the search process**

---

[5] The review team contacted authors of 100 studies, requesting additional information to facilitate the computation of the effect sizes or to get clarity on the quantitative results or intervention details; 34 authors replied to the request.

```
┌─────────────────────────────────────────────────┐
│ Search in over 70 sources using search terms in │
│ EN, SP, FR, GE, PO, including:                  │
│  ▪ 12 specialized databases                     │
│  ▪ 11 general databases                         │
│  ▪ 35 websites, including institutional and     │
│    conference websites                          │
│  ▪ 5 dissertations, theses and grey literature  │
│    databases and gateways                       │
│  ▪ 9 other reviews and meta-analyses            │
└─────────────────────────────────────────────────┘
                      ↓
┌──────────────────────────────────────┐    ┌──────────────────────────────────┐
│ Primary and complementary search:    │ ➡  │ Duplicates in EndNote: 3,742     │
│ 32,117 records                       │    │ records                          │
└──────────────────────────────────────┘    └──────────────────────────────────┘
                      ↓
┌──────────────────────────────────────┐    ┌──────────────────────────────────┐
│ Title/abstract review of 28,375      │ ➡  │ 27,234 records excluded          │
│ records                              │    │                                  │
└──────────────────────────────────────┘    └──────────────────────────────────┘
                      ↓
┌──────────────────────────────────────┐
│ Full text review of 1,141 records    │
└──────────────────────────────────────┘
                      ↓
┌──────────────────────────────────────┐    ┌──────────────────────────────────┐
│ 86 reports selected                  │ ➡  │ Review of 6,782 records          │
│                                      │    │ identified through reference     │
│                                      │    │ lists, citation tracking, hand   │
│                                      │    │ searching of key journals and    │
│                                      │    │ contacting authors and experts   │
└──────────────────────────────────────┘    └──────────────────────────────────┘
                      ↓                                     ↓
┌──────────────────────────────────────┐    ┌──────────────────────────────────┐
│ Final selection: 113 reports         │ ⬅  │ 27 reports selected              │
└──────────────────────────────────────┘    └──────────────────────────────────┘
            ↓              ↓
┌──────────────┐  ┌──────────────┐            ┌──────────────────────────────────┐
│ 31 countries │  │ 87 programmes│  ➡         │ 107 interventions                │
└──────────────┘  └──────────────┘            └──────────────────────────────────┘
```
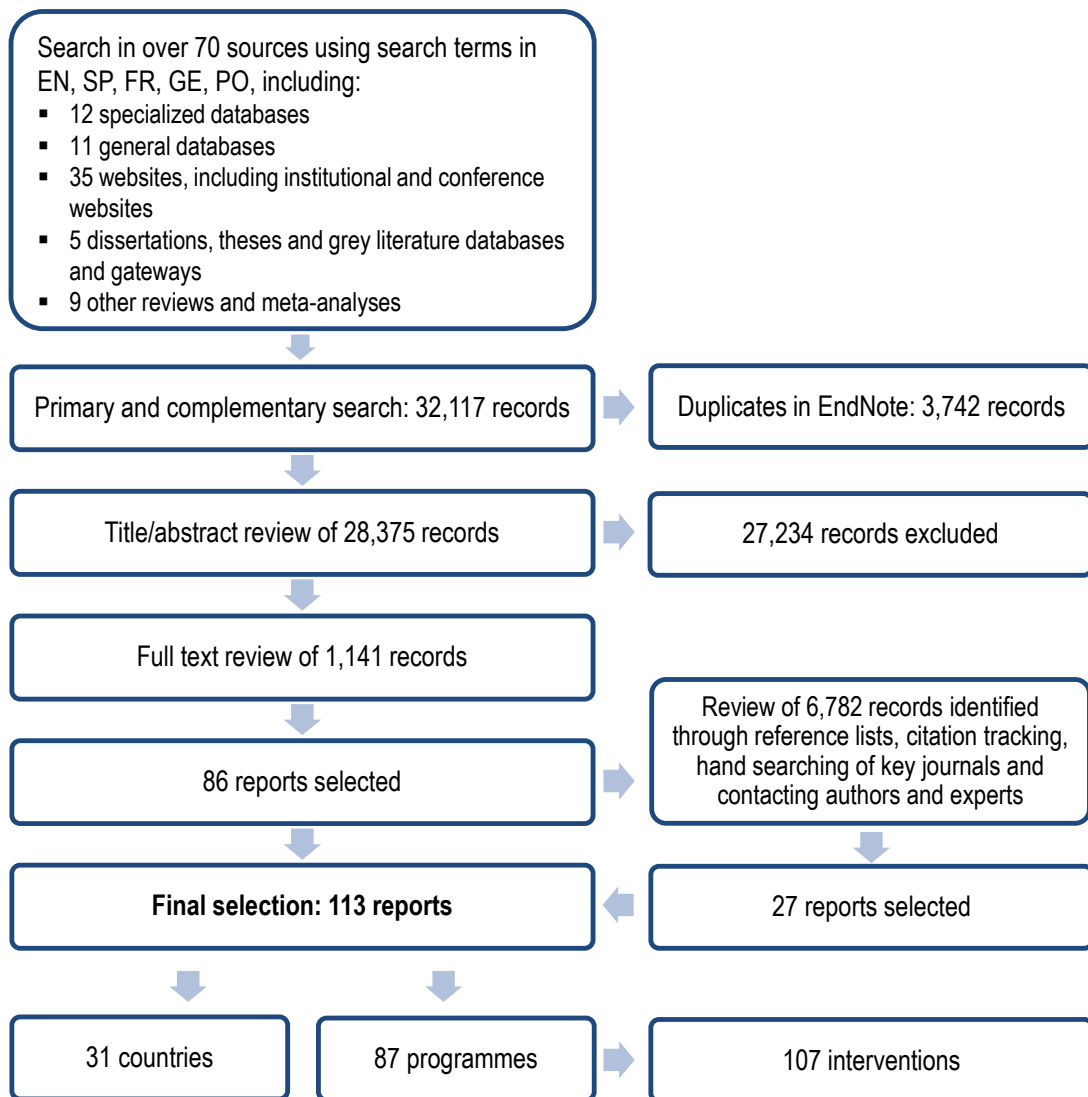
## 4. Study and intervention characteristics

### 4.1 Characteristics of included studies

Table 2 provides an overview of the 113 studies included in the database. The extensive search effort was successful in identifying unpublished studies (Panel (c)). Around one-third of the studies are from peer-reviewed journals, with the remainder split between technical reports from implementing organizations and working papers. Other reviews, in comparison, have a much lower share of unpublished reports; the Systematic Review team believes that the detection of grey literature is one of the strengths of this Review.

Table 2 shows the recent surge of research in the field (Panel (b)). Almost half of the studies in our sample were published after 2010, with 21 studies published in 2014 alone. Most of the latter are working papers.

**Table 2. Study characteristics**

|   |   | n | % |   |   | n | % |
|---|---|---|---|---|---|---|---|
| a) | **Country income level** |   |   | e) | **Timing of Evaluation Follow-up** |   |   |
|   | High-income country | 65 | 58% |   | Follow-up <= 1 year | 58 | 51% |
|   | Low and middle-income country | 48 | 42% |   | Follow-up > 1 year | 71 | 63% |
|   |   |   |   |   |   |   |   |
| b) | **Year of Publication** |   |   | f) | **Sub-group analysis in addition to the overall analysis** |   |   |
|   | 1991-2000 | 14 | 12% |   | Gender Disaggregated | 56 | 50% |
|   | 2001-2005 | 20 | 18% |   | Low-income participants | 4 | 4% |
|   | 2006-2010 | 27 | 24% |   | Education level of participants | 13 | 12% |
|   | 2011-2014 | 52 | 46% |   |   |   |   |
|   |   |   |   | g) | **Outcome Category** |   |   |
| c) | **Type of Publication** |   |   |   | Employment | 98 | 87% |
|   | Peer-Reviewed Journal | 41 | 36% |   | Earnings | 91 | 81% |
|   | Working Paper | 28 | 25% |   | Business Performance | 10 | 9% |
|   | Evaluation / Technical Report | 30 | 27% |   |   |   |   |
|   | Other (Book / Dissertation) | 14 | 12% | h) | **Main Intervention** |   |   |
|   |   |   |   |   | Skills Training | 74 | 65% |
| d) | **Evaluation Design** |   |   |   | Entrepreneurship Promotion | 12 | 11% |
|   | Experimental | 53 | 47% |   | Employment Services | 11 | 10% |
|   | Natural Experiment | 11 | 10% |   | Subsidized Employment | 17 | 15% |
|   | Quasi-experimental | 50 | 44% |   | Unspecified | 9 | 8% |

Note: N=113. Reports may not be exclusive across the different typologies in this table. E.g. One study may estimate multiple outcomes or look into more than one intervention type.

In contrast to other systematic reviews, we seem to find a large share of experimental studies in the form of randomized control trials (RCTs). Many of the results from these RCTs have been published quite recently (66 per cent after 2010) and, hence, are not included in previous reviews. Figure 2 shows the increase in rigorous evidence that is covered in this Systematic Review for the first time. Furthermore, before 2011, most RCTs in the sample were conducted in Organisation for Economic Co-operation and Development (OECD) countries, while the last five years have seen a remarkable increase in RCTs in developing countries. Most notably, in 2014, 12 out of 15 RCTs included in this review were from non-OECD countries; 5 of them evaluated youth employment programs in Sub-Saharan Africa.

Table 2 also provides an overview of the types of outcomes that are evaluated by these studies (Panel (g)). Three out of every four studies in the sample report results for several outcomes. Employment probability[6] is, by far, the most commonly measured and reported outcome within our set of studies, with more than 88 studies providing an estimate of program impact on employment probability. Another 35 studies estimate the effect of an intervention on hours worked. Since only

---

[6] Note employment outcomes include other outcomes aside from probability of employment, such as hours worked.

entreneurship promotion interventions measured business performance outcomes, we have limited observations for these outcomes across the sample.

**Figure 2. Increasing body of evidence**
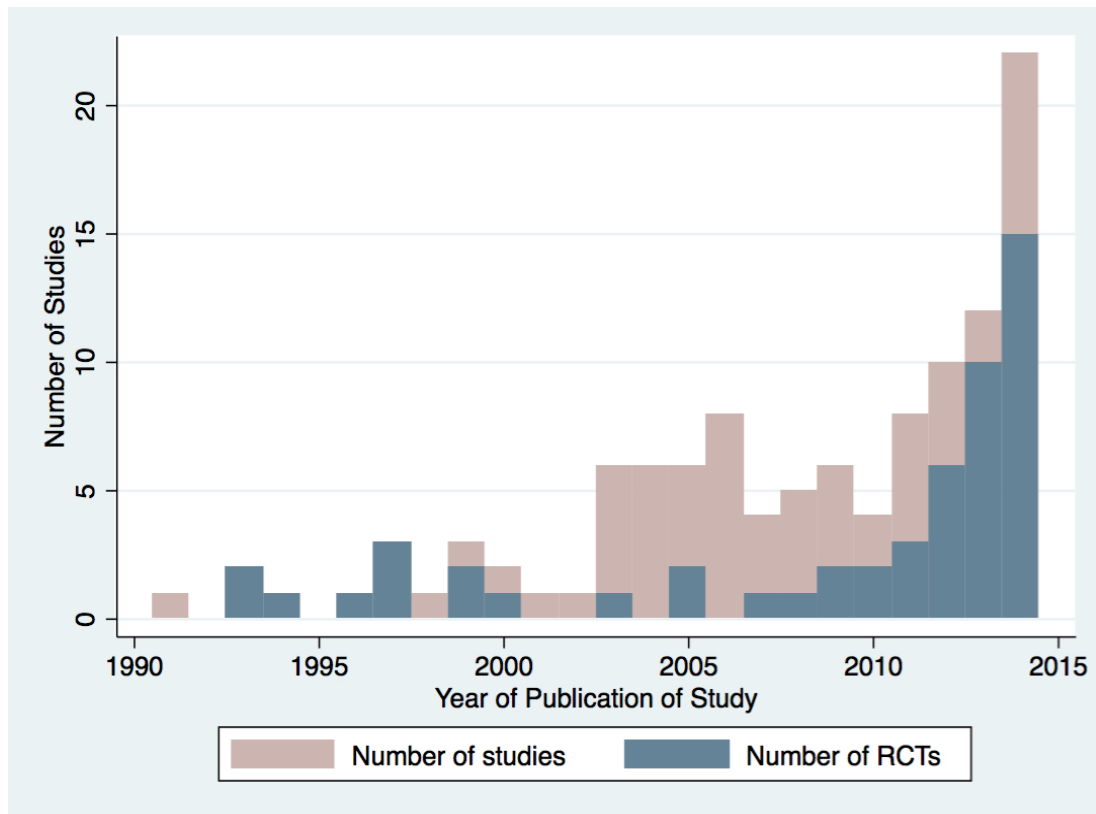


In relation to the evaluation features, 39 studies provide impact estimates at multiple time points. In addition, 71 studies measured changes in outcomes of interest over 12 months after treatment exposure (Panel e)). These longer-term effects are estimated primarily across skills training interventions.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

### 4.2 Characteristics of evaluated interventions

Some of the 87 youth employment programs covered in this review consist of several interventions. For example, a single program may have different tracks – one track providing job placement services and another providing skills training. To provide evidence of which interventions and combinations have been shown to work best, these different types are evaluated separately in the meta-analysis.

Table 3 provides an overview of the characteristics of the 107 interventions in the sample. Many of the interventions provide a combination of different components, whose distribution is reported in the table, as well as the distribution by main category of intervention.

**Table 3. Intervention characteristics**

| | n | % | | n | % |
|---|---|---|---|---|---|
| **a) Main category** | | | **d) Country income level** | | |
| Skills Training | 55 | 51% | High-income country | 60 | 56% |
| Entrepreneurship Promotion | 15 | 14% | Low and middle-income country | 47 | 44% |
| Employment Services | 10 | 9% | | | |
| Subsidized Employment | 21 | 20% | **e) Intervention Region** | | |
| Unspecified | 6 | 6% | OECD | 56 | 52% |
| | | | Sub-Sahara Africa | 15 | 14% |
| **b) Has Component** | | | Europe and Central Asia | 4 | 4% |
| Skills Training | 68 | 64% | Latin America and Caribbean | 22 | 21% |
| Entrepreneurship Promotion | 17 | 16% | Middle East and North Africa | 6 | 6% |
| Employment Services | 40 | 37% | South Asia | 4 | 4% |
| Subsidized Employment | 25 | 23% | | | |
| | | | **f) Scale of Intervention** | | |
| **c) Combinations** | | | National | 59 | 55% |
| Skills Training Only | 32 | 30% | Regional | 21 | 20% |
| Entrepreneurship Promotion Only | 14 | 13% | Local or pilot | 30 | 28% |
| Employment Services Only | 9 | 8% | | | |
| Subsidized Employment Only | 12 | 11% | **g) Intervention features** | | |
| Skills Training & Entrepreneurship Promotion | 1 | 1% | **Target group:** | | |
| Skills Training & Employment Services | 24 | 22% | Women only | 16 | 15% |
| Skills Training & Subsidized Employment | 8 | 7% | Unemployed at intervention start | 48 | 45% |
| Entrepreneurship Promotion & Employment Services | 1 | 1% | Low-Income/Disadvantaged Youth | 45 | 42% |
| Employment Services & Subsidized Employment | 3 | 3% | **Implemented with participation of:** | | |
| Skills Training & Employment Services & more | 3 | 3% | Government | 75 | 70% |
| | | | Private Sector | 63 | 59% |
| | | | NGO/Non-profit | 37 | 35% |
| | | | Mulitlateral organisation | 11 | 10% |

Note: N=107.

Panel (c) in Table 3 splits the interventions by the various combinations of intervention components. It can be seen that over 30 per cent of interventions provide a combination of different components for participants. Among single-component interventions, only skills training programs constitute a similarly large share (30 per cent).

The 113 studies included in this review cover a wide range of countries from all major world regions. As in previous reviews of ALMPs (e.g., Card et al. 2010 and Betcherman et al. 2007), a large share of the interventions that are evaluated have been implemented in OECD countries. Another large share of studies comes from Latin America and the Caribbean, where many countries have experimented with ALMPs – in particular, skills training – and started to evaluate their impact early in

the 1990s using quasi-experimental and experimental designs. As mentioned above, we are able to include a large number of recent ALMP evaluations from Sub-Saharan Africa (15 out of a global total of 105), which contrasts with the number of evaluations found in other developing and emerging regions, such as Europe and Central Asia and South Asia, with four interventions each, respectively. There is no evidence from programs implemented in East Asia and the Pacific. In regard to scale, most interventions have a national coverage. In 30 cases, the evaluations looked into localized interventions implemented as pilots (Panel (f)).

A close examination of program targeting led to the identification of 16 interventions (15 per cent) that served only young women and 45 (42 per cent) that focused exclusively on low-income and disadvantaged youth. Two interventions in 5 targeted unemployed youth. When it comes to implementation, most interventions had public and private entities delivering services, ranging from the provision of in-classroom training to internships agreements and coaching. Non-Governmental Organizations (NGOs) appear as implementers in about one-third of the evaluated interventions in the sample (Panel (g)).

**Skills training interventions:** Education and skills are considered core factors in determining the outcomes for young people in the labor market (Biavaschi et al. 2012), As a result, skills training programs are the most widely used labor market intervention for young people worldwide (Kluve 2010, Betcherman et al. 2007) and are increasingly delivered as complements to other labor market measures (Betcherman et al. 2007; Fares & Puerto 2009). Over half of youth interventions have some skills training components. About 50 percent of the included skills training interventions are in high-income countries, 35 per cent in middle-income countries, and 12 per cent in low-income countries. The objective of these skills training programs is to develop employment-relevant skills of jobseekers to improve their chances of finding stable employment.

Skills training interventions for youth are largely multifaceted with regard to the services provided, skills taught, and settings in which they are taught. About 40 per cent of the interventions complement training with other services. Combinations have paired training with employment services (31 interventions), subsidized employment (8 interventions), and entrepreneurship promotion (4 interventions). While most interventions offer technical skills, non-technical skills (or soft skills) are increasingly embedded in the package, reflecting employers' growing demand for these abilities (Cunningham, Sanchez-Puerta, & Wuermli 2010; Youth Employment Network & IYF 2009). Our sample has 29 interventions that incorporated soft skills within the package.

A significant portion of the interventions relied on a results-based approach for the service providers to improve outcomes for beneficiaries. Of these, 18 out of 32 skills training interventions with available information on payment systems relied on a results- or performance-based approach. In

these cases, a central agency provides traditionally bonus payments to the service providers, contingent on the employment or earnings outcomes of beneficiaries.

About one-third of programs applied participant profiling to match the services provided with participant needs. Profiling involves incorporating a mechanism in the program design to proactively assure that the specific constraints faced by the individuals within target group are being mitigated. This is generally implemented through (1) tailored services for each participant based on an individual assessment done by modeling or interviews by caseworkers or (2) increased emphasis on detailed targeting so a program's services match the needs of its beneficiaries. The Adolescent Girls Employment Initiative (AGEI) in Nepal, for instance, provided technical and life skills training with a comprehensive incentive scheme. Training providers were selected through a competitive bidding process and offered a bonus payment based on the number of trainees who "gainful" employment six months after completing the training and a second bonus for the share of participants who met pre-specified vulnerability criteria and were successfully placed in employment (World Bank 2015).

Of the skills training interventions in lower- and middle-income countries in our sample, almost one-fourth are located in Latin America and the Caribbean and follow the "Jovenes" model.[7] The model, piloted in the 1990s in Chile, combines in-classroom and on-the-job training in a demand-driven fashion. The design ensures private sector involvement in the definition of training content, securing correspondence between the skills taught and those demanded by the productive sector. Moreover, implementation is demand-driven through a stringent, competitive bidding process for the selection of training providers and incentive payments schemes based on trainee outcomes.

**Entrepreneurship promotion interventions:** Programs that focus on forming or supporting youth entrepreneurs are comparatively rare in the dataset, comprising 15 interventions. These interventions are primarily found in low-income (e.g., Uganda and Liberia) and upper-middle-income economies (Peru, Colombia, Tunisia, Bosnia, and Herzegovina).

Similar to skills training interventions, the entrepreneurship promotion interventions included in the sample apply multipronged approaches, often providing technical and non-technical skills training in entrepreneurship in combination with other services. Two-thirds of the evaluated entrepreneurship-oriented interventions fall into this group, most offering a combination of business skills training, business advisory services (including mentoring), and access to finance. An example is the Women's Income Generation Support (WINGS) program in Uganda, which applies this specific combination. Only a small proportion of interventions in our sample (fewer than 3) provided access to finance through credits or grants in a single-pronged manner, or solely business skills training.

---

[7] These programs were implemented in Argentina, Chile, Dominican Republic, Colombia, Peru, and Panama.

**Employment services interventions:** The sample contains 10 employment service interventions embedded in 9 different programs that almost always deliver job counseling, job-search assistance, and/or mentoring services, sometimes complemented by job placement services and/or financial assistance. The only intervention that focuses solely on financial assistance for job search is a subsidized transportation experiment by Franklin (2014) conducted in Addis Ababa, Ethiopia. In contrast to other main intervention types, the employment service interventions examined in this review exhibit a trend toward single-pronged approaches, mostly the provision of job counseling, job-search assistance, and/or mentoring services.

While only 10 interventions in our sample have the main intervention category of employment services, 40 interventions have employment service components. These are most frequently observed as sub-components of either skills training or subsidized employment interventions.

Except for 3 regional or pilot studies that are set in lower- or middle-income countries (Ethiopia, India, Jordan), employment services for youth have so far been evaluated mostly in high-income countries (Finland, France, Germany, Portugal, the U.S.), where they are typically implemented by public employment agencies and operate on a national scale (all interventions are found in high-income/OECD countries).

**Subsidized employment interventions:** The evaluations included in this review contain a heterogeneous sample of 21 subsidized employment interventions that consider reductions in employer social security contributions, reductions in employer labor/wage costs, direct subsidy payments to the individual (e.g., through vouchers), and labor-intensive infrastructure development projects, or any combination of these services, as the predominant category of intervention. The interventions in our sample have been implemented exclusively in high-income economies (Australia, Canada, Chile, France, Germany, Sweden, and the U.S.) and upper-middle income economies (Jordan, South Africa, Tunisia, and Turkey).

Of the subsidized employment interventions, 9 out of 21 are multipronged interventions combining different employment services, such as skills training with job mentoring, all within the framework of a subsidized job. The remaining 12 interventions provide standalone subsidized employment opportunities, such as the Targeted Jobs Tax Credit (TJTC) program in the U.S. In terms of the intervention setting, the majority of the subsidized employment interventions in our sample take place within private enterprises (for example, the Swedish employer-paid payroll tax or the French Contrat Jeune en Entreprise).

## 5. Empirical analysis

## 5.1 Effect size computations and meta-analysis methods

We use two measures created from the treatment effects reported in the evaluations to enable us to both summarize their results succinctly as well as to conduct a multivariate analysis to inform what drives the evaluation results. The two measures are the standardized mean difference (SMD) effect size statistic and a binary variable holding the value of one if a treatment effect is positive and statistically significant (PSS).[8]

The SMD captures the relative magnitude of the treatment effect in a way that is unit-less and hence comparable across outcomes and studies. It is the ratio of the treatment effect (ATT, ITT, or LATE) for a specific outcome relative to the standard deviation of that outcome within the evaluation sample used to estimate the treatment effect. The true effect size ($\theta$) is the mean difference between the treatment ($\mu_t$) and control groups ($\mu_c$) as a proportion of the standard deviation of the outcome variables:

(1) $\theta = \frac{\mu_t - \mu_c}{\sigma}$

The most intuitive form of estimating $\theta$ is applying Cohen's d (Cohen, 1988) defined by

(2) $d = \frac{\overline{Y}_t - \overline{Y}_c}{S_p} = \frac{D}{S_p}$

Where $\overline{Y}_t$ is the mean outcome of the treatment group and $\overline{Y}_c$ that of the control group. The numerator of d captures the treatment effect and is often reported as a treatment effect parameter estimate, such as an average treatment effect on treated (ATT), intention-to-treat effect (ITT), or local average treatment effect (LATE), rather than as differences in means; thus we use D to denote a treatment effect estimate. The denominator of $d$, $S_p$, is the pooled standard deviation from standard deviations of the treatment and control groups and is equivalent to

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

Where $n_c$ and $n_t$ are the sample sizes of the control and treatment groups, respectively, and $S_c$ and $S_t$ are the sample standard deviations of the control and treatment groups, respectively. While $d$ is an intuitive estimator for $\theta$, research shows that $d$ has a bias and overestimates the absolute value of $\theta$ in small samples (Hedge, 1981). For this reason we use a small sample size adjusted estimator referred to as Hedge's g, which is given by

(4) $g = d \left( 1 - \left( \frac{3}{4(n_t + n_c) - 9} \right) \right)$

The standard error of Hedge's d is given by

---

[8] We consider a treatment effect as statistically significant if it has a p-value from a two-tailed test of less than 0.05.

$$(5) \quad SE_g = \sqrt{\left[\frac{n_t+n_c}{n_c*n_t} + \frac{g^2}{2*(n_c+n_t)}\right]}$$

A challenge encountered in the data extraction was the limited information available to compute $g$. Standard deviations for the treatment, control, and total sample groups were often missing, even after attempting to correspond with authors in attempts to acquire this information. In such cases, the standard deviation of the outcome variable was approximated using the formula from Borenstein et al. (2009b)

$$(6) \quad S_p = SE * \sqrt{\frac{n_c*n_t}{n_c+n_t}}$$

where SE is the Standard Error of a means test (e.g. standard error of the regression coefficient estimate).

### 5.2 Creating an effect size per intervention

Most studies provide more than one impact estimator, but in order to arrive at summary effect sizes and aggregated effect sizes we have to combine them to arrive at a single effect size estimate per outcome for each program. Estimating summary effect sizes[9] requires careful estimation to avoid permitting a single group of evaluation survey respondents to influence the aggregate disproportionately. For example, it is common that a treatment effect is reported in a study for the entire, pooled evaluation sample and subsequently reported for sub-groups of the same sample, such as gender. The median number of treatment effect estimates per study is 12, with some reports providing more than 100 estimates. In such instances, there can be a multitude of treatment effects reported for the same group where there is no a priori reason to give preference to one measure over another.

In these scenarios it is possible to mitigate the disproportionate influence on the aggregate effect sizes. One way is by identifying a set of effect sizes that are derived from the same independent group of participants and then, where applicable, selecting the effect sizes for this group where it is possible to establish a preference, for example, keeping only pooled estimates and discarding of sub-group estimates when appropriate. By dropping some of the effect sizes derived from the sample, we remove redundancy.[10,11]

---

[9] These can include intervention level, study level, and other types of aggregates.

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

[11] For the purpose of brevity, we do not include the guidelines we used to drop effect sizes within each group here. This is available upon request.

Once redundant effect sizes are removed, it is likely to still be left with multiple effect sizes for each independent group without clear justification for dropping some over others. In order to arrive at overall effect sizes so as to have one effect size per intervention, we apply the method for combining effect sizes from the same independent population suggested by Borenstein et al. (2009a). The approach is as follows: Let $g_{ij}$ and $SE_{g,ij}$ be the $i^{th}$ effect size, where $i = (1, .., m)$, and its standard error, respectively, for the sample population (e.g. intervention) identified by j. To arrive at a single combined effect size for intervention we take a simple average:

(7) $g_j = \frac{1}{m} \sum_{i=1}^{m} g_{ij}$

And calculate the standard error of $g_j$ by

(8) $SE_{g,j} = \sqrt{\left(\frac{1}{m}\right)^2 \left(\sum_{i=1}^{m} SE_{g,i}^2 + \sum_{i \neq k} \rho_{ik} SE_{g,ij} SE_{g,kj}\right)}$

Where $\rho_{ik}$ is the correlation coefficient between $g_{ij}$ and $g_{kj}$. [12]

**5.3 Creating aggregate effect sizes for groups of interventions:**

With one effect size per intervention, we are able to create aggregate effect sizes for different groupings of interventions as well as an aggregate effect size for the whole sample. Given the breadth of interventions included in our sample it is likely that not all interventions have an identical effect size but, rather, that each intervention's true effect size ($\theta_i$) deviates from the true aggregate effect size for the overall group it belongs to. Furthermore, each observed effect size, estimated by Hedge's g, contains a sampling error and there g will either be less than or greater than $\theta_i$. This can be expressed as

(9) $g_i = \mu + \zeta_i + \varepsilon_i = \theta_i + \varepsilon_i$

Where $\mu$ *is the true aggregate effect size for the group as a whole*, $\zeta_i$ is the deviation of the true effect size of intervention *i* from the group's aggregate effect, and $\varepsilon_i$ is the sampling error. In order to estimate the true aggregate effect size for the group as a whole ($\mu$), equation (9) is estimated using a random-effects regression. Moreover, in order to obtain the most accurate estimate of $\mu$, a weighted random effects regression is applied where the weights are each study's inverse variance[13].

---

[12] Ideally we would estimate $\rho_{ik}$ from the data. However, due to the lack sufficient number of observations an assumption on $\rho_{ik}$ was required. The assumption of $\rho_{ik} = 0$ would likely overestimate precision, while the assumption of $\rho_{ik} = 1$ will likely underestimate precision. We take the more conservative assumption that $\rho_{ik} = 1 \ \forall \ (i,j) \ where \ i \neq k$.

[13] Note that the study's variance corresponds to the term in equation 8 squared.

In the case of PSS, we provide a weighted average of the effect sizes such that each study within the group carries equal weight for the aggregates. As with Hedge's g, the PSS average is based on independent groups created by methods described in Section 5.2.

### 5.4 Multivariate analysis

The multivariate analysis estimates regressions of the effect size on covariates of interest. Building on the results from the earlier expansive version of this Systematic Review (see Kluve et al. 2016), in this paper we conduct an analysis equivalent to what is found in that version of the review, applying random effects inverse-variance weighted random effect regressions on the SMD,[14] and add two more estimation methods: probit regressions of positive and statistically significance estimator and weighted least square regressions with clustered standard errors on Hedge's g.

For analysis of the PSS indicator, $I_{pss,i}$ we estimate the following probit model via Maximum Likelihood

$$(10)\ \text{Prob}(I_{pss,i} = 1 | X_i) = \Phi(X_i \delta)$$

Where $\Phi(.)$ is the Cumulative Distribution Function (CDF) for the standard normal distribution, $X_i$ a vector of the covariates of interest, and $\delta$ the vector of parameter being estimated. The covariates include intervention characteristics, outcome characteristics, and study characteristics. The results from this model are reported as marginal effects.

For the analysis of the Hedge's g, we estimate the following weighted least squares (WLS) regression

$$(11)\ Y_{ij} = X_{ij}\delta + \varepsilon_{ij}$$

Where $Y_{ij}$ is effect size *i* extracted from study *j* and $X_{ij}$ are the relevant covariates values for the sample (or sub-sample) used in estimating. $Y_{ij}$. For all regressions we used the fully disaggregated effect sizes to retain the variation with respect to covariates. We weight the regressions by the inverse of the number of effect size observations contributed by each intervention and cluster standard errors at the intervention level.

It should be noted that the three multivariate methods that are being applied represent different strengths and weaknesses. When taken together, these multivariate methods provide us with a clear view of the findings. The inverse-variance weighted random effects regressions enable more precise estimates from the studies in the meta-analysis to have greater influence on covariate coefficients. However, since these regressions are not weighted by the number of observations that each

---

[14] For a discussion and description of the inverse-variance weighted random effects regressions, see Kluve et al. (2016).

intervention contributes, and we are using fully disaggregated data to enable greatest variation of covariates, it is possible that the interventions that contributed greater number of estimates and sub-group estimates have an unduly disproportionate influence on the regression results. Moreover, as discussed above, there is a very wide range of estimates that each study contributes. On the other hand, the WLS regressions, weighted by the inverse of the number of observations contributed from each intervention, assures that no single intervention has a disproportionate amount of influence on results. Moreover, in the WLS regressions the errors are also being clustered at the intervention level. Given that groups of estimates were taken from different interventions, clustering the errors provides conservative estimates in the likelihood that the error structure contains heteroscedasticity at the intervention level. In contrast to the random effects regression, the WLS regressions are limited in that effect sizes based on estimates with great precision are handled the same as those of lesser precision. Finally, the probit models, while being more general by only capturing the positive direction and statistical significance of estimates, permit us to use a much broader sample since SMDs and Hedge's g require information that is often not available in studies. Only about two-thirds of the estimates that have information to be included in a probit regression also have enough information to calculate or impute an SMD or Hedge's g. Probit regressions on a binary variable similar to the one we apply are at times criticized in that statistical significance is driven by sample size. In order to mitigate this effect, we include sample size as a covariate. Further, we also include the weights and cluster errors as in the weighted least squares regressions, following the same rationale discussed above for the WLS regressions. Card et al. (2015) find that there is effectively no difference in their results derived from meta-analytical models based on effect sizes and models based on sign and significance.

## 6. Empirical results

### 6.1 Results for aggregate effect sizes

The review of the 113 primary studies led to the identification and coding of 3,402 treatment effects. We have the information necessary to construct the PSS indicator variable for a total of 2,932 effect sizes and to compute Hedge's g in 2,059 cases. The number of effect sizes is substantially higher than in other systematic reviews. This, to some extent, is the result of intensive efforts to acquire missing information from authors; primary studies only provided the required information to compute Hedge's g for roughly 13 per cent of reported treatment effects.

Taking all the intervention summary effect sizes, combined to have one per intervention by applying equations (7) and (8), we estimated the inverse variance weighted random effects model shown in equation (11) and created the aggregate PSS for the pooled sample as well for several sub-groups of interventions and outcomes separately. These are all shown in Table 4.

Overall the results from the random effects model show that youth employment interventions have a positive and statistically significant effect but that this does not apply to all sub-groups of

interventions. Slightly more than one-third of the impact estimates are positive and statistically significant. Across the different dimensions captured in Table 4 we observe that the percent of positive statistical significance estimates in this approximate range with a few exceptions that are above forty percent (subsidized employment interventions and some of the design features). The aggregate random effects model (equation 9) is consistent with this, with an overall estimate for Hedge's g of .04 and a 95 percent confidence interval from .02 to .06 standard deviations. Although most subgroups have an aggregate effect size estimate that is positive and statistically significant, we did not find evidence that programs in which the main intervention focus is employment services or subsidized employment (along with interventions whose single focus was not specified) have an effect size that is statistically different from zero. The precision (i.e. the standard error) for these three types of intervention – employment services, subsidized employment, focus not specified – did not hold extreme or outlying values relative to the estimates of other sub-groups, suggesting that the low magnitude of their effects was the reason for the generally insignificant effect sizes. Card et al. (2015) confirm this for employment services using a meta-analysis sample of youth and adults, at all time horizons (short-term, medium-term, and long term impacts).

There is a significant contrast between entrepreneurship promotion interventions and skills training interventions with respect to the unconditional aggregate effect sizes. While both types of interventions have similar proportions of positive and statistically significant impact estimates (.37), the magnitude of aggregate effect size for entrepreneurship programs overshadows that for skills programs (.12 compared to .05) while having a much lower level of precision (standard error of .08 compared to .02). Nonetheless it seems – from the perspective of unconditional estimates – that entrepreneurship promotion programs have greater treatment effect magnitudes. It should be noted that entrepreneurship programs and skills training programs are often implemented in different contexts. There is much variation for skills training programs, partly due to the significant number of skills programs in our sample, in terms of target populations, scale, implementers, and location, among other factors, while entrepreneurship programs in our sample are more homogeneous and have a tendency of being implemented at a smaller scale, targeting poor and disadvantaged populations, and more frequently taking place in lower-income countries. Hence, this difference in effect size magnitudes should be interpreted with caution and requires testing with conditional analysis, which we do below in the multivariate regressions.

In terms of outcomes, the aggregate estimates indicate analogous findings for earnings and employment results, as both have a similar proportion of impact estimates being positive and significant (.32 and .37) as well as effect size (both at .04). Breaking down these into more specific outcomes (where we have sufficiently large number of observations), we find that the impacts of

youth employment programs tend to be the largest on the probability of employment.[15] Youth employment program effects on other employment outcomes, which include labor force participation, unemployment duration, quality of employment,[16] and hours worked, carry smaller values and drive the aggregate effect size for employment outcomes down. On the other hand, we observe similar magnitudes (.03 and .04) for impacts on income and hourly wages (or equivalent), which are the two most common types of earnings outcomes.

The aggregate effect sizes show that several elements of the program design may have a significant impact on the outcomes of programs, specifically including 1) a focus on features that will increase the likelihood that participants will finish and/or perform better in the programs (engagement), 2) proactively using information about individual participants to direct them to the services that best fit their constraints (profiling), and 3) providing service providers with incentives based on results. The contrast is starkest with respect to the proportion of the estimates in evaluations that were positive and statistically significant. Programs with engagement mechanisms had over twice the proportion of positive and statistically significant estimates (.41 compared to .17) while implementing profiling and providing service providers incentives led to having increases of 13 percent and 11 percent, respectively, in large proportion of positive and statistically significant evaluation estimates. However, as with other unconditional estimates discussed in this section, without conditional (multivariate) analysis the weight placed on these differences is limited, and they cannot be generalized, since often profiling and engagement mechanisms are more highly associated with some programs over others. For example, subsidized employment programs by their nature include financial incentives to continue participation, since there is a direct link between payment and attendance.

---

[15] To provide a more representative picture we have combined employment and unemployment probabilities.

[16] Quality of employment captures intervention effects on outcomes such as attaining a fixed contract and receiving benefits.

### Table 4. Effect size summary

**Summary effect size and positive significance estimates**

| | Percent Positive & Statistically Significant | Aggregate* Hedge's G | 95% Confience Interval (G) | |
| --- | --- | --- | --- | --- |
| | | | Lower Bound | Upper Bound |
| **Main category of program** | | | | |
| Skill Training | 0.37 | 0.05 | 0.02 | 0.07 |
| Entrepreneurship Prom. | 0.37 | 0.12 | 0.04 | 0.19 |
| Employment Services | 0.17 | 0.00 | -0.03 | 0.02 |
| Subsidized Employment | 0.41 | 0.02 | -0.01 | 0.05 |
| Unspecified | 0.30 | 0.04 | -0.03 | 0.10 |
| | | | | |
| **Outcomes** | | | | |
| Earnings Outcomes (combined) | 0.32 | 0.04 | 0.02 | 0.05 |
| Employment Outcomes (combined) | 0.37 | 0.04 | 0.03 | 0.06 |
| Specific outcomes: | | | | |
| Employment Probability | 0.39 | 0.06 | 0.04 | 0.08 |
| Number of hours/days worked | 0.26 | 0.03 | 0.00 | 0.06 |
| Income | 0.35 | 0.04 | 0.02 | 0.06 |
| Salary/Wage | 0.35 | 0.03 | 0.02 | 0.05 |
| | | | | |
| **Design features** | | | | |
| Offered extra services: | | | | |
| Yes | 0.37 | 0.05 | 0.01 | 0.08 |
| No | 0.34 | 0.04 | 0.02 | 0.05 |
| Profiled participants: | | | | |
| Yes | 0.41 | 0.06 | 0.02 | 0.10 |
| No | 0.28 | 0.04 | 0.01 | 0.06 |
| Participant engagment mechanism : | | | | |
| Yes | 0.41 | 0.04 | 0.03 | 0.06 |
| No | 0.17 | 0.03 | 0.00 | 0.06 |
| Incentives to service providers | | | | |
| Yes | 0.43 | 0.06 | 0.03 | 0.09 |
| No | 0.32 | 0.05 | 0.01 | 0.08 |
| | | | | |
| **Type of participant** | | | | |
| Male | 0.35 | 0.04 | 0.02 | 0.06 |
| Female | 0.31 | 0.08 | 0.04 | 0.11 |
| Low income/disadvantaged: | | | | |
| Yes | 0.35 | 0.06 | 0.03 | 0.10 |
| No | 0.34 | 0.03 | 0.01 | 0.05 |
| | | | | |
| **Country** | | | | |
| High Income Countries | 0.34 | 0.02 | 0.00 | 0.03 |
| Low & middle income (combined) | 0.37 | 0.09 | 0.06 | 0.12 |
| Low income | 0.34 | 0.15 | 0.10 | 0.21 |
| Middle income | 0.38 | 0.06 | 0.03 | 0.10 |
| Sub-saharan Africa | 0.31 | 0.14 | 0.09 | 0.19 |
| Latin American & Caribbean | 0.50 | 0.10 | 0.05 | 0.15 |
| | | | | |
| **Total** | **0.35** | **0.04** | **0.02** | **0.06** |

*Notes: All aggregated measures estimated are based on independent groups. Aggregate Hedge's G represents estimate of μ from random effects inverse variance weighted regression (see equation 9 above). For variable definitions see Appendix.*

Finally, the aggregate estimates provide initial evidence that youth employment programs in certain macroeconomic contexts and targeting certain populations have better evaluation results in general. These differences present themselves, in varying degrees, in terms of income level and country income, where programs that target low-income populations have more positive effects and programs in high-income countries have lower effects, as compared to those in low- and middle-income countries.

## 6.2 Multivariate meta-regression results

The main objective of this analysis is to look at the effect of different covariates on labor market outcomes, focusing on both the probability of having a positive statistically significant impact and the average standardized mean difference (or effect size) as measured by Hedge's g and SMD and applying the three methods discussed above: weighted least squares, (invariance-variance weighted) random effects, and probit regressions. We first analyze the effects of these covariates on aggregate labor market outcomes – both employment and earnings combined. We then look separately at aggregate labor market outcomes for high-income countries, and low- and middle-income countries. Finally, we estimate separate regressions for the effects on employment and earnings outcomes.

In terms of covariates we focus on four dimensions: the type of program (main intervention category); the country context; individual characteristics of the participants; and general information about program design, including whether it is accompanied by extra services.[17] In addition, we add variables to control for different types of evaluation/publication (RCT, sample size, peer-reviewed journal, estimated unadjusted differences in means, and follow-up after one year).[18] Controlling for sample size is key for estimating the probability of a positive statistically significant impact, since it accounts for larger studies finding statistically significant impacts due to more observations. The results from the pooled regressions are presented in Tables 5 through 7. For brevity Tables 8 to 19, which contain the separate multivariate results for high- and low-income countries and earning and employment outcomes are included in Appendix 3.

### 6.2.1 Pooled regressions

In the estimation results for the pooled sample there is no strong pattern of specific program types or combinations of interventions systematically outperforming others. Some programs, such as a subsidized employment intervention, may at times have smaller effect sizes or a lower likelihood of having a positive statistically significant impact. In general, it therefore appears that there are no consistent statistically significant differences between types of programs across our estimation methods (see Tables 5 to 7). Note at this point, however, that subsequent stratification by country

---

[17] This refers to having services in a category of intervention aside for the main intervention category.
[18] Detailed definitions of select variables are provided in the Appendix.

income group (see section 6.2.2) does reveal impact heterogeneity by intervention type. In particular, skills training and entrepreneurship interventions produce the greatest impacts in low- and middle income countries based on the random effects regressions. The result here suggests that some of the unconditional differences in aggregate effect sizes found in the prior section may not hold when accounting for important contextual features. Following our theory of change, we expect that the effect of a given intervention on labor market outcomes (for youths) will depend on beneficiaries' characteristics and country context, and thus this result is not surprising. For some population groups, for instance, employment services are likely to be the most relevant intervention, whereas for others wage subsidies are more important in creating positive outcomes.

At the same time, the results support the stylized fact observed in the prior section that combining alternative interventions (having extra services) increases the likelihood of success of a given program (And again we preview that this finding is particularly pronounced for the low- and middle income country group, see below). Thus, once controlling for relevant covariates, offering services that complement the main intervention can increase the magnitude of the effect from .05 to .09 standard deviations (Tables 5 and 6) and the probability of success by 14 to 21 percent (Table 7). Again, this is consistent with our theory of change. Most population groups are likely to face multiple constraints affecting their likelihood of getting a job, the types of jobs they get, and associated earnings. The Economic Empowerment of Adolescent Girls (EPAG) in Liberia, which showed strong evaluation results, combined six months of classroom-based training followed by employment services through six months of follow-up support in entering wage employment or starting a business. Also having a positive impact, the Teenage Parent Demonstration in the U.S. aimed to address multiple constraints on youth employment outcomes. Mandatory for teenage mothers receiving welfare, it provided a wide array of services that were employment oriented including enrollment in alternative education programs, participation in job training, job search guidance, and employment. Programs that only address one of these constraints are less likely to have an impact than those that are multi-pronged. It is not possible, however, to identify the one specific multi-component combination that always works; the types of interventions that are needed seem to be specific to the individual and the country context. Even within the Teenage Parent Demonstration, implementation in some sites included additional specific services like transportation stipends which were not broadly applied.

The pooled results also suggest that, systematically, programs implemented in low- and middle-income countries have been more likely to succeed and have larger effect sizes than programs implemented in high-income countries. For instance, the likelihood of success in high-income countries can be between 21 to 32 percent lower than in low-middle income countries. One explanation is that differences in performance are explained by differences in the severity of the constraints facing beneficiaries to access jobs and improve their earnings. High-income countries

could be dealing, on average, with population groups that are harder to serve or for whom there are fewer job opportunities relative to those facing beneficiaries in low- and middle-income countries. Moreover, the larger impacts can be associated with low- and middle-income countries having a low starting point. We also conjecture that programs in the latter set of countries tend to be newer and might have benefited from better designs and technical innovations.

Regarding beneficiaries, we do not find differences in program performance based on their age or gender. However, we find that programs that focus on vulnerable workers – either low-income workers or youth at risk – have larger effect sizes (by a value in the range of .05 to .07, see Tables 5 and 6), although we do not observe consistent effects on the likelihood of having a positive and statistically significant impact. This is consistent with the aggregates discussed above (Table 4). While the difference in the likelihood of having a positive effect was small, with low-income (more vulnerable) participants slightly higher at .35 versus .34, the magnitude of the aggregate effect size was twice that of programs not targeting low-income and disadvantaged participants. The most probable explanation for this result is that programs that deal with targeted groups of vulnerable beneficiaries are better able to target the necessary interventions, not only in terms of the types of interventions offered but also their content and/or design.

Another important result from the analysis is that program design and implementation matter. As discussed in the previous section, the majority of youth employment programs have not been able to have statistically significant positive impacts. This, most likely, does not imply that the programs are not needed, but instead that there are problems in terms of design and implementation. The programs might not have positive results because they are not using the right combination of interventions, given the constraints facing beneficiaries. They also may be affected by implementation challenges, e.g.: was the program implemented as designed? Did the program participant group reflect the targeted population? Did participants complete all components?

Although our set of variables to control for program design and implementation is limited,[19], we are able to capture effects regarding three elements that international experiences suggest are key determinants of program success: profiling systems, systems to monitor and follow up on beneficiaries, and incentives for providers.

---

[19] See Appendix 2 for the definitions of select indicators.

**Table 5. Weighted Least Squares Hedge's g regressions: pooled sample**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| <u>Main intervention category</u> <u>(base=skills training)</u> | | | | | | |
| Entr. Prom. | 0.032 | 0.012 | 0.050 | 0.053 | 0.026 | 0.055 |
| | [0.62] | [0.23] | [0.97] | [1.37] | [0.59] | [1.43] |
| Empl. Serv. | -0.064** | -0.041* | 0.023 | 0.074** | 0.000 | 0.026 |
| | [2.57] | [1.82] | [0.62] | [2.07] | [0.01] | [0.57] |
| Subs. Empl. | -0.036 | -0.013 | -0.071* | 0.017 | -0.059 | 0.007 |
| | [1.28] | [0.48] | [1.91] | [0.48] | [1.15] | [0.19] |
| Unspecified | 0.002 | 0.024 | -0.002 | -0.045 | -0.013 | -0.122 |
| | [0.04] | [0.61] | [0.03] | [0.78] | [0.24] | [1.61] |
| <u>Evaluation features</u> | | | | | | |
| Log Evaluation Sample Size | | -0.015* | -0.012 | -0.020** | -0.004 | -0.004 |
| | | [1.94] | [1.31] | [2.38] | [0.51] | [0.53] |
| Publication Peer-Reviewed | | -0.033 | -0.042* | -0.059** | -0.064** | -0.072*** |
| | | [1.45] | [1.86] | [2.23] | [2.49] | [2.88] |
| Eval. design:RCT | | -0.029 | -0.056 | -0.048* | -0.035 | -0.030 |
| | | [0.96] | [1.61] | [1.74] | [1.29] | [1.27] |
| High income country | | -0.042 | -0.106*** | -0.133*** | -0.088*** | -0.135*** |
| | | [1.50] | [3.41] | [5.52] | [2.81] | [4.80] |
| <u>Program design features</u> | | | | | | |
| Additional services | | | -0.007 | 0.051** | 0.024 | 0.067** |
| | | | [0.23] | [2.07] | [0.83] | [2.33] |
| Participant profiling | | | 0.062* | 0.074*** | 0.069* | 0.053 |
| | | | [1.80] | [2.80] | [1.93] | [1.57] |
| Participant engagement mechanism | | | 0.092** | 0.073*** | 0.063** | 0.088*** |
| | | | [2.60] | [2.92] | [2.09] | [3.15] |
| Incentives for service providers | | | 0.067* | 0.046 | 0.084** | 0.026 |
| | | | [1.78] | [1.36] | [2.43] | [0.75] |
| Program has soft skills training | | | | | -0.005 | 0.031 |
| | | | | | [0.11] | [0.90] |
| <u>Outcome characteristics</u> | | | | | | |
| Employment outcome (base=earnings outcome) | | | | 0.012 | | 0.001 |
| | | | | [0.65] | | [0.04] |
| Estimated unadjusted difference in means | | | | -0.105* | | -0.189** |
| | | | | [1.75] | | [2.24] |
| Measured over one year after exit from program | | | | 0.088*** | | 0.102*** |
| | | | | [3.66] | | [4.13] |
| <u>Target/evaluation group</u> | | | | | | |
| Low income / disadvantaged | | | | 0.061** | | 0.018 |
| | | | | [2.47] | | [0.59] |
| Male (base=male and female combined) | | | | -0.041 | | -0.036 |
| | | | | [1.68] | | [1.44] |
| Female (base=male and female combined) | | | | 0.012 | | 0.027 |
| | | | | [0.51] | | [1.26] |
| Younger Participants | | | | -0.025 | | 0.048 |
| | | | | [0.77] | | [1.32] |
| <u>Type of implementer (base=Private and public sector joinly implement)</u> | | | | | | |
| Government only | | | | | 0.007 | -0.023 |
| | | | | | [0.19] | [0.49] |
| Private sector only | | | | | 0.023 | 0.050 |
| | | | | | [0.39] | [1.30] |
| Constant | 0.072*** | 0.223*** | 0.137 | 0.127 | 0.066 | -0.003 |
| | [3.99] | [2.92] | [1.53] | [1.63] | [0.91] | [0.03] |
| R2 | 0.02 | 0.06 | 0.17 | 0.28 | 0.15 | 0.33 |
| N | 2,001 | 2,001 | 1,369 | 1,024 | 1,250 | 962 |
| Number of studies: | 96 | 96 | 70 | 58 | 62 | 51 |
| Number of interventions: | 97 | 97 | 58 | 42 | 51 | 37 |

30

Notes: Regressions weighted by inverse of number of observations coming from each intervention and errors clustered at the intervention level. Marginal effects evaluated at the variable means reported for probit regressions. For variable definitions see appendix.

**Table 6. Random Effects SMD regressions: pooled sample**

| Main intervention category (base=skills training) | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Entr. Prom. | 0.037** | 0.014 | 0.001 | 0.076*** | 0.061*** | 0.097*** |
|  | [2.15] | [0.81] | [0.06] | [3.36] | [2.74] | [3.87] |
| Empl. Serv. | -0.027*** | -0.009 | -0.003 | 0.115*** | -0.008 | 0.067** |
|  | [2.78] | [0.91] | [0.24] | [6.10] | [0.36] | [2.25] |
| Subs. Empl. | -0.031*** | -0.003 | -0.047*** | 0.007 | -0.045** | -0.018 |
|  | [3.72] | [0.29] | [3.08] | [0.36] | [2.32] | [0.67] |
| Unspecified | -0.040*** | -0.029** | -0.046*** | -0.064*** | -0.026* | -0.049** |
|  | [3.09] | [2.18] | [3.12] | [3.50] | [1.70] | [2.33] |
| **Evaluation features** | | | | | | |
| Log Evaluation Sample Size |  | -0.010*** | -0.006** | -0.015*** | -0.004 | -0.003 |
|  |  | [4.83] | [2.22] | [4.24] | [1.58] | [0.67] |
| Publication Peer-Reviewed |  | 0.001 | -0.015 | -0.012 | -0.022** | -0.019** |
|  |  | [0.07] | [1.56] | [1.40] | [2.25] | [2.14] |
| Eval. design:RCT |  | -0.013* | -0.019* | -0.050*** | -0.071*** | -0.076*** |
|  |  | [1.74] | [1.86] | [4.56] | [6.20] | [6.24] |
| High income country |  | -0.037*** | -0.084*** | -0.137*** | -0.093*** | -0.152*** |
|  |  | [4.72] | [7.22] | [11.77] | [7.25] | [10.80] |
| **Program design features** | | | | | | |
| Additional services |  |  | -0.001 | 0.061*** | 0.063*** | 0.096*** |
|  |  |  | [0.09] | [5.03] | [4.14] | [5.81] |
| Participant profiling |  |  | 0.032*** | 0.048*** | 0.058*** | 0.023 |
|  |  |  | [2.79] | [3.60] | [4.64] | [1.31] |
| Participant engagement mechanism |  |  | 0.064*** | 0.072*** | 0.058*** | 0.107*** |
|  |  |  | [5.12] | [6.30] | [4.61] | [8.09] |
| Incentives for service providers |  |  | 0.031*** | 0.043*** | 0.065*** | 0.058*** |
|  |  |  | [2.96] | [3.93] | [5.48] | [4.32] |
| Program has soft skills training |  |  |  |  | -0.025* | -0.003 |
|  |  |  |  |  | [1.79] | [0.23] |
| **Outcome characteristics** | | | | | | |
| Employment outcome (base=earnings outcome) |  |  |  | -0.020*** |  | -0.018*** |
|  |  |  |  | [2.85] |  | [2.60] |
| Estimated unadjusted difference in means |  |  |  | -0.009 |  | -0.018 |
|  |  |  |  | [0.65] |  | [1.33] |
| Measured over one year after exit from program |  |  |  | 0.046*** |  | 0.050*** |
|  |  |  |  | [5.51] |  | [5.93] |
| **Target/evaluation group** | | | | | | |
| Low income / disadvantaged |  |  |  | 0.071*** |  | 0.047*** |
|  |  |  |  | [5.55] |  | [3.11] |
| Male (base=male and female combined) |  |  |  | -0.007 |  | 0.002 |
|  |  |  |  | [0.68] |  | [0.15] |
| Female (base=male and female combined) |  |  |  | -0.005 |  | 0.000 |
|  |  |  |  | [0.55] |  | [0.06] |
| Younger Participants |  |  |  | -0.013 |  | 0.010 |
|  |  |  |  | [1.52] |  | [1.18] |
| **Type of implementer (base=Private and public sector joinly implement)** | | | | | | |
| Government only |  |  |  |  | -0.026* | -0.042** |
|  |  |  |  |  | [1.93] | [2.13] |
| Private sector only |  |  |  |  | 0.036** | 0.033* |
|  |  |  |  |  | [2.07] | [1.88] |
| Constant | 0.048*** | 0.159*** | 0.107*** | 0.099*** | 0.095*** | 0.003 |
|  | [12.66] | [8.77] | [3.97] | [2.65] | [3.46] | [0.07] |
| R2 | . | . | . | . | . | . |
| N | 2,059 | 2,000 | 1,369 | 1,024 | 1,250 | 962 |
| Number of studies: | 98 | 96 | 70 | 58 | 62 | 51 |
| Number of interventions: | 97 | 97 | 58 | 42 | 51 | 37 |

**Table 7. Positive & statistically significant probit regressions: pooled sample**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category (base=skills training) | | | | | | |
| Entr. Prom. | 0.028 | 0.013 | 0.042 | 0.054 | -0.007 | 0.002 |
| | [0.25] | [0.12] | [0.35] | [0.42] | [0.08] | [0.03] |
| Empl. Serv. | -0.139 | -0.123 | -0.048 | 0.067 | -0.366** | -0.277* |
| | [1.45] | [1.34] | [0.39] | [0.45] | [2.27] | [1.84] |
| Subs. Empl. | 0.035 | -0.106 | -0.235*** | -0.053 | -0.342*** | -0.195 |
| | [0.42] | [1.37] | [2.60] | [0.38] | [3.17] | [1.61] |
| Unspecified | -0.044 | 0.094 | 0.021 | 0.092 | 0.063 | -0.231 |
| | [0.46] | [1.01] | [0.14] | [0.49] | [0.43] | [1.28] |
| Evaluation features | | | | | | |
| Log Evaluation Sample Size | | 0.034* | 0.035* | 0.028 | 0.061*** | 0.067*** |
| | | [1.81] | [1.72] | [1.00] | [3.79] | [2.82] |
| Publication Peer-Reviewed | | -0.052 | -0.050 | -0.048 | -0.033 | -0.074 |
| | | [0.71] | [0.64] | [0.54] | [0.43] | [1.03] |
| Eval. design:RCT | | -0.075 | -0.095 | -0.054 | -0.119 | -0.020 |
| | | [1.16] | [1.22] | [0.74] | [1.60] | [0.27] |
| High income country | | -0.094 | -0.214** | -0.324*** | -0.208** | -0.288*** |
| | | [1.37] | [2.32] | [3.05] | [2.57] | [3.39] |
| Program design features | | | | | | |
| Additional services | | | -0.050 | 0.011 | 0.140* | 0.214** |
| | | | [0.58] | [0.14] | [1.67] | [2.49] |
| Participant profiling | | | 0.168** | 0.196*** | 0.173** | 0.086 |
| | | | [2.36] | [2.58] | [2.42] | [1.16] |
| Participant engagement mechanism | | | 0.273*** | 0.222** | 0.228** | 0.241** |
| | | | [2.68] | [2.09] | [2.51] | [2.52] |
| Incentives for service providers | | | 0.121 | 0.161 | 0.107 | 0.006 |
| | | | [1.30] | [1.30] | [1.36] | [0.06] |
| Program has soft skills training | | | | | -0.226** | -0.201** |
| | | | | | [2.50] | [2.10] |
| Outcome characteristics | | | | | | |
| Employment outcome (base=earnings outcome) | | | | -0.014 | | -0.049 |
| | | | | [0.27] | | [1.08] |
| Estimated unadjusted difference in means | | | | 0.050 | | -0.218 |
| | | | | [0.34] | | [1.46] |
| Measured over one year after exit from program | | | | 0.187*** | | 0.253*** |
| | | | | [2.86] | | [4.22] |
| Target/evaluation group | | | | | | |
| Low income / disadvantaged | | | | 0.077 | | 0.102 |
| | | | | [0.77] | | [1.02] |
| Male (base=male and female combined) | | | | -0.047 | | -0.091 |
| | | | | [0.74] | | [1.60] |
| Female (base=male and female combined) | | | | -0.001 | | -0.042 |
| | | | | [0.01] | | [0.85] |
| Younger Participants | | | | -0.026 | | 0.128 |
| | | | | [0.24] | | [1.34] |
| Type of implementer (base=Private and public sector joinly implement) | | | | | | |
| Government only | | | | | -0.019 | 0.000 |
| | | | | | [0.17] | [0.00] |
| Private sector only | | | | | 0.224* | 0.299** |
| | | | | | [1.87] | [2.51] |
| N | 2,932 | 2,058 | 1,441 | 1,068 | 1,323 | 1,007 |
| Number of studies: | 104 | 94 | 67 | 55 | 60 | 49 |
| Number of interventions: | 105 | 98 | 58 | 42 | 51 | 37 |

Notes: see Table 5 notes.

32

There is strong evidence that programs that profile beneficiaries and/or have systems in place to monitor/follow beneficiaries and keep them in the program are more likely to succeed and have larger effect sizes. Profiling, for example, allows program managers to better understand and respond to the needs/constraints facing different groups of beneficiaries. A program proactively taking information from participants to enable them to succeed, such as the Programa de Capacitación Jóvenes con Futuro (JCF) in Colombia or the Galpao Program in Brazil, uses information about participant aptitudes to place them in the type of training where candidates may be most likely to succeed. However, profiling does not necessarily imply that each individual needs to have a differentiated treatment. Instead, the ability to group beneficiaries into broad categories – from those who require minimal support to those who are hard to serve – seems to be critical for the performance of the program. The Adolescent Girls Employment Initiative (AGEI) in Nepal applied an innovative approach by incorporating a results-based system whereby training providers received different bonus payments for successfully placing participants from specific vulnerable populations in "gainful" employment.

Similarly, the continuous follow-up and engagement of beneficiaries is important for program performance. Following up beneficiaries is not only necessary to assess whether a given intervention is delivering the expected results, but also to obtain timely feedback in terms of whether adjustments to the intervention are required, both in the composition and intensity of different services. In general, this requires having in place adequate monitoring and evaluation systems. For example, the Women's Income Generation Support (WINGS) in Uganda, which focused on entrepreneurship promotion activities, required that staff maintain close supervision of business activities for the first few business cycles and provide advice on meeting market challenges and implementing sound business practices. While continuous follow-ups and monitoring also address the problem of beneficiary drop-out, providing incentives for beneficiaries to stay in the program can also be used for this purpose. The Satya/Pratham program in India, which provided young women with specific skills training, required beneficiaries to deposit Rs 50 per month for continuing in the program. This required participants to be ready to commit a total of Rs 300 for the entire duration of the training program with a promise that upon program completion, they would be repaid Rs 350.

The evidence regarding incentives for providers continues to support what we saw in the aggregate group statistics (Table 4) but is less persuasive regarding its influence on the proportion of results that are positive and statistically significant. In principle, programs that pay service providers based on results and performance are more likely to have positive impacts. There are different examples of how contracts and payment systems can be structured in this way. The 2008 Employment Package, a Turkish subsidized employment program, combined a payroll tax subsidy to employers for

newly hired employees with cuts in their social security payments of employers. For Contrat Jeune en Entreprise in France, a subsidized employment program with a negligible impact based on its evaluation results, firms were entitled to claim a subsidy whenever they hired an eligible young worker on an open-ended contract. Unfortunately, the information about contracting and payment system in our sample of programs is quite sparse. We are thus only able to distinguish between programs that provide some type of incentive system to providers explicitly mentioned in the program materials and programs that do not have such an incentive.

Another important finding in terms of program design is that training programs that focus on soft or non-cognitive skills may not be the silver bullet that many expected them to be. Several studies have emphasized the role that non-cognitive skills have in determining labor market outcomes. Programs such as the Jovenes in Latin America became well known, in part, because of their focus on training that aims to improve non-cognitive skills. Our results suggest, however, that other things being equal, programs that include training in socio-emotional and behavioral skills do not necessarily do better than other programs. On the contrary, once we control for key design features, these programs seem to have been less likely to achieve positive outcomes.

The interpretation of the result is that, if programs are not set up to address the needs of beneficiaries through good profiling and follow-up systems, appropriate contracting and payments systems, and a diversified package of interventions, simply adding a soft skills training component is unlikely to make a difference. This may be due to a broad set of non-technical skills components that have been applied that have made it difficult to determine empirically which are the most likely to have positive impacts, and for which target group. The JOBSTART program in the U.S. offers an illustration of a training program with soft skills components and limited impacts. JOBSTART applied an intensive exposure model through which it provided school dropouts training which included work-readiness, life, and communication skills. Another example is Entra 21 in Latin America, which was implemented in several countries by different organizations and had a design that incorporated "life skills." However, "life skills" was defined broadly, ranging from writing resumés and job search assistance to reproductive health (Alzua 2007).

Finally, the results suggest that programs implemented by the private sector alone, as opposed to joint public-private implementation or sole implementation by the government, seem to lead to moderately larger gains (as shown in the random effect regressions and probit regressions, Tables 6 and 7). The estimates show that programs solely implemented by the private sector have greater outcomes by .03 to .04 standard deviations, compared to jointly implemented programs (Table 6) and an increased chance of having a positive and significant effect of 20 to 30 percentage points (Table 7). The main interpretation of this result is that programs managed by the private sector may be more likely to have built-in incentives to respond to the needs of employers and job seekers.

In terms of the type of evaluation/publication, the estimation results do not suggest many strong patterns. The only variable that correlates significantly is the time of the follow-up survey relative to the baseline survey: program evaluations that estimate impacts over one year after the intervention are more likely to identify significant results (a finding in line with Card et al. 2010, 2015) This indicates that, in most cases, ALMPs do not have immediate effects either on employment rates or on earnings. Time is required to properly implement the program and, subsequently, to deliver the necessary services. Also as anticipated, the sample size is positively correlated with the likelihood of positive statistical significance.

### 6.2.2 Differences by country income level

Stratifying the sample by country income group addresses a key dimension of impact heterogeneity of youth interventions. Specifically, we look at two groups, high-income countries and low-middle income countries, and present empirical results for both. Clearly, making comparisons between high- and middle-low income countries is challenging, because sample sizes and the number of programs within each group is reduced. At the same time, the types of interventions in the groups of countries are markedly different. For instance, in high-income countries employment services and subsidized employment are more frequent. Middle-low income countries, on the other hand, tend to focus more on entrepreneurship programs. Many of the result patterns reported by country group below are consistent with those found in the pooled results, so we focus on highlighting those for which we observe relevant heterogeneity between the two country groups or results deviating from the pooled results. The regression results are found in Tables 8 to 13 in Appendix 3.

One result from looking at country groups is that, whereas impact differences between program types is not very strong overall, some pronounced intervention patterns emerge. In some specifications for high-income countries, for instance, we see that wage subsidies tend to be less successful than other programs, which is consistent with other reviews of this type of program (see Almeida 2012). In contrast, for low- and middle-income countries, some specifications indicate that entrepreneurship programs fare better, as well as skills training programs. Moreover, for low- and middle-income countries, many of the studies indicate that programs that add extra services to the main intervention tend to do better, whereas this pattern is not pronounced for high-income countries (Tables 10 and 11).

There are also different findings in terms of program design and implementation. In high income countries there is no conclusive evidence about the role/importance of profiling but, on the other hand, we observe evidence that providing incentives to service providers increase both the proportion of positive evaluation estimates and the effect size magnitude (Tables 8 to 10). In contrast, the evidence that incentives to service providers have a large impact is tenuous in the results for low and middle income countries. For high income countries we also find that programs implemented

solely by the private sector or public sector are no better or, in some specifications, achieve lower outcomes than those jointly implemented in a public-private partnership. The result for low- and middle-income countries reverse this pattern and show there is an advantage in outcomes for those programs that were solely implemented by the private sector, which have greater impacts between .07 and .11 standard deviations (Tables 11 and 12) and an increased probability of having a positive and statistically significant outcome of between 38 and 56 percent. In low income countries there is evidence about the importance of profiling and follow-up and engagement systems, albeit no consistent evidence regarding the role of incentives for providers.

### 6.2.3 Differentiated effects on employment and earnings

Most of the results discussed above in terms of the role of different types of interventions, the type of beneficiaries, and program design hold when we look separately at effects on employment and earnings, although there were some differences regarding how youth employment programs affect earnings and employment outcomes. The results for the separate regressions for employment and earnings outcomes are found in Tables 14 to 19 in Appendix 3.

In the case of employment outcomes, there is also some evidence suggesting that entrepreneurship programs have tended to do better, but only in terms of the magnitude of the impact. One explanation can be that in the case of programs supporting self-employment and/or small scale entrepreneurship the creation of jobs is more directly under the control of the program, relative to programs that need to connect individuals to wage employment. The important question, however, is how sustainable these jobs are in the long term. They can be created more easily but their retention rate can also be lower. Unfortunately, our analysis cannot address this question.

There is strong evidence that subsidized employment programs are less effective than skills training programs in increasing participant incomes, while there is no available evidence for either type of intervention for employment outcomes. This is a reasonable result, since the constraints that skills training programs address are the limits to available human capital and the lack of participant skill sets that are demanded by the private sector, and hence earn a greater wage, while subsidized employment programs focus on providing participants the opportunity to enter the job market, where they may build networks and develop a career. For example, the Ninaweza Youth Empowerment Program in Kenya targeted young women in informal settlements in Nairobi and included class-room training and employment support. While the program demonstrated overall positive impacts in both employment and earnings, the evaluation found earnings effects to be statistically significant more often than employment effects. While for some treatment groups there was no statistically significant impact on employment, the authors find it was actually the type of employment that made a difference between treatment and control. The higher human capital provided through technology training led to program beneficiaries more often attaining full-time positions and employment in more productive

modern sectors, while control group participants were more often employed as casual laborers. In many cases, subsidized employment programs often focus directly on reducing unemployment, such as the 2008 Employment Package (Turkey) which was created in order to increase the formal employment of youth, with the ancillary goals of both improving employment rates and reducing the prevalence of workers who are not registered by their employer with a social security institution.

While most results regarding program design features were consistent with those found in the pooled regression, we observe that incorporating well-defined participant engagement mechanisms into the design plays a strong role in improving earnings outcomes. While the effects of engagement mechanisms are most statistically significant in both magnitude (effect size) and predicting the percent of positive and statistically significant impact estimates, the effect is statistically significant across specifications for earnings. This may demonstrate that providing strong guidance to program participants through case workers or mentors, which is among the most common form of monitoring implemented in employment programs, allows participants to take advantage of better opportunities in the market, such as higher income employment or employment where their skills are more highly valued.


**7. Conclusion**

Labor market prospects for youth are a cause of concern for policymakers worldwide. As a result, many programs have been implemented to bring youth into the labor market, connect them to jobs, increase their earnings, and/or help them set up a business. However, the majority of these programs have not been properly evaluated and therefore there has been, to date, limited information available about the types of interventions that work and the reasons why.

This paper aims to improve our understanding of the effectiveness of youth employment programs, focusing on skills training, entrepreneurship promotion, employment services, and subsidized employment. To this end, we identified all relevant empirical studies with rigorous evaluations produced over the last ten years. We created a database with 113 studies and coded information about impacts on employment and earnings, beneficiaries, and program design and implementation. Our analysis looks at Standardized Mean Difference by programs and models of having positive and statistically significant evaluation estimates.

The results of the analysis show that, overall, youth programs have positive effect sizes. However, just above one third of the programs in our database display statistically significant positive effects. The interpretation of this result is not that youth employment programs, in most cases, do not work. Instead, much of the difference in performance seems to be related to design and implementation factors, as well as the characteristics of the country and population of beneficiaries.

In general, programs have been more successful in middle- and low-income countries. This may be because the programs' investments more strongly affect target populations that are very vulnerable (low-skilled, low-income). We also conjecture that because these programs are more recent they might have benefited from innovations in design and implementation. There are, however, no systematic differences in performance by the age or gender of the beneficiary.

We do not find that certain types of programs, or combinations of programs, systematically outperform others. For instance, training programs do not necessarily perform better than employment services, and combining training with wage subsidies does not always improve the chances of success. Some programs, however, seem to be better at affecting certain outcomes than others. There is some evidence that entrepreneurship programs can do better at increasing employment rates than other programs, perhaps because by design the intervention directly "creates" a job for each beneficiary. At the same time, employment services and wage subsidies are less likely to succeed if the goal is to increase earnings.

We find evidence that programs that integrate multiple interventions are more likely to have a positive impact, in particular in low-middle income countries. Hence, while there is no specific combination of services that always works, programs that add complementary services to the main intervention, regardless of what those are, tend to do better. The interpretation is that the success of youth employment programs rests on their ability to respond to multiple needs/constraints facing a heterogeneous group of beneficiaries. In other words, the efficient portfolio of services is specific to the population of beneficiaries. Programs that target multiple categories of beneficiaries are likely to need multiple portfolios of interventions. One implication is that successful youth employment programs will need to be able to offer a comprehensive set of interventions, from training, to counseling, intermediation, and income support.

We also find evidence about the importance of profiling and individualized follow-up and monitoring systems in determining program performance. Consistent with the finding above, programs that are able to profile beneficiaries are also able to better respond to their needs. Profiling does not necessarily imply having services tailored to each individual. Instead, it often involves being able to group beneficiaries in broad categories, from those requiring minimal support to the most disadvantaged or hard-to-reach. Efficient follow-up systems and incentives to keep youth in the program are, not surprisingly, also critical for success. This often implies having in place robust monitoring and evaluation systems.

Evidence about the importance of incentive systems for services providers is also positive, although weaker. At least conceptually, programs that pay providers based on performance are more likely to achieve their objectives. It remains unclear, however, what the best types of contracting and payments systems are and how these need to be adjusted depending on the context. Unfortunately, in

our analysis, we are unable to capture these differences in design and therefore only code whether a given program offers some type of incentive at all. Still, in our full specifications, we find that incentives for providers are associated with a higher chance of success and larger size effects.

Finally, we find that programs solely implemented/managed by the private sector lead to moderately higher gains. This suggests that this type of program implementation may better at providing built-in incentives to program providers to respond to the needs of employers and job-seekers.

The findings of this systematic review also bring to light the importance of including information on programs costs in impact evaluations. Our findings provide insights into mechanisms and designs features that may improve youth employment program performance, but the general unavailability of standardized information on program costs limits our ability to make absolute statements about the efficient allocation of resources available for improving outcomes. The infrequent presentation of standardized program costs measures alongside impact evaluation results may be the largest gap in our knowledge of what works for youth employment programs.

**References**

3ie (2013). Impact Evaluation Glossary. Available at: http://ridie.3ieimpact.org/downloads/Impact%20Evaluation%20Glossary.pdf [13 Oct. 2014].

Almeida, Rita, Jere Behrman, David Robalino (2012). The right skills for the job? Rethinking training policies for workers. Human Development Perspectives. Washington, DC: World Bank.

Almeida, R., L. Orr, David Robalino (2014). Wage subsidies in developing countries as a tool to build human capital: Design and implementation issues. IZA *Journal of Labor Policy*, 3, 12.

Baird, S., F.H.G. Ferreira, F. H. G., B. Özler, M. Woolcock (2013). Relative effectiveness of conditional and unconditional cash transfers for schooling outcomes in developing countries: A systematic review. Campbell Systematic Reviews, 8.

Bell, D., D. Blanchflower (2009). What should be done about rising unemployment in the UK? IZA Discussion Paper No. 4040.

Betcherman, G., A. Dar, K. Olivas (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., M. Godfrey, S. Puerto, S. F. Rother, F. A. Stavreska (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., W. Eichhorst, C. Giulietti, M.J. Kendzia, A. Muravyev, J. Pieters, N. Rodriguez Planas, R. Schmidl, R., K.F. Zimmermann (2012). Youth unemployment and vocational training. IZA Discussion Paper No. 6890.

Boone, J., J.C. van Ours (2004). Effective Active Labor Market Polices. IZA Discussion Paper No. 1335.

Bonnal, L., D. Fougere, A. Sérandon (1997). Evaluating the impact of French employment policies on individual labor market histories. *Review of Economic Studies* 64, 4, 683–713.

Borenstein, M., H. Cooper, L.V. Hedges, J.C. Valentine (2009a). Effect sizes for continuous data. In H. Cooper, V. Hedges, & J. C. Valentine (Eds.). *The handbook of research synthesis and meta-analysis*, 221–235. New York, NY: Russell Sage Foundation.

Borenstein, M., L.V. Hedges, J.P.T. Higgins, H.R. Rothstein (2009b). *Introduction to Meta-Analysis.* Hoboken, NJ: John Wiley & Sons, Ltd.

Brewer, L. (2013). Enhancing youth employability: What? Why? and How? Guide to core work skills. International Labour Office, Skills and Employability Department. Geneva: ILO.

Brodaty, T., B. Crepon, D. Fougere (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.

Campbell Collaboration. (2014). Campbell Systematic Reviews: Policies and Guidelines: Supplement 1. DOI: 10.4073/csrs.2014.1 Available at http://www.campbellcollaboration.org/resources/research/the_production.php [14 Oct. 2014]. Oslo, Norway: Campbell Collaboration.

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

Card, D., J. Kluve, A. Weber (2015). What works? A meta analysis of recent active labor market program evaluations, NBER Working Paper 21431.

Cho, Y., M. Honorati (2013). Entrepreneurship programs in developing countries: A meta regression analysis. World Bank Policy Research Working Paper No. 6402. Washington, DC: World Bank.

ho, Y., D.N. Margolis, D. Newhouse, D. Robalino (2012). Labor Markets in Low and Middle-Income Countries: Trends and Implications for Social Protection and Labor Policies. Human Development Social Protection Discussion Paper Series. Washington, DC: World Bank.

Cho, Y., D. Robalino, S. Watson (2014). Supporting self-employment and small-scale entrepreneurship: Potential programs to improve livelihoods for vulnerable workers. Washington, DC: World Bank.

Cunningham, W., L. McGinnis, R.G. Verdú, C. Tesliuc, D. Verner (2008). Youth at risk in Latin America and the Caribbean. Washington, DC: World Bank.

Cunningham, W., M.L. Sanchez-Puerta, A. Wuermli (2010). Active Labor Market Programs for Youth: A Framework to Guide Youth Employment Interventions. Washington, DC: World Bank.

Dar, A., Z. Tzannatos (1999). Active labor market programs: A review of the evidence from evaluations. Social Protection Discussion Paper Series No. 9901. Washington, DC: World Bank.

eeks, J. J., D.G. Altman, M.J. Bradburn (2008). Statistical methods for examining heterogeneity and combining results from several studies in meta-analysis. *Systematic Reviews in Health Care: Meta-Analysis in Context*, second edition, 285–312. Hoboken, NJ: John Wiley and Sons.

Doucouliagos, C.H., T.D. Stanley (2009). Publication selection bias in minimum-wage research? A meta-regression analysis. *British Journal of Industrial Relations* 47(2): 406–428.

Doucouliagos, C. H., T.D. Stanley, M. Giles (2012). Are estimates of the value of a statistical life exaggerated?, *Journal of Health Economics* 31(1): 197-206.

Duvendack, M., R. Palmer-Jones, J.G. Copestake, L. Hooper, Y. Loke, N. Rao (2011). What is the evidence of the impact of microfinance on the well-being of poor people? London: EPPI-Centre, Social Science Research Unit, Institute of Education, University of London.

Duvendack, M., J.G. Hombrados, R. Palmer-Jones, H. Waddington (2012). Assessing 'what works' in international development: meta-analysis for sophisticated dummies. *Journal of Development Effectiveness* 4(3), 456-471.

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

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

Eichhorst, W., U. Rinne (2015). An Assessment of the Youth Employment Inventory and Implications for Germany's Development Policy. IZA Research Report No. 67.

Fares, J., O.S. Puerto (2009). Towards comprehensive training. Washington, DC: World Bank.

Fields, G.S. (2011). Labor market analysis for developing countries. *Labor Economics* 18, S16-S22.

Franco, A., N. Malhotra, G. Simonovits (2014). Publication bias in the social sciences: Unlocking the file drawer. *Science* 345(6203), 1502–1505.

Galasso, E., M. Ravallion, M., A. Salvia (2001). Assisting the transition from workfare to work: Argentina's proempleo experiment. Washington, DC: World Bank.

Gindling, T.H., D. Newhouse (2012). Self-employment in the developing world. World Bank Policy Research Working Paper No. 6201. Washington, DC: World Bank.

Grimm, M., A.L. Paffhausen (2014). Do interventions targeted at micro-entrepreneurs and small and medium-sized firms create jobs? A systematic review of the evidence for low and middle income countries. *Labor Economics* 32 (2015) 67–85.

Grosh, M., C. del Ninno, E. Tesliuc, A. Ouerghi (2008). For protection and promotion: The design and implementation of effective safety nets. Washington, DC: World Bank.

Heckman, J.J., R.J. LaLonde, J.A. Smith (1999). The economics and econometrics of active labor market programs, in Ashenfelter, O. and D. Card (eds.), *Handbook of Labor Economics 3*. Amsterdam: Elsevier.

Hedges, L., I. Olkin (1985). *Statistical models for meta-analysis*. New York: Academic Press.

Hedges, LV, T.D. Pigott (2001). The power of statistical tests in meta-analysis. Psychological Methods, 6, 203-217.IEG. (2013). World Bank and IFC support for youth employment programs. Washington, DC: World Bank.

International Labour Office (2012). Global Employment Trends for Youth 2012. Geneva: ILO, 2012.

International Labour Office (2015a). Global Employment Trends for Youth 2015. Scaling up investments in decent jobs for youth. Geneva: ILO, 2015.

International Labour Office (2015b). Small and medium-sized enterprises and decent and productive employment creation. Report IV to the International Labour Conference, 104th Session, 2015. Geneva: ILO.

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

Jameel Poverty Action Lab (J-PAL) (2013). J-PAL Youth Initiative Review Paper. Cambridge, MA: Abdul Latif Jameel Poverty Action Lab.

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

Kluve, J. (2006). The effectiveness of European active labor market policy. IZA Discussion Paper No. 2018.

Kluve, J. (2010). The effectiveness of European active labor market programs. *Labor Economics* 17, 6, 904–918.

Kluve, J., H. Lehmann, C.M. Schmidt (1999). Active Labor Market Policies in Poland: Human Capital Enhancement, Stigmatization or Benefit Churning. IZA Discussion Paper No. 30.

Kluve, J., S. Puerto, D. Robalino, J.M. Romero, F. Rother, J. Stoeterau, F. Weidenkaff, M. Witte (2014). Protocol: Interventions to improve labor market outcomes of youth: a systematic review of training, entrepreneurship promotion, employment services, mentoring, and subsidized employment interventions. Oslo, Norway: Campbell Collaboration.

Kluve, J., S. Puerto, D. Robalino, J.M. Romero, F. Rother, J. Stoeterau, F. Weidenkaff, M. Witte (2016, forthcoming). Interventions to improve labor market outcomes of youth: a systematic review of

training, entrepreneurship promotion, employment services, and subsidized employment interventions. Oslo, Norway: Campbell Collaboration.

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

Kritikos, A. (2014). Entrepreneurs and their impact on jobs and economic growth. IZA World of Labor 2014: 8.

Leroy, J. L., P. Gadsden, P., M. Guijarro, M. (2012). The impact of daycare programs on child health, nutrition and development in developing countries: A systematic review. *Journal of Development Effectiveness* 4, 3, 472–496.

Lieuw-Kie-Song, M., K. Philip, M. Tsukamoto, M. Van Imschoot (2010). Toward the right to work: Innovations in Public Employment Programs (IPEP). Geneva: ILO.

Lipsey, M.W., D.B. Wilson (2001). Practical meta-analysis. Applied Social Research Methods Series, Volume 49. SAGE Publications.

Majurin, E. (2014). Start and improve your business (SIYB): implementation guide. Geneva: ILO.

Mishra, A., D. Cameron (2013). Register of Impact Evaluation Published Studies (RIEPS) Protocol. Washington, DC: International Initiative for Impact Evaluation.

Nataraj, S., F. Perez-Arce, K.B. Kumar, S.V. Srinivasan (2013). The impact of labor market regulation on employment in low-income countries: A meta-analysis. *Journal of Economic Surveys*.

OECD. (2013). Glossary of statistical terms, Active Labor Market Programs. Available at: http://stats.oecd.org/glossary/detail.asp?ID=28 [13 Oct. 2014].

O'Leary, J.C. (1998). Evaluating The Effectiveness of Active Labor Programs in Poland. Upjohn Institute Technical Report No.98-012; June 1998.

Pasali, S. (2015) Youth Employment in Africa: What Works and Why? Mimeo.

Piza, C., T. Cravo, L. Taylor, L. Gonzalez, I. Musse, I. Furtado, A.C. Sierra, S. Abdelnour (2016). The Impact of Business Support Services for Small and Medium Enterprises on Firm Performance in Low- and Middle-Income Countries: A Systematic Review. Campbell Systematic Reviews 2016:1.

Robalino, D., D. Margolis, F. Rother, D. Newhouse, M. Lundberg (2013). Youth employment: a human development agenda for the next decade. Social Protection and Labor discussion paper ; no. 1308. Washington, DC: World Bank Group.

Schoof, U. (2006). Stimulating youth entrepreneurship: Barriers and incentives to enterprise start-ups by young people. SEED Working Paper No. 76. Geneva: ILO.

Sparreboom, T., A. Staneva (2014). Is education the solution to decent work for youth in developing economies?: identifying qualifications mismatch from 28 school-to-work transition surveys. Geneva: ILO.

Stanley, T.D. (2013). Does economics add up? An introduction to meta-regression analysis. *European Journal of Economics and Economic Policies: Intervention* 10, 2, 207–220.

Tripney, J., J. Hombrados, M. Newman, K. Hovish, C. Brown, K. Steinka-Fry, E. Wilkey (2013). Technical and vocational education and training (TVET) interventions to improve the employability and employment of young people in low- and middle-income countries: A systematic review. Campbell Systematic Reviews 9.

Valerio, A., B. Parton, A. Robb, A. (2014). Entrepreneurship Education and Training Programs around the World: Dimensions for Success. Washington, DC: World Bank.

van Lieshout, S., M. Sievers, M. Aliyev, M. (2012). Start and Improve Your Business Global Tracer Study 2011. Geneva: ILO.

Waddington, H., J.G. Hombrados (2012). Tool to assess risk of bias and internal validity of social experiments and quasi-experiments. Mimeo, 3ie.

Walther, A., A. Pohl, A. (2005). Thematic Study on Policy Measures Concerning Disadvantaged Youth. Study commissioned by European Commission, DG Employment and Social Affairs in framework of Community Action Program to Combat Social Exclusion 2002–2006. Brussels: European Commission.

World Bank. (2015). The role of skills training for youth employment in Nepal: an impact evaluation of the employment fund. Adolescent Girls Initiative (AGI); February 2015. Results Series. Washington, DC: World Bank Group.

YEN & IYF (2009). Private sector demand for youth labor in Ghana and Senegal – Ghana and Senegal study findings. Washington, DC: Youth Employment Network and International Youth Foundation, World Bank.

**Appendix 1: List of included evaluations**

Adoho, Franck, Shubha Chakravarty, Dala Korkoyah Jr., Mattias Lundberg, Afia Tasneem (2014). The impact of an adolescent girls employment program: The EPAG project in Liberia. Washington, DC: World Bank Policy Research Working Paper 6832.

Aedo, Cristián, Marcelo Pizarro Valdivia (2004). Rentabilidad económica del programa de capacitación laboral de jóvenes Chile Joven. Report.

Aedo, Cristián, Sergio Nuñez (2004). The Impact of Training Policies in Latin America and the Caribbean: The Case of Programa Joven. Washington, DC: Inter-American Development Bank, Research Department.

Ahmed, Ali, Shubha Chakravarty, Mattias Lundberg, Plamen Nikolov (2014). The Role of Training Programs for Youth Employment in Nepal: Interim Impact Evaluation Report on the Employment Fund.

Almeida, Rita K., Mahdi Barouni, Stefanie Brodmann, Rebekka Grun, Patrick Premand (2012). Entrepreneurship Training and Self-Employment among University Graduates: Evidence from a Randomized Trial in Tunisia. Washington, DC: World Bank.

Alvares de Azevedo, Thomaz, Jeff Davis, Munene Charles (2013). Testing What Works In Youth Employment: Evaluating Kenya's Ninaweza Program. Volume 1: A Summative Report and Volume 2: Appendices. Global Partnership for Youth Employment (GPYE).

Alzuá, María L., Pablo Brassiolo (2006). The Impact of Training policies in Argentina: An Evaluation of Proyecto Joven. Washington, DC: Inter-American Development Bank, Office of Evaluation and Oversight (OVE).

Alzúa, María L., Paula Nahirñak, Belissario Alvarez de Toledo (2007). Evaluation of Entra 21 using quantitative and qualitative data. Q-Squared Working Paper (41).

Alzúa, María L., Guillermo Cruces, Carolina Lopez Erazo (2013). Youth training programs beyond employment. Evidence from a randomized controlled trial.

Arandarenko, Mihail, Aleksandra Nojkovic, Marko Vladisavljevic (2014). Impact Evaluation of Employment Programs Targeting Disadvantaged Youth in Serbia. Belgrade, Serbia: Foundation for the Advancement of Economics.

Attanasio, Orazio, Adriana Kugler, Costas Meghir (2011). Subsidizing Vocational Training for Disadvantaged Youth in Colombia: Evidence from a Randomized Trial. *American Economic Journal: Applied Economics* 3, 3, 188–220.

Bampasidou, Maria (2012). Unbundling the degree effect in a job training program for disadvantaged youths. 3569526 Ph.D., University of Florida.

Bampasidou, Maria, Carlos A. Flores, Alfonso Flores-Lagunes, Daniel J. Parisian (2014). The Role of Degree Attainment in the Differential Impact of Job Corps on Adolescents and Young Adults. Factors Affecting Worker Well-being: The Impact of Change in the Labor Market Emerald Group Publishing Limited, *Research in Labor Economics* 40, 113–156.

Bandiera, Oriana, Niklas Buehren, Robin Burgess, Markus Goldstein, Selim Gulesci, Imran Rasul, Munshi Sulaiman (2014). Women's empowerment in action: evidence from a randomized control trial in Africa.

Barza, Rachael (2011). Essays in Labor Economics. Ph.D., Harvard University.

Blanco, German, Carlos A. Flores, Alfonso Flores-Lagunes (2013a). Bounds on Average and Quantile Treatment Effects of Job Corps Training on Wages. *Journal of Human Resources* 48, 3, 659–701.

Blanco, German, Carlos A. Flores, Alfonso Flores-Lagunes (2013b). The Effects of Job Corps Training on Wages of Adolescents and Young Adults. *American Economic Review* 103, 3, 418–22.

Blasco, Jaume, David Casado, Jordi Sanz, Federico A. Todeschini (2014). Evaluación del impacto de los programas PCPI del Servicio de Empleo de Catalunya. *Revista de evaluación de programas y políticas públicas/Journal of public programs and policy evaluation* (*REPPP*) 2, 187–217.

Blattman, Christopher, Jeannie Annan, Eric Green, Julian Jamison (2013). Building Women's Economic and Social Empowerment Through Enterprise. An Experimental Assessment of the Women's Income Generating Support (WINGS) Program in Uganda. –New Haven, CT: Innovations for Poverty Action (IPA).

Blattman, Christopher, Eric P. Green, Julian Jamison, Jeannie Annan (2014). Employing and empowering marginalized women: A randomized trial of microenterprise assistance.

Blattman, Christopher, Nathan Fiala, Sebastian Martinez (2013). Generating Skilled Self-Employment in Developing Countries: Experimental Evidence from Uganda. *Quarterly Journal of Economics* forthcoming.

Bloom, Howard S., Larry L. Orr, Stephen H. Bell, George Cave, Fred Doolittle, Winston Lin, Johannes M. Bos (1997). The Benefits and Costs of JTPA Title II-A Programs: Key Findings from the National Job Training Partnership Act Study. *Journal of Human Resources* 32, 3, 549–576.

Blundell, Richard, Monica Costa Dias, Costas Meghir, John Van Reenen (2004). Evaluating the Employment Impact of a Mandatory Job Search Program. *Journal of the European Economic Association* 2, 4, 569–606.

Brodaty, Thomas (2007). La politique active de l'emploi en faveur des jeunes: Les dispositifs ont-ils touche leur cible? (Are French Youth Employment Active Labor Market Policies Well Targeted? with English summary). *Annales d'Economie et de Statistique* 85, 3–40.

Broecke, Stijn (2013). Tackling graduate unemployment in North Africa through employment subsidies: A look at the SIVP program in Tunisia. IZA *Journal of Labor Policy* 2, 9, 1–19.

Bruhn, Miriam, Bilal Zia (2013). Stimulating managerial capital in emerging markets: the impact of business training for young entrepreneurs. *Journal of Development Effectiveness* 5, 2, 232–266.

Calero, Carla, Carlos Henrique Corseuil, Veronica Gonzales, Jochen Kluve, Yuri Soares (2014). Can arts-based interventions enhance labor market outcomes among youth? Evidence from a randomized trial in Rio de Janeiro.

Caliendo, Marco, Steffen Künn, Ricarda Schmidl (2011). Fighting Youth Unemployment: The Effects of Active Labor Market Policies. Bonn, Germany: Forschungsinstitut zur Zukunft der Arbeit (IZA)/Institute for the Study of Labor.

Cansino Muñoz-Repiso, José Manuel, Antonio Sánchez Braza (2011). Effectiveness of Public Training Programs Reducing the Time Needed to Find a Job/Eficacia de los programas públicos de formación en la reducción del tiempo necesario para encontrar un empleo. *Estudios de Economía Aplicada* 29, 1, 1–26.

Card, David, Pablo Ibarrarán, Ferdinando Regalia, David Rosas-Shady, Yuri Soares (2011). The Labor Market Impacts of Youth Training in the Dominican Republic. *Journal of Labor Economics* 29, 2, 267–300.

Carling, Kenneth, Laura Larsson (2005). Does early intervention help the unemployed youth? *Labor Economics* 12, 3, 301–319.

Cave, George, Hans Bos, Fred Doolittle, Cyril Toussaint (1993). Jobstart: Final Report on a Program for School Dropouts.

Centeno, Luis, Mario Centeno, Alvaro A. Novo (2009). Evaluating job-search programs for old and young individuals: Heterogeneous impact on unemployment duration. *Labour Economics* 16, 1, 12–25.

Centeno, Mário, Álvaro A. Novo (2006). Avaliação Do Impacto De Um Programa De Apoio À Procura De Emprego. *Boletim Económico*, Banco de Portugal Verão.

Chang, Young Eun, Aletha C. Huston, Danielle A. Crosby, Lisa A. Gennetian (2007). The Effects of Welfare and Employment Programs on Children's Participation in Head Start. *Economics of Education Review* 26, 1, 17–32.

Charpail, Christine, Tristan Klein, Serge Zilberman (2005). Évaluation des politiques d'emploi: la deuxième génération des panels des bénéficiaires. DARES Document d'Etudes No. 95.

Chen, Xuan (2013). Partial identification of average treatment effects in program evaluation: Theory and applications. Ph.D., Economics, University of Miami.

Cho, Yoon Y., Davie Kalomba, Ahmed Mushfiq Mobarak, Victor Orozco (2013). Gender Differences in the Effects of Vocational Training: Constraints on Women and Drop-Out Behavior. Washington, DC: World Bank.

Chong, Alberto, Jose Galdo (2006). Training Quality and Earnings: The Effects of Competition on the Provision of Public-Sponsored Training Programs. Mimeo, Washington, DC:  Inter-American Development Bank.

Chong, Alberto, Jose Galdo (2012). Does the quality of public-sponsored training programs matter? Evidence from bidding processes data. *Labor Economics* 19, 970–986.

Chong, Alberto, Jose Galdo, Jaime Saavedra (2008). Informality and productivity in the labor market in Peru." *Journal of Economic Policy Reform* 11, 4, 229–245.

Corseuil, Foguel, Gonzaga, Pontual Ribeiro (2014). Youth Turnover in Brazil: Job and Worker Flows and an Evaluation of a Youth-Targeted Training Program.

Costa Dias, Monica, Hidehiko Ichimura, Gerard J. van den Berg (2013). Treatment Evaluation With Selective Participation and Ineligibles. *Journal of the American Statistical Association* 108, 502, 441–455.

Crépon, Bruno, Esther Duflo, Marc Gurgand, Roland Rathelot, Philippe Zamora (2013). Do Labor Market Policies have Displacement Effects? Evidence from a Clustered Randomized Experiment. *Quarterly Journal of Economics*.

Crépon, Bruno, Esther Duflo, Elise Huillery, William Parienté, Juliette Seban (2014). Les effets du dispositif d'accompagnement à la création d'entreprise CréaJeunes: résultats d'une expérience contrôlée Rapport d'évaluation.

De Giorgi, Giacomo (2005). The New Deal for Young People Five Years On. *Fiscal Studies* 26, 3, 371–383.

Díaz, Juan J., Miguel Jaramillo (2006). An evaluation of the Peruvian 'youth labor training program' – Projoven. Working Paper: OVE/WP-10/06. Washington, DC: Inter-American Development Bank.

Egebark, Johan, Niklas Kaunitz (2014). Do payroll tax cuts raise youth employment? Stockholm, Sweden: Research Institute of Industrial Economics (IFN).

Ehlert, Christoph, Jochen Kluve, Sandra Schaffner (2012). Temporary work as an active labor market policy: Evaluating an innovative program for disadvantaged youths. *Economics Bulletin* 32, 2, 1765–1773.

Elías, Víctor, Fernanda Ruiz Núñez, Ricardo D. Cossa, D. Bravo (2004). An Econometric Cost-Benefit Analysis of Argentina's Youth Training Program. Washington, DC: Inter-American Development Bank.

Espinoza Peña, Henry (2010). Impact evaluation of a job-training program for disadvantaged youths: The case of Projoven. Ph.D., Maastricht University.

Fein, David, Erik Beecroft, John D. Blomquist (1994). The Ohio Transitions To Independence Demonstration: Final Impacts For Jobs And Work Choice. Cambridge, MA: Abt Associates.

Fiala, Nathan (2014). Stimulating Microenterprise Growth: Results from a Loans, Grants and Training Experiment in Uganda.

Flores-Lagunes, Alfonso, Arturo Gonzalez, and Todd Neumann (2010). Learning but Not Earning? The Impactof Job Corps Training for Hispanic Youth. *Economic Inquiry* 48, 3, 651–667.

Franklin, Simon (2014). Job Search, Transport Costs and Youth Unemployment: Evidence from Urban Ethiopia.

Frölich, Markus, Martin Huber (2014). Direct and Indirect Treatment Effects: Causal Chains and Mediation Analysis with Instrumental Variables. IZA Discussion paper No. 8280.

Frumento, Paolo, Fabrizia Mealli, Barbara Pacini, Donald B. Rubin (2012). Evaluating the effect of training on wages in the presence of noncompliance, nonemployment, and missing outcome data. *Journal of the American Statistical Union* 107, 498, 450–466.

Galdo, Jose, Miguel Jaramillo, Veronica Monatalva (2008). Household Wealth and Heterogeneous Impacts of a Market-Based Training Program: the Case of PROJOVEN in Peru.

General Accounting Office (1996). Job Training Partnership Act: Long-term earnings and employment outcomes. New York, NY: United Nations General Accounting Office.

Gong, Tao (2005). Two essays on youths' labor supply. Ph.D., Economics and Finance, Middle Tennessee State University.

Griffith, James (2001). An approach to evaluating school-to-work initiatives: post-secondary activities of high school graduates of work-based learning. *Journal of Vocational Education and Training* 53, 1, 37–60.

Groh, Matthew, Nandini Krishnan, David McKenzie, Tara Vishwanath (2012). Soft skills or hard cash? The impact of training and wage subsidy programs on female youth employment in Jordan. Policy Research Working Paper Series: 6141. Washington, DC: World Bank.

Groh, Matthew, David McKenzie, Nour Shammout, Tara Vishwanath (2014). Testing the Importance of Search Frictions, Matching, and Reservation Prestige Through Randomized Experiments in Jordan. Impact Evaluation series. Policy Research working paper no. WPS 7030. Washington, DC: World Bank.

Hämäläinen, Kari, Ulla Hämäläinen, Juha Tuomala (2014). The Labor Market Impacts of Youth Guarantee Evidence from a Finnish Reform.

Hall, Georgia S. (2000). The impact of a school-to-work program ProTech: A study of post-high school outcomes. Ph.D., Boston College.

Heckman, James J., Hidehiko Ichimura, Petra E. Todd (1997). Matching As an Econometric Evaluation Estimator: Evidence from Evaluating a Job Training Program. *Review of Economic Studies* 64, 4, 605–654.

Heckman, James J., Jeffrey Smith (1999). The Pre-Program Earnings Dip and the Determinants of Participation in a Social Program: Implications for Simple Program Evaluation Strategies. *Economic Journal* 109, 457, 313–348.

Heckman, James J., Jeffrey Smith (2000). The Sensitivity of Experimental Impact Estimates (Evidence from the National JTPA Study). In *Youth Employment and Joblessness in Advanced Countries*, David G. Blanchflower and Richard B. Freeman (Eds.), 331–356. Cambridge, MA: National Bureau of Economic Research.

Hicks, Joan Hamory, Michael Kremer, Isaac Mbiti, Edward Miguel (2013). Vocational Education in Kenya: Evidence from A Randomized Evaluation Among Youth.

Hollenbeck, Kevin, W.J. Huang (2006). Net Impact and Benefit-Cost Estimates of the Workforce Development System in Washington State. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Hollenbeck, Kevin, Richard J. Willke (1991). The Employment and Earnings Impacts of the Targeted Jobs Tax Credit. Staff Working Papers. Kalamazoo, MI: W.E. Upjohn Institute for Employment Research.

Ibarrarán, Pablo, David Rosas (2007). Impact Evaluation of a Labor Training Program in Panama.

Ibarrarán, Pablo, Laura Ripani, Bibiana Taboada, Juan M. Villa, Brígida García (2014). Life Skills, Employability and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design. IZA *Journal of Labor & Development* 3, 10.

Jackson, Russell H., R. Malené Dixon, Ann McCoy, Carol Pistorino, Paul Zador, Cynthia Thomas, John Lopdell, Juanita Lucas-McLean, Frank Bennici, Andy Sum, Neeta Fogg, Ron D'Amico, Andrew Weigand, Lee Bruno (2007). Youth Opportunity Grant Initiative: Impact and synthesis report. Washington, DC: U.S. Department of Labor – Decision Information Resources.

Jaramillo, Miguel, Sandro Parodi (2003). Jóvenes emprendedores. Lima: Instituto Apoyo.

Jaramillo, Miguel, Sandro Parodi (2005). Evaluación Experimental del Programa "Formación de Líderes Empresariales en el Departamento de Huancavelica."

Jensen, Robert (2012). Do Labor Market Opportunities Affect Young Women's Work and Family Decisions? Experimental Evidence from India. *Quarterly Journal of Economics* 127, 2, 753–792.

Klasen, Stephan, Carlos Villalobos Barría (2014). The Impact of SENAI's vocational training program on employment, wages, and mobility in Brazil: What lessons for Sub Saharan Africa? Manchester, UK: International Research Initiative on Brazil and Africa (IRIBA).

Knight, G. (2002). Evaluation of the Australian Wage Subsidy Special Youth Employment and Training Program, SYETP. Ph.D. thesis, University of Sydney.

Kornfeld, Robert, Howard S. Bloom (1999). "Measuring Program Impacts on Earnings and Employment: do unemployment insurance wage reports from employers agree with surveys of individuals?" *Journal of Labor Economics* 17, 1, 168–197.

Larsson, Laura (2003). Evaluation of Swedish Youth Labor Market Programs. *Journal of Human Resources* 38, 4, 891–927.

Lee, David S. (2009). Training, Wages, and Sample Selection: Estimating Sharp Bounds on Treatment Effects. *Review of Economic Studies* 76, 1071–1102.

Levinsohn, James, Neil Rankin, Gareth Roberts, Volker Schoer (2014). Wage subsidies and youth employment in South Africa: Evidence from a randomised control trial.

Maitra, Pushkar, and Subha Mani (2013). "Learning and earning: evidence from a randomized evaluation in India." Discussion Paper Series 2, Fordham University, Department of Economics.

Maynard, Rebecca, Walter Nicholson, Anu Rangarajan (1993). Breaking the cycle of poverty: The effectiveness of mandatory services for welfare-dependent teenage parents. Mathematica Policy Research.

McClanahan, Wendy S., Cynthia L. Sipe, Thomas J. Smith (2004). Enriching Summer Work: An Evaluation of the Summer Career Exploration Program. Public/Private Ventures.

Meager, Nigel, Peter Bates, Marc Cowling (2003). An evaluation of business start-up support for young people. *National Institute Economic Review* 186, 59–72.

Mensch, Barbara S., Monica J. Grant, Mary P. Sebastian, Paul C. Hewett, Dale Huntington (2004). The Effect of a Livelihoods Intervention in an Urban Slum in India: Do Vocational Counseling and Training Alter the Attitudes and Behavior of Adolescent Girls? In Working Paper No. 194: New York: Population Council.

Millenky, Megan, Dan Bloom, Sara Muller-Ravett, Joseph Broadus (2011). Staying on course: Three-year results of the National Guard Youth Challenge Evaluation. MDRC.

Miller, Cynthia, Kristin E. Porter, Fannie M. Tseng, Yasuyo Abe, Johannes M. Bos. 2005. The Challenge of Repeating Success in a Changing World – Final Report on the Center for Employment Training Replication Sites.

Neumark, David, and Donna Rothstein. 2006. School-to-Career Programs and Transitions to Employment and Higher Education. *Economics of Education Review* 25, 4, 374–393.

Ñopo, Hugo, Jaime Saavedra (2003). Evaluación del impacto de mediano plazo de Projoven: Resultados de las mediciones realizadas a los seis, doce y dieciocho meses de culminado el programa. GRADE report.

Ñopo, Hugo, Miguel Robles, Jaime Saavedra (2008). Occupational training to reduce gender segregation: The impacts of ProJoven. *Revista Economía* 62, 33–54.

Pessoa e Costa, Sofia, Stéphane Robin (2009). An Illustration of the Returns to Training Programs: The Evaluation of the "Qualifying Contract" in France. Université catholique de Louvain, Institut de Recherches Economiques et Sociales (IRES).

Quint, Janet C., Johannes M. Bos, Denise F. Polit (1997). New chance: Final report on a comprehensive program for young mothers in poverty and their children. New York, NY: Manpower Demonstration Research Corporation.

Richardson, James (1998). "Do Wage Subsidies Enhance Employability? Evidence from Australian Youth." Center for Economic Performance Working Paper No: 387. London School of Economic and Political Science, London.

Rodríguez-Planas, Núria (2012). Longer-Term Impacts of Mentoring, Educational Services, and Learning Incentives: Evidence from a Randomized Trial in the United States. *American Economic Journal: Applied Economics* 4, 4, 121–39.

Roger, Muriel, Philippe Zamora (2011). Hiring young, unskilled workers on subsidized open-ended contracts: a good integration program? Oxford Review of Economic Policy 27, 2, 380-–396.

Steiner, Roberto, Paula Acosta, Norberto Rojas, Natalia Millán, Carlos Castañeda, José L. González (2010). Evaluación de Impacto del Programa Jóvenes Rurales Emprendedores Del Servicio Nacional De Aprendizaje – Sena.

Saniter, Nils Peter (2014). Educational interventions and labor market success. Berlin, Germany: Freie Universität Berlin.

Santa María, Mauricio, Paula Acosta, Alexander Estacio, Natalia Millán, Mauricio Olivera, Mónica Parra Torrado, Carlos Alberto Castañeda, Yanira Marcela Oviedo, Carlos Andrés Pérez (2009). Evaluación de Impacto de Tres Programas y Línea de Base de un Programa del Servicio Nacional de Aprendizaje – SENA. Informe Final. Fedesarrollo – Centro de Investigación Ecónomica y Social.

Santa María, Mauricio, Mauricio Olivera, Paula Acosta, Tatiana Vásquez, Anwar Rodríguez (2009). Evaluación de Impacto Programa Jóvenes con Futuro. Working Paper. Fedesarrollo.

Santiago Consultores Asociados (1999). Evaluacion Ex-post Chile Joven Fase II. Report.

Schneider, Hilmar, Arne Uhlendorff, Klaus F. Zimmermann (2011). Ökonometrie vs. Projektdesign: Lehren aus der Evaluation eines Modellprojekts zur Umsetzung des Workfare-Konzepts. (Econometrics vs. Project Design: Lessons from the Evaluation of a Workfare Project. With English summary.). IZA DP No. 5599.

Schochet, Peter, Sheena McConnell, John Burghardt (2003). National Job Corps Study: Findings using administrative earnings records data – Final Report. Mathematica Policy Research.

Schochet, Peter, John Burghardt, Sheena McConnell (2008). Does Job Corps Work? Impact Findings from the National Job Corps Study. *American Economic Review* 98, 5, 1864–86.

SENCE (2008). Evaluación de Impacto del Programa de Formación en Oficios para Jóvenes de Escasos Recursos, Ejecución 2005 – Informe Final.

Mónica Riquelme Peña, Margarita Cifuentes, Bárbara Morales, Álvaro Espinoza (2006). Informe Final Preliminar Estudio: Evaluación de Impacto Programa Aprendices. SENCE.

Universidad de Chile (2012). Evaluación de Impacto del Programa de Subsidio al Empleo Joven. Centro de Microdatos, Departamento de Economía, Universidad de Chile.

van Gameren, Edwin (2010). Evaluación de impacto del programa de apoyo al empleo. Centro de Estudios Económicos El Colegio de México.

Webb, Matthew, Arthur Sweetman, Casey Warman (2014). How Targeted is Targeted Tax Relief? Evidence from the Unemployment Insurance Youth Hires Program. Queen's University, Department of Economics.

Wilkinson, David (2003). New Deal for Young People: evaluation of unemployment flows. Research Discussion Paper 15. Policy Studies Institute.

Zhang, Junni L., Donald B. Rubin, Fabrizia Mealli (2009). Likelihood-Based Analysis of Causal Effects of Job-Training Programs Using Principal Stratification. *Journal of the American Statistical Association* 104, 485, 166–176.

# Appendix 2. Variable definitions

| | |
|---|---|
| **Employment outcomes** | Employment Probability, Unemployment Probability, Participation Rate, Hours Worked (or equivalent) , Unemployment Duration, and Quality of Employment (e.g. contract, fixed, benefits). |
| **Earnings outcomes** | Earnings, Income, Household Income, Consumption, Salary , or hourly wage (or equivalent). |
| **Incentives for service providers** | Funders of the project provide implementors payments (or bonuses) contingent on participant performance. Performance can refer to the outcomes of participants (getting a job after program) or performance within program such as completion or higher test scores. |
| **Incentives for participants** | Most frequently payments or living stipends that are contingent on attendance. This also include incentives for participation such as taking away welfare or unemployment benefits. |
| **Monitoring** | Direct one on one supervision of program participants often through scheduled visits with a case manage or mentor. |
| **Participant profiling** | Profiling is the identification of individual factors that represent a risk in the labor market and assigning appropriate services based on this. Collecting the information to asses the main contraints (risks) in the labor market include caseworker mentor discretion, detailed screening, eligibility rules, or statistical profiling among others. A program is considered to profile if it uses information to assign particpants specific services among an array of services offered or to determine the intensity of services (duration, etc.). A key element of this is that the program collects information and proactively assigns services. |
| **Exra services** | The program offers services that fall into more than one of the four main categories (skills training, employment services, subsidized employment, and entrepreneurship promotion). A commonn case is offerring skills training combined with employment services. |
| **Low-income/ disadvantaged populations** | This includes low income, disadvantaed, at risk, and vulnerable youth and is based on targetting described in the evaluations. |
| **Country income groups** | Based on the World Bank's country and lending group definitions for the 2015 fiscal year. See http://data.worldbank.org/about/country-and-lending-groups. |
| **Unadjusted mean differences** | The evaluation did not apply covariate adjustment in estimating treatement effect that was used. |
| **Younger participants** | Codes a younger participant sample if the approximate mean age of the evaluation sample is below 22 years. |
| **Implementers** | Government implementers include local, regional, national governments, and multilateral organizations. Private sector implemeters consist of private firms, foundations, and NGOs. |
| **Soft skills training** | Training in non-cognitive, socio-emotional, and/or life skills. |
| **Unspecified categroy programs** | Program balances equally among more than one service category (skills traning, entrepreneurship promotion, subsidized employment, employment services) and the evaluation information does not list any single service as predominant. |

# Appendix 3. Separate multivariate results for high- and low-income countries, earnings outcomes, and employment outcomes.

**Table 8. Weighted Least Squares Hedge's g regressions: high income country sample**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Main intervention category (base=skills training)** | | | | | | |
| Entr. Prom. | -0.074*** | -0.076*** | | | | |
| | [2.72] | [3.00] | | | | |
| Empl. Serv. | -0.061* | -0.049 | 0.057 | 0.182** | 0.231** | 0.301** |
| | [1.99] | [1.52] | [0.72] | [2.60] | [2.32] | [2.70] |
| Subs. Empl. | -0.035 | -0.020 | -0.026 | 0.035 | 0.023 | 0.092 |
| | [0.90] | [0.53] | [0.45] | [1.06] | [0.48] | [1.36] |
| Unspecified | 0.020 | 0.019 | 0.006 | -0.036 | -0.001 | -0.193* |
| | [0.47] | [0.49] | [0.09] | [0.64] | [0.03] | [1.83] |
| **Evaluation features** | | | | | | |
| Log Evaluation Sample Size | | -0.020** | -0.013 | -0.025** | -0.010* | -0.001 |
| | | [2.16] | [1.32] | [2.21] | [2.01] | [0.07] |
| Publication Peer-Reviewed | | -0.046 | -0.089 | -0.152*** | -0.148*** | -0.220** |
| | | [1.47] | [1.53] | [2.95] | [3.19] | [2.83] |
| Eval. design:RCT | | -0.049 | 0.007 | 0.049 | -0.011 | 0.057 |
| | | [1.52] | [0.17] | [1.11] | [0.36] | [0.58] |
| **Program design features** | | | | | | |
| Additional services | | | 0.021 | 0.105* | 0.031 | 0.120 |
| | | | [0.38] | [2.06] | [0.66] | [1.55] |
| Participant profiling | | | 0.007 | 0.071* | 0.047 | 0.050 |
| | | | [0.17] | [1.94] | [1.08] | [0.99] |
| Participant engagement mechanism | | | 0.073 | 0.238*** | 0.108 | 0.294* |
| | | | [1.25] | [3.62] | [1.63] | [2.00] |
| Incentives for service providers | | | 0.100 | 0.170*** | 0.146** | 0.098* |
| | | | [1.49] | [4.56] | [2.76] | [1.76] |
| Program has soft skills training | | | | | 0.045 | -0.012 |
| | | | | | [0.67] | [0.05] |
| **Outcome characteristics** | | | | | | |
| Employment outcome (base=earnings outcome) | | | | 0.015 | | 0.011 |
| | | | | [0.70] | | [0.64] |
| Estimated unadjusted difference in means | | | | -0.041 | | -0.181*** |
| | | | | [1.05] | | [2.96] |
| Measured over one year after exit from program | | | | 0.096** | | 0.073 |
| | | | | [2.51] | | [1.54] |
| **Target/evaluation group** | | | | | | |
| Low income / disadvantaged | | | | 0.044 | | 0.127 |
| | | | | [0.88] | | [0.53] |
| Male (base=male and female combined) | | | | -0.011 | | -0.009 |
| | | | | [0.52] | | [0.48] |
| Female (base=male and female combined) | | | | -0.060* | | -0.035 |
| | | | | [2.02] | | [1.45] |
| Younger Participants | | | | 0.014 | | 0.080* |
| | | | | [0.31] | | [1.96] |
| **Type of implementer (base=Private and public sector joinly implement)** | | | | | | |
| Government only | | | | | -0.110*** | -0.201 |
| | | | | | [3.35] | [0.94] |
| Private sector only | | | | | -0.166*** | -0.045 |
| | | | | | [3.48] | [0.36] |
| Constant | 0.053* | 0.235** | 0.015 | -0.281* | -0.076 | -0.470* |
| | [1.98] | [2.46] | [0.12] | [2.02] | [0.63] | [2.05] |
| R2 | 0.02 | 0.08 | 0.08 | 0.46 | 0.25 | 0.60 |
| N | 1,300 | 1,300 | 792 | 511 | 705 | 480 |
| Number of studies: | 54 | 54 | 38 | 30 | 35 | 27 |
| Number of interventions: | 53 | 53 | 28 | 18 | 25 | 16 |

Notes: see Table 5 notes.

**Table 9. Random Effects SMD regressions: high income country sample**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| **Main intervention category (base=skills training)** | | | | | | |
| Entr. Prom. | -0.031 | -0.04 | | | | |
| | [0.75] | [0.96] | | | | |
| Empl. Serv. | -0.013 | -0.012 | -0.043** | 0.153*** | -0.011 | 0.372*** |
| | [1.21] | [1.01] | [2.20] | [3.90] | [0.13] | [3.15] |
| Subs. Empl. | -0.022** | -0.018 | -0.029 | 0.029 | -0.024 | -0.031 |
| | [2.36] | [1.46] | [1.43] | [1.21] | [0.33] | [0.53] |
| Unspecified | -0.025* | -0.028** | -0.033** | -0.078*** | -0.033* | -0.084*** |
| | [1.78] | [1.97] | [1.96] | [3.40] | [1.88] | [3.24] |
| **Evaluation features** | | | | | | |
| Log Evaluation Sample Size | | -0.006** | -0.006 | -0.012** | -0.007* | -0.011** |
| | | [2.52] | [1.62] | [2.54] | [1.87] | [2.11] |
| Publication Peer-Reviewed | | -0.002 | -0.02 | -0.054*** | -0.038** | -0.048*** |
| | | [0.22] | [1.28] | [3.61] | [2.48] | [3.08] |
| Eval. design:RCT | | -0.017* | 0.014 | 0.005 | -0.040* | -0.086*** |
| | | [1.83] | [0.91] | [0.26] | [1.76] | [2.81] |
| **Program design features** | | | | | | |
| Additional services | | | -0.051** | 0.128*** | 0.012 | 0.078* |
| | | | [2.57] | [4.09] | [0.46] | [1.88] |
| Participant profiling | | | -0.003 | 0.028 | 0.041** | 0.018 |
| | | | [0.19] | [1.27] | [2.06] | [0.52] |
| Participant engagement mechanism | | | 0.037 | 0.195*** | 0.034 | 0.074 |
| | | | [1.48] | [5.28] | [1.07] | [1.35] |
| Incentives for service providers | | | -0.027 | 0.102*** | 0.018 | -0.027 |
| | | | [1.64] | [4.25] | [0.66] | [0.63] |
| Program has soft skills training | | | | | -0.042 | -0.459*** |
| | | | | | [0.60] | [2.90] |
| **Outcome characteristics** | | | | | | |
| Employment outcome (base=earnings outcome) | | | | -0.009 | | -0.007 |
| | | | | [1.00] | | [0.81] |
| Estimated unadjusted difference in means | | | | -0.040** | | -0.047*** |
| | | | | [2.30] | | [2.75] |
| Measured over one year after exit from program | | | | 0.071*** | | 0.070*** |
| | | | | [6.39] | | [6.64] |
| **Target/evaluation group** | | | | | | |
| Low income / disadvantaged | | | | 0.092*** | | 0.491*** |
| | | | | [2.78] | | [3.55] |
| Male (base=male and female combined) | | | | 0.006 | | 0.007 |
| | | | | [0.51] | | [0.60] |
| Female (base=male and female combined) | | | | -0.029** | | -0.023* |
| | | | | [2.37] | | [1.92] |
| Younger Participants | | | | -0.008 | | 0.008 |
| | | | | [0.78] | | [0.85] |
| **Type of implementer (base=Private and public sector joinly implement)** | | | | | | |
| Government only | | | | | -0.067*** | -0.434*** |
| | | | | | [3.55] | [3.19] |
| Private sector only | | | | | 0.01 | -0.205** |
| | | | | | [0.20] | [2.00] |
| Constant | 0.031*** | 0.094*** | 0.102** | -0.299*** | 0.109 | 0.018 |
| | [6.65] | [4.28] | [2.37] | [3.27] | [1.07] | [0.12] |
| R2 | . | . | . | . | . | . |
| N | 1,299 | 1,299 | 792 | 511 | 705 | 480 |
| Number of studies: | 54 | 54 | 38 | 30 | 35 | 27 54 |
| Number of interventions: | 53 | 53 | 28 | 18 | 25 | 16 |

Notes: see Table 5 notes.

**Table 10. Positive & statistically significant probit regressions: high income country sample**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category | | | | | | |
| Entr. Prom. | -0.330** | -0.186 | | | | |
| | [2.19] | [1.31] | | | | |
| Empl. Serv. | -0.138 | -0.159 | 0.025 | 0.801*** | -0.977*** | 0.809*** |
| | [1.10] | [1.37] | [0.16] | [5.30] | [3.46] | [14.71] |
| Subs. Empl. | 0.026 | -0.208** | -0.292** | -0.289** | -1.216*** | -1.391*** |
| | [0.25] | [2.16] | [2.53] | [2.32] | [5.69] | [15.33] |
| Unspecified | -0.045 | 0.083 | -0.089 | -0.081 | 0.02 | -0.750*** |
| | [0.44] | [1.01] | [0.79] | [0.78] | [0.15] | [18.62] |
| Evaluation features | | | | | | |
| Log Evaluation Sample Size | | 0.025 | 0.040** | 0.06 | 0.048*** | -0.001 |
| | | [1.34] | [2.23] | [1.55] | [3.57] | [0.04] |
| Publication Peer-Reviewed | | -0.143* | -0.371*** | -0.203* | -0.256** | -0.022 |
| | | [1.69] | [3.05] | [1.88] | [2.32] | [1.02] |
| Eval. design:RCT | | -0.157* | 0.036 | 0.08 | -0.084 | -0.088*** |
| | | [1.95] | [0.45] | [0.88] | [0.80] | [2.70] |
| Program design features | | | | | | |
| Additional services | | | 0.171 | 1.070*** | 0.175 | 1.436*** |
| | | | [1.42] | [6.86] | [1.46] | [19.07] |
| Participant profiling | | | 0.002 | -0.104 | 0.150* | -0.669*** |
| | | | [0.02] | [1.56] | [1.77] | [14.64] |
| Participant engagement mechanism | | | 0.528** | | 0.339 | |
| | | | [2.06] | | [1.20] | |
| Incentives for service providers | | | 0.285*** | 1.116*** | 0.333*** | 0.496*** |
| | | | [2.75] | [9.88] | [2.76] | [9.08] |
| Program has soft skills training | | | | | -0.967*** | -2.364*** |
| | | | | | [4.14] | [10.55] |
| Outcome characteristics | | | | | | |
| Employment outcome | | | | -0.071 | | -0.053* |
| | | | | [1.62] | | [1.90] |
| Estimated unadjusted difference in | | | | 0 | | -0.089*** |
| | | | | [0.16] | | [11.80] |
| Measured over one year after exit | | | | 0.012 | | -0.047 |
| | | | | [0.10] | | [0.39] |
| Target/evaluation group | | | | | | |
| Low income / disadvantaged | | | | 0.829*** | | 3.019*** |
| | | | | [4.60] | | [18.18] |
| Male (base=male and female | | | | 0.006 | | 0.01 |
| | | | | [0.18] | | [0.72] |
| Female (base=male and female | | | | -0.04 | | -0.028 |
| | | | | [0.99] | | [0.79] |
| Younger Participants | | | | 0.088* | | -0.032 |
| | | | | [1.70] | | [1.19] |
| Type of implementer (base=Private | | | | | | |
| Government only | | | | | -0.097 | -0.971*** |
| | | | | | [0.89] | [7.45] |
| Private sector only | | | | | 0.027 | -0.063 |
| | | | | | [0.18] | [0.65] |
| Constant | | | | | | |
| N | 1,715 | 1,334 | 845 | 492 | 756 | 459 |
| Number of studies: | 60 | 54 | 37 | 29 | 34 | 26 |
| Number of interventions: | 60 | 54 | 28 | 18 | 25 | 16 |

Notes: see Table 5 notes.

**Table 11. Weighted Least Squares Hedge's g regressions: low- and middle-income country sample**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category | | | | | | |
| Entr. Prom. | 0.037 | 0.036 | 0.031 | 0.041 | 0.001 | 0.036 |
| | [0.61] | [0.57] | [0.56] | [1.08] | [0.02] | [1.13] |
| Empl. Serv. | -0.048 | -0.039 | 0.03 | -0.015 | -0.134 | -0.207*** |
| | [1.06] | [0.82] | [0.54] | [0.25] | [1.54] | [2.99] |
| Subs. Empl. | -0.018 | -0.01 | -0.109* | 0.106* | -0.065 | 0.224*** |
| | [0.50] | [0.29] | [1.87] | [1.92] | [0.85] | [4.34] |
| Evaluation features | | | | | | |
| Log Evaluation Sample Size | | -0.007 | -0.018 | -0.036*** | 0.011 | -0.029 |
| | | [0.50] | [1.10] | [3.10] | [0.73] | [1.22] |
| Publication Peer-Reviewed | | -0.012 | 0.000 | 0.036 | -0.009 | 0.100* |
| | | [0.53] | [0.01] | [1.20] | [0.22] | [1.99] |
| Eval. design:RCT | | -0.013 | -0.084 | -0.114*** | -0.045 | -0.141*** |
| | | [0.23] | [1.59] | [3.99] | [1.22] | [4.90] |
| Program design features | | | | | | |
| Additional services | | | 0.022 | 0.098*** | 0.125 | 0.183** |
| | | | [0.37] | [3.52] | [1.50] | [2.24] |
| Participant profiling | | | 0.105*** | 0.094*** | 0.059 | 0.052 |
| | | | [2.80] | [3.17] | [1.38] | [1.51] |
| Participant engagement mechanism | | | 0.118** | 0.067** | 0.072* | 0.077 |
| | | | [2.39] | [2.40] | [1.87] | [1.70] |
| Incentives for service providers | | | 0.03 | -0.05 | -0.001 | -0.128*** |
| | | | [0.48] | [1.13] | [0.01] | [3.79] |
| Program has soft skills training | | | | | -0.037 | 0.008 |
| | | | | | [0.72] | [0.11] |
| Outcome characteristics | | | | | | |
| Employment outcome | | | | 0.026 | | 0.021 |
| | | | | [1.02] | | [0.78] |
| Estimated unadjusted difference in | | | | 0.159*** | | 0.171*** |
| | | | | [4.97] | | [4.61] |
| Measured over one year after exit | | | | 0.032* | | 0.02 |
| | | | | [1.84] | | [1.34] |
| Target/evaluation group | | | | | | |
| Low income / disadvantaged | | | | 0.104*** | | 0.108** |
| | | | | [3.36] | | [2.38] |
| Male (base=male and female | | | | -0.092*** | | -0.054 |
| | | | | [3.28] | | [1.44] |
| Female (base=male and female | | | | 0.012 | | 0.039* |
| | | | | [0.68] | | [1.90] |
| Younger Participants | | | | -0.068** | | 0.001 |
| | | | | [2.17] | | [0.05] |
| Type of implementer (base=Private | | | | | | |
| Government only | | | | | 0.061 | 0.001 |
| | | | | | [0.97] | [0.02] |
| Private sector only | | | | | 0.110* | 0.039 |
| | | | | | [2.04] | [0.56] |
| Constant | 0.092*** | 0.149 | 0.191 | 0.314*** | -0.013 | 0.263 |
| | [3.90] | [1.13] | [1.27] | [3.13] | [0.12] | [1.72] |
| R2 | 0.01 | 0.01 | 0.23 | 0.26 | 0.20 | 0.31 |
| N | 701 | 701 | 577 | 513 | 545 | 482 |
| Number of studies: | 42 | 42 | 32 | 28 | 27 | 24 |
| Number of interventions: | 44 | 44 | 30 | 24 | 26 | 21 |

Notes: see Table 5 notes.

**Table 12.  Random Effects SMD regressions: low- and middle-income country sample**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category | | | | | | |
| Entr. Prom. | 0.017 | 0.02 | 0.048** | 0.050* | 0.047** | -0.002 |
| | [1.01] | [1.09] | [2.10] | [1.90] | [1.98] | [0.06] |
| Empl. Serv. | -0.032 | -0.033 | 0.012 | -0.029 | -0.092* | -0.226*** |
| | [0.96] | [0.98] | [0.30] | [0.67] | [1.71] | [3.58] |
| Subs. Empl. | 0.003 | 0.007 | -0.077*** | 0.084* | -0.060** | 0.153*** |
| | [0.13] | [0.29] | [2.89] | [1.65] | [2.13] | [2.63] |
| Evaluation features | | | | | | |
| Log Evaluation Sample Size | | -0.024*** | -0.020*** | -0.030*** | 0.004 | 0.009 |
| | | [5.08] | [2.99] | [3.70] | [0.55] | [0.77] |
| Publication Peer-Reviewed | | -0.003 | 0.025* | 0.070*** | 0.021 | 0.050** |
| | | [0.21] | [1.76] | [3.95] | [1.21] | [2.38] |
| Eval. design:RCT | | -0.021 | -0.074*** | -0.108*** | -0.090*** | -0.090*** |
| | | [1.56] | [4.99] | [5.08] | [5.42] | [4.02] |
| Program design features | | | | | | |
| Additional services | | | 0.049** | 0.078*** | 0.195*** | 0.290*** |
| | | | [2.16] | [2.95] | [5.43] | [5.62] |
| Participant profiling | | | 0.110*** | 0.126*** | 0.067*** | 0.094*** |
| | | | [5.81] | [5.38] | [3.07] | [3.67] |
| Participant engagement mechanism | | | 0.048** | 0.052** | 0.054** | 0.015 |
| | | | [2.29] | [2.06] | [2.37] | [0.50] |
| Incentives for service providers | | | 0.051*** | -0.021 | 0.022 | -0.061** |
| | | | [2.63] | [0.70] | [1.03] | [2.03] |
| Program has soft skills training | | | | | -0.074*** | -0.094*** |
| | | | | | [3.55] | [2.92] |
| Outcome characteristics | | | | | | |
| Employment outcome | | | | -0.032*** | | -0.030** |
| | | | | [2.71] | | [2.48] |
| Estimated unadjusted difference in | | | | 0.106 | | 0.116 |
| | | | | [1.01] | | [1.17] |
| Measured over one year after exit | | | | -0.012 | | -0.002 |
| | | | | [0.85] | | [0.16] |
| Target/evaluation group | | | | | | |
| Low income / disadvantaged | | | | 0.075*** | | 0.056** |
| | | | | [3.33] | | [2.00] |
| Male (base=male and female | | | | -0.063*** | | -0.028 |
| | | | | [3.62] | | [1.52] |
| Female (base=male and female | | | | 0 | | 0.021 |
| | | | | [0.00] | | [1.50] |
| Younger Participants | | | | -0.013 | | 0.043** |
| | | | | [0.77] | | [2.05] |
| Type of implementer (base=Private | | | | | | |
| Government only | | | | | 0.033 | 0.105** |
| | | | | | [0.82] | [2.09] |
| Private sector only | | | | | 0.073*** | 0.106*** |
| | | | | | [3.28] | [3.08] |
| Constant | 0.088*** | 0.280*** | 0.198*** | 0.302*** | 0.053 | 0.024 |
| | [14.71] | [6.85] | [3.11] | [3.86] | [0.80] | [0.27] |
| R2 | . | . | . | . | . | . |
| N | 760 | 701 | 577 | 513 | 545 | 482 |
| Number of studies: | 44 | 42 | 32 | 28 | 27 | 24 |
| Number of interventions: | 44 | 44 | 30 | 24 | 26 | 21 |

Notes: see Table 5 notes.

**Table 13. Positive & statistically significant probit regressions: low- and middle-income country sample**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category | | | | | | |
| Entr. Prom. | 0.079 | 0.078 | 0.053 | 0.038 | -0.024 | -0.05 |
| | [0.61] | [0.62] | [0.37] | [0.44] | [0.31] | [0.93] |
| Empl. Serv. | -0.136 | -0.123 | 0.166 | 0.125 | -0.402* | -0.515** |
| | [1.08] | [0.94] | [1.08] | [0.77] | [1.83] | [2.52] |
| Subs. Empl. | 0.057 | 0.026 | -0.181 | 0.372** | -0.162 | 0.094 |
| | [0.46] | [0.25] | [1.34] | [2.06] | [1.32] | [0.52] |
| Evaluation features | | | | | | |
| Log Evaluation Sample Size | | 0.047 | 0.022 | -0.052 | 0.119*** | 0.117** |
| | | [1.30] | [0.49] | [1.41] | [2.77] | [2.27] |
| Publication Peer-Reviewed | | 0.124 | 0.141 | 0.189** | 0.094 | 0.068 |
| | | [1.35] | [1.36] | [2.03] | [0.93] | [0.71] |
| Eval. design:RCT | | -0.045 | -0.141 | -0.290*** | -0.124 | -0.241*** |
| | | [0.42] | [1.31] | [3.32] | [1.42] | [2.80] |
| Program design features | | | | | | |
| Additional services | | | -0.094 | 0.053 | 0.407*** | 0.848*** |
| | | | [0.66] | [0.56] | [2.74] | [4.82] |
| Participant profiling | | | 0.161* | 0.129 | 0.137* | 0.249*** |
| | | | [1.65] | [1.62] | [1.68] | [4.36] |
| Participant engagement mechanism | | | 0.365*** | 0.284*** | 0.246** | 0.144* |
| | | | [2.83] | [3.05] | [2.35] | [1.81] |
| Incentives for service providers | | | 0.09 | -0.122 | -0.071 | -0.189* |
| | | | [0.55] | [0.86] | [0.57] | [1.87] |
| Program has soft skills training | | | | | -0.373*** | -0.574*** |
| | | | | | [3.16] | [3.75] |
| Outcome characteristics | | | | | | |
| Employment outcome (base=earnings | | | | 0.107 | | 0.054 |
| | | | | [1.45] | | [0.73] |
| Measured over one year after exit from | | | | 0.168** | | 0.255*** |
| | | | | [2.18] | | [3.39] |
| Target/evaluation group | | | | | | |
| Low income / disadvantaged | | | | 0.217** | | -0.086 |
| | | | | [1.96] | | [0.53] |
| Male (base=male and female combined) | | | | -0.097 | | 0.001 |
| | | | | [1.40] | | [0.01] |
| Female (base=male and female | | | | 0.043 | | 0.034 |
| | | | | [0.70] | | [0.74] |
| Younger Participants | | | | -0.176 | | 0.129 |
| | | | | [1.22] | | [0.92] |
| Type of implementer (base=Private and | | | | | | |
| Government only | | | | | 0.110 | 0.031 |
| | | | | | [0.69] | [0.20] |
| Private sector only | | | | | 0.388*** | 0.560*** |
| | | | | | [3.00] | [2.98] |
| N | 1,217 | 724 | 596 | 530 | 567 | 502 |
| Number of studies: | 44 | 40 | 30 | 26 | 26 | 23 |
| Number of interventions: | 45 | 44 | 30 | 24 | 26 | 21 |

Notes: see Table 5 notes.

**Table 14. Weighted Least Squares Hedge's g regressions: employment outcomes only**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| <u>Main intervention category</u> | | | | | | |
| Entr. Prom. | 0.086* | 0.082 | 0.121** | 0.126*** | 0.150*** | 0.117** |
| | [1.67] | [1.57] | [2.02] | [3.02] | [2.92] | [2.22] |
| Empl. Serv. | -0.065** | -0.048* | 0.041 | 0.103** | 0.018 | 0.056 |
| | [2.15] | [1.74] | [1.20] | [2.38] | [0.31] | [1.08] |
| Subs. Empl. | -0.03 | -0.007 | -0.042 | 0.03 | -0.046 | 0.019 |
| | [1.01] | [0.26] | [1.11] | [0.86] | [0.92] | [0.53] |
| Unspecified | 0.00 | 0.005 | 0.003 | -0.036 | 0.01 | -0.098 |
| | [0.01] | [0.12] | [0.05] | [0.58] | [0.19] | [1.18] |
| <u>Evaluation features</u> | | | | | | |
| Log Evaluation Sample Size | | -0.015** | -0.012 | -0.022** | -0.009 | 0.00 |
| | | [2.15] | [1.52] | [2.51] | [1.55] | [0.05] |
| Publication Peer-Reviewed | | -0.022 | -0.060*** | -0.087*** | -0.097*** | -0.097*** |
| | | [0.85] | [2.82] | [3.66] | [4.22] | [4.05] |
| Eval. design:RCT | | -0.005 | -0.028 | -0.044 | -0.046* | -0.029 |
| | | [0.17] | [1.12] | [1.52] | [1.86] | [1.20] |
| High income country | | -0.02 | -0.077*** | -0.114*** | -0.071** | -0.127*** |
| | | [0.64] | [2.80] | [4.26] | [2.23] | [3.86] |
| <u>Program design features</u> | | | | | | |
| Additional services | | | -0.003 | 0.045* | 0.038 | 0.072** |
| | | | [0.09] | [1.91] | [1.34] | [2.48] |
| Participant profiling | | | 0.053 | 0.083*** | 0.088** | 0.075* |
| | | | [1.67] | [3.02] | [2.52] | [2.03] |
| Participant engagement mechanism | | | 0.03 | 0.051* | 0.001 | 0.081** |
| | | | [0.86] | [1.79] | [0.03] | [2.53] |
| Incentives for service providers | | | 0.085** | 0.064 | 0.096*** | 0.05 |
| | | | [2.32] | [1.62] | [2.77] | [1.25] |
| Program has soft skills training | | | | | -0.053 | 0.008 |
| | | | | | [1.12] | [0.25] |
| <u>Outcome characteristics</u> | | | | | | |
| Estimated unadjusted difference in | | | | -0.093 | | -0.189** |
| | | | | [1.58] | | [2.23] |
| Measured over one year after exit | | | | 0.082*** | | 0.102*** |
| | | | | [3.32] | | [3.79] |
| <u>Target/evaluation group</u> | | | | | | |
| Low income / disadvantaged | | | | 0.051* | | 0.021 |
| | | | | [1.78] | | [0.65] |
| Male (base=male and female | | | | -0.025 | | -0.027 |
| | | | | [1.18] | | [1.19] |
| Female (base=male and female | | | | 0.013 | | 0.015 |
| | | | | [0.50] | | [0.74] |
| Younger Participants | | | | -0.037 | | 0.041 |
| | | | | [1.24] | | [1.19] |
| <u>Type of implementer (base=Private</u> | | | | | | |
| Government only | | | | | -0.029 | -0.056 |
| | | | | | [0.80] | [1.17] |
| Private sector only | | | | | 0.037 | 0.072* |
| | | | | | [0.63] | [1.77] |
| Constant | 0.072*** | 0.198*** | 0.144* | 0.155 | 0.166** | -0.05 |
| | [3.80] | [2.81] | [1.80] | [1.69] | [2.31] | [0.57] |
| R2 | 0.04 | 0.06 | 0.16 | 0.36 | 0.22 | 0.43 |
| N | 1,331 | 1,331 | 905 | 619 | 798 | 568 |
| Number of studies: | 86 | 86 | 63 | 52 | 56 | 45 |
| Number of interventions: | 91 | 91 | 53 | 38 | 47 | 33 |

Notes: see Table 5 notes.

**Table 15. Random Effects SMD regressions: employment outcomes only**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category | | | | | | |
| Entr. Prom. | 0.047* | 0.034 | -0.005 | 0.119*** | 0.075** | 0.145*** |
| | [1.79] | [1.26] | [0.15] | [3.52] | [2.39] | [4.30] |
| Empl. Serv. | -0.029* | 0.015 | 0.008 | 0.121*** | -0.001 | 0.077** |
| | [1.92] | [0.90] | [0.45] | [5.20] | [0.04] | [2.10] |
| Subs. Empl. | -0.030** | 0.015 | -0.025 | 0.028 | -0.026 | 0 |
| | [2.34] | [1.03] | [1.28] | [1.21] | [0.96] | [0.00] |
| Unspecified | -0.050*** | -0.035* | -0.045** | -0.065*** | -0.023 | -0.053** |
| | [2.69] | [1.82] | [2.25] | [2.77] | [1.12] | [1.96] |
| Evaluation features | | | | | | |
| Log Evaluation Sample Size | | -0.023*** | -0.008** | -0.016*** | -0.005 | 0 |
| | | [7.02] | [2.21] | [3.74] | [1.49] | [0.11] |
| Publication Peer-Reviewed | | -0.02 | -0.029** | -0.036*** | -0.052*** | -0.054*** |
| | | [1.63] | [2.09] | [3.12] | [3.62] | [4.57] |
| Eval. design:RCT | | -0.026** | 0.001 | -0.02 | -0.061*** | -0.052*** |
| | | [2.41] | [0.09] | [1.50] | [4.12] | [3.49] |
| High income country | | -0.034*** | -0.079*** | -0.115*** | -0.089*** | -0.125*** |
| | | [2.92] | [5.07] | [7.71] | [5.36] | [7.54] |
| Program design features | | | | | | |
| Additional services | | | -0.022 | 0.040*** | 0.063*** | 0.077*** |
| | | | [1.38] | [2.75] | [3.37] | [4.03] |
| Participant profiling | | | 0.028* | 0.065*** | 0.059*** | 0.053** |
| | | | [1.94] | [3.94] | [3.82] | [2.56] |
| Participant engagement mechanism | | | 0.074*** | 0.059*** | 0.055*** | 0.087*** |
| | | | [4.12] | [4.08] | [3.04] | [5.32] |
| Incentives for service providers | | | 0.023 | 0.047*** | 0.064*** | 0.061*** |
| | | | [1.59] | [3.40] | [4.16] | [3.76] |
| Program has soft skills training | | | | | -0.042** | -0.027 |
| | | | | | [2.05] | [1.52] |
| Outcome characteristics | | | | | | |
| Estimated unadjusted difference in | | | | -0.014 | | -0.037* |
| | | | | [0.67] | | [1.85] |
| Measured over one year after exit | | | | 0.047*** | | 0.050*** |
| | | | | [4.10] | | [4.55] |
| Target/evaluation group | | | | | | |
| Low income / disadvantaged | | | | 0.054*** | | 0.039** |
| | | | | [3.35] | | [2.15] |
| Male (base=male and female | | | | -0.009 | | -0.002 |
| | | | | [0.70] | | [0.18] |
| Female (base=male and female | | | | 0 | | 0.004 |
| | | | | [0.02] | | [0.33] |
| Younger Participants | | | | -0.025** | | 0.004 |
| Type of implementer (base=Private | | | | | | |
| Government only | | | | | -0.040** | -0.084*** |
| | | | | | [2.16] | [3.53] |
| Private sector only | | | | | 0.060*** | 0.036* |
| | | | | | [2.59] | [1.70] |
| Constant | 0.063*** | 0.274*** | 0.112*** | 0.084* | 0.108*** | -0.024 |
| | [10.54] | [9.96] | [3.04] | [1.83] | [2.84] | [0.50] |
| R2 | . | . | . | . | . | . |
| N | 1,382 | 1,330 | 905 | 619 | 798 | 568 |
| Number of studies: | 88 | 86 | 63 | 52 | 56 | 45 |
| Number of interventions: | 91 | 91 | 53 | 38 | 47 | 33 |

Notes: see Table 5 notes.

**Table 16. Positive & statistically significant probit regressions: employment outcomes only**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| _Main intervention category_ | | | | | | |
| Entr. Prom. | 0.101 | 0.101 | 0.148 | 0.206 | 0.231** | 0.137 |
| | [1.01] | [1.08] | [1.00] | [1.29] | [2.03] | [1.24] |
| Empl. Serv. | -0.195* | -0.192* | 0.016 | 0.104 | -0.253 | -0.123 |
| | [1.76] | [1.84] | [0.12] | [0.57] | [1.58] | [0.68] |
| Subs. Empl. | 0.032 | -0.101 | -0.188* | -0.113 | -0.297** | -0.243 |
| | [0.36] | [1.14] | [1.86] | [0.75] | [2.49] | [1.51] |
| Unspecified | -0.075 | 0.036 | 0.046 | 0.052 | 0.11 | -0.23 |
| | [0.77] | [0.34] | [0.34] | [0.26] | [0.77] | [1.06] |
| _Evaluation features_ | | | | | | |
| Log Evaluation Sample Size | | 0.038** | 0.038* | 0.013 | 0.059*** | 0.056* |
| | | [2.14] | [1.83] | [0.38] | [3.50] | [1.86] |
| Publication Peer-Reviewed | | -0.029 | -0.048 | -0.096 | -0.065 | -0.089 |
| | | [0.37] | [0.59] | [1.00] | [0.76] | [0.94] |
| Eval. design:RCT | | 0.003 | -0.034 | -0.013 | -0.109 | 0.002 |
| | | [0.05] | [0.44] | [0.15] | [1.40] | [0.02] |
| High income country | | -0.064 | -0.217** | -0.293** | -0.228** | -0.316*** |
| | | [0.80] | [2.24] | [2.38] | [2.56] | [2.78] |
| _Program design features_ | | | | | | |
| Additional services | | | -0.018 | 0.052 | 0.204** | 0.257*** |
| | | | [0.20] | [0.58] | [2.00] | [2.67] |
| Participant profiling | | | 0.138* | 0.150* | 0.180** | 0.075 |
| | | | [1.92] | [1.66] | [2.56] | [0.80] |
| Participant engagement mechanism | | | 0.183 | 0.113 | 0.121 | 0.195 |
| | | | [1.45] | [0.92] | [1.05] | [1.58] |
| Incentives for service providers | | | 0.178* | 0.225 | 0.123 | 0.05 |
| | | | [1.82] | [1.60] | [1.38] | [0.45] |
| Program has soft skills training | | | | | -0.336*** | -0.233** |
| | | | | | [3.04] | [2.15] |
| _Outcome characteristics_ | | | | | | |
| Estimated unadjusted difference in | | | | 0.049 | | -0.196 |
| | | | | [0.28] | | [1.08] |
| Measured over one year after exit | | | | 0.152** | | 0.248*** |
| | | | | [1.98] | | [3.22] |
| _Target/evaluation group_ | | | | | | |
| Low income / disadvantaged | | | | -0.039 | | 0.025 |
| | | | | [0.34] | | [0.20] |
| Male (base=male and female | | | | -0.054 | | -0.133 |
| | | | | [0.66] | | [1.57] |
| Female (base=male and female | | | | 0.066 | | -0.017 |
| | | | | [0.72] | | [0.27] |
| Younger Participants | | | | -0.015 | | 0.119 |
| | | | | [0.13] | | [1.08] |
| _Type of implementer (base=Private_ | | | | | | |
| Government only | | | | | -0.204* | -0.19 |
| | | | | | [1.84] | [1.03] |
| Private sector only | | | | | 0.195 | 0.264* |
| | | | | | [1.53] | [1.96] |
| R2 | . | . | . | . | . | . |
| N | 1,983 | 1,299 | 927 | 640 | 821 | 590 |
| Number of studies: | 91 | 83 | 59 | 48 | 53 | 42 |
| Number of interventions: | 98 | 92 | 53 | 38 | 47 | 33 |

Notes: see Table 5 notes.

**Table 17. Weighted Least Squares Hedge's g regressions: earnings outcomes only**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category | | | | | | |
| Entr. Prom. | -0.025 | -0.041 | 0.015 | 0.016 | -0.025 | -0.006 |
| | [0.44] | [0.76] | [0.34] | [0.48] | [0.67] | [0.16] |
| Empl. Serv. | -0.113*** | -0.071** | -0.005 | 0.057 | -0.073 | 0.006 |
| | [3.81] | [2.35] | [0.14] | [1.07] | [1.60] | [0.10] |
| Subs. Empl. | -0.120*** | -0.081*** | -0.148*** | -0.097 | -0.151** | -0.101 |
| | [3.47] | [3.01] | [2.87] | [1.39] | [2.63] | [1.55] |
| Unspecified | -0.051 | 0.025 | 0.013 | -0.04 | -0.01 | -0.069 |
| | [1.05] | [0.48] | [0.19] | [0.71] | [0.18] | [1.40] |
| Evaluation features | | | | | | |
| Log Evaluation Sample Size | | -0.019* | -0.017 | -0.042*** | -0.008 | -0.038*** |
| | | [1.72] | [1.30] | [2.79] | [0.71] | [2.77] |
| Publication Peer-Reviewed | | -0.093** | -0.024 | -0.004 | -0.028 | -0.011 |
| | | [2.05] | [0.59] | [0.09] | [0.73] | [0.31] |
| Eval. design:RCT | | -0.064 | -0.118** | -0.122** | -0.096 | -0.106* |
| | | [1.45] | [2.15] | [2.12] | [1.63] | [2.02] |
| High income country | | -0.088** | -0.123*** | -0.169*** | -0.118** | -0.185*** |
| | | [2.38] | [2.74] | [3.79] | [2.34] | [3.79] |
| Program design features | | | | | | |
| Additional services | | | 0.049 | 0.089* | 0.079 | 0.109* |
| | | | [0.98] | [1.70] | [1.35] | [1.90] |
| Participant profiling | | | 0.029 | -0.023 | -0.012 | -0.05 |
| | | | [0.64] | [0.53] | [0.21] | [0.86] |
| Participant engagement | | | 0.110*** | 0.102*** | 0.095*** | 0.092** |
| | | | [2.92] | [3.37] | [2.73] | [2.64] |
| Incentives for service providers | | | 0.036 | 0.000 | 0.029 | -0.005 |
| | | | [0.86] | [0.01] | [0.71] | [0.11] |
| Program has soft skills training | | | | | 0.024 | 0.078* |
| | | | | | [0.70] | [1.88] |
| Outcome characteristics | | | | | | |
| Estimated unadjusted difference | | | | 0.051 | | 0.045 |
| | | | | [0.91] | | [0.74] |
| Measured over one year after exit | | | | 0.052* | | 0.057** |
| | | | | [1.96] | | [2.06] |
| Target/evaluation group | | | | | | |
| Low income / disadvantaged | | | | 0.091** | | 0.014 |
| | | | | [2.62] | | [0.40] |
| Male (base=male and female | | | | -0.064* | | -0.060* |
| | | | | [1.88] | | [1.93] |
| Female (base=male and female | | | | -0.033 | | -0.026 |
| | | | | [1.14] | | [0.97] |
| Younger Participants | | | | 0.004 | | 0.054 |
| | | | | [0.08] | | [1.13] |
| Type of implementer | | | | | | |
| Government only | | | | | 0.06 | 0.088 |
| | | | | | [1.27] | [1.52] |
| Private sector only | | | | | 0.071 | 0.074 |
| | | | | | [1.46] | [1.52] |
| Constant | 0.110*** | 0.320*** | 0.229* | 0.380*** | 0.149 | 0.343*** |
| | [4.03] | [2.98] | [1.98] | [2.73] | [1.43] | [2.77] |
| R2 | 0.06 | 0.17 | 0.33 | 0.37 | 0.32 | 0.42 |
| N | 670 | 670 | 464 | 405 | 452 | 394 |
| Number of studies: | 74 | 74 | 55 | 47 | 51 | 44 |
| Number of interventions: | 77 | 77 | 46 | 36 | 42 | 33 |

Notes: see Table 5 notes.

**Table 18. Random Effects SMD regressions: earnings outcomes only**

| | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category | | | | | | |
| Entr. Prom. | 0.022 | 0.006 | 0.056* | 0.064* | 0.045 | 0.054 |
| | [1.04] | [0.27] | [1.92] | [1.95] | [1.42] | [1.37] |
| Empl. Serv. | -0.041*** | -0.044*** | 0.000 | 0.091** | -0.020 | 0.080 |
| | [3.00] | [3.13] | [0.01] | [2.51] | [0.61] | [1.57] |
| Subs. Empl. | -0.034*** | -0.040*** | -0.074*** | -0.016 | -0.125*** | -0.048 |
| | [3.63] | [3.05] | [3.02] | [0.48] | [4.07] | [1.03] |
| Unspecified | -0.011 | 0.016 | -0.016 | -0.059* | -0.017 | -0.058* |
| | [0.56] | [0.82] | [0.73] | [1.79] | [0.81] | [1.75] |
| Evaluation features | | | | | | |
| Log Evaluation Sample Size | | 0.001 | -0.005 | -0.020*** | -0.001 | -0.016** |
| | | [0.47] | [1.15] | [3.11] | [0.23] | [2.41] |
| Publication Peer-Reviewed | | 0.029*** | 0.018 | 0.026* | 0.013 | 0.022 |
| | | [3.24] | [1.45] | [1.87] | [1.07] | [1.59] |
| Eval. design:RCT | | -0.026*** | -0.088*** | -0.107*** | -0.115*** | -0.129*** |
| | | [2.64] | [5.33] | [5.29] | [5.74] | [5.91] |
| High income country | | -0.039*** | -0.097*** | -0.155*** | -0.115*** | -0.194*** |
| | | [3.89] | [5.58] | [7.52] | [5.34] | [7.27] |
| Program design features | | | | | | |
| Additional services | | | 0.082*** | 0.106*** | 0.108*** | 0.148*** |
| | | | [4.04] | [4.90] | [3.78] | [4.74] |
| Participant profiling | | | 0.037* | 0.012 | 0.024 | -0.015 |
| | | | [1.92] | [0.49] | [0.93] | [0.45] |
| Participant engagement mechanism | | | 0.053*** | 0.092*** | 0.081*** | 0.122*** |
| | | | [3.05] | [4.54] | [4.29] | [5.20] |
| Incentives for service providers | | | 0.069*** | 0.036* | 0.078*** | 0.049** |
| | | | [4.53] | [1.77] | [3.82] | [2.00] |
| Program has soft skills training | | | | | 0.005 | 0.036 |
| | | | | | [0.27] | [1.60] |
| Outcome characteristics | | | | | | |
| Estimated unadjusted difference in | | | | 0.005 | | 0.003 |
| | | | | [0.27] | | [0.14] |
| Measured over one year after exit | | | | 0.047*** | | 0.053*** |
| | | | | [3.79] | | [4.08] |
| Target/evaluation group | | | | | | |
| Low income / disadvantaged | | | | 0.086*** | | 0.062** |
| | | | | [3.57] | | [2.43] |
| Male (base=male and female | | | | 0.001 | | 0.000 |
| | | | | [0.04] | | [0.01] |
| Female (base=male and female | | | | -0.006 | | -0.008 |
| | | | | [0.45] | | [0.58] |
| Younger Participants | | | | -0.002 | | 0.017 |
| Type of implementer (base=Private | | | | | | |
| Government only | | | | | 0.031 | 0.038 |
| | | | | | [1.40] | [1.06] |
| Private sector only | | | | | 0.028 | 0.035 |
| | | | | | [0.99] | [1.11] |
| Constant | 0.037*** | 0.067*** | 0.089** | 0.141** | 0.062 | 0.100 |
| | [9.37] | [2.88] | [2.37] | [2.18] | [1.41] | [1.49] |
| R2 | . | . | . | . | . | . |
| N | 677 | 670 | 464 | 405 | 452 | 394 |
| Number of studies: | 75 | 74 | 55 | 47 | 51 | 44 |
| Number of interventions: | 77 | 77 | 46 | 36 | 42 | 33 |

Notes: see Table 5 notes.

**Table 19. Positive & statistically significant probit regressions: earnings outcomes only**

|  | (1) | (2) | (3) | (4) | (5) | (6) |
|---|---|---|---|---|---|---|
| Main intervention category |  |  |  |  |  |  |
| Entr. Prom. | -0.068 | -0.087 | 0.012 | 0.001 | -0.075 | -0.215* |
|  | [0.53] | [0.64] | [0.09] | [0.01] | [0.75] | [1.91] |
| Empl. Serv. | -0.229 | -0.197 | -0.325** | -0.229 | -0.631*** | -0.629*** |
|  | [1.36] | [1.24] | [2.22] | [1.28] | [4.15] | [3.50] |
| Subs. Empl. | -0.304** | -0.324*** | -0.373*** | 0.000 | -0.653*** | -0.295** |
|  | [2.35] | [2.68] | [2.87] | [0.00] | [5.75] | [2.46] |
| Unspecified | -0.080 | 0.042 | -0.015 | -0.088 | 0.004 | -0.152 |
|  | [0.60] | [0.24] | [0.07] | [0.35] | [0.02] | [0.96] |
| Evaluation features |  |  |  |  |  |  |
| Log Evaluation Sample Size |  | 0.027 | 0.013 | -0.047 | 0.048** | 0.018 |
|  |  | [0.99] | [0.45] | [1.37] | [1.97] | [0.66] |
| Publication Peer-Reviewed |  | -0.092 | 0.038 | 0.044 | 0.015 | -0.011 |
|  |  | [0.91] | [0.37] | [0.49] | [0.25] | [0.23] |
| Eval. design:RCT |  | -0.081 | -0.214** | -0.265*** | -0.288*** | -0.225*** |
|  |  | [0.87] | [2.24] | [3.04] | [3.80] | [3.70] |
| High income country |  | -0.085 | -0.069 | -0.251** | -0.088 | -0.249*** |
|  |  | [0.99] | [0.61] | [2.15] | [0.93] | [2.82] |
| Program design features |  |  |  |  |  |  |
| Additional services |  |  | -0.020 | 0.058 | 0.219** | 0.254*** |
|  |  |  | [0.17] | [0.61] | [2.18] | [3.11] |
| Participant profiling |  |  | 0.177* | 0.187** | 0.206** | 0.200** |
|  |  |  | [1.73] | [1.99] | [1.97] | [2.40] |
| Participant engagement mechanism |  |  | 0.218** | 0.241*** | 0.150** | 0.209*** |
|  |  |  | [2.03] | [2.68] | [1.97] | [3.08] |
| Incentives for service providers |  |  | 0.021 | -0.070 | 0.025 | -0.049 |
|  |  |  | [0.18] | [0.48] | [0.33] | [0.54] |
| Program has soft skills training |  |  |  |  | -0.293*** | -0.207*** |
|  |  |  |  |  | [3.68] | [3.33] |
| Outcome characteristics |  |  |  |  |  |  |
| Estimated unadjusted difference in |  |  |  | -0.084 |  | -0.209 |
|  |  |  |  | [0.47] |  | [1.41] |
| Measured over one year after exit |  |  |  | 0.261*** |  | 0.278*** |
|  |  |  |  | [3.69] |  | [4.45] |
| Target/evaluation group |  |  |  |  |  |  |
| Low income / disadvantaged |  |  |  | 0.182 |  | 0.045 |
|  |  |  |  | [1.28] |  | [0.44] |
| Male (base=male and female | |  |  | -0.058 |  | -0.045 |
|  |  |  |  | [0.75] |  | [0.85] |
| Female (base=male and female | |  |  | -0.119 |  | -0.131** |
|  |  |  |  | [1.37] |  | [2.41] |
| Younger Participants |  |  |  | -0.069 |  | 0.076 |
|  |  |  |  | [0.66] |  | [1.21] |
| Type of implementer (base=Private |  |  |  |  |  |  |
| Government only |  |  |  |  | 0.217* | 0.285** |
|  |  |  |  |  | [1.81] | [2.27] |
| Private sector only |  |  |  |  | 0.255** | 0.315*** |
|  |  |  |  |  | [2.37] | [3.34] |
| R2 | . | . | . | . | . | . |
| N | 949 | 759 | 514 | 428 | 502 | 417 |
| Number of studies: | 84 | 77 | 57 | 48 | 53 | 45 |
| Number of interventions: | 80 | 78 | 47 | 37 | 43 | 34 |

Notes: see Table 5 notes.

# Appendix 4. List of included interventions.

| Program Name | Country | Main intervention category | Impact evaluaton reference(s) |
|---|---|---|---|
| Apprenticeship Training Program and Entrepreneurship Support for Vulnerable Youth | Malawi | Skills training | Cho et al. (2013) |
| Becate | Mexico | Skills training | Van Gameren (2010) |
| Chile Joven | Chile | Skills training | Aedo et al. (2004) |
| | | | Consultores (1999) |
| Contrat de Qualification | France | Skills training | Charpail et al. (2005) |
| | | | Pessao e Costa and Robin (2009) |
| Economic Empowerment of Adolescent Girls (EPAG) | Liberia | Skills training | Adoho et al. (2014) |
| Employment Fund | Nepal | Skills training | Ahmed et al. (2014) |
| Empowerment and Livelihood for Adolescents (ELA) | Uganda | Skills training | Bandiera et al. (2014) |
| Entra 21 | Argentina | Skills training | Alzua et al. (2007) |
| | | | Alzua et al. (2013) |
| Entra 21 | Brazil | Skills training | Alzua et al. (2007) |
| Formacion Tecnica y Tecnologica (FT&T) | Colombia | Skills training | Santa Maria et al. (2009) |
| Formación en Oficios para Jóvenes de Escasos Recursos | Chile | Skills training | Sence (2008) |
| Further training (medium to long-term) (FT) | Germany | Skills training | Caliendo et al. (2011) |
| Galpao | Brazil | Skills training | Kluve et al. (2014) |
| German Apprenticeship Programme | Germany | Skills training | Clark and Fahr (2002) |
| JTPA | United States | Skills training | Bloom et al. (1997) |
| | | | GAO (1996) |
| | | | Heckman and Smith (1999) |
| | | | Heckman and Smith (2000) |
| | | | Heckman et al. (1997) |
| | | | Kornfeld and Bloom (1999) |
| Job Corps | United States | Skills training | Bampasidou (2012) |
| | | | Bampasidou et al. (2014) |
| | | | Blanco et al. (2011) |
| | | | Blanco et al. (2013) |
| | | | Chen (2013) |
| | | | Flores-Lagunes et al. (2010) |
| | | | Frumento et al. (2012) |
| | | | Frölich and Huber (2014) |
| | | | Lee (2009) |
| | | | Schochet et al. (2003) |
| | | | Schochet et al. (2008) |
| | | | Zhang et al. (2009) |

| Program Name | Country | Main intervention category | Impact evaluaton reference(s) |
| --- | --- | --- | --- |
| JobStart | United States | Skills training | Cave et al. (1993) |
| Jordan New Opportunities for Women (NOW) | Jordan | Skills training | Groh et al. (2012) |
| Jovenes En Accion | Colombia | Skills training | Attanasio et al. (2011) |
| Labour Market Training | Sweden | Skills training | Larsson (2003) |
| Lei do Aprendiz | Brazil | Skills training | Corseuil et al. (2014) |
| Livelihoods Training for Adolescent Living | India | Skills training | Mensch et al. (2004) |
| Mandatory internships(Germany) | Germany | Skills training | Saniter (2014) |
| New Chance | United States | Skills training | Chang et al. (2007) |
|  |  |  | Quint et al. (1997) |
| Ohio Transitions To Independence Demonstration | United States | Skills training | Fein et al. (1994) |
| Preparatory training (PT) | Germany | Skills training | Caliendo et al. (2011) |
| Procajoven | Panama | Skills training | Ibarraran and Rosas (2007) |
| Programa Juventud y Empleo | Dominican Republic | Skills training | Card et al. (2011) |
|  |  |  | Ibarraran et al. (2014) |
| Programa de Escuelas taller y Casas de Oficio | Spain | Skills training | Cansino Muñoz-Repiso and Sanchez Braza (2011) |
| Programa de capacitación Jóvenes con Futuro (JCF) | Colombia | Skills training | Santa Maria et al. (2009) |
| Programes de Qualificació Professional Inicial (PQPI) | Spain | Skills training | Blasco et al. (2014) |
| Projoven | Peru | Skills training | Chong and Galdo (2006) |
|  |  |  | Chong and Galdo (2012) |
|  |  |  | Chong et al. (2008) |
|  |  |  | Diaz and Jaramillo (2006) |
|  |  |  | Espinoza Peña (2010) |
|  |  |  | Galdo et al. (2008) |
|  |  |  | Ñopo and Saavedra (2003) |
|  |  |  | Ñopo et al. (2008) |
| Quantum Opportunity Program (QOP) | United States | Skills training | Rodriguez-Planas (2012) |
| Satya/Pratham programme | India | Skills training | Maitra and Mani (2014) |
| School-to-Work Opportunities Act (STWOA) | United States | Skills training | Gong (2005) |
|  |  |  | Griffith (2001) |
|  |  |  | Neumark and Rothstein (2006) |
|  |  |  | Page (2012) |
| Secondary Career and Technical Education programme | United States | Skills training | Hollenbeck and Huang (2006) |
| Senai Vocational Training | Brazil | Skills training | Klasen and Villalobos Barria (2014) |
| Short-term training (STT) | Germany | Skills training | Caliendo et al. (2011) |
| Summer Career Exploration Programme | United States | Skills training | McClanahan et al. (2004) |
| Technical and Vocational Voucher Programme (TVVP) | Kenya | Skills training | Hicks et al. (2013) |
| Utvecklingsgarantin (UVG) | Sweden | Skills training | Carling and Larsson (2005) |

| Program Name | Country | Main intervention category | Impact evaluaton reference(s) |
|---|---|---|---|
| Workforce Investment Act (WIA) | United States | Skills training | Hollenbeck and Huang (2006) |
| Youth Employment and Migration(YEM) | Serbia | Skills training | Arandarenko et al. (2014) |
| Proyecto Joven | Argentina | Skills training | Aedo and Nuñez (2004) |
| | | | Alzua and Brassiolo (2006) |
| | | | Elias et al. (2004) |
| Calificacion De Jovenes Creadores De Microempresas | Peru | Entrepreneurship promotion | Parodi (2003) |
| Créa Jeunes | France | Entrepreneurship promotion | Crepon et al. (2014) |
| Economic Empowerment of Adolescent Girls (EPAG) | Liberia | Entrepreneurship promotion | Adoho et al. (2014) |
| Formacion Empresarial De La Juventud (Project JUMP) | Peru | Entrepreneurship promotion | Parodi (2003) |
| Formacion de Lideres Empresariales | Peru | Entrepreneurship promotion | Jaramillo and Parodi (2005) |
| Jóvenes Rurales Emprendedores | Colombia | Entrepreneurship promotion | Rojas et al. (2010) |
| Partner Microcredit Foundation Experiment | Bosnia and Herzegovina | Entrepreneurship promotion | Bruhn and Zia (2013) |
| Start and Improve Your Business (SIYB) programme | Uganda | Entrepreneurship promotion | Fiala (2014) |
| The Prince's Trust | United Kingdom | Entrepreneurship promotion | Meager et al. (2003) |
| Turning Theses into Enterprises | Tunisia | Entrepreneurship promotion | Almeida et al. (2012) |
| Women's Income Generation Support (WINGS) | Uganda | Entrepreneurship promotion | Blattman et al. (2013) |
| | | | Blattman et al. (2014) |
| Youth Opportunities Programme (YOP) | Uganda | Entrepreneurship promotion | Blattman et al. (2013) |
| Arbeit Sofort! | Germany | Employment services | Schneider et al. (2011) |
| BPO recruiting services | India | Employment services | Jensen (2012) |
| Counseling and Job Placement for Young Graduate Job Seekers | France | Employment services | Crepon et al. (2013) |
| Finnish Vocational Labour Market Training (LMT) | Finland | Employment services | Hämäläinen' et al. (2014) |
| Franklin Subsidized Transport Experiment | Ethiopia | Employment services | Franklin (2014) |
| Job search assistance (JS) | Germany | Employment services | Caliendo et al. (2011) |
| Jordan New Opportunities for Women 2.0 (NOW) | Jordan | Employment services | Groh et al. (2014) |
| Mandatory visits to job info. centers(Germany) | Germany | Employment services | Saniter (2014) |
| Programa Inserjovem | Portugal | Employment services | Centeno and Novo (2006) |
| | | | Centeno et al. (2009) |
| School-to-Work Opportunities Act (STWOA) | United States | Employment services | Gong (2005) |
| | | | Neumark and Rothstein (2006) |
| Contrat Jeune en Entreprise | France | Subsidized employment | Roger and Zamora (2011) |
| JUMP wage subsidies (JWS) | Germany | Subsidized employment | Caliendo et al. (2011) |
| Job creation schemes (JCS) | Germany | Subsidized employment | Caliendo et al. (2011) |
| Jordan New Opportunities for Women (NOW) | Jordan | Subsidized employment | Groh et al. (2012) |
| SGB III wage subsidies (WS) | Germany | Subsidized employment | Caliendo et al. (2011) |

| Program Name | Country | Main intervention category | Impact evaluaton reference(s) |
|---|---|---|---|
| School-to-Work Opportunities Act (STWOA) | United States | Subsidized employment | Gong (2005) |
| | | | Neumark and Rothstein (2006) |
| Special Youth Employment and Training Programme (SYETP) | Australia | Subsidized employment | Knight (2002) |
| | | | Richardson (1998) |
| Stage d'Initiation à la Vie Professionnelle (SIVP) | Tunisia | Subsidized employment | Broecke (2013) |
| Stages de formation | France | Subsidized employment | Brodaty (2007) |
| Stages d'Initiation à la Vie Professionnelle (SIVP) | France | Subsidized employment | Brodaty (2007) |
| Subsidio al Empleo Joven | Chile | Subsidized employment | Universidad de Chile (2012) |
| Swedish employer-paid payroll tax | Sweden | Subsidized employment | Egebark and Kaunitz (2014) |
| Targeted Jobs Tax Credit (TJTC) | United States | Subsidized employment | Hollenbeck and Willke (1991) |
| Travaux d'Utilité Collective (TUC) | France | Subsidized employment | Brodaty (2007) |
| Youth Hires | Canada | Subsidized employment | Webb et al. (2014) |
| Youth Practice | Sweden | Subsidized employment | Costa Dias et al. (2013) |
| | | | Larsson (2003) |
| Youth Wage Subsidies for South Africa | South Africa | Subsidized employment | Levinsohn et al. (2014) |
| 2008 Employment Package | Turkey | Unspecified | Barza (2011) |
| ALMP for disadvantaged youth in Germany | Germany | Unspecified | Ehlert et al. (2012) |
| Apprentices Hiring Programme | Chile | Unspecified | Sence (2006) |
| CET replication sites | United States | Unspecified | Miller et al. (2005) |
| Jordan New Opportunities for Women (NOW) | Jordan | Unspecified | Groh et al. (2012) |
| National Guard Youth Challenge Programme | United States | Unspecified | Millenky et al. (2011) |
| New Deal For The Young Unemployed | United Kingdom | Unspecified | Blundell et al. (2004) |
| | | | De Georgi (2005) |
| | | | Wilkinson (2003) |
| Ninaweza Youth Empowerment Programme | Kenya | Unspecified | Alvares de Azevedo et al. (2013) |
| School-to-Work Opportunities Act (STWOA) | United States | Unspecified | Hall (2000) |
| Teenage Parent Demonstration | United States | Unspecified | Maynard et al. (1993) |
| Youth Opportunity Grant Initiative | United States | Unspecified | Jackson et al. (2007) |

*Notes: There can be more than one intervention per program. When estimates were provided for intervention with no single category of focus programs were placed under unspecified.*