

# Adult Training Methodology

Authors: Matias Busso, Julián Cristia, and Julián Messina



## INCLUSION PROTOCOL

After determining the pool of studies to be analyzed from the search protocol, we dropped all those papers that appeared twice in the dataset. Then, we focused on the title, abstract and the whole study in that order. The filters we analyzed here timing and language, population, intervention, methods, and measures, in that order.

The timing criteria exclude studies published before 1985 and the language criteria include studies only in Spanish or English. The population criteria include individuals older than 15 years old, from any country. We also include interventions that targeted only one gender (for example, females only).

In general, the intervention criteria involve whether or not the program seeks to improve skills to generate improvement in labor market outcomes. These programs are targeted to individuals seeking employment or employed (i.e., not self-employment). The type of skills developed can be job-specific or socioemotional. If programs provide classroom training and other services in addition (e.g., job search assistance, placement), we will include them. **We will exclude the following programs:**

- i. Programs that involve formal education.
- ii. Programs that seek to improve labor market outcomes through other means (i.e., job search assistance, subsidized employment in the private sector, public employment, threats of sanctions for those receiving public subsidies).
- iii. Programs that do not provide any classroom training.
- iv. Programs for entrepreneurs, that is, programs with the objective of increasing self-employment or successfully managing a business.

Another criterion regarding the intervention is that it should be contrasted with a “natural” status quo. If there are two treatment arms compared between each other and there is not a status quo control (representative of “business as usual”), we exclude the evaluation. When the evaluation includes even a small classroom training component, we include it. If the treated individuals in the evaluation choose which training to take and we cannot identify who choose which type of training, we also include the evaluation.

In regards to the methods criteria, we exclude evaluations that use only post cross-sectional data and that try to tackle selection using regressions or matching. On the other hand, we include them if the methods used are one of the following: Randomized Control Trial (RCT), Regression Discontinuity Design (RDD), difference-in-difference or Instrumental Variables (IV). The specific requirements are detailed in [Table 1](#).

We only include papers that measure outcomes related with improvement in market labor conditions. When the outcomes reported in the paper are for all populations (employed and not employed), we required at least one the following: employment, earnings or formality. When the outcome is only for those employed, we required outcomes related to earnings, log earnings or wages (per hour). The excluded outcomes are blue-collar employment, job satisfaction. If the paper has at least one of these outcomes, we keep it; otherwise, we dropped the paper from the study.

•• TABLE 1. Specific Requirements for Methods Included

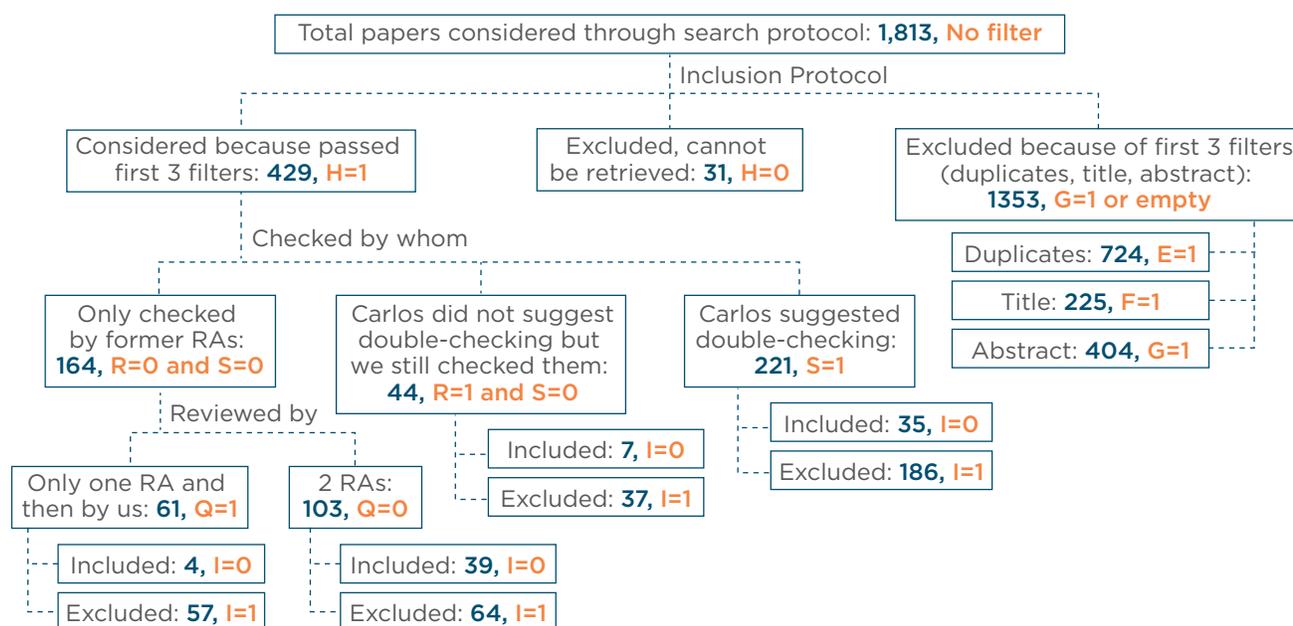
Method	Specific requirements
RCT	It must include a follow-up sample for the treatment group not targeted to compliers.
RDD	Study shows balance in covariates.
	Graph shows no jump of the density distribution for the forcing variable at the threshold.
	Includes convincing argument indicating that procedure to assign units is not prone to manipulation.
Difference-in-difference	Diff-in-diff specification or controlling for baseline outcome.
Difference-in-difference + PSM	Diff-in-diff specification or controlling for baseline outcome and constructing the control group using PSM.
Instrumental variables	F-test for first stage above 10.
	Convincing argument of exogeneity of instrument.

Figure 1 summarizes who read each paper from the Excel Effect sizes and the sheet Papers. In total there are 85 papers included.

Figure 2 explains the holistic evaluations, where there are in total 26+1 holistic papers and 58 job-training papers.

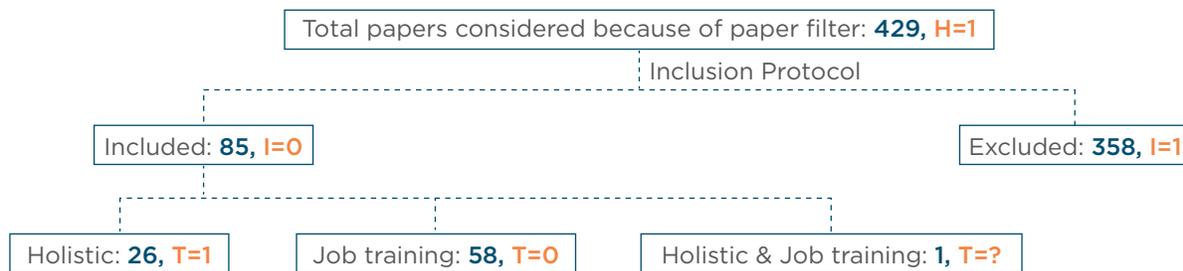
A holistic evaluation is an intervention that groups several programs under the label of training (Perry 2007 i, Michalopoulos 2004 i). In addition, the program has several treatment arms, and the effects are for the whole program; we cannot identify which was related to classroom training (Kornfeld 1999, Nordlund 2011). Finally, another option is, that after several other options, training was offered. (Black 2003 i, Graversen 2008 i, Giorgi 2005 i, Rosholm 2008 iii).

•• FIGURE 1. Stages and Papers Checked



Note: Column E=duplicated papers, F= Title, G=Abstract, H= Retrieved, I=Included/Excluded, Q=2<sup>nd</sup> Stage Retrieval, R= 2<sup>nd</sup> Stage Carlos, S=Review

FIGURE 2. Holistic Papers



Note 1: Column E=duplicated papers, F= Title, G=Abstract, H= Retrieved, I=Included/Excluded, Q=2<sup>nd</sup> Stage Retrieval, R= 2<sup>nd</sup> Stage Carlos, S=Review, T= Is it a holistic evaluation  
 Note 2: Fitzenberger 2016 has two evaluations: one is holistic and the other involves job training.

## CALCULATING STANDARDIZED MEAN DIFFERENCES FOR EFFECTS RELATED TO WAGE

In order to compare effects across studies when the effect sizes are reported under different metrics, we use the Standardized Mean Difference (SMD): the effect size divided by its standard deviation. There are two ways to report the wage related effects in our sample: difference in means and the  $\hat{\beta}$  parameter from a regression.

### Case I: Difference in means

We can estimate the SMD  $d$  as

$$d = \frac{\bar{y}_t - \bar{y}_c}{S_{within}} \left(1 - \frac{3}{4df - 1}\right)$$

where the factor correction of the SMD follows Hedges'  $g$ , a correction for the SMD used in small samples.  $\bar{y}_t$  and  $\bar{y}_c$  are the means for the treated and control group (with sample sizes  $n_t$  and  $n_c$ ).  $S_{within}$  is the within-group standard deviation, pooled across groups, measured at the end line.  $df=n_t+n_c-2$  is the degrees of freedom.

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

The variance of the SMD  $d$  is:

$$V_d = \left(\frac{n_t + n_c}{n_t n_c} + \frac{d^2}{2(n_t + n_c)}\right) \left(1 - \frac{3}{4df - 1}\right)^2$$

### Case II: Regression

Following Kluve (2017) the SMD  $d$  can be calculated from the quotient of the regression coefficient  $\hat{\beta}$  and the standard deviation of the effect  $\hat{\sigma}$  (not the standard error).

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

Due to the lack of information about the standard deviation of the regression residual, we approximate  $\hat{\sigma}$  as  $S_p$  from the standard error of the coefficient  $\hat{\beta}$ .

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

The variance of  $d$  is  $V_d$  and it depends on the SMD ( $d$ ) itself, the t-statistic  $t$  and the degrees of freedom of the regression coefficient which is in general  $v=n-k$ . In this case we stated  $k=2$  because it is the minimum number of coefficients estimated.  $d$  and  $V_d$  are not sensible to the value of  $k$  in our sensitivity analysis.

$$V_d = \frac{d^2}{v - 2} \left(\frac{v}{t^2} + v c^2(v) - v + 2\right)$$

Where  $\frac{1}{c(v)} = \sqrt{\frac{v \Gamma(\frac{v}{2} - 0.5)}{\Gamma(\frac{v}{2})}}$  and  $\Gamma()$  is the gamma function. When the study has a large  $n$ ,  $c(v)=1$

<sup>1</sup> Skemer 2017 reports a 90% coefficient interval and instead of 1.96 it is 1.64.

Some papers report de confidence interval, in which case we used the formula

$$IC = \hat{\beta} \pm 1.96^1 \times SE$$

to recover  $SE$ . When the paper reported the t-stat we used the formula

$$\frac{\hat{\beta}}{SE} = t$$

## GETTING THE FOREST PLOT

Evaluations usually report several estimates regarding the effect of the intervention on a given outcome (in this case, employment or wages). For example, the effects for an evaluation can be reported at 6, 9, 12 and 15 months. In order to compare these effects on the same population and same treatment, we follow Borenstein (2009)

$$\bar{Y} = \frac{1}{m} \left( \sum_j Y_j \right)$$

$$V_{\bar{Y}} = \left( \frac{1}{m} \right)^2 \left( \sum_{j=i}^m V_i + \sum_{j \neq k} r_{jk} \sqrt{V_j} \sqrt{V_k} \right)$$

where  $m$  is in the number of effects in different periods,  $j$  is each effect,  $Y_j$  is the effect of the evaluation (either SMD or change in probability of employment)  $V_j$  is the variance of  $Y_j$ , and  $r_{jk}$  is the correlation among different periods  $j$  and  $k$  of the evaluation. Following the *Methodology for Skills-Bank* we assume a 0.5 correlation.

There are some<sup>2</sup> evaluations without results for the whole population; instead, there are either men or women separated. We group these outcomes without correlation since they are different subsamples. Also, there are no differences in the sample size of male and female subsamples.

Borenstein (2009) uses a fixed effect meta-analysis in order to combine these subgroups. The effect size is  $\bar{Y}$ , and it is a weighted average of each effect on population, male or female.

$$\bar{Y} = \frac{\sum W_i Y_i}{\sum W_i}, i = \{M, F\}$$

where the weight  $W_i$  is the inverse of the variance of

$$Y_i: W_i = \frac{1}{V_{Y_i}}$$

The variance of  $\bar{Y}$ ,  $V_{\bar{Y}}$  is

$$V_M = \frac{1}{\sum^k W_i}$$

Once we have one effect per evaluation, these evaluations are combined in a meta-analysis using random effects because all studies come from different populations and different methods.

Additionally, in order to summarize one effect per intervention, we selected from each intervention the most aggregated effect. For example, if there are effects for male, female, and pooled sample we use only the pooled sample. If there are effects for several cities or regions and an aggregated effect, we use the latter. This work depended largely on how each paper reported each intervention and was done case by case.

<sup>2</sup> Total of 13 evaluations: Acevedo 2017 i, Acevedo 2017 ii, Adoho 2014 i, Biewen 2014 i, Biewen 2014 ii, Fitzenberger 2000, GAO 1996, Groh 2016, Honorati 2015, Maitra 2017, De Hoop 2018, Mueser 2007, Stenberg 2015.

Copyright © 2022 Inter-American Development Bank. This work is licensed under a Creative Commons IGO 3.0 Attribution-Non-Commercial-NoDerivatives (CC-IGO BY-NC-ND 3.0 IGO) license (<https://creativecommons.org/licenses/by-nc-nd/3.0/igo/legalcode>) and may be reproduced with attribution to the IDB and for any non-commercial purpose. No derivative work is allowed.

Any dispute related to the use of the works of the IDB that cannot be settled amicably shall be submitted to arbitration pursuant to the UNCITRAL rules. The use of the IDB's name for any purpose other than for attribution, and the use of IDB's logo shall be subject to a separate written license agreement between the IDB and the user and is not authorized as part of this CC-IGO license.

Note that link provided above includes additional terms and conditions of the license.

The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the Inter-American Development Bank, its Board of Directors, or the countries they represent.

