Session 4a: Regression Analysis in CIE

C4ED – EUTF

October 2023

Center for Evaluation and Development

European Commission

Welcome to the Training Workshop on Counterfactual Impact Evaluation (CIE)

The material of this workshop was produced with the financial support of the European Union. Its contents are the sole responsibility of C4ED and do not necessarily reflect the views of the European Union

Center for Evaluation and Development

European Commission

Introduction

Communication during the training

MUTE BUTTON

QUESTIONS

FEEDBACK

Communication during the training

MUTE BUTTON

QUESTIONS

FEEDBACK

Communication during the training

MUTE BUTTON

QUESTIONS

FEEDBACK

Asking Questions

- Please post your questions in the chat room
- Like  the questions of others, so we know they are particularly relevant for you as well
- Carolin will read out all questions and we will answer these at once
- Use the longer breaks to ask more questions

Communication during the training

MUTE BUTTON

QUESTIONS

FEEDBACK

Center for Evaluation and Development

Asking Questions

European Commission

- Please make suggestions
- Feel free to share your comments
- More feedback and questions (especially for the Q&A session)

Day 1 Agenda

10:00 – 10:15	Welcome and Q&A on day 1
10:15 – 11:05	**Session 4a**: Regression analysis for CIE
11:05 – 11:20	Break (15 minutes)
11:20 – 11:40	**Session 4b**: Case Study presentation
11:40 – 11:50	**Q&A**
11:50 – 12:45	**Session 5a**: Qualitative analysis and triangulation for CIE
12:45 – 13:45	Lunch (60 minutes)
13:45 – 14:05	**Session 5b:** Case Study Presentation
14:05 – 14:50	**Session 6a:** Understanding and using outputs of CIE
14:50 – 15:00	Break (10 minutes)
15:00 – 15:40	Session 6b: Breakout session
15:40 – 16:05	Final quiz
16:05 – 16:15	Closing Day 2

Center for Evaluation and Development

European Commission

Session 4a: Regression analysis for CIE

Statistical Testing – Recap

- We want to test a hypothesis about the *population* of interest
- However, in practice we usually cannot gather information on the whole population → we use a *sample*
- 2 sources of uncertainty:

1. From trying to draw conclusions for the *population* based on *partial information* – i.e., sample data
2. From the process of *choosing/selecting* the sample itself

→Inferential statistics (confidence intervals, t-test, etc.) explicitly account for both sources of uncertainty when testing hypotheses

Center for Evaluation and Development

Statistical Testing – Recap

European Commission

- Steps of a statistical test:

Define the null hypothesis → Calculate the test statistic → Set the confidence level and the critical value → Conclude

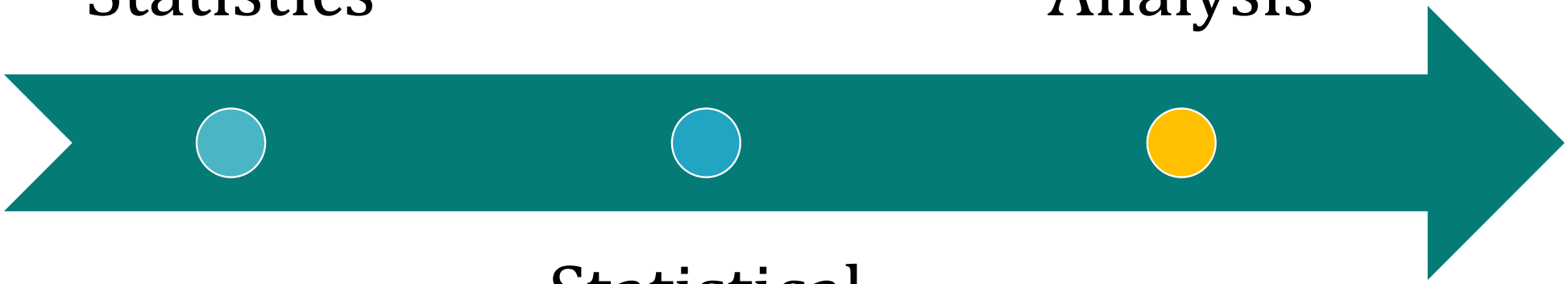
- In practice, you can use the *p-value* to conclude
  - p-value = probability of being wrong when *rejecting* the null hypothesis
  - p-value is low enough → High confidence to reject the null hypothesis
- Good news: you need only know the null hypothesis and the p-value, and you can read/interpret a test result!
- t-test to compare means is not enough to establish causality → Need to account for other factors that could explain the difference!

Center for Evaluation and Development

European Commission

Descriptive Statistics

Regression Analysis



Statistical Testing

Regression Analysis in CIE
Overview

- What is a regression?

→Quantify the relationship between variables

- Statistical testing in regressions

- Why use regression analysis?

→Controlling for confounding variables

- Multivariate regressions

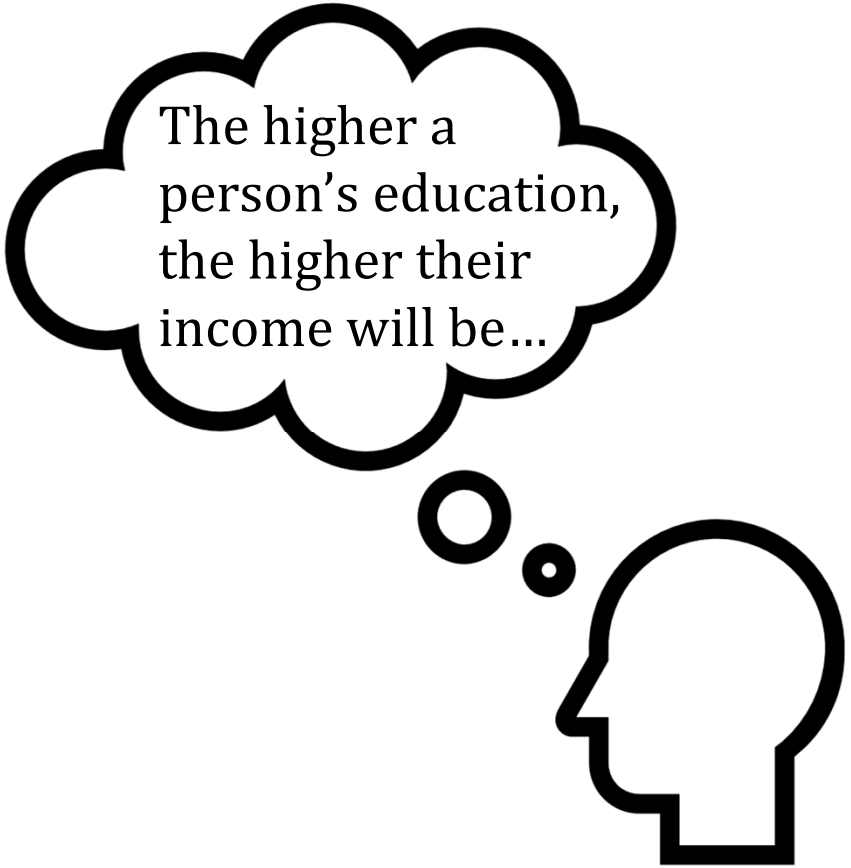
- Regression Analysis and causality

What is Regression Analysis?

Regression Analysis in CIE

What is regression analysis? – Example setup

- In the previous session, we considered the link between education and income
- For ease of exposure, we used a simplified measure of education – i.e., completion of primary school
- Now, we still focus on income and education, but let's assume we measure the latter by the *number of years of education completed*

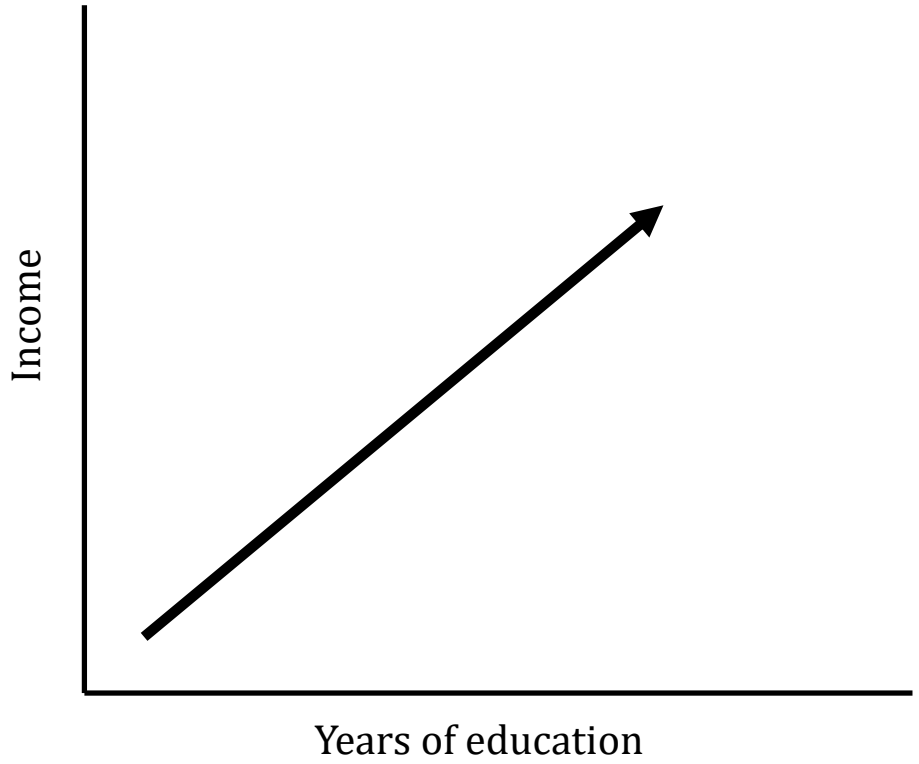


Regression Analysis in CIE

What is regression analysis?

- We focus on income and education measured as *number of years of education completed*
- *Regression analysis* allows to investigate the relationship between income and years of education

→"If I stay in school for one more year, how much more income can I expect?"



Regression Analysis in CIE

What is regression analysis?

- We select a sample of people and ask them how many years of school they have completed and their monthly income

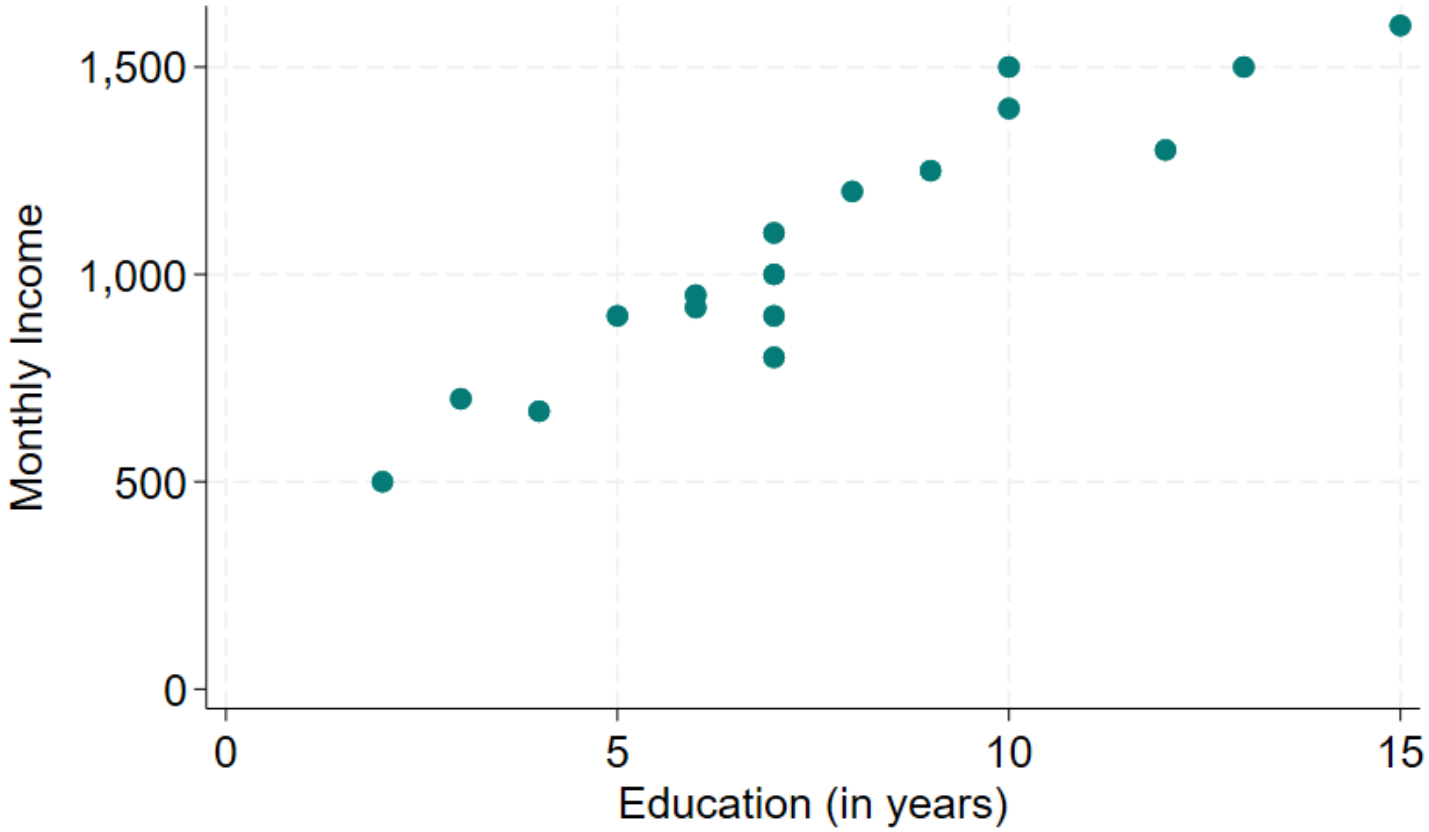
- Let's visualize the data on a *scatter plot*

Observation	Education (in years)	Monthly Income
Person 1	7	1000
Person 2	2	500
Person 3	3	700
Person 4	7	1100
Person 5	8	1200
Person 6	7	900
Person 7	7	800
Person 8	12	1300
Person 9	10	1400
Person 10	4	670
Person 11	6	920
Person 12	13	1500
Person 13	10	1500
Person 14	5	900
Person 15	15	1600
Person 16	6	950
Person 17	9	1250

Regression Analysis in CIE

What is regression analysis?

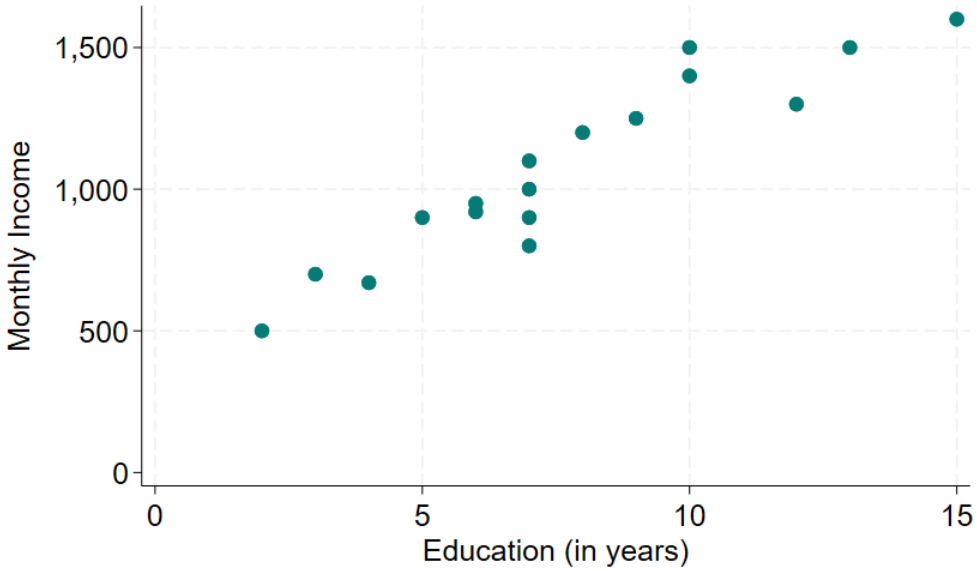
Scatter plot = each point represents a person's years of education (x-axis) and monthly income (y-axis)



Regression Analysis in CIE

What is regression analysis?

- Visually, there appears to be some positive relationship between years of education and monthly income – i.e., we tend to observe higher values of income for higher values of education



- Regression analysis allows us to:

1. *Quantify* the relationship between two variables – i.e., to measure by *how much* income increases when completing one extra year of education
2. Gauge whether the relationship is due to chance or whether it is *significant* → statistical testing

Regression Analysis in CIE

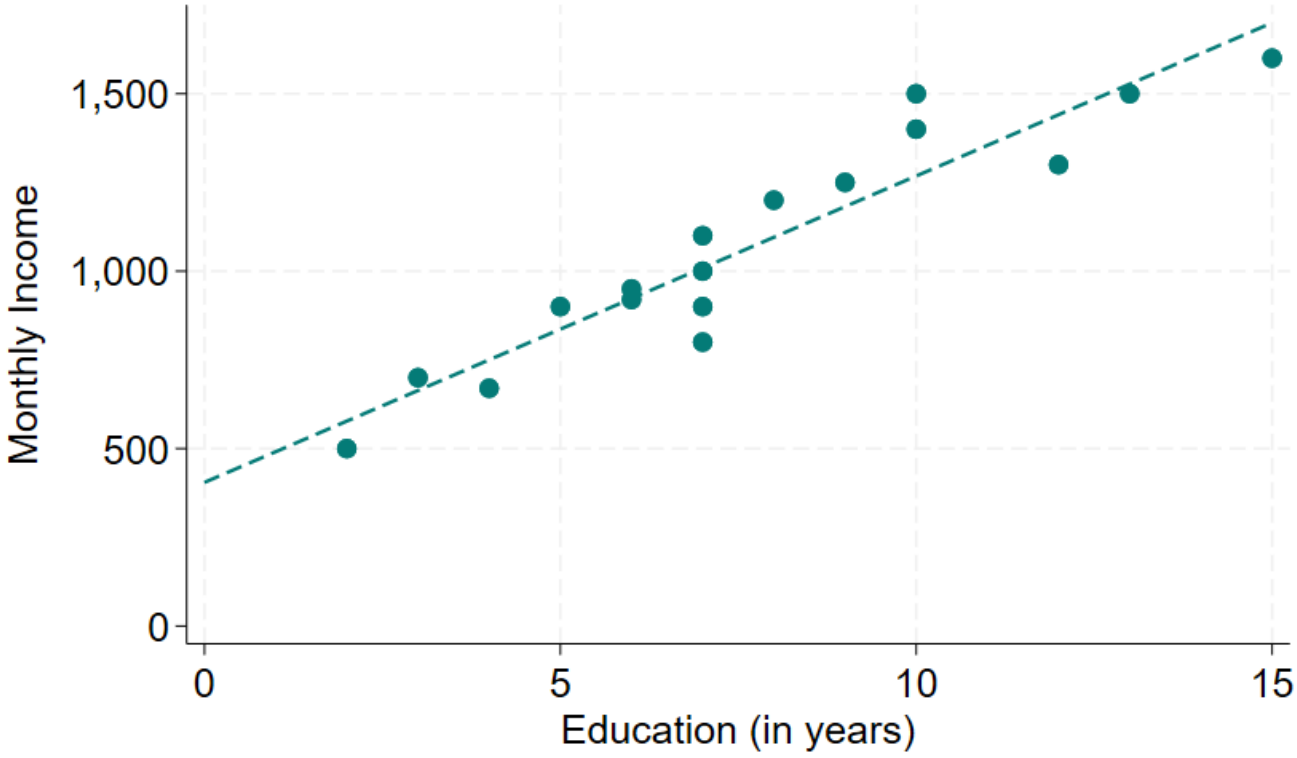
What is regression analysis?

- To quantify the relationship between income and education, we can fit a line to our data points

- You may be wondering:

➢Why a *line*?

➢Why *this particular* line?



Center for Evaluation and Development

Regression – Why a *line*?

Regression Analysis in CIE

Regression – Why a line?

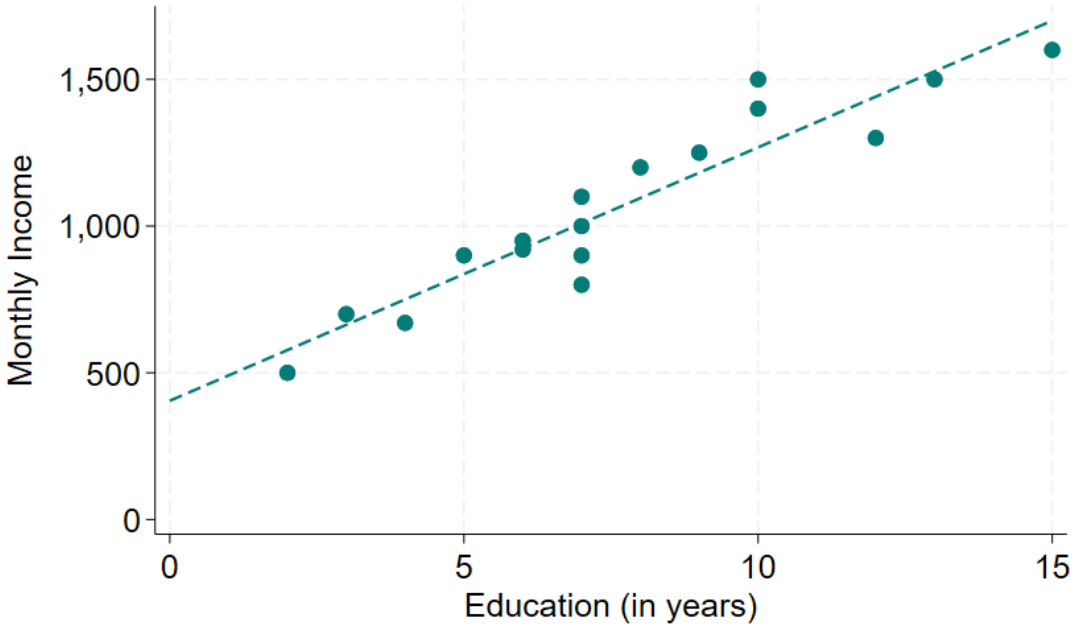
- Why a *line*?

➢ The line is the *simplest*, most intuitive mathematical tool available to describe the relationship between two variables

➢ Recall from your maths lessons:

$$y = a + bx$$

➢ When x increases by 1 unit, y changes by b units



Regression Analysis in CIE

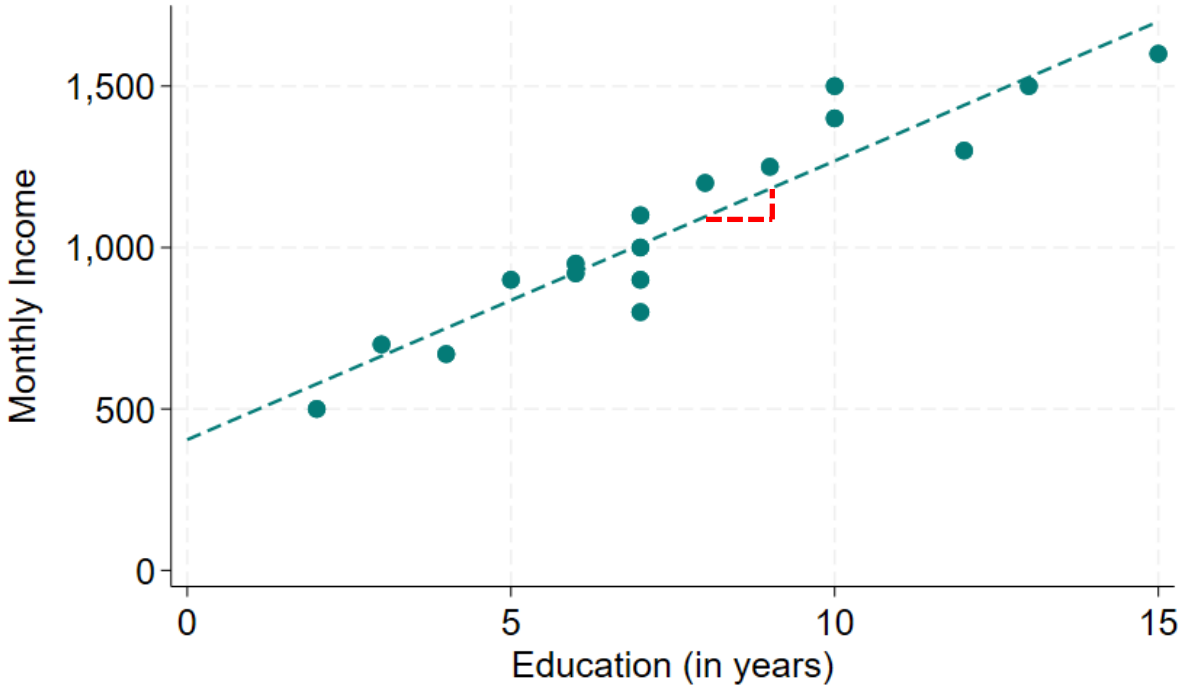
Regression – Why a line?

- *Lines* are intuitive → $y = a + bx$

➢In our case:

This is what we're interested in

Monthly income = Intercept + (Slope of the line x Years of education)



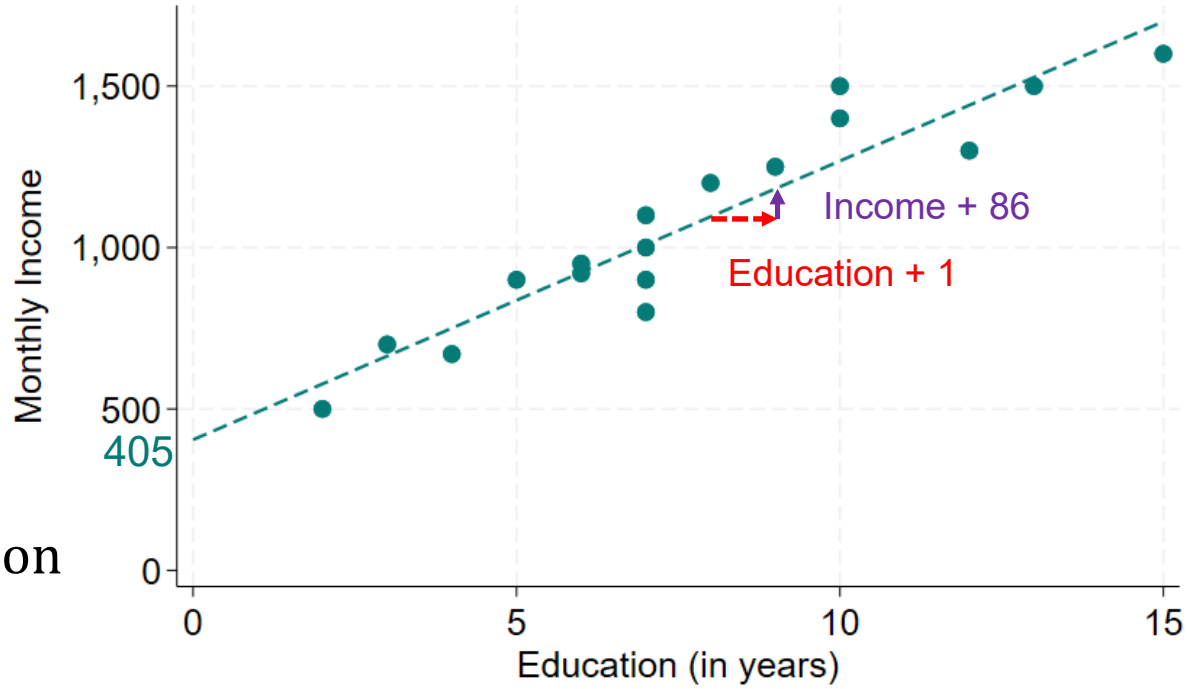
Regression Analysis in CIE

Regression – Why a line?

- *Lines* are intuitive

Slope of the line = by how much monthly income increases when years of education increase by 1

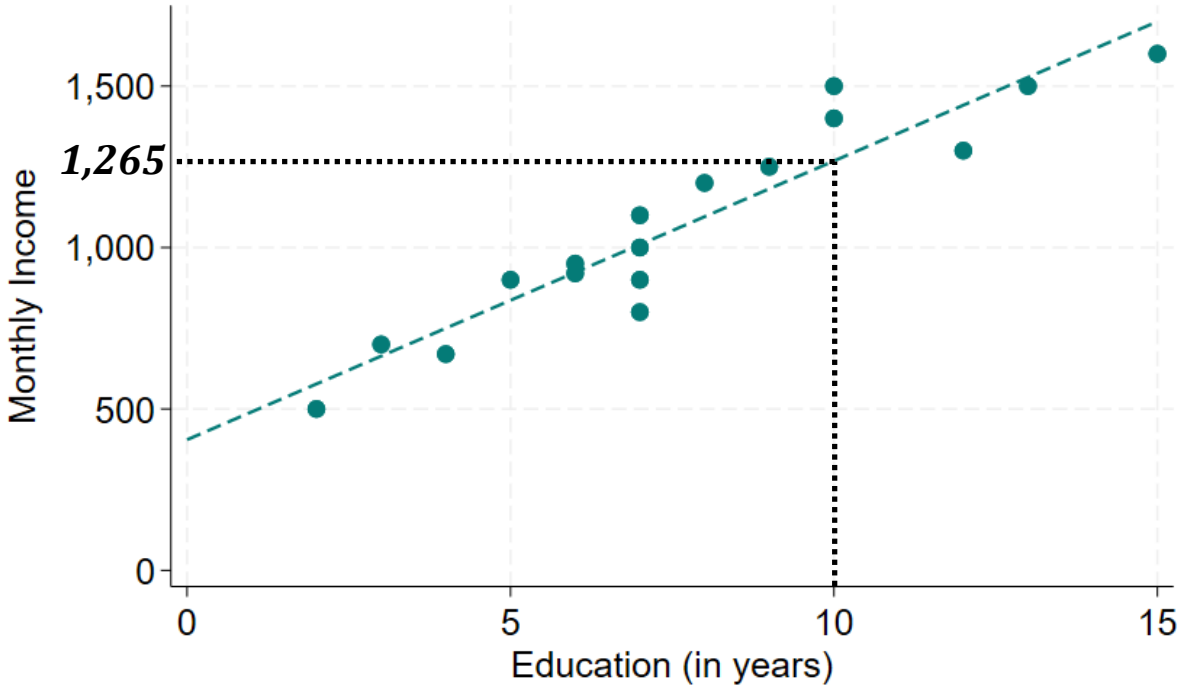
Monthly income =
405 + 86 x Years of education



Regression Analysis in CIE

Regression – Why a line?

- Monthly income = 405 + 86 x Years of education
- Can you predict how much monthly income a person with 10 years of education can expect to earn?



Center for Evaluation and Development

Regression – Why *this particular* line?

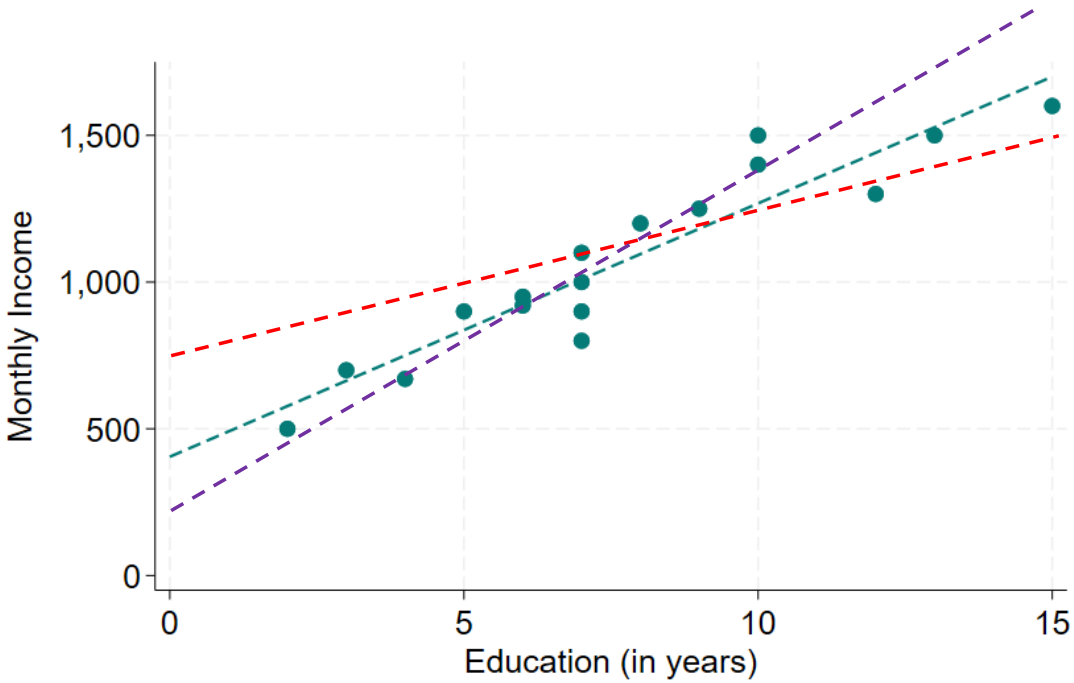
Regression Analysis in CIE

Regression – Why this particular line?

- Why *this particular* line?

➢Why not **this one**?

➢Or **this one**?



Regression Analysis in CIE

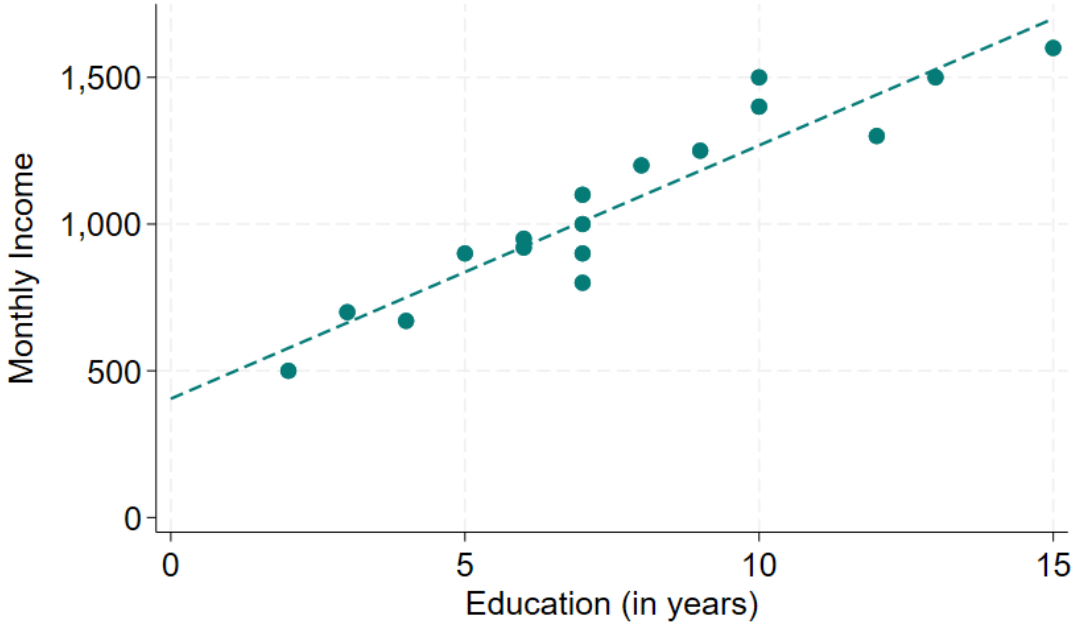
Regression – Why this particular line?

- Why *this particular* line?

- We did not choose this line randomly

→It is the *best* line we can fit – i.e., this is the line that comes *closest to all data points*

- Technically, we find it using **Ordinary Least Squares (OLS)**
- In regression analysis, OLS will always find the *best line* to fit your data (under some conditions)



Regression Analysis in CIE

The regression line – Recap

- We want to explore the relationship between a variable *y* – e.g., income – and a variable *x* – e.g., education
- *y* and *x* take many values → the easiest way to represent their relationship is a *line* ⇒ $y = a + bx$
- b = slope of the line → This is what interests us
- *Regression analysis* = find values for a and b
- The OLS line is the *best* regression line – i.e., it comes closest to all data points (under some conditions)

→What to make of the results from regression analysis?

Interpretation and Statistical Testing in Regression Analysis

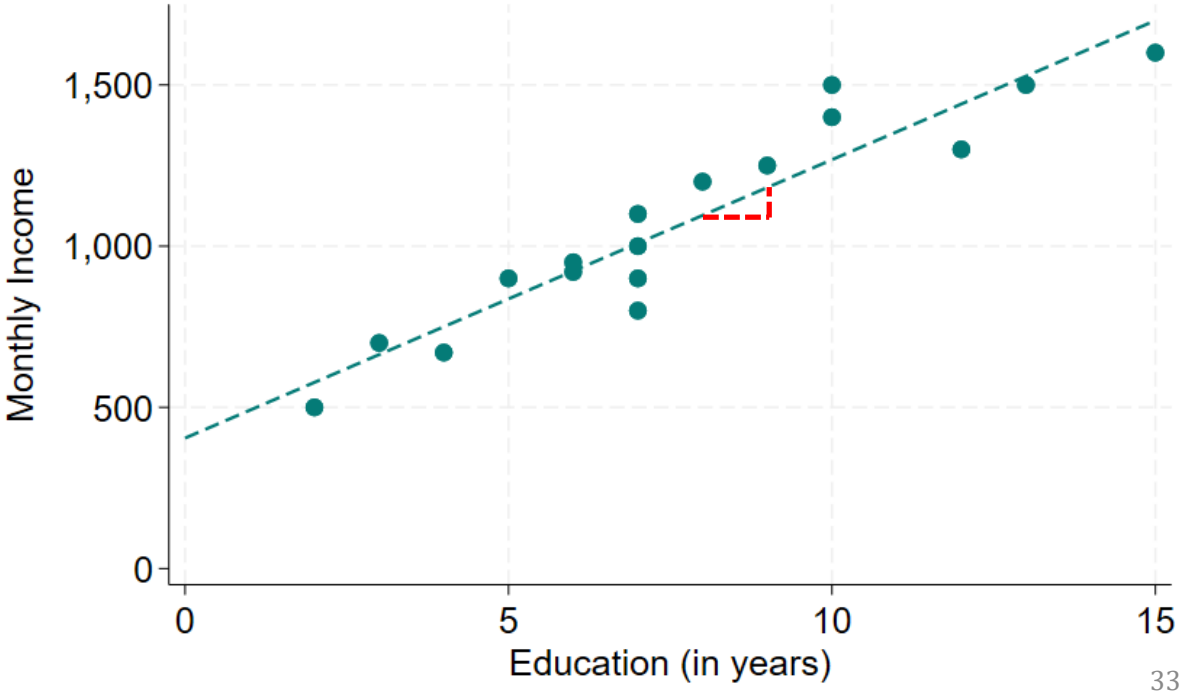
Regression Analysis in CIE

Regression line – Interpretation

- Line slopes upwards → Positive relationship
- Line slopes downwards → Negative relationship
- Line is horizontal → No relationship

- But we are using *sample data* to try to learn about the *population* →how confident are we about the results from this regression?

➢We need a statistical test



Regression Analysis in CIE

Regression line – Statistical Test

- We represent the relationship of interest by a line:

  Monthly Income = **a** + **b** x Years of education

- Regression analysis aims to find values for parameters **a** and **b** → we call these values ***estimates***
- The result:

  Monthly Income = **405** + **86** x Years of education

→How reliable are these results/estimates?

- Recall the steps of statistical testing:
  - Null hypothesis
  - Calculate test statistic
  - Calculate the p-value (or set the level of confidence and find the critical value)
  - Set the significance level and use the p-value to conclude (or compare the test statistic and the critical value)

Regression Analysis in CIE

Regression line – Statistical Test

Monthly Income = **a** + **b** x Years of education

Estimates: **a** = 405, **b** = 86

- What question do we want to ask here?

→How confident can we be that the positive relationship between income and education is true in the population?

- What null hypothesis could we formulate to test this?

[TIP: Recall that Null hypotheses are usually framed in terms of "no effect/no difference"]

→Null hypothesis: "There is no relationship between income and education in the population".

→In other words: "The true value of parameter **b** in the population is 0."

Regression Analysis in CIE

Regression line – Statistical Test

Monthly Income = **a** + **b** x Years of education

Estimates: **a** = 405, **b** = 86

- Null hypothesis: "The true value of parameter **b** in the population is 0."
- Now that we have the null hypothesis, all we need is the test **p-value**
- In this example, p-value = 0.000
- What do you decide for the test?

→We can reject the null hypothesis at the 1% significance level – i.e., with at least 99% confidence

- What do you conclude?

→Parameter **b** is *statistically significantly different from 0*

→Hence, there is a *statistically significant* positive relationship between income and education in the population of interest

Regression Analysis in CIE

Statistical Tests in regressions – Remarks

- *Remark 1*: We focussed on parameter **b** – the slope of the regression line – because it is what interests us most, but we can carry out the same test for all the estimated parameters in the regression – e.g., we can test whether parameter **a** is truly equal to 0 in the population.

- *Remark 2*: For regressions, most statistical software provide you with the p-value for the significance test of each estimated parameter.

- *Remark 3*: Technically, the test used here is a **t-test** – i.e., the same family of tests we used to compare means between two groups – which allows to test whether the parameter is equal to any specific value, not just 0.
  - ➢ E.g., Null hypothesis: "The true value of parameter **b** in the population is 100."
  - ➢ With our data, p-value = 0.12 → What do you conclude?

Regression Analysis in CIE

What we covered so far…

- **What is regression analysis?**
- Regression analysis allows us to:

1. *Quantify* the relationship between two variables – by how much does income increase when years of education increase by 1?
2. Gauge whether the relationship is due to chance or whether it is *statistically significant* → statistical testing

- We finally get to the big question:

→***Why*** use regression analysis?

Center for Evaluation and Development

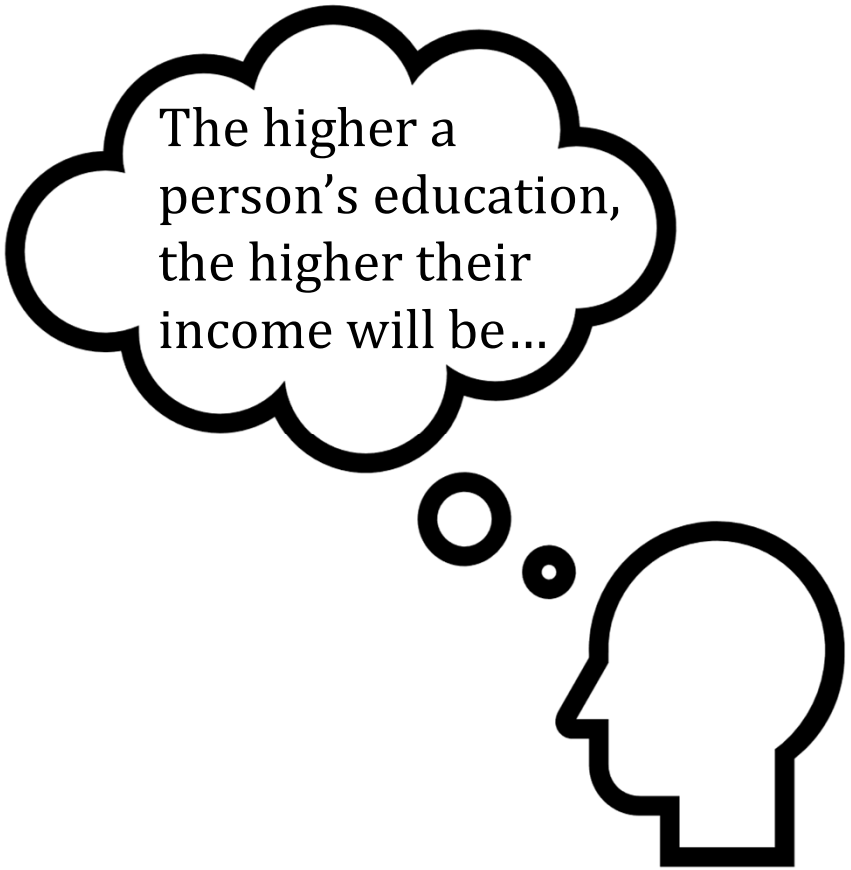
Why use Regression Analysis?

Regression Analysis in CIE
Why use regression analysis?

- In the previous session, we considered the link between education – i.e., completing primary school – and income
- We tested our hypothesis using sample data
- We used a **t-test** to compare the (sample) average income in both groups

→Null hypothesis: "The difference in income between those who completed primary education and those who did not is 0."

→p-value = 0.0319



Regression Analysis in CIE

Why use regression analysis?

→Null hypothesis: "The difference in income between those who completed primary education and those who did not is 0."

→ p-value = 0.0319

- We rejected the null hypothesis with 95% confidence, and concluded that the difference in income between the two groups is *statistically significant* – i.e., we are highly confident that this difference really exists in the population

- However, we cannot say that completing primary education *completely* explains the difference in income, or that it directly causes it

Regression Analysis in CIE

Why use regression analysis? – Confounders

- We cannot say that completing primary education *completely* explains the difference in income, or that it directly causes it
- There could be other important factors driving the difference

➢ *Remark 1*: such factors are called **confounders** or **confounding variables**, because they blur/disturb the relationship of interest between education and income

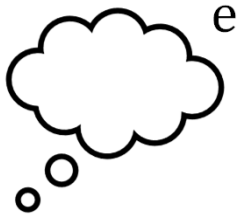
➢ *Remark 2*: **confounders** cause problems because they are correlated with both education *and* income at the same time

Can you think of examples of confounders in the case of education and income?

Regression Analysis in CIE
Confounders – Example

- Along with information on education and income, we asked people in our sample to report the average income of their parents

→How/Why could parental income be a confounder in the relationship between education and income?



Observation	Education (in years)	Monthly Income	Parental Income
Person 1	7	1,000	900
Person 2	2	500	800
Person 3	3	700	900
Person 4	7	1,100	2,000
Person 5	8	1,200	1,800
Person 6	7	900	1,500
Person 7	7	800	1,600
Person 8	12	1,300	900
Person 9	10	1,400	1,100
Person 10	4	670	850
Person 11	6	920	750
Person 12	13	1,500	1,900
Person 13	10	1,500	2,500
Person 14	5	900	1,350
Person 15	15	1,600	1,600
Person 16	6	950	950
Person 17	9	1,250	1,050

Regression Analysis in CIE
Confounders – Example

- How/Why could parental income be a confounder in the relationship between education and income?

→If parental income is correlated with both education and income

- Let's plot the data and fit regression lines to see if that might be the case

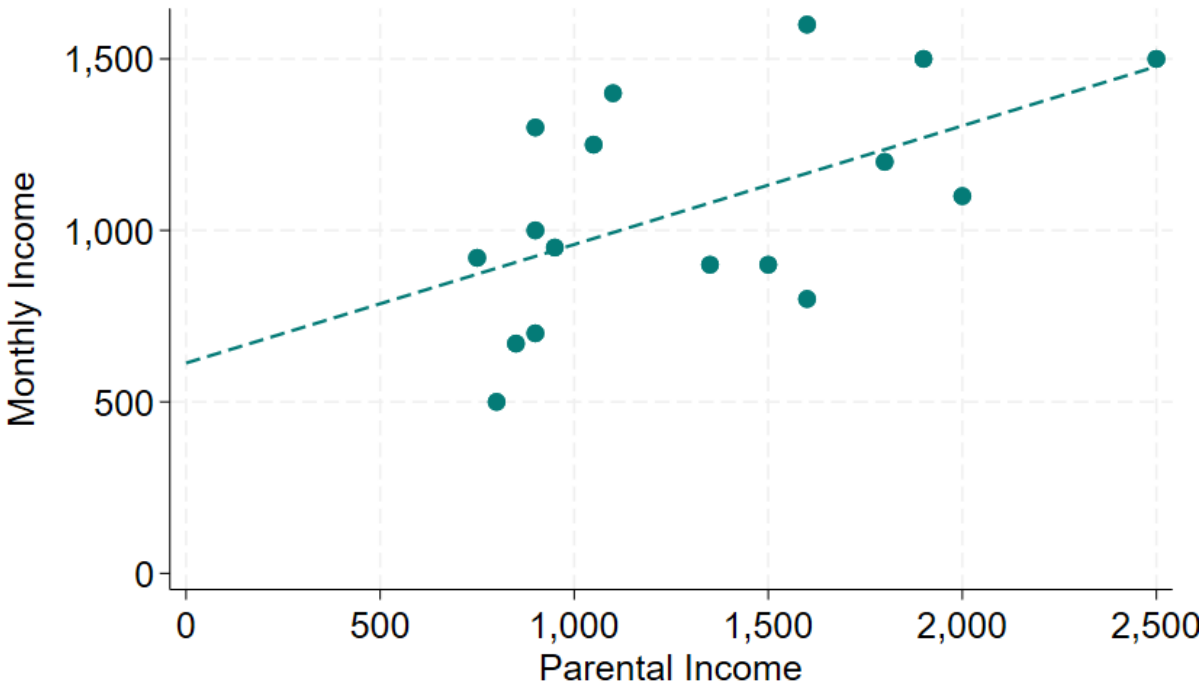
Regression Analysis in CIE
Confounders – Example

- Monthly income and parental income

- Line slopes upward → positive relationship
- Estimated value of slope parameter = 0.35
- p-value = 0.022

→Conclusion?

→There appears to be a *significant* relationship between income and parental income in the population



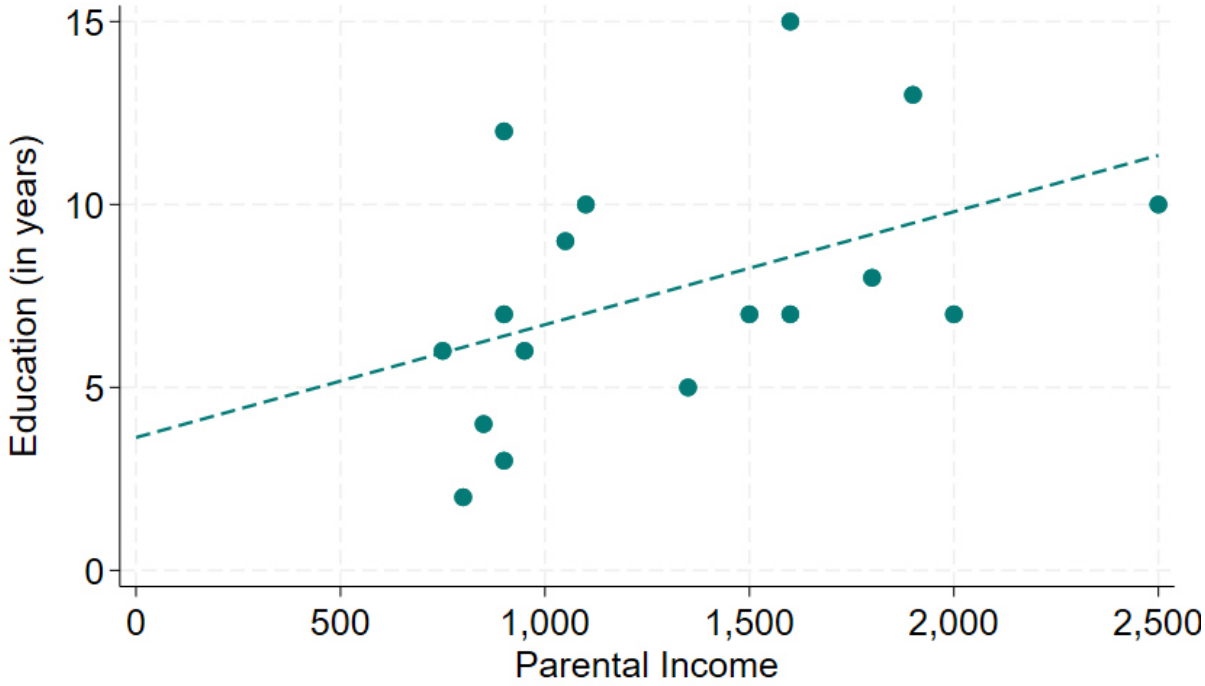
Regression Analysis in CIE
Confounders – Example

- Education and parental income

- Line slopes upward → positive relationship
- Estimated value of slope parameter = 0.003
- p-value = 0.068

→Conclusion?

→There appears to be a *significant* relationship between education and parental income in the population



Regression Analysis in CIE
Confounders – Example

- Our first regression analysis allowed us to conclude there was a significant positive relationship between education and income

- However, we just saw that *parental income* is also significantly correlated with both income and education
- In addition, we have good explanations (theoretical/conceptual) why this correlation should exist

→We should probably account for parental income in our regression analysis

Regression Analysis in CIE

Why use regression analysis? – Confounders

- We want to account for parental income in our regression analysis. But wait…

- Up to now, we only talked about *bivariate* regressions – i.e., the relationship between 2 variables only e.g., income and education

- So, how do we account for a *third* variable in our analysis?

→Multivariate regressions

Multivariate Regressions

Regression Analysis in CIE
Multivariate Regressions

- Bivariate regression → We try to explain *y* as a linear function of *x*

$$y = a + bx$$

- In our example: Monthly income = **a** + **b** x Education
- Multivariate regression → Try to explain *y* as a linear function of *multiple x*'s – i.e., there can be more than one *x* variable!

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 \ldots$$

- In our example, we can simply add parental income in the regression:

Monthly income = **a** + $\mathbf{b_1}$ x Education + $\mathbf{b_2}$ x Parental Income

Regression Analysis in CIE

Multivariate Regressions – Remarks

$$y = a + b_1x_1 + b_2x_2 + b_3x_3...$$

Monthly income = **a** + $\mathbf{b_1}$ x Education + $\mathbf{b_2}$ x Parental Income

- The intuitions for bivariate regressions carry over to multivariate regressions:
  - Positive/Negative slope parameter indicates a positive/negative relationship
  - You can test the statistical significance of each parameter with a t-test
  - Ordinary Least Squares (OLS) give the best linear approximation
- Note: unfortunately, with more than 2 variables, we cannot use 2-dimension visualization tools (scatter plot, regression line)

Regression Analysis in CIE

Multivariate Regressions – Remarks (cont'd)

$$y = a + b_1x_1 + b_2x_2 + b_3x_3...$$

Monthly income = **a** + $\mathbf{b_1}$ x Education + $\mathbf{b_2}$ x Parental Income

- In principle, you can add as many *x* variables as you want.
- In practice, this is limited by:
  - Sample size (degrees of freedom)
  - Theory/Concepts → do not add variables just because you can!

Regression Analysis in CIE

Multivariate Regressions – Example

- Our first (bivariate) regression yielded – p-value in parentheses:

Monthly income = **405** + **86** x Education
(0.000)

- Now, let's include parental income

→We say we **control** for confounding variables

- New results:

Monthly income = **324** + **80** x Education + **0.1** x Parental Income
(0.000) (0.115)

→What can you say about the new results?

Regression Analysis in CIE
Multivariate Regressions – Example

- New results – **controlling** for parental income:

  Monthly income = **324** + **80** x Education + **0.1** x Parental Income
  (0.000) (0.115)

- Income and education: Positive relationship, strongly significant.
  → *Remark*: the estimated parameter value is *lower* than before
- Income and parental income: Positive relationship, not significant.
  → *Remark*: parameter value on "parental income" in bivariate regression was larger and significant (see Slide 45)

→ The goal of multivariate regressions is to understand how variables behave *jointly/in combination*, instead of just by pairs

Regression Analysis in CIE

Regression and confounders – Another example

- Along with information on education, income, and parental income, we have data on gender

→How/Why could gender be a confounder in the relationship between education and income?

Observation	Education (in years)	Monthly Income	Gender
Person 1	7	1000	Male
Person 2	2	500	Female
Person 3	3	700	Female
Person 4	7	1100	Female
Person 5	8	1200	Male
Person 6	7	900	Female
Person 7	7	800	Male
Person 8	12	1300	Male
Person 9	10	1400	Male
Person 10	4	670	Female
Person 11	6	920	Female
Person 12	13	1500	Male
Person 13	10	1500	Male
Person 14	5	900	Female
Person 15	15	1600	Male
Person 16	6	950	Male
Person 17	9	1250	Female

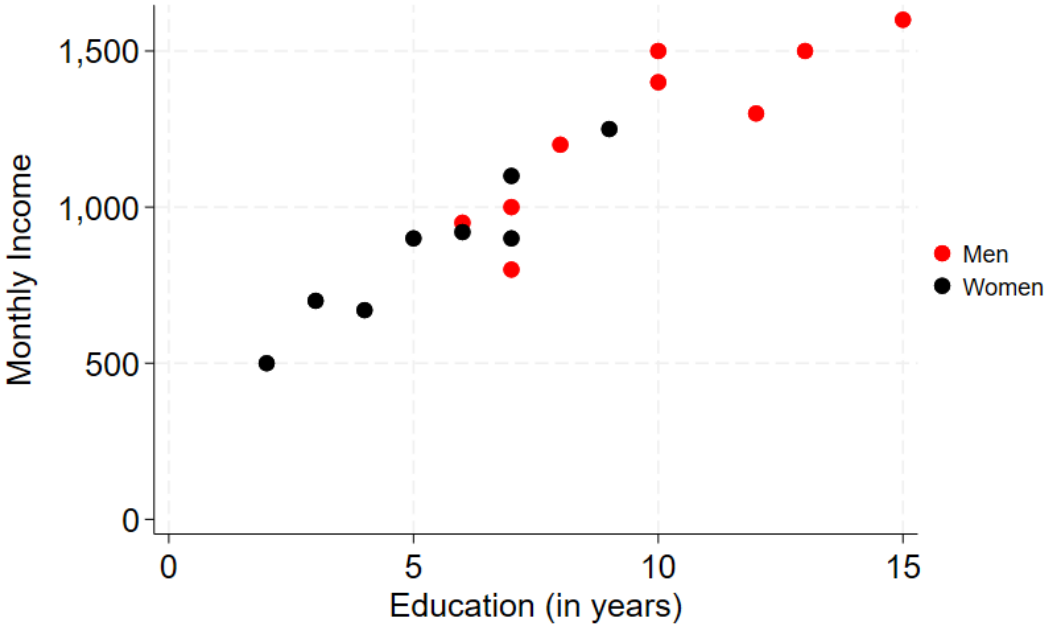
Regression Analysis in CIE

Regression and confounders – Another example

- Income, education and gender
- Visually, it seems men experience high education/high income combinations more often than women
- Let's see what the data say:

Variable	Men	Women	p-value (t-test)
Income	1,250	867.5	0.009
Education	9.8	5.4	0.005

→Is gender potentially a confounding variable?



Regression Analysis in CIE

Regression and confounders – Another example

- Let's control for gender as well:

Monthly income = **324** + **80** x Education + **0.1** x Parental Income – **0.93** x Gender
(0.000) (0.129) (0.990)

Where Gender = 1 if male, 0 if female

- Results are very similar to those controlling for parental income only

→Including "gender" does not add valuable information

- Guiding principle in regression analysis → **parsimony**

➢Include *x* variables that should be meaningful *conceptually*, or that the data indicate as potential confounders

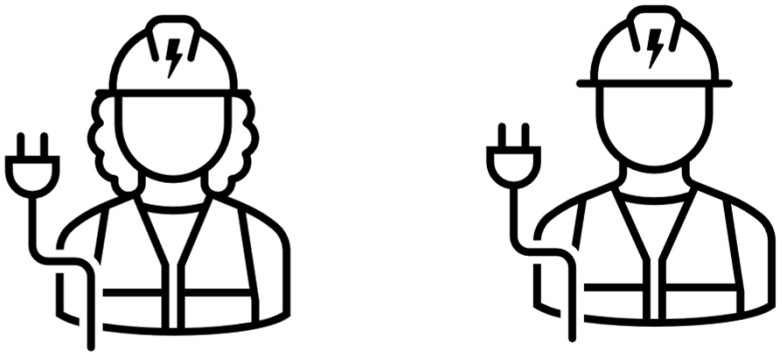
➢If results are the same/very stable with and without a given variable → remove it from your analysis

Regressions, Confounders and Causality

Regression Analysis in CIE

- Let's now return to our example from Day 1

- Client designs and implements a vocational training programme at TVET centers aimed at young people
- Overarching aim: Economically empower disadvantaged youth to engage in employment and livelihood strategies



Regression Analysis in CIE

- You are asked to quantify the effect of the vocational training programme on income
- You want to use regression analysis

→What factors may you wish to control for that may affect the relationship between the programme and income?

→Where/How could you get this data?

Regression Analysis in CIE

- You take a large, representative random sample of young people in the area and conduct a rigorous data collection, focussing on accurately measuring income.
- You collect other key information such as gender, age, parental income, education level, previous employment, etc. that you will use as control variables in your analysis.
- You conduct a regression analysis and find that taking part in the program increased average monthly income by 300 USD and you are confident the finding is not down to chance (p-value = 0.01)

→Are you confident to tell the client they should expand the program as you have found evidence it will be an effective investment?

Regression Analysis in CIE

Regressions, Confounders and Causality

- Regression analysis allows us to quantify the relationship between the vocational program and income, and to conduct statistical tests to be confident about the results
- In addition, regressions are flexible and allow to control for confounders – i.e., factors that may "blur" the relationship between the programme and income

Caution

- But… Regressions and statistical tests are not sufficient

→They can only include data that are available!

→What about data from people who don't participate in the programme?

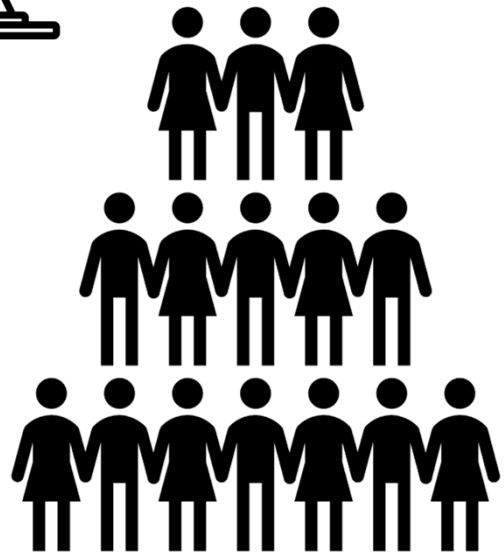
→What about factors that *cannot be measured* – e.g., motivation?

Regression Analysis in CIE

Regressions, Confounders and Causality

- At this point, we go full circle and come back to **CIE designs**

- CIE designs allow us to remove / minimize / account for differences in things we can and cannot observe between participants and non-participants, so that we can ***attribute*** differences in outcomes to the ***causal effect*** of the program



Center for Evaluation and Development

European Commission

END OF SESSION 4a

Regression Analysis in CIE

Appendix

Center for Evaluation and Development

APPENDIX
Multivariate Regressions – Digression on F-test

European Commission

$$y = a + b_1x_1 + b_2x_2 + b_3x_3...$$

- In regressions, we saw that simple t-tests allow you to test whether a parameter is significant – i.e., statistically different from 0
  - Can do a t-test on b_1, a t-test on b_2, a t-test on b_3, etc.
  - When you do a t-test on, say, b_1, you cannot say anything about b_2 or b_3 → it is an *individual* test
- For *multivariate regressions* (more than one *x*), statistical software usually report a p-value for another test – the **F-test**
  - Null hypothesis of F-test: "All estimated parameters are simultaneously equal to 0."
  - In other words, you test whether b_1=0 and b_2=0 and b_3=0 etc. at the same time → we refer to the F-test as **joint hypothesis testing** – i.e., we test multiple hypotheses at the same time (here we test *jointly* the null hypotheses of all the individual t-tests)
  - When you reject the null hypothesis of the F-test (i.e., if the p-value is low enough), you conclude that your regression model has *some explanatory power* – that is, the *x* variables you chose do explain (part of) the variation in *y*
  - If you cannot reject the F-test null, it means your regression model does not explain anything, and you need to think again about your analysis.

Session 4b: Case Study

Impact Evaluation of the STEDE Programme – Ethiopia

C4ED – EUTF

October 2023

Outline

- Project description
- Evaluation design
- Data collection
- Descriptive Statistics
- (Preliminary) Results from Regression analysis

Center for Evaluation and Development

European Commission

Project Description

Project Overview

- **STEDE** = Strengthened Socio-Economic Development and Better Employment Opportunities for Refugees and Host Communities in Fafan Zone, Somali region, Ethiopia.
- **Overarching goal:** improving the economic inclusion of refugees by expanding access to financial resources for both refugees and host communities
  - Aims to support 54,000 individuals from refugee and host communities
  - Started in May 2019 and expected to last 44 months – i.e., until January 2023

Intervention of interest (for evaluation)

- **Intervention:** expanding Village Savings and Loans Associations (VSLAs); special focus on women and youth!
- **Aim:** establish 318 VSLAs, where each VSLA consists of 20 individuals (average) ➔ 7,620 beneficiaries
- **Rationale:** Improve refugee and host communities' access to saving and loan services and improve their economic wellbeing

Implementation Area

Region: Somali region, Ethiopia
Implementation zone(s):
Fafan - one of nine zones of the Somali region
Implementation Woreda(s):
→**Refugee camps:** Awbarre, Sheder, and Kebribeyah
→**Host:** Awbarre, Kebribayah, and Jijiga (proximity to the camps)

Livelihood:
→**Host**: Agro-pastoralists and small business owners
→**Refugee:** support from NGOs and small business

Center for Evaluation and Development

European Commission

Evaluation Design

Key Evaluation Question

- For the purpose of this presentation, we will focus on a single Evaluation Question:

To what extent did the STEDE programme change resilience and livelihoods for beneficiaries?

Evaluation method: Quantitative & Qualitative

Dimensions: Economic indicators; Mechanisms and perceived barriers to seize economic opportunities; Income diversification; Food security and coping strategies; Self-perceptions

- Let's focus on the *quantitative* aspects, and on economic outcomes relating to *income and employment*

Evaluation Design

- **To what extent did the STEDE programme change resilience and livelihoods for beneficiaries?**
- **Treatment** = Formation of VSLAs
- **Clustered Design** =
  - Cluster = Group of units → here they are *communities* defined geographically
  - Treatment is assigned/implemented at the level of *communities* rather than individuals

  →Mitigates the risk of social unrest resulting from differential treatment within one community
  - Clusters (communities) in the treatment group will receive support to form VSLAs
- **Aim of the evaluation** = Measure the impact of VSLAs on employment outcomes and income

Evaluation Design (cont'd)

- How to measure impact?

→Compare outcomes in clusters (communities) that receive support to form VSLAs to outcomes in clusters that do not

- Challenge?

→The selection of treatment and comparison communities ⇒ we need to be able to ***attribute*** differences in outcomes to the program

- Solution?

→Ideally: first identify eligible communities that *could* receive the program, then *randomize* treatment to have a rigorous (clustered) RCT design – i.e., choose at random the *clusters* that receive support to form VSLAs

Evaluation Design (cont'd)

- However, randomization was not feasible (operational constraints)

  →We still have program (treatment) and non-program (comparison) clusters in the sample, they are just not allocated to treatment randomly

- Plan B?

  →Use a quasi-experimental approach, and use **regression analysis** to *control for confounding variables*/ensure comparability of the treatment and comparison groups

Center for Evaluation and Development

European Commission

Data

Sampling and Tools

- Host communities
  - →70 clusters (here, villages) selected for the study ⇒ 35 clusters in treatment group, 35 clusters in comparison group
  - →Target = 1,050 households (15 to 18 households per cluster)
- Refugee camps
  - →6 clusters (here, blocks/sections) ⇒ 3 clusters in treatment group, 3 clusters in comparison group
  - →Target = 720 households (40 per cluster)
- Data collection tools (questionnaire)
  - →Focus on key outcome variables – e.g., employment, income
  - →Collect extra information on potentially meaningful/confounding variables – e.g., gender, age, education, household size, assets, housing, etc.

Data Collection

- 2 rounds of data collection:
  - Baseline – i.e., before the intervention starts
  - Endline – i.e., to measure key outcomes *after* beneficiaries have received the program

	Baseline	Endline
Date	November 2021	May/June 2023
Total interviews completed	1,840	1,821 (98.96%)
Total in Host communities	*1,120*	*1,103 (98.48%)*
Total in Refugee camps	*720*	*718 (99.72%)*

Center for Evaluation and Development

European Commission

Descriptive Statistics

Descriptive Statistics
General Remarks

- CIE reports often present descriptive statistics in so-called **balance tables** that look like this:

	(1) Full sample	(2) Treatment	(3) Control	(4) (2)-(3) (p-value)
Age	36.7	36.5	36.9	-0.4
	(11.3)	(11.1)	(11.4)	(0.48)
Female respondent	0.72	0.74	0.70	0.04*
	(0.45)	(0.44)	(0.46)	(0.07)

- It presents the sample mean and standard deviation for the full sample, the treatment group, and the comparison group (sometimes Column (1) is omitted).
- In addition, it usually presents the results of a **t-test** of the equality of means between the treatment and comparison groups – here we report the difference between the group means and the p-value of the test.
- **Aim**: present descriptive stats (mean and SD) and see at a glance if the two groups seem to differ too greatly in terms of some characteristics – i.e., check "balance" between the groups

Center for Evaluation and Development

Baseline characteristics

Which variables show ***imbalances*** between the two groups?

	(1) Full sample	(2) Treatment	(3) Control	(4) (2)-(3) (p-value)
Age	36.7	36.5	36.9	-0.4
	(11.3)	(11.1)	(11.4)	(0.48)
Female respondent	0.72	0.74	0.70	0.04*
	(0.45)	(0.44)	(0.46)	(0.07)
Female HH head	0.29	0.30	0.25	0.05**
	(0.45)	(0.46)	(0.43)	(0.03)
Married	0.85	0.84	0.88	-0.04**
	(0.36)	(0.37)	(0.33)	(0.03)
# of Household members	7.5	7.6	7.4	0.1
	(3.0)	(3.1)	(2.8)	(0.35)
# of Children	3.4	3.6	3.2	0.5***
	(3.1)	(3.2)	(2.9)	(0.00)
Education Status				
Never enrolled	0.60	0.53	0.71	-0.18***
	(0.49)	(0.50)	(0.46)	(0.00)
Primary level or Informal education	0.30	0.36	0.21	0.14***
	(0.46)	(0.48)	(0.41)	(0.00)
At least completed primary education	0.10	0.12	0.08	0.04**
	(0.30)	(0.32)	(0.27)	(0.01)
Observations	***1,936***	***1,129***	***692***	

*Note: *** p<0.01; ** p<0.05; * p<0.1*

Outcome Variables – Employment and Income

- We report the same for outcome variables measured at endline.

→What do the t-tests show?

	(1) Full sample	(2) Treatment	(3) Control	(4) (2)-(3) (p-value)
Has a stable job	0.21	0.25	0.13	0.12***
	(0.41)	(0.44)	(0.34)	(0.00)
Self-employed in stable job	0.14	0.16	0.10	0.06***
	(0.35)	(0.37)	(0.30)	(0.00)
Interested in starting own business	0.85	0.84	0.87	-0.03*
	(0.35)	(0.36)	(0.33)	(0.08)
Beneficiary interested in investing in their business	0.76	0.79	0.68	0.11*
	(0.43)	(0.41)	(0.47)	(0.07)
Average income over last 6 months	4,138	4,071	4,357	-286
	(5,177)	(4,353)	(7,262)	(0.65)
Annualized monthly employment income	4,360	4,439	4,095	344
	(4,118)	(4,070)	(4,29)	(0.53)
Observations	***1,821***	***1,129***	***692***	

*Note: *** p<0.01; ** p<0.05; * p<0.1*

Center for Evaluation and Development

European Commission

(Preliminary) Results

Data Analysis Plan

- **Goal**: isolate the *causal* relationship between the outcomes of interest and the treatment
- **Regression equation:**

$$Outcome = a + \boldsymbol{b}.Treatment + d.Confounders + error$$

Where Treatment = 1 if individual lives in an area where VSLA formation was encouraged, 0 if not; Confounders = Baseline characteristics shown earlier

- We're interested in parameter **b**
- The following tables report the estimated value of **b** and the p-value from its significance test

Data Analysis Plan

- **Regression equation:**

$$Outcome = a + \boldsymbol{b}.Treatment + d.Confounders + error$$

- **Option 1**

→Simple multivariate regression including all potential confounding variables

- **Option 2**

→Similar to Option 1 but explicitly account for *clustering*

→Having clusters changes *standard errors*, and hence changes results of statistical tests!

- **Option 3**

→Complex approach that mixes regression analysis with *matching* – i.e., the actual quasi-experimental method used in this evaluation

Results – Simple Regression vs. Clustering

Can you spot any differences between the two approaches?

Outcome variable	(1) Simple regression	(2) Regression with clustering
Has a stable job	0.119***	0.118***
	(0.00)	(0.00)
Self-employed in stable job	0.061***	0.061***
	(0.00)	(0.01)
Interested in starting own business	-0.042**	-0.042*
	(0.02)	(0.06)
Beneficiary interested in investing in their business	0.057***	0.057***
	(0.00)	(0.01)
Average income over last 6 months	-327.53	-325.04
	(0.64)	(0.73)
Annualized monthly employment income	133.50	136.15
	(0.82)	(0.84)

Note: p-values in parentheses.
**** p<0.01; ** p<0.05; * p<0.1*

Option 3 – Regression and Matching

- Matching deals with ***selection on observables***

  →That is, it assumes that once all measured/observed confounding factors have been accounted for, the differences in outcomes between treatment and comparison groups can be attributed to the causal effect of the program

- The intuition is similar to multivariate regression controlling for confounders
- However, *matching* is better suited to estimate *causal impacts* (it solves some technical limitations of regressions)
- **The intuition of matching** = For each unit in the treatment group, find a unit in the comparison group that is as similar as possible based on observed characteristics. In practice, you may match one or several treatment units to one or several comparison units, with different ways to assess "similarity".
- The actual analysis essentially uses regressions *adjusted* with matching

Results – Regression + Matching

Does using *matching* substantially change the results?

Outcome variable	(1) Simple regression	(2) Regression with clustering	(3) Regression + Matching
Has a stable job	0.119***	0.118***	0.123***
	(0.00)	(0.00)	(0.00)
Self-employed in stable job	0.061***	0.061***	0.063***
	(0.00)	(0.01)	(0.00)
Interested in starting own business	-0.042**	-0.042*	-0.037**
	(0.02)	(0.06)	(0.03)
Beneficiary interested in investing in their business	0.057***	0.057***	0.058***
	(0.00)	(0.01)	(0.00)
Average income over last 6 months	-327.53	-325.04	137.48
	(0.64)	(0.73)	(0.86)
Annualized monthly employment income	133.50	136.15	490.56
	(0.82)	(0.84)	(0.35)

Note: p-values in parentheses.
**** p<0.01; ** p<0.05; * p<0.1*

Concluding remarks

- Regressions are a powerful, flexible tool, but they are *only one of the tools* available for analysis
- Regression must be used in conjunction with careful design and sampling
- In practice, only rigorous RCTs allow (relatively) simple analyses based on regressions only
- Quasi-experimental evaluations designs (difference-in-differences, matching, regression discontinuity) require more sophisticated analytical skills, to account for certain design specificities such as e.g., clustering
- Here, we presented *preliminary* results only → Conclusions are not final yet
- In addition, we presented only the quantitative aspect, but the evaluation also has a rich *qualitative* investigation that will be crucial to contextualize and deepen the quantitative results!

Center for Evaluation and Development

European Commission

END OF SESSION 4b

Case Study on Mixed Method Analysis

Impact Evaluation of the RISE programme implemented by GIZ in Uganda

---

October, 2023

Center for Evaluation and Development

giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

European Commission

EU Emergency Trust Fund for Africa

Outline

1. **The RISE programme**
2. **Impact Evaluation design**
3. **Mixed-Methods Analysis**
4. **Complementarity and Triangulation**

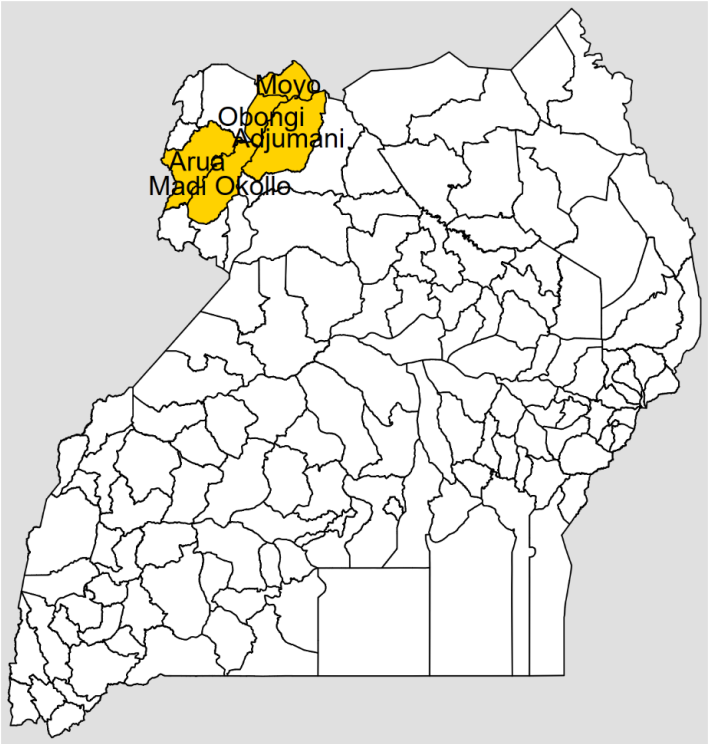
The RISE Programme

Background

- Implementing Partner: GIZ Uganda & Norwegian Refugee Council (NRC)
- Period: 2021-2024
- Districts: Adjumani, Arua, Madi-Okollo, Obongi and Moyo

Focus on 3 components:

1. strengthening local authorities' ability to deliver public services
2. improving employment opportunities and non-agricultural income
3. increasing income from agricultural activities.



Background

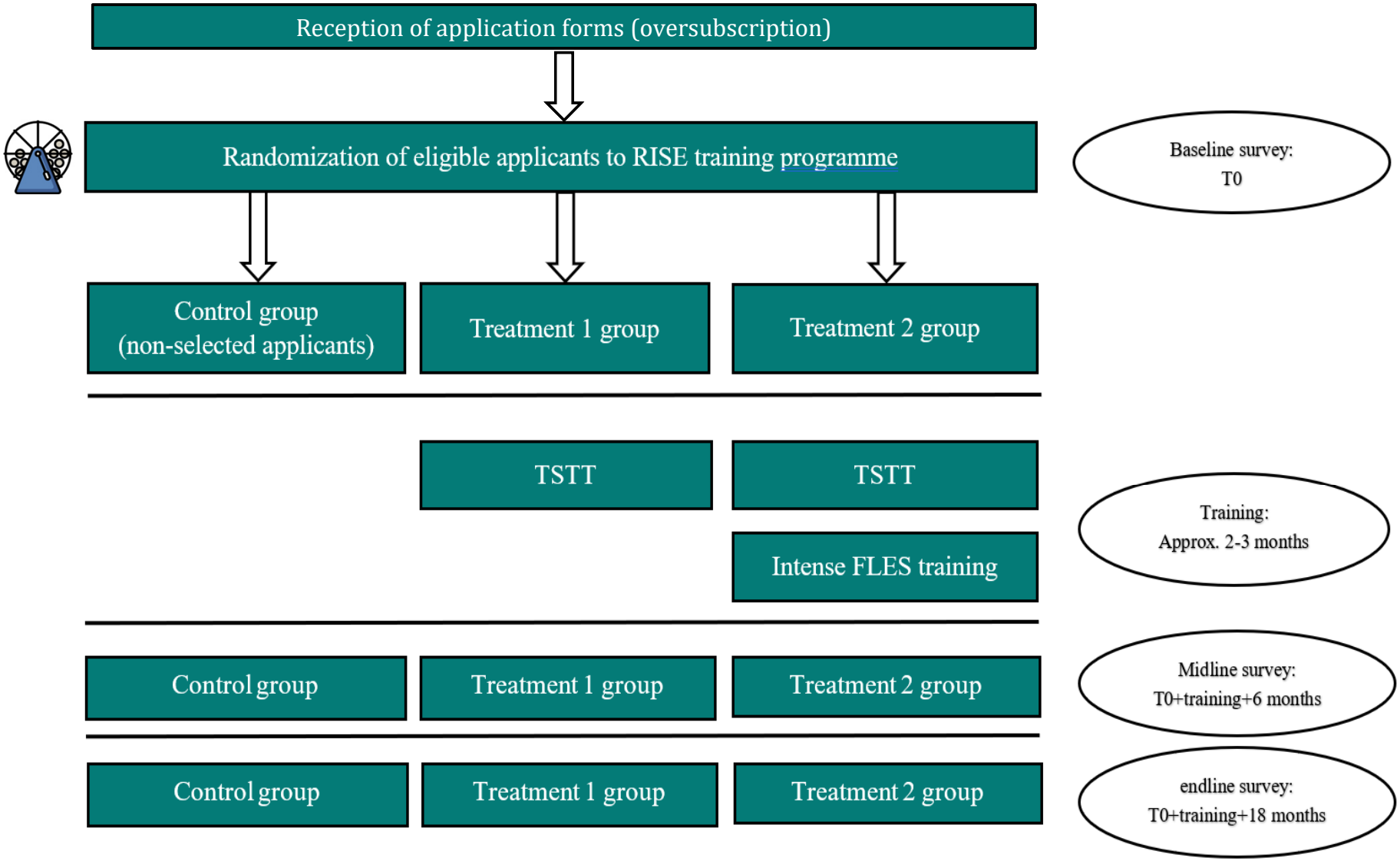
- The IE focuses on component 2: **improve employment opportunities and non-agricultural income** *of* ***young people*** *in* ***host and refugee communities***

Activities

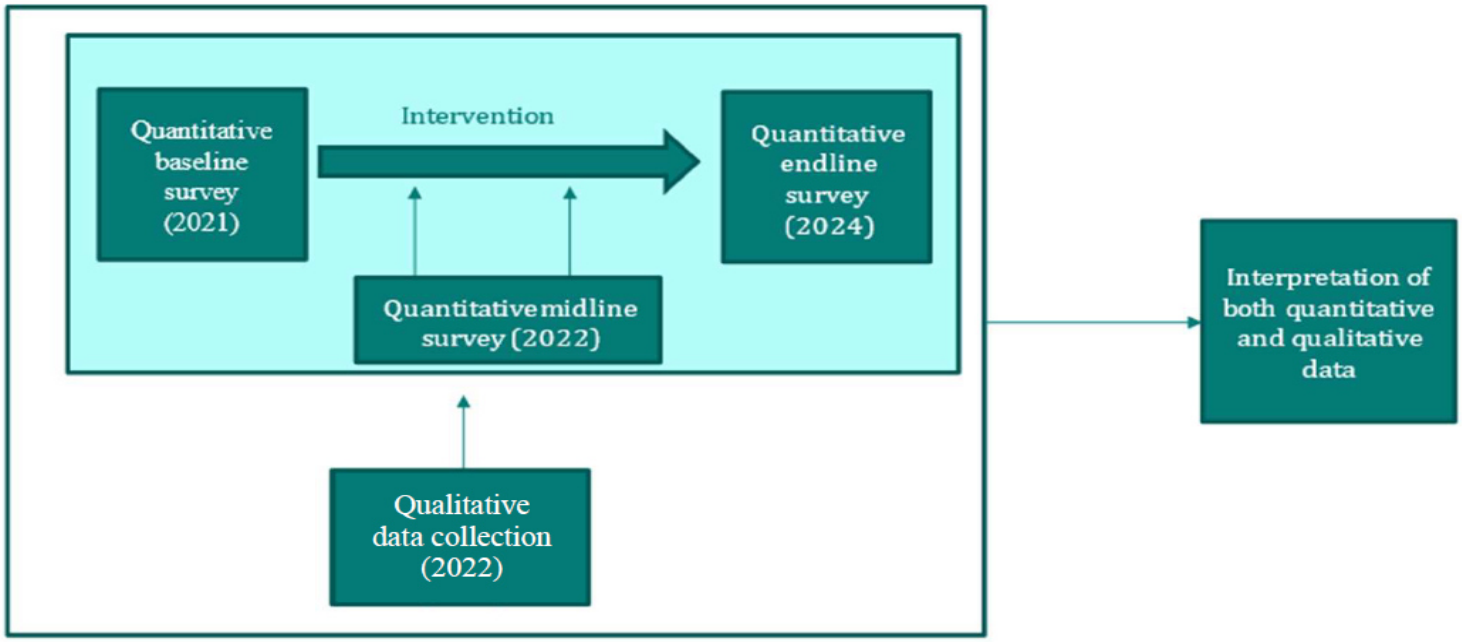
- Train 2,000 youth (70% female and 50% refugees) in 3 cohorts
- Two complementary trainings:
  1. Technical Short-Term training (TSTT) – 2.5 months (2,000 youth)
  2. Intense Financial Literacy and Entrepreneurial Skills (FLES) training - 2 weeks (1,000 youth from TSTT)

RISE Impact Evaluation Design

IE design: Randomized Control Trial



Mixed methods design



Source: C4ED adaptation of the concept developed by (Creswell, 2006)

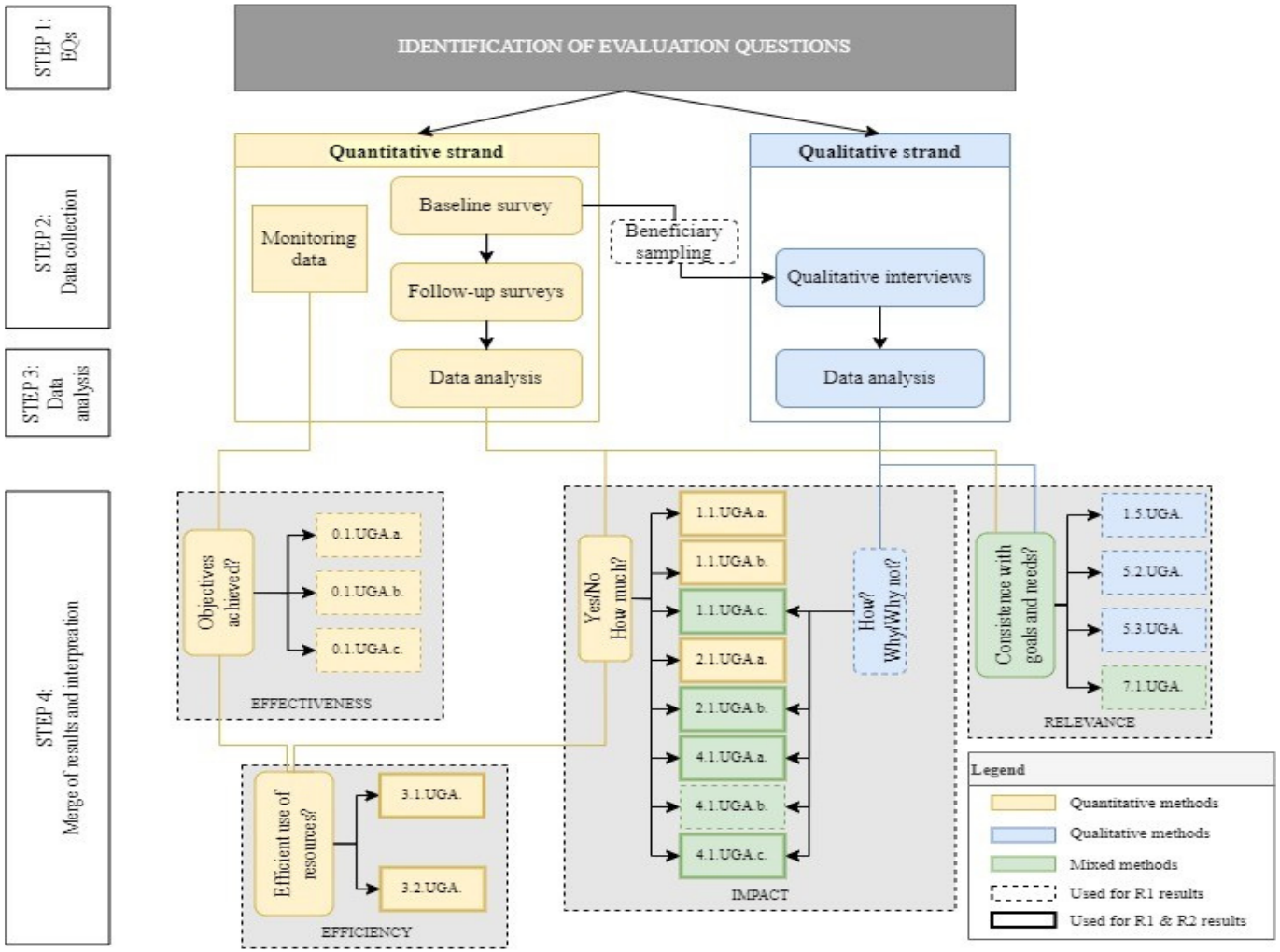
Evaluation Questions

EQ	Main question	Theme addressed
EQ0	**Programme-specific monitoring EQ**	**Enrollment**
EQ1	To what extent did EUTF interventions contribute to **employment, job creation, and skills**?	**Employment, job creation & skills**
EQ2	To what extent did EUTF interventions change **resilience and livelihoods** for beneficiaries?	**Resilience and livelihoods**
EQ3	Which were the most **cost-effective EUTF support** options to enhance employability?	**Cost-effectiveness analysis**
EQ4	What other **intended and unintended outcomes** (e.g. mobility, migration, migration intentions, employment policies and reforms) did EUTF interventions contribute to?	**Intended & unintended outcomes**
EQ5	How did EUTF interventions **include and promote different vulnerable groups** such as youths, women, refugees, IDPs, migrants and host communities alike through its activities?	**Inclusion of vulnerable groups**

EQs in MM design

Evaluation design poses EQs which are:

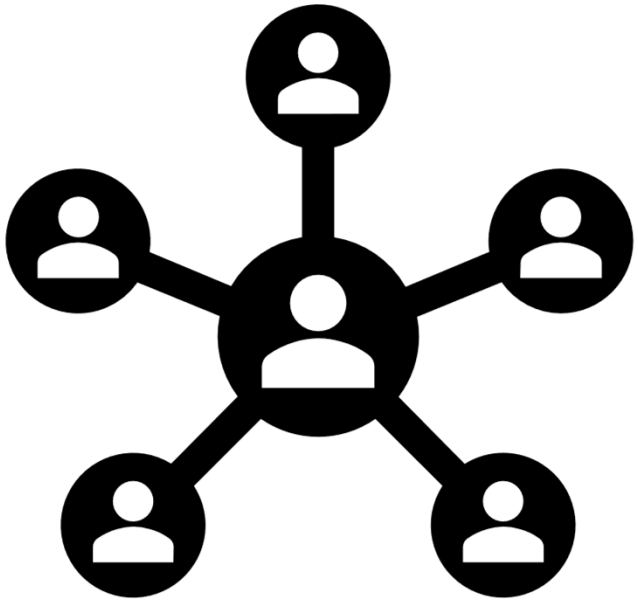
1. exclusively addressed by qualitative means (1.5.UGA; 5.2.UGA; and 5.3.UGA)
2. Exclusively addressed by quantitative means (0.1, 1.1.UGA.a; 1.1.UGA.b; 2.1.UGA.a. and EQs 3)
3. Both quantitative and qualitative means (1.1.UGA.b; 2.1.UGA.b; 4.1.UGA and 7.1.UGA.).



Mixed Methods Analysis

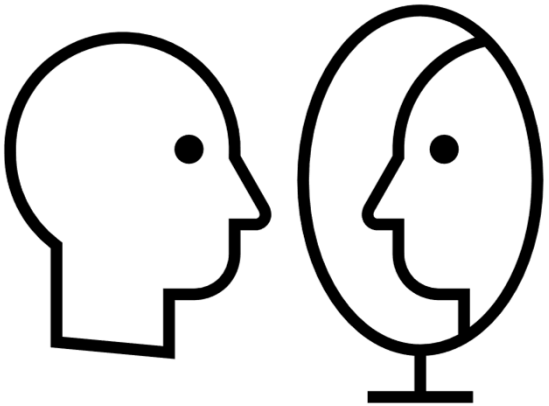
Planning mixed methods analysis

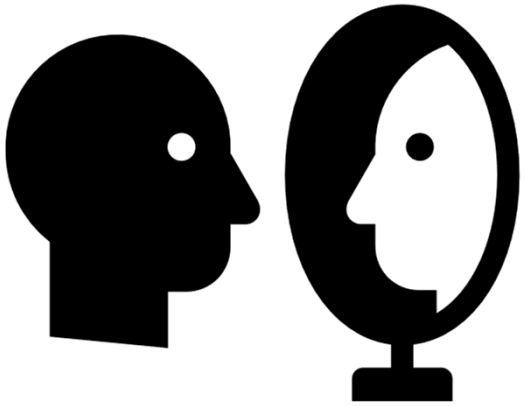
- **Post-data collection and cleaning**
  - Quantitative and quantitative teams meet and:
    - Discuss expectations of each approach
    - Reflect on the EQs and sub-questions and if there were any changes in indicators and means of verification
    - Outline questions and indicators to start with
    - Set a timeline for analysis and reflection meetings



Reflections on the evaluation questions, and design

- ➢ Some questions the team should ask itself
- ❑ Does the plan for analysis at design still fit?
- ❑ If not, what changed in the course of the evaluation?
- ❑ How can we integrate these changes into our analysis?





Reflection of indicators

- Answering sub-questions 0.1 UGA.b. and 1.1.UGA.c uses both quantitative and qualitative approaches.
- At design EQ 0.1 was to be answered quantitatively, but later qualitative explanatory data was required to explain some results.
- Qualitative indicators preempt reasons how, why, or why not results identified under quantitative indicators were achieved/not achieved.

Evaluation Matrix

Sub-question	Approach	Indicators	Tool	DAC
0.1.UGA.b. Did the RISE programme select and train the intended number of youth?	Quant	Number of applicants selected Number of individuals trained Number of women trained Number of refugees trained	Monitoring data	Effectiveness
1.1.UGA.c. What effects does the RISE programme have on employability?	Quant.	Received a job offer Searched for employment Received a job offer related to trade Self-perceived employability score	Youth questionnaire	Impact
	Qual.	Opportunities and barriers for finding/maintaining (decent) wage employment Influence (perception of) of the programme for finding employment Self perceived employability Support perceived as the most useful for professional development	IDI KII FGD Life stories	Impact

Timeline

Timeline interval

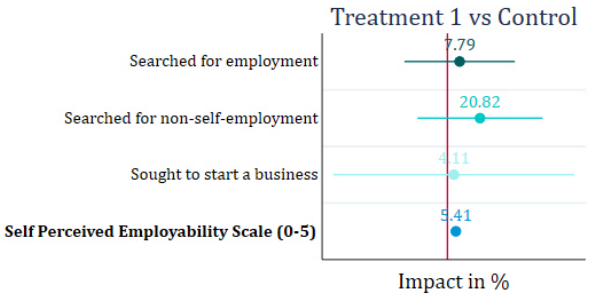
Week 1	First meeting General planning
Week 3	Reflecting on the codes and indicators
Week 5	Preliminary results and gaps from codes and matching
Week 7	Preliminary results and gaps from themes and regressions analysis
Week 9	Consolidating results for all EQs and sub-questions
Week 11	Structuring report
Week 11- 15	Report writing

Common practices

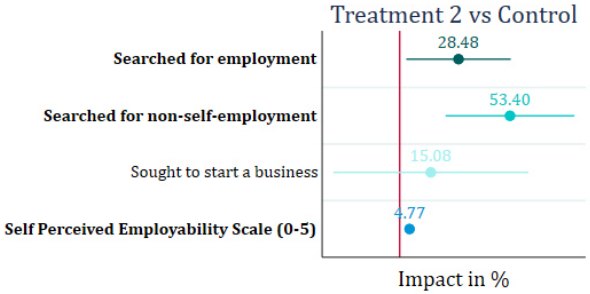
- Project Lead sets up and moderates the meeting
- Discuss a reasonable timeline depending on the project – the amount of data, the complexity of analysis, the deadline for the deliverables
- Regular meetings/check-in are important

Systematizing Evidence: Quantitative analysis

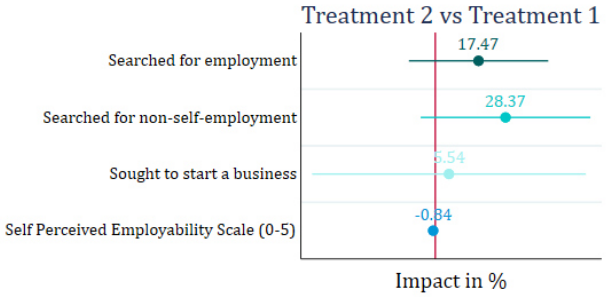
Impacts of the RISE programme



→ Impact of being selected for TSTT



→ Impact of being selected for TSTT and FLES



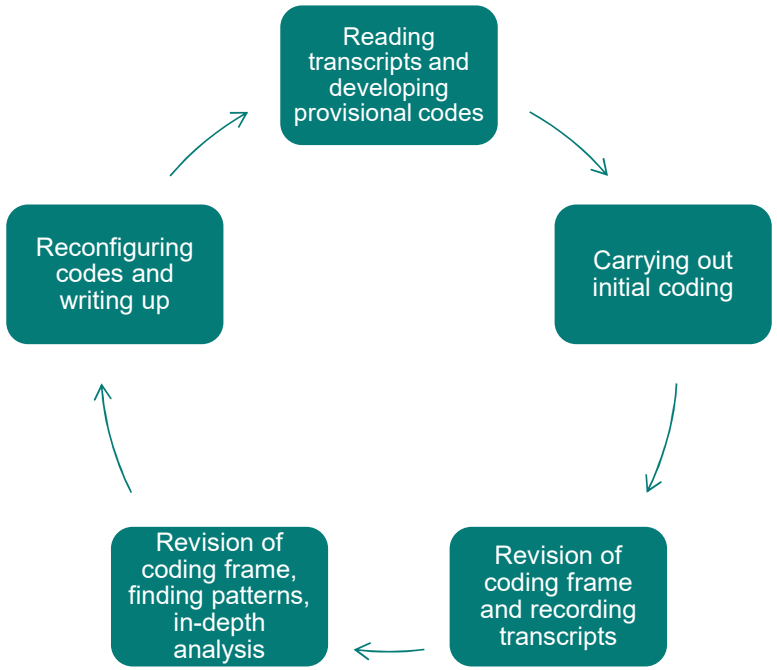
→ Impact of being selected for FLES

Sytematizing Evidence: Qualitative analysis

Coding tree

- Code System
  - Code System
    - Training received
    - Employment and livelihoods
      - Improved income
      - Diversified livelihoods
      - Opportunities+barriers for finding/maintaining (decent)wage (+)
        - Tools/start-up kits
        - Skills (mis)match
    - Quality of training
      - Trainers
      - Training facility
      - Training resources
      - Training is adapted to beneficiaries' needs and wishes
    - Social cultural hindrances and opportunities
      - gender
      - Refugee status

Iterative Coding



Complementarity & Triangulation

Demonstrated with preliminary findings

Disclaimer

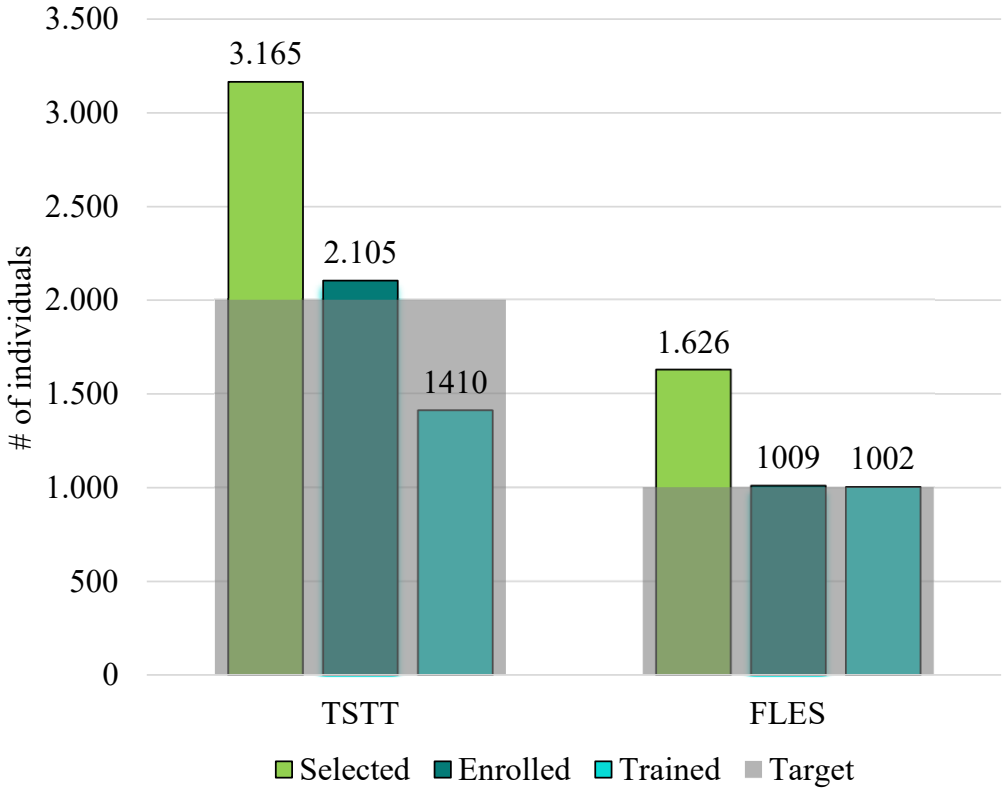
The following presentation contains preliminary findings and insights based on the data collected and analyzed up to this point.

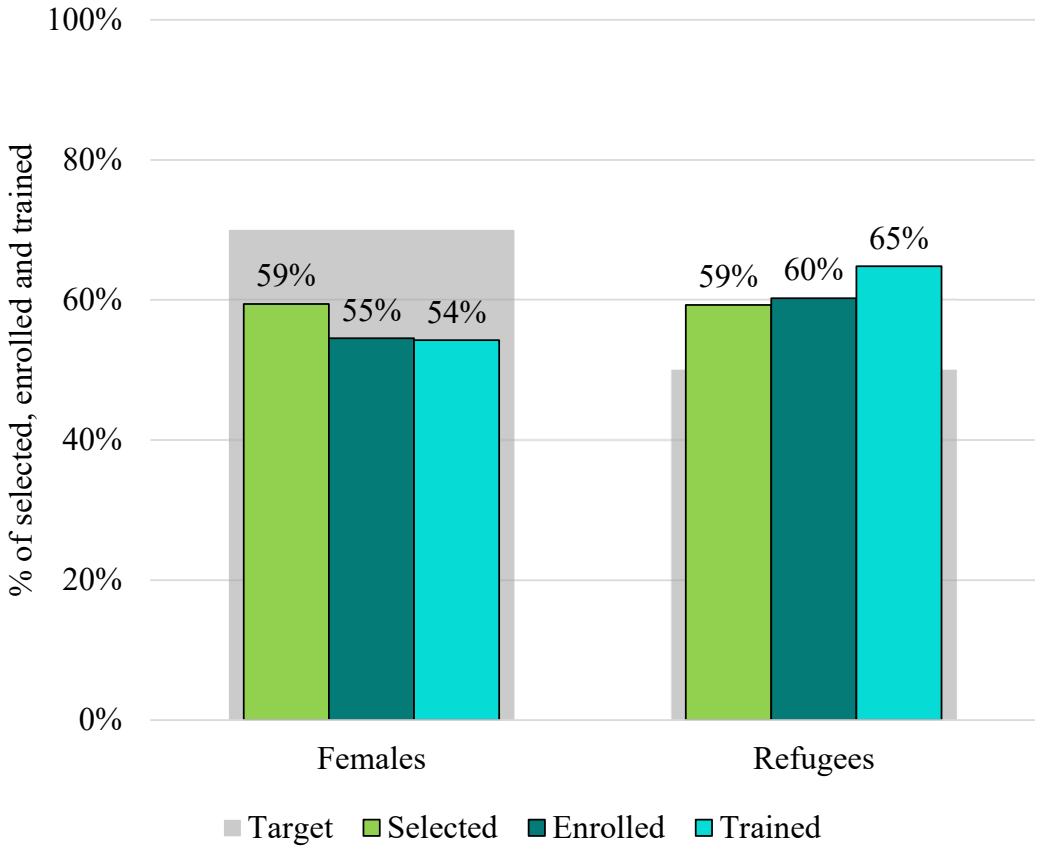
The following findings are a snapshot of an ongoing research process, and adjustments or revisions may occur as we continue to refine our analysis.

Did the RISE programme select and train the intended number of youth?

Complementarity/explanatory

Selection, enrollment and training (quant)





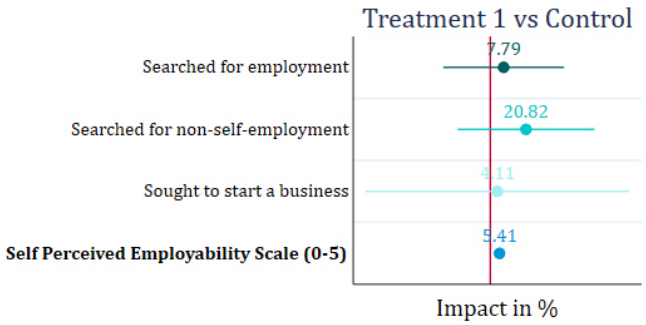
Main reasons for no-shows and drop-outs (qual)

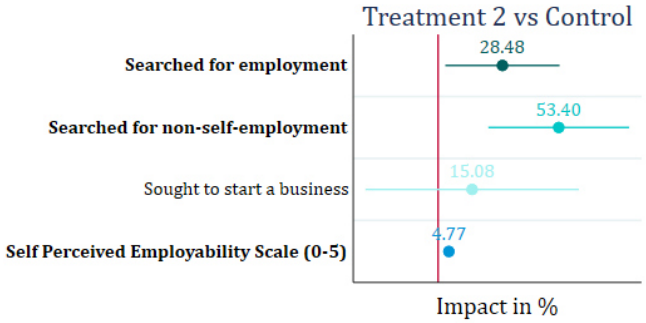
- Communication on outcome of the selection
- Competition with other NGOs and institutions (with “start-up kits”)
- Disappointment with content of training (cohort 1)
- Inadequate training tools
- Household and professional obligations
- Cost of training (transportation, food…)

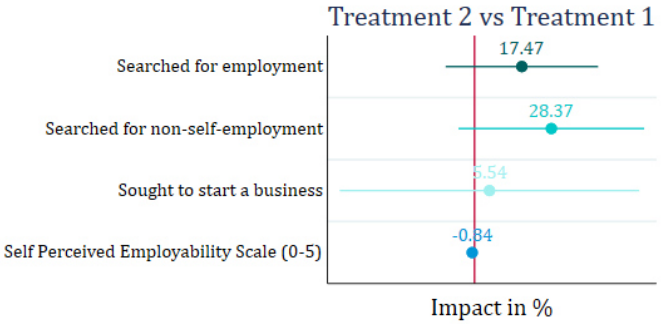
What effects did the RISE programme have on employability?

Complementarity

Impact on employment search (quant)







- TSTT is the main driver to improve the perception of being employable.
- The programme has stimulated job search among male beneficiaries.
- While female beneficiaries are also more confident that they can find a job, they are not more likely to search for one.

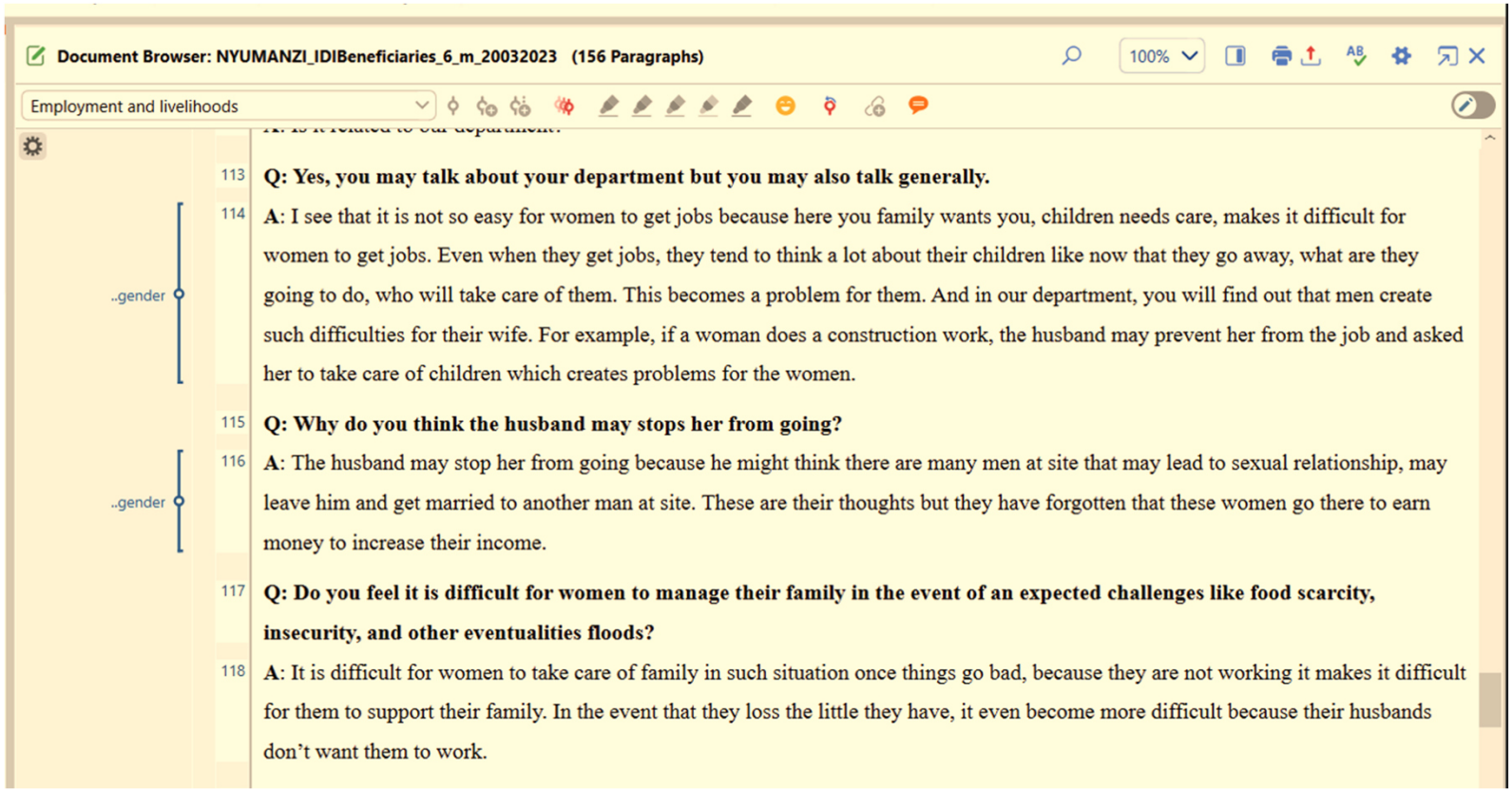
Reasons why and how (qual)

- Limited start-ups

Lack of start-up kits

- More males looking for jobs

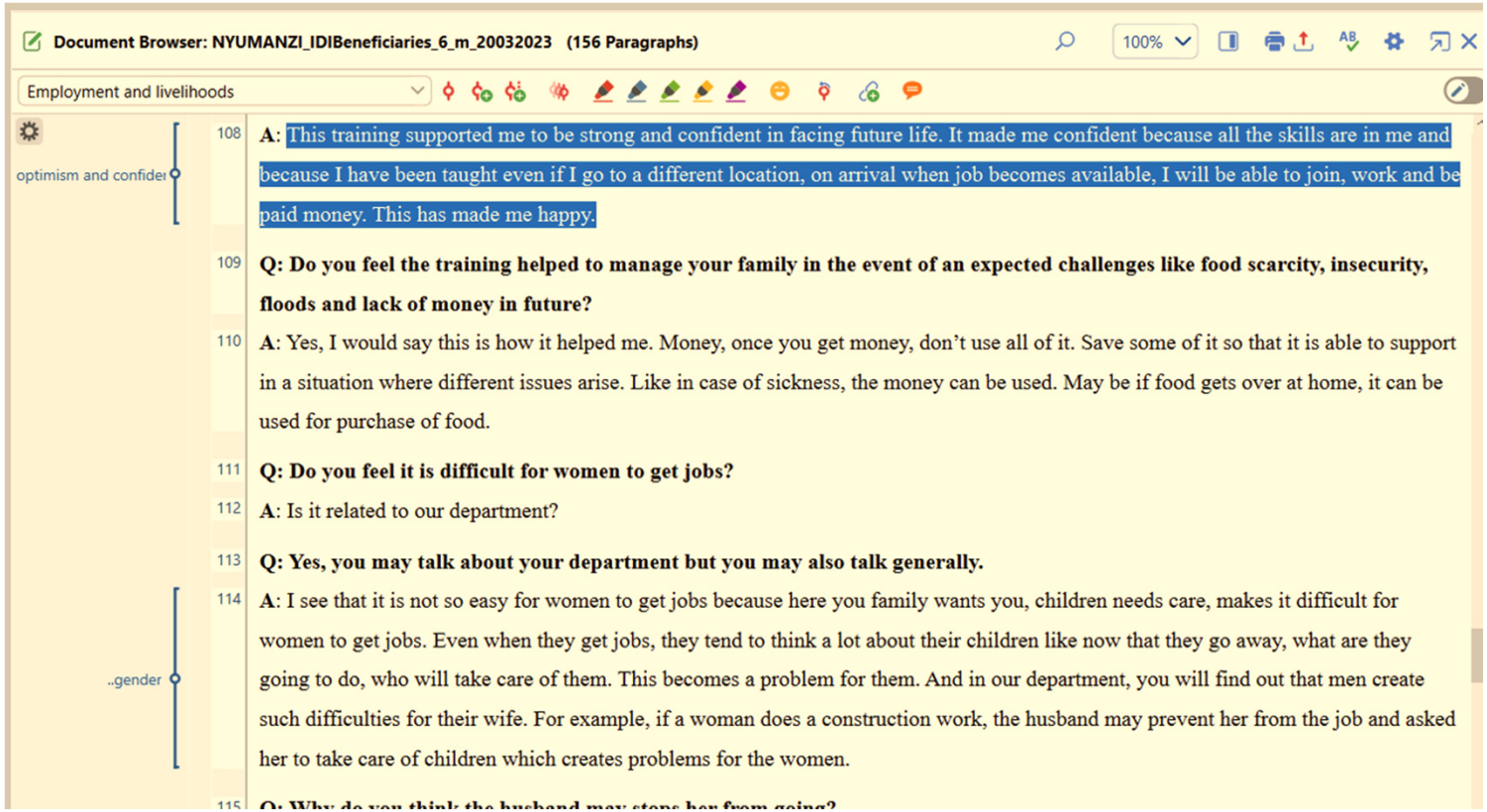
Social cultural barriers limiting women in the job search – women prioritising domestic work, decision dependent on the husbands

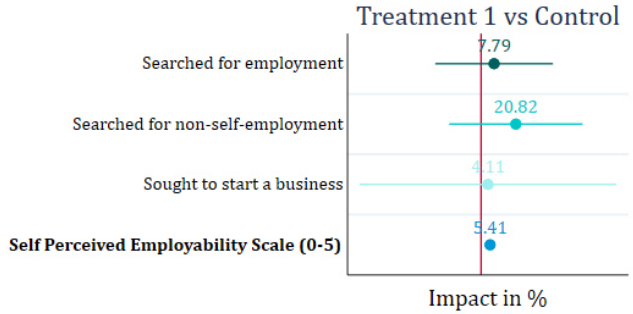


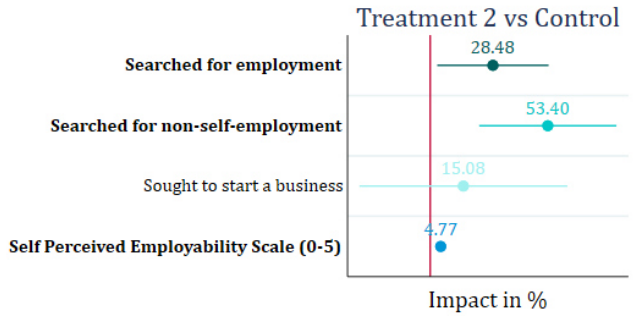
What effects did the RISE programme have on employability?

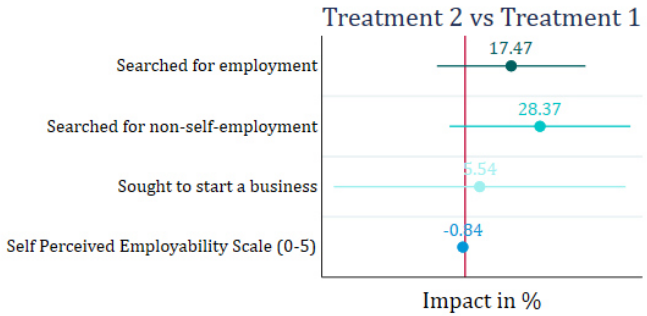
Triangulation

Triangulation









giz Deutsche Gesellschaft für Internationale Zusammenarbeit (GIZ) GmbH

Thank you

Q&A

European Commission

EU Emergency Trust Fund for Africa

Center for Evaluation and Development

European Commission

Session 5a: Qualitative Analysis and Triangulation for CIE

C4ED – EUTF

September 2023

Center for Evaluation and Development

Outline:

- Overview of qualitative research and uses
- Basic approaches for collecting qualitative data
- Methods and tools used for qualitative data analysis
- How can qualitative data be used to answer questions that quantitative cannot?
- How can qualitative data be used to validate findings from quantitative research?

Center for Evaluation and Development

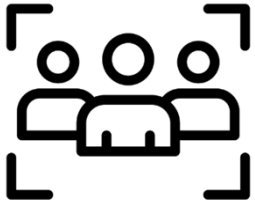
Qualitative research and its uses:

Qualitative research is a **systematic** research approach that focuses on:

- ✓ *Understanding and interpreting the complex and nuanced aspects of human experiences, behaviours, and phenomena,*
- ✓ *Involves gathering and analysing non-numerical data, such as textual, visual, or audio information, through methods such as interviews, observations or case studies;*
- ✓ *Qualitative research aims to uncover underlying meanings, patterns, and contextual insights, often using inductive reasoning to generate theories or hypotheses from the collected data.*

Center for Evaluation and Development

Why qualitative research in evaluation in CIE?



Understand the views, experiences and motivations of beneficiaries, implementers and stakeholders in greater depth.

Understand the processes and mechanisms by which impacts occur – How and why?

- Investigate if a project had any unintended (both positive and negative) consequences

Questions about meaning and motivation examine how a particular behavior or action is understood, or how people make sense of their circumstances.

Center for Evaluation and Development

Why qualitative research in evaluation?

Addresses evaluation questions that are not tackled by the quantitative research methods

Complements the quantitative research on common evaluation questions

- By providing additional information that cannot be covered with surveys and tests.
- By including hard-to-reach respondents that are not “counted” in households

Analyses underlying social, cultural, political and economic aspects of change

Center for Evaluation and Development

Timing for qualitative research in CIE

E.g. EUTF Ethiopia	Before: to clarify certain conditions, patterns for better planning or later-on specifications for quantitative approach and questionnaire design
E.g. Most of the EUTF CIE projects	In parallel to quantitative impact evaluation: to complement and triangulate findings
E.g. EUTF Uganda – GIZ - RISE Project	After: to further and deeper investigate and explain phenomena

Center for Evaluation and Development

Main methods

Interviews

Focus group discussions

Observations

Center for Evaluation and Development

Key Informant Interview (KII)

KIIs involve interviewing people who have particularly informed perspectives on an aspect of the program being evaluated.

Interviewers frame open ended questions, probe for information and take notes, which are elaborated on later.

Why do we use them?

- When it is important **to gain an understanding of the perspectives and motivations of stakeholders and partners** regarding an activity or project in order to explain the implementation process, shortcomings and successes of an intervention.
- When **generating recommendations** is the key purpose

Center for Evaluation and Development

In-Depth Interview (IDI)

IDIs involve interviewing beneficiaries of the interventions under study or people with a specific relevant profile (farmers, child labourers, GNB survivors, female entrepreneurs…)

Why do we use them?

- To understand their personal experiences with the topic at hand.
- To understand what worked and what didn't work and why
- To generate recommendations and lessons learned

Center for Evaluation and Development

Focus Group Discussions (FGDs)

FGD is used as a group interview designed to explore people's attitudes. It can be used to find out what issues are of most concern for a community or group when little or no information is available.

Why do we use them?

- They are helpful for adding meaning and understanding to existing knowledge or getting at the “why” and “how” of a topic.
- To understand collective attitudes.

How do we use them?

- The group has a specific discussion topic.
- The group has a facilitator.
- The group composition and discussion is carefully planned to create an environment where participants feel safe and can speak freely.

Any questions so far?

Center for Evaluation and Development

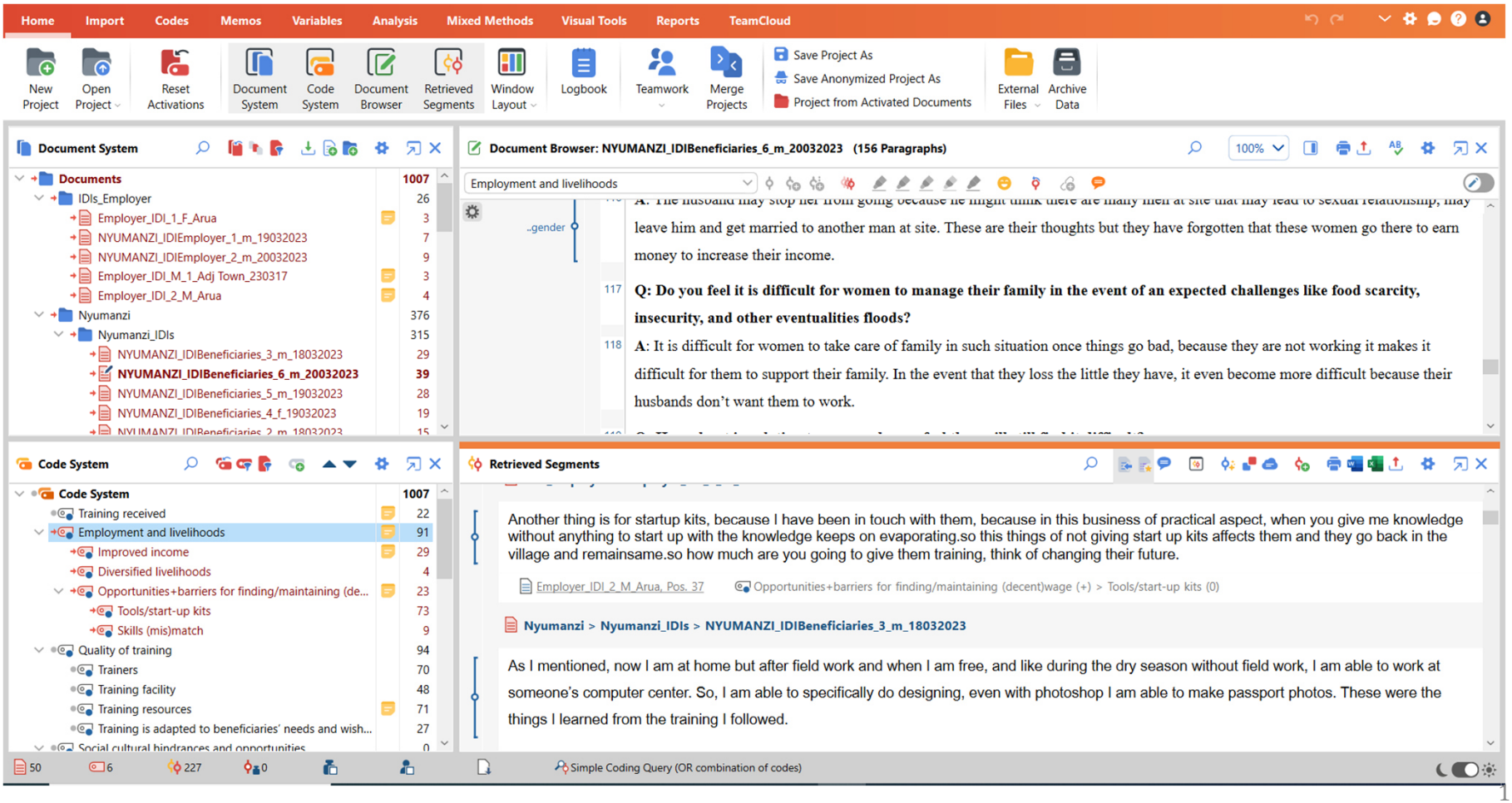
Tools and methods used for analyzing qualitative data:

➢ Coding

This Photo by Unknown Author is licensed under CC BY

- ✓ Qualitative data analysis software such as NVivo, MAXQDA, or ATLAS.ti,
- ✓ **Open Coding:** identify and label meaningful segments or "codes" within the data. These codes represent concepts, themes, or patterns in the text.
- ✓ **Axial Coding:** Organize and connect the open codes into categories or themes. This stage helps in making sense of the data and understanding relationships between codes.
- ✓ **Selective Coding:** Develop a more refined and focused coding scheme by selecting the most important categories or themes to analyse in-depth.

Center for Evaluation and Development



Center for Evaluation and Development

➢ Approaches:

✓ **Content Analysis:** Content analysis is a systematic approach for identifying and categorizing textual or visual data. Researchers identify keywords, phrases, or elements and code them according to predefined categories. This method is often used for analysing large volumes of data, such as interviews, or textual documents.

✓ **Thematic Analysis:** Thematic analysis involves identifying, analysing, and reporting patterns or themes within the data. Researchers read through transcripts or notes, code segments of text, and group codes into overarching themes. This method helps in organizing and summarizing the data into meaningful categories.

✓ **Grounded Theory:** Grounded theory is a method for developing theories or conceptual frameworks from qualitative data. Researchers begin with an open mind, without preconceived theories, and continually analyse data to identify concepts and relationships. This approach aims to generate new theories or models based on the data itself.

Center for Evaluation and Development

How can qualitative data complement quantitative data

Qualitative data can be very useful in addressing questions that quantitative data alone cannot fully answer.

✓ **Exploration of Context:** Qualitative data can help provide context to quantitative findings. It allows researchers to understand the "**how**" and "**why**" behind quantitative results. For example, if a quantitative study finds that a certain program's effectiveness decreased, qualitative data can help uncover the reasons behind this change.

✓ **Rich Descriptions:** Qualitative data offers rich, detailed descriptions of phenomena. It can capture nuances, complexities, and subtleties that quantitative data might oversimplify. This is particularly useful when studying social or cultural phenomena.

Center for Evaluation and Development

- ✓ **Understanding Motivations and Perceptions:** Qualitative research can delve into people's motivations, attitudes, beliefs, and perceptions. This is crucial for understanding why individuals or groups behave in certain ways, which may not be evident from quantitative data alone.

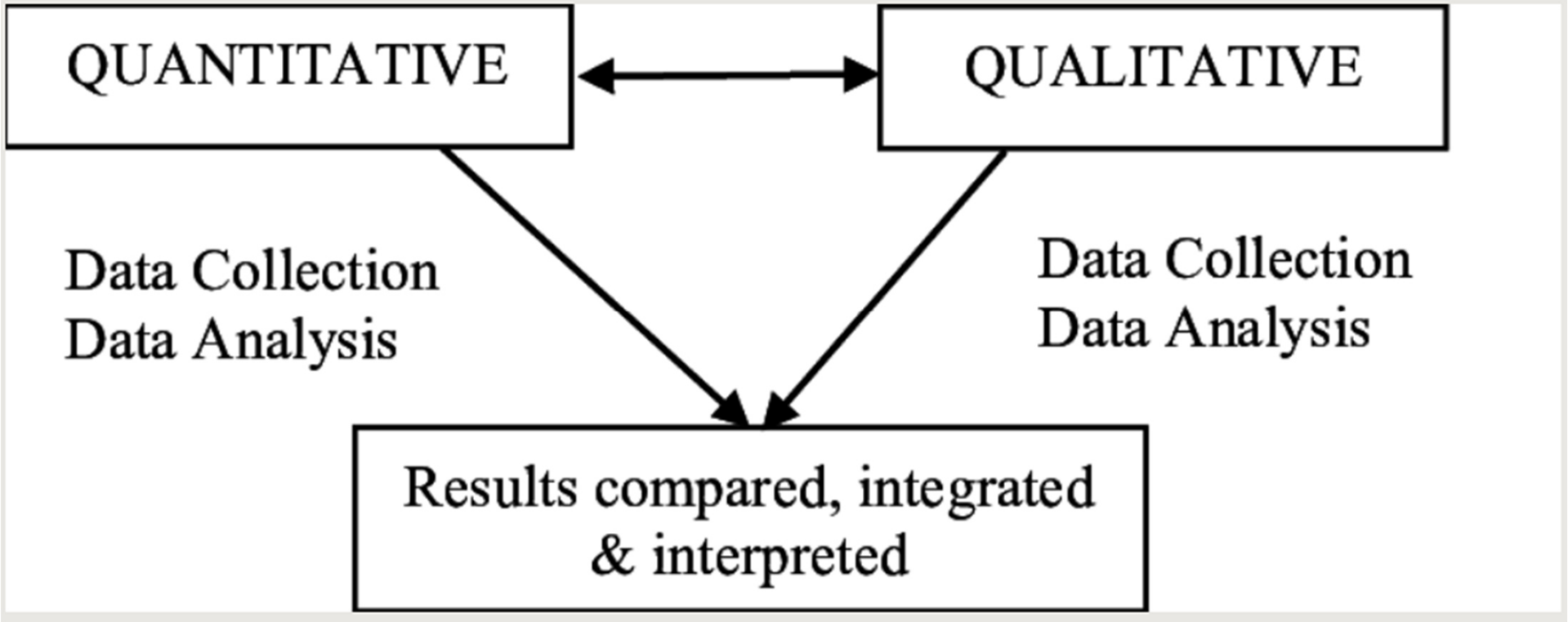
- ✓ **Generating Hypotheses:** Qualitative data can be used to generate hypotheses or research questions for subsequent quantitative studies. It can help identify patterns and variables that warrant further investigation using quantitative methods.

- ✓ **Triangulation:** Combining qualitative and quantitative data (triangulation) can improve the validity and reliability of research findings. If qualitative and quantitative data converge on similar conclusions, it strengthens the overall research argument.

- ✓ **Identifying Unanticipated Factors:** Qualitative data can uncover unexpected factors or variables that researchers may not have considered in their quantitative analysis. This can lead to new research directions and insights.

Center for Evaluation and Development

Triangulating quantitative data with qualitative data



Center for Evaluation and Development

Convergence:

- ✓Look for convergence or agreement between quantitative and qualitative findings.
- ✓When quantitative data and qualitative data independently point to similar conclusions or patterns, it strengthens the validity of those conclusions.

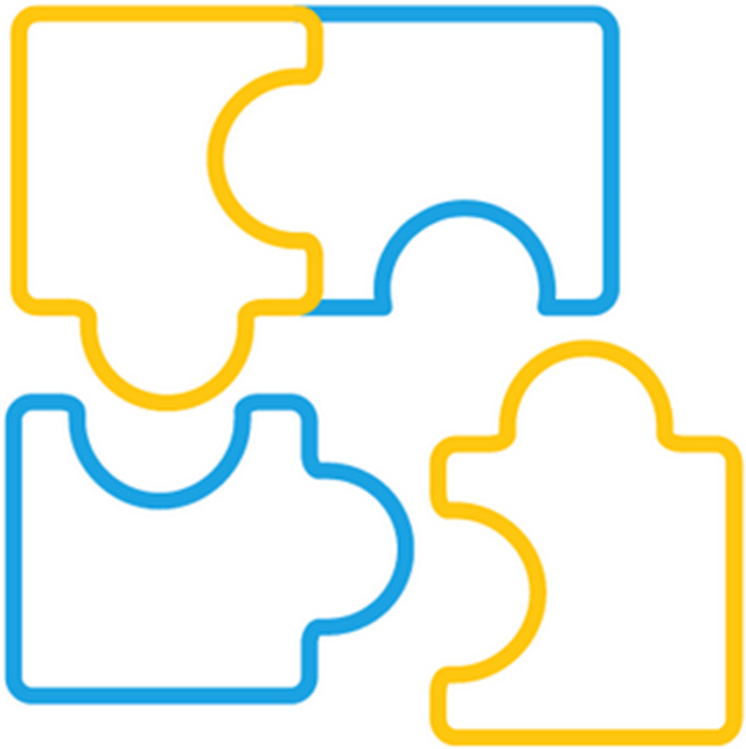
Center for Evaluation and Development

Validation:

✓If the quantitative data show a correlation between two variables, qualitative interviews can provide insights into why this relationship exists.

Center for Evaluation and Development

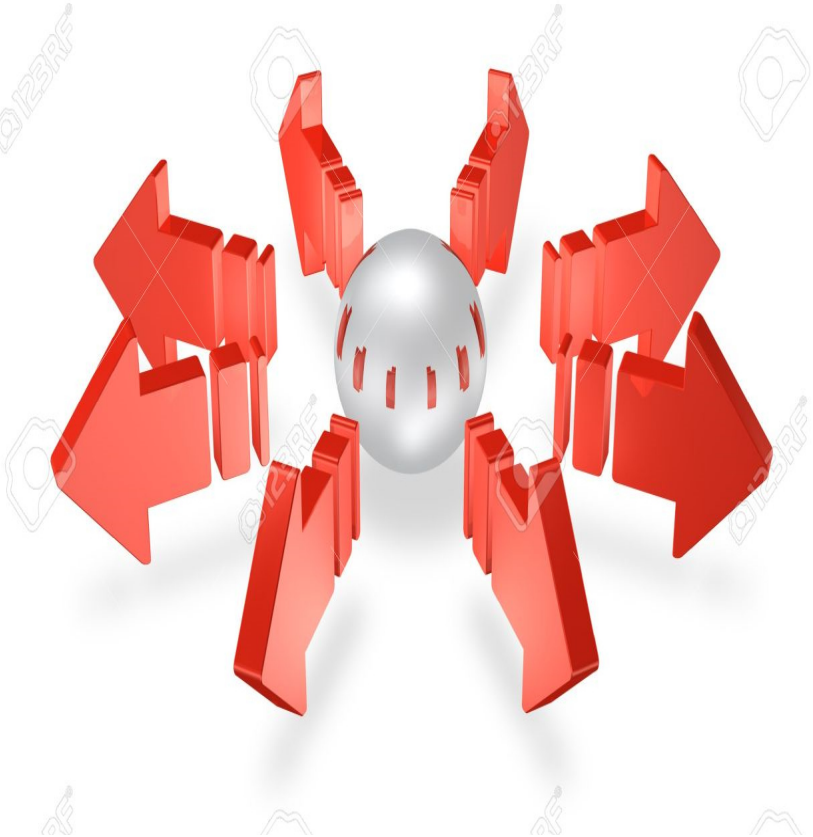
Complementarity:



- ✓Quantitative data may provide a broad overview, while qualitative data can provide depth and context.
- ✓Qualitative data can be used to explore the "**why**" and "**how**" behind quantitative results.

Center for Evaluation and Development

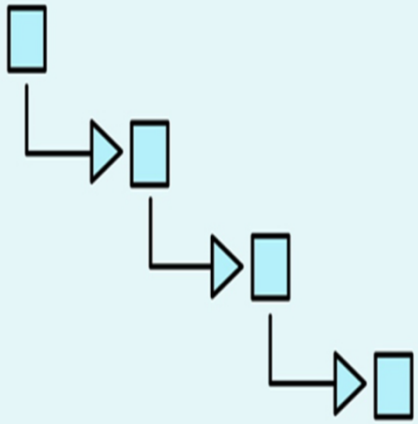
Expansion:



✓Qualitative data can expand or extend quantitative findings by uncovering unexpected insights or nuances that were not captured in the initial quantitative analysis.

Center for Evaluation and Development

Sequential Phases:

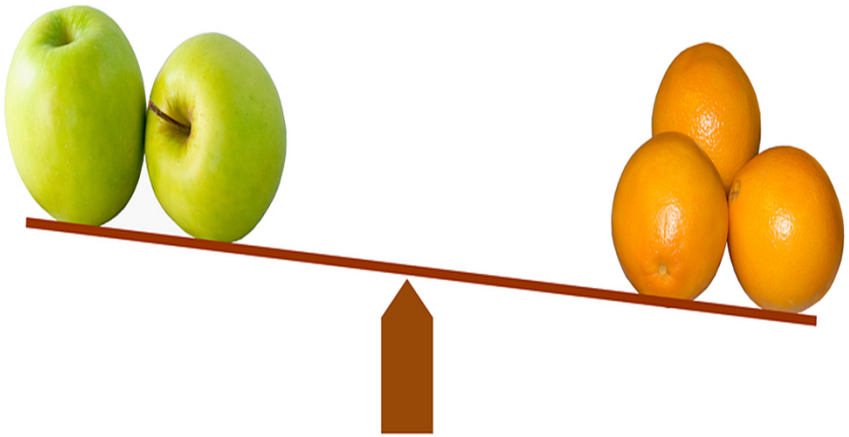


- ✓ Quantitative data collection and analysis in one phase of the study, followed by qualitative data collection and analysis in a subsequent phase.

- ✓ The two sets of findings are integrated during the interpretation phase to build a more comprehensive picture of the research question.

Center for Evaluation and Development

Comparison:



- ✓ Compare and contrast quantitative and qualitative results to
  - ↳ Identify areas of agreement, disagreement, or complementarity
  - ↳ Identify areas where further research or exploration is needed.

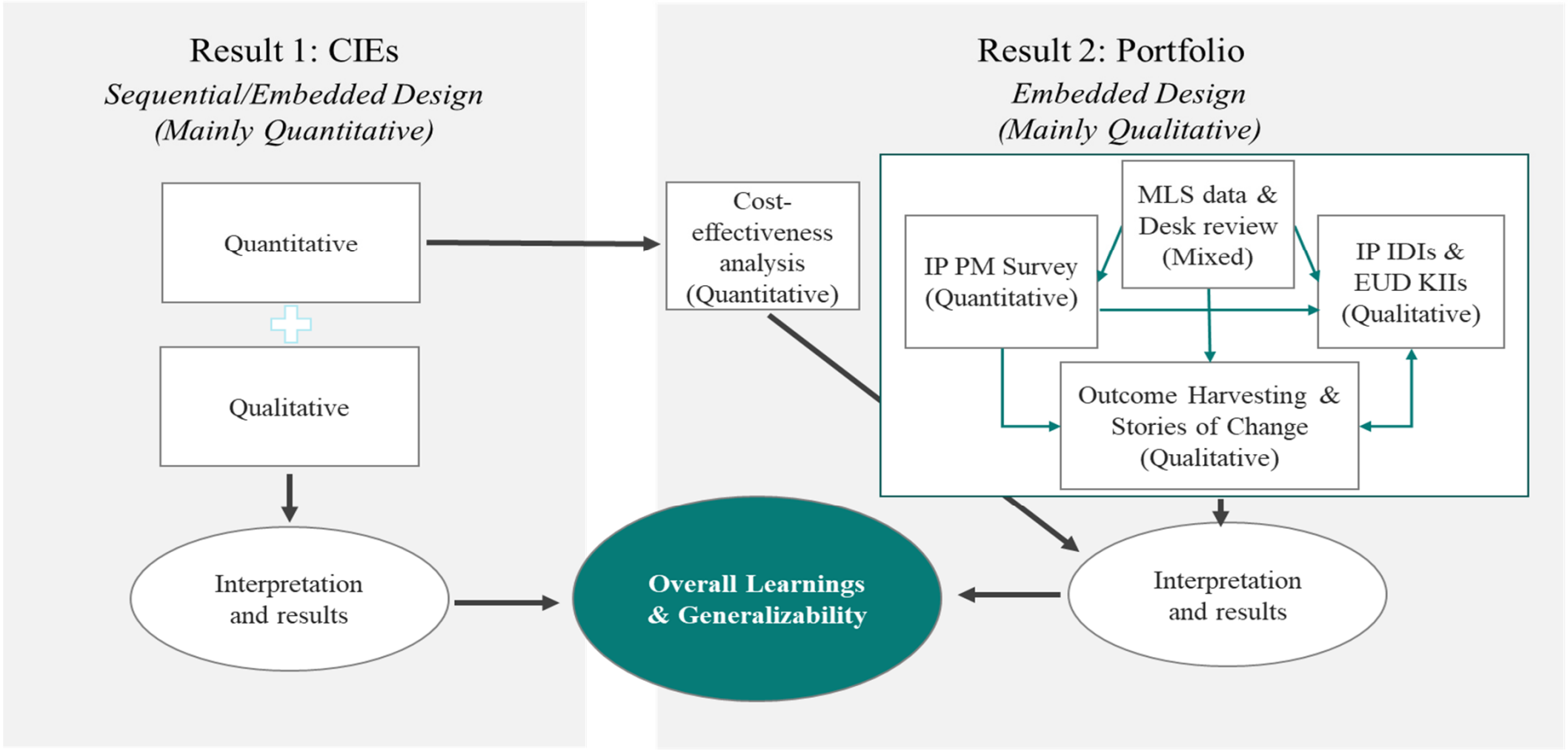
Center for Evaluation and Development

Priority Sequencing:

- ✓Either quantitative or qualitative data may take priority in the analysis.
  - ↳If quantitative data are considered foundational, qualitative findings may be used to help explain or contextualize the quantitative results.

Center for Evaluation and Development

Triangulation design used by C4ED FOR EUTF portfolio evaluation



Center for Evaluation and Development

In summary

Aspect	Qualitative Research	Quantitative Research
Research Objective	Explore and understand phenomena, generate theories, and gain in-depth insights.	Measure and quantify variables, test hypotheses, establish patterns.
Data Type	Non-numeric, textual, narrative, visual, and descriptive data.	Numeric, statistical, structured, and quantifiable data.
Data Collection Methods	In-depth interviews, focus groups, participant observation, content analysis, and observations.	Surveys, experiments, observations, content analysis, and structured data collection instruments.
Sample Size	Smaller, often purposeful or non-probabilistic sampling.	Larger, often randomly selected samples, using probabilistic sampling methods.
Data Analysis Methods	Qualitative coding, thematic analysis, constant comparative analysis, and narrative synthesis.	Statistical analysis, inferential analysis, descriptive statistics, and hypothesis testing.

Center for Evaluation and Development

Researcher's Role	More subjective, as researchers often interact with participants and interpret data.	More objective, as researchers aim for neutrality and minimal interference.
Contextualization	Emphasizes context, culture, and the subjective experiences of participants.	Context may be controlled or minimized to focus on variables.
Validity and Reliability	Focus on trustworthiness, credibility, and rigor.	Focus on validity, reliability, and replicability.
Generalizability	Findings are context-specific and not easily generalizable to broader populations.	Findings aim for broader generalizability and the ability to make predictions.
Timeframe	Often a longer data collection process; results may take time.	Typically more structured and efficient, with faster results.

Center for Evaluation and Development

Thank you for listening!

Questions?

Session 6: Understanding and using outputs of CIE

C4ED – EUTF

October 2023

Center for Evaluation and Development

European Commission

- In the previous sessions over the last three years, we have discussed step-by-step the method of conducting a CIE

- Overarching goal of CIEs are to expand the evidence base on programmes, policies or intervention
  - Does it work?
  - Why?
  - Why not?

Center for Evaluation and Development

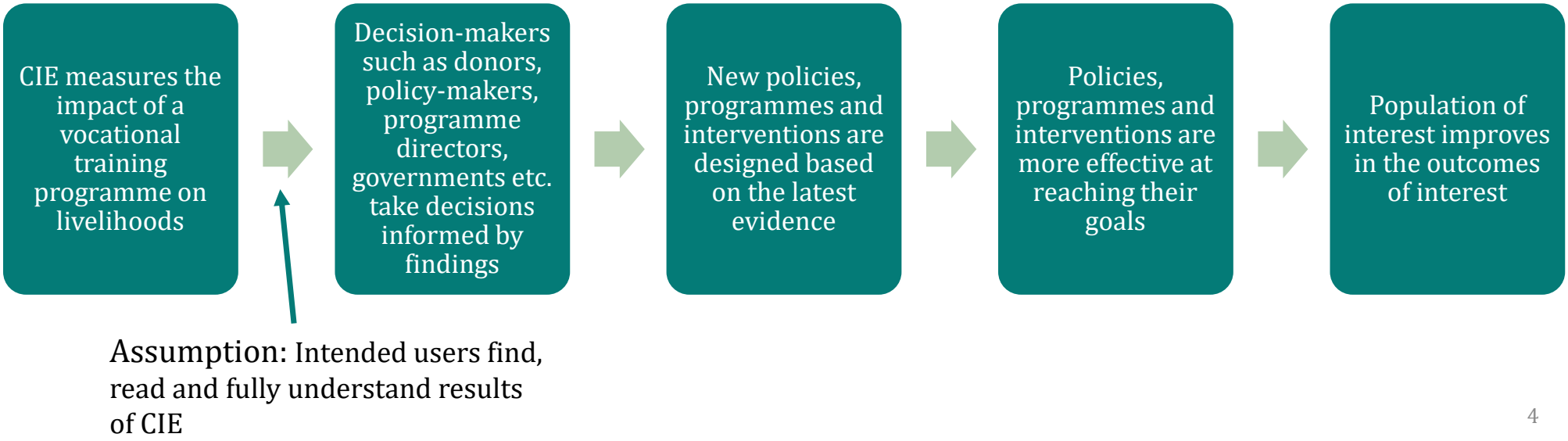
European Commission

- To help increase the impact of CIEs we can map out a basic theory of change
- Just as we do for an intervention itself

Center for Evaluation and Development

European Commission

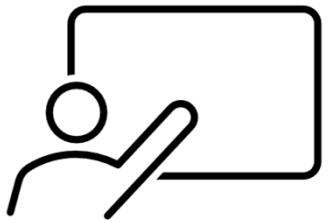
- To help increase the impact of CIEs we can map out a basic theory of change
- Just as we do for an intervention itself

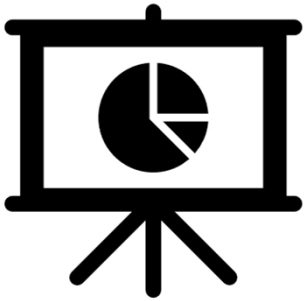


Center for Evaluation and Development

European Commission

- Dissemination of CIE results:
  - Workshops
  - Presentations
  - Report
  - Academic Journal article
  - Briefs





Center for Evaluation and Development

European Commission

- Throughout these sessions we have covered:
  - Descriptive Statistics
  - Statistical Tests for differences between groups
  - Regression models

- The outputs of these items should be displayed in CIE outputs

- Conclusions and recommendations should flow from these outputs

Center for Evaluation and Development

European Commission

- **Descriptive Statistics**
- What are the characteristics of the participants of a programme?
  - Age
  - Gender
  - Education Level
  - Urban/Rural
  - Income Level

- Does this fit the context that you – as a user – is interested in?
- Generalizability?

	(1) Full sample	(2) Treatment	(3) Comparison
Principal sociodemographic characteristics			
Age	25.0	25.2	24.8
	(4.3)	(4.3)	(4.4)
Female	0.50	0.52	0.47
	(0.50)	(0.50)	(0.50)
Returnee	0.09	0.12	0.05
	(0.29)	(0.33)	(0.22)
Not married	0.85	0.85	0.85
	(0.36)	(0.36)	(0.36)
Education			
Primary or lower	0.06	0.06	0.07
	(0.24)	(0.24)	(0.25)
Junior Secondary	0.20	0.17	0.23
	(0.40)	(0.37)	(0.42)
Senior Secondary	0.48	0.50	0.45
	(0.50)	(0.50)	(0.50)
Tertiary/vocational	0.25	0.26	0.23
training	(0.43)	(0.44)	(0.42)
Higher level	0.02	0.01	0.03
	(0.14)	(0.10)	(0.17)
Region of birth and residence			
Born in Greater Banjul	0.56	0.54	0.58
Area	(0.50)	(0.50)	(0.49)
Residence in Greater	0.69	0.64	0.74
Banjul Area	(0.46)	(0.48)	(0.44)
Other characteristics			
Parents' income	4,767	4,627	4,923
	(5,603)	(5,810)	(5,370)
# of adults in HH	5.1	5.3	5.0
	(3.5)	(3.9)	(3.1)
# of children in HH	5.0	4.9	5.2
	(4.7)	(5.1)	(4.2)
Interview score			
Average interview score	15.5	16.4	14.3
	(2.7)	(2.1)	(2.8)
Observations	1,402	760	642

Center for Evaluation and Development

European Commission

- Average age of participants in the study: 25
  - Using mean average
- 50% of the study participants are female
- Recall: Measures of central tendency!

	(1) Full sample	(2) Treatment	(3) Comparison
Principal sociodemographic characteristics			
Age	25.0	25.2	24.8
	(4.3)	(4.3)	(4.4)
Female	0.50	0.52	0.47
	(0.50)	(0.50)	(0.50)
Returnee	0.09	0.12	0.05
	(0.29)	(0.33)	(0.22)
Not married	0.85	0.85	0.85
	(0.36)	(0.36)	(0.36)
Education			
Primary or lower	0.06	0.06	0.07
	(0.24)	(0.24)	(0.25)
Junior Secondary	0.20	0.17	0.23
	(0.40)	(0.37)	(0.42)
Senior Secondary	0.48	0.50	0.45
	(0.50)	(0.50)	(0.50)
Tertiary/vocational	0.25	0.26	0.23
training	(0.43)	(0.44)	(0.42)
Higher level	0.02	0.01	0.03
	(0.14)	(0.10)	(0.17)
Region of birth and residence			
Born in Greater Banjul	0.56	0.54	0.58
Area	(0.50)	(0.50)	(0.49)
Residence in Greater	0.69	0.64	0.74
Banjul Area	(0.46)	(0.48)	(0.44)
Other characteristics			
Parents' income	4,767	4,627	4,923
	(5,603)	(5,810)	(5,370)
# of adults in HH	5.1	5.3	5.0
	(3.5)	(3.9)	(3.1)
# of children in HH	5.0	4.9	5.2
	(4.7)	(5.1)	(4.2)
Interview score			
Average interview score	15.5	16.4	14.3
	(2.7)	(2.1)	(2.8)
Observations	1,402	760	642

Center for Evaluation and Development

European Commission

- Recall: Good practice to include measures of dispersion in results
- Typically standard deviations within descriptive statistics

	(1) Full sample	(2) Treatment	(3) Comparison
Principal sociodemographic characteristics			
Age	25.0	25.2	24.8
	(4.3)	(4.3)	(4.4)
Female	0.50	0.52	0.47
	(0.50)	(0.50)	(0.50)
Returnee	0.09	0.12	0.05
	(0.29)	(0.33)	(0.22)
Not married	0.85	0.85	0.85
	(0.36)	(0.36)	(0.36)
Education			
Primary or lower	0.06	0.06	0.07
	(0.24)	(0.24)	(0.25)
Junior Secondary	0.20	0.17	0.23
	(0.40)	(0.37)	(0.42)
Senior Secondary	0.48	0.50	0.45
	(0.50)	(0.50)	(0.50)
Tertiary/vocational	0.25	0.26	0.23
training	(0.43)	(0.44)	(0.42)
Higher level	0.02	0.01	0.03
	(0.14)	(0.10)	(0.17)
Region of birth and residence			
Born in Greater Banjul	0.56	0.54	0.58
Area	(0.50)	(0.50)	(0.49)
Residence in Greater	0.69	0.64	0.74
Banjul Area	(0.46)	(0.48)	(0.44)
Other characteristics			
Parents' income	4,767	4,627	4,923
	(5,603)	(5,810)	(5,370)
# of adults in HH	5.1	5.3	5.0
	(3.5)	(3.9)	(3.1)
# of children in HH	5.0	4.9	5.2
	(4.7)	(5.1)	(4.2)
Interview score			
Average interview score	15.5	16.4	14.3
	(2.7)	(2.1)	(2.8)
Observations	1,402	760	642

Center for Evaluation and Development

European Commission

- It's important to know how the participants in a programme may differ from non-participants
- Recall from Session 3a – we can use statistical tests
- Compare the means
- Check likelihood that this is down to chance

	(1) Full sample	(2) Treatment	(3) Comparison	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age	25.0	25.2	24.8	0.4*
	(4.3)	(4.3)	(4.4)	(0.07)
Female	0.50	0.52	0.47	0.05*
	(0.50)	(0.50)	(0.50)	(0.05)
Returnee	0.09	0.12	0.05	0.07***
	(0.29)	(0.33)	(0.22)	(0.00)
Not married	0.85	0.85	0.85	0.00
	(0.36)	(0.36)	(0.36)	(0.88)
Education				
Primary or lower	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.25)	(0.48)
Junior Secondary	0.20	0.17	0.23	-0.06***
	(0.40)	(0.37)	(0.42)	(0.01)
Senior Secondary	0.48	0.50	0.45	0.06**
	(0.50)	(0.50)	(0.50)	(0.03)
Tertiary/vocational training	0.25	0.26	0.23	0.03
	(0.43)	(0.44)	(0.42)	(0.19)
Higher level	0.02	0.01	0.03	-0.02***
	(0.14)	(0.10)	(0.17)	(0.01)
Region of birth and residence				
Born in Greater Banjul Area	0.56	0.54	0.58	-0.04
	(0.50)	(0.50)	(0.49)	(0.14)
Residence in Greater Banjul Area	0.69	0.64	0.74	-0.10***
	(0.46)	(0.48)	(0.44)	(0.00)
Other characteristics				
Parents' income	4,767	4,627	4,923	-296
	(5,603)	(5,810)	(5,370)	(0.49)
# of adults in HH	5.1	5.3	5.0	0.3
	(3.5)	(3.9)	(3.1)	(0.11)
# of children in HH	5.0	4.9	5.2	-0.2
	(4.7)	(5.1)	(4.2)	(0.35)
Interview score				
Average interview score	15.5	16.4	14.3	2.1***
	(2.7)	(2.1)	(2.8)	(0.00)
Observations	1,402	760	642	

Significance stars: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Center for Evaluation and Development

European Commission

- Difference between treatment and comparison is 0.4 years
  - Participants are, on average, almost half a year older than non-participants

	(1) Full sample	(2) Treatment	(3) Comparison	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age	25.0	25.2	24.8	0.4*
	(4.3)	(4.3)	(4.4)	(0.07)
Female	0.50	0.52	0.47	0.05*
	(0.50)	(0.50)	(0.50)	(0.05)
Returnee	0.09	0.12	0.05	0.07***
	(0.29)	(0.33)	(0.22)	(0.00)
Not married	0.85	0.85	0.85	0.00
	(0.36)	(0.36)	(0.36)	(0.88)
Education				
Primary or lower	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.25)	(0.48)
Junior Secondary	0.20	0.17	0.23	-0.06***
	(0.40)	(0.37)	(0.42)	(0.01)
Senior Secondary	0.48	0.50	0.45	0.06**
	(0.50)	(0.50)	(0.50)	(0.03)
Tertiary/vocational training	0.25	0.26	0.23	0.03
	(0.43)	(0.44)	(0.42)	(0.19)
Higher level	0.02	0.01	0.03	-0.02***
	(0.14)	(0.10)	(0.17)	(0.01)
Region of birth and residence				
Born in Greater Banjul Area	0.56	0.54	0.58	-0.04
	(0.50)	(0.50)	(0.49)	(0.14)
Residence in Greater Banjul Area	0.69	0.64	0.74	-0.10***
	(0.46)	(0.48)	(0.44)	(0.00)
Other characteristics				
Parents' income	4,767	4,627	4,923	-296
	(5,603)	(5,810)	(5,370)	(0.49)
# of adults in HH	5.1	5.3	5.0	0.3
	(3.5)	(3.9)	(3.1)	(0.11)
# of children in HH	5.0	4.9	5.2	-0.2
	(4.7)	(5.1)	(4.2)	(0.35)
Interview score				
Average interview score	15.5	16.4	14.3	2.1***
	(2.7)	(2.1)	(2.8)	(0.00)
Observations	1,402	760	642	

Significance stars: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Center for Evaluation and Development

European Commission

- In the brackets we have the p-value
- P-value of 0.07 translates as having 93% confidence that your results are not simply down to chance

	(1) Full sample	(2) Treatment	(3) Comparison	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age	25.0	25.2	24.8	0.4*
	(4.3)	(4.3)	(4.4)	(0.07)
Female	0.50	0.52	0.47	0.05*
	(0.50)	(0.50)	(0.50)	(0.05)
Returnee	0.09	0.12	0.05	0.07***
	(0.29)	(0.33)	(0.22)	(0.00)
Not married	0.85	0.85	0.85	0.00
	(0.36)	(0.36)	(0.36)	(0.88)
Education				
Primary or lower	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.25)	(0.48)
Junior Secondary	0.20	0.17	0.23	-0.06***
	(0.40)	(0.37)	(0.42)	(0.01)
Senior Secondary	0.48	0.50	0.45	0.06**
	(0.50)	(0.50)	(0.50)	(0.03)
Tertiary/vocational training	0.25	0.26	0.23	0.03
	(0.43)	(0.44)	(0.42)	(0.19)
Higher level	0.02	0.01	0.03	-0.02***
	(0.14)	(0.10)	(0.17)	(0.01)
Region of birth and residence				
Born in Greater Banjul Area	0.56	0.54	0.58	-0.04
	(0.50)	(0.50)	(0.49)	(0.14)
Residence in Greater Banjul Area	0.69	0.64	0.74	-0.10***
	(0.46)	(0.48)	(0.44)	(0.00)
Other characteristics				
Parents' income	4,767	4,627	4,923	-296
	(5,603)	(5,810)	(5,370)	(0.49)
# of adults in HH	5.1	5.3	5.0	0.3
	(3.5)	(3.9)	(3.1)	(0.11)
# of children in HH	5.0	4.9	5.2	-0.2
	(4.7)	(5.1)	(4.2)	(0.35)
Interview score				
Average interview score	15.5	16.4	14.3	2.1***
	(2.7)	(2.1)	(2.8)	(0.00)
Observations	1,402	760	642	

Significance stars: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.

Center for Evaluation and Development

European Commission

- The stars next to a value provide information on the level of confidence in the difference

- The more stars – the higher degree of confidence

	(1) Full sample	(2) Treatment	(3) Comparison	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age	25.0	25.2	24.8	0.4*
	(4.3)	(4.3)	(4.4)	(0.07)
Female	0.50	0.52	0.47	0.05*
	(0.50)	(0.50)	(0.50)	(0.05)
Returnee	0.09	0.12	0.05	0.07***
	(0.29)	(0.33)	(0.22)	(0.00)
Not married	0.85	0.85	0.85	0.00
	(0.36)	(0.36)	(0.36)	(0.88)
Education				
Primary or lower	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.25)	(0.48)
Junior Secondary	0.20	0.17	0.23	-0.06***
	(0.40)	(0.37)	(0.42)	(0.01)
Senior Secondary	0.48	0.50	0.45	0.06**
	(0.50)	(0.50)	(0.50)	(0.03)
Tertiary/vocational training	0.25	0.26	0.23	0.03
	(0.43)	(0.44)	(0.42)	(0.19)
Higher level	0.02	0.01	0.03	-0.02***
	(0.14)	(0.10)	(0.17)	(0.01)
Region of birth and residence				
Born in Greater Banjul Area	0.56	0.54	0.58	-0.04
	(0.50)	(0.50)	(0.49)	(0.14)
Residence in Greater Banjul Area	0.69	0.64	0.74	-0.10***
	(0.46)	(0.48)	(0.44)	(0.00)
Other characteristics				
Parents' income	4,767	4,627	4,923	-296
	(5,603)	(5,810)	(5,370)	(0.49)
# of adults in HH	5.1	5.3	5.0	0.3
	(3.5)	(3.9)	(3.1)	(0.11)
# of children in HH	5.0	4.9	5.2	-0.2
	(4.7)	(5.1)	(4.2)	(0.35)
Interview score				
Average interview score	15.5	16.4	14.3	2.1***
	(2.7)	(2.1)	(2.8)	(0.00)
Observations	1,402	760	642	

Significance stars: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

Center for Evaluation and Development

European Commission

- Can we conclude from our results whether there is a difference in age between participants and non-participants?

	(1) Full sample	(2) Treatment	(3) Comparison	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age	25.0	25.2	24.8	0.4*
	(4.3)	(4.3)	(4.4)	(0.07)
Female	0.50	0.52	0.47	0.05*
	(0.50)	(0.50)	(0.50)	(0.05)
Returnee	0.09	0.12	0.05	0.07***
	(0.29)	(0.33)	(0.22)	(0.00)
Not married	0.85	0.85	0.85	0.00
	(0.36)	(0.36)	(0.36)	(0.88)
Education				
Primary or lower	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.25)	(0.48)
Junior Secondary	0.20	0.17	0.23	-0.06***
	(0.40)	(0.37)	(0.42)	(0.01)
Senior Secondary	0.48	0.50	0.45	0.06**
	(0.50)	(0.50)	(0.50)	(0.03)
Tertiary/vocational training	0.25	0.26	0.23	0.03
	(0.43)	(0.44)	(0.42)	(0.19)
Higher level	0.02	0.01	0.03	-0.02***
	(0.14)	(0.10)	(0.17)	(0.01)
Region of birth and residence				
Born in Greater Banjul Area	0.56	0.54	0.58	-0.04
	(0.50)	(0.50)	(0.49)	(0.14)
Residence in Greater Banjul Area	0.69	0.64	0.74	-0.10***
	(0.46)	(0.48)	(0.44)	(0.00)
Other characteristics				
Parents' income	4,767	4,627	4,923	-296
	(5,603)	(5,810)	(5,370)	(0.49)
# of adults in HH	5.1	5.3	5.0	0.3
	(3.5)	(3.9)	(3.1)	(0.11)
# of children in HH	5.0	4.9	5.2	-0.2
	(4.7)	(5.1)	(4.2)	(0.35)
Interview score				
Average interview score	15.5	16.4	14.3	2.1***
	(2.7)	(2.1)	(2.8)	(0.00)
Observations	1,402	760	642	

Significance stars: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$

Center for Evaluation and Development

European Commission

- What can we say based on what this line in our table tells us?

	(1) Full sample	(2) Treatment	(3) Comparison	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age	25.0	25.2	24.8	0.4*
	(4.3)	(4.3)	(4.4)	(0.07)
Female	0.50	0.52	0.47	0.05*
	(0.50)	(0.50)	(0.50)	(0.05)
Returnee	0.09	0.12	0.05	0.07***
	(0.29)	(0.33)	(0.22)	(0.00)
Not married	0.85	0.85	0.85	0.00
	(0.36)	(0.36)	(0.36)	(0.88)
Education				
Primary or lower	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.25)	(0.48)
Junior Secondary	0.20	0.17	0.23	-0.06***
	(0.40)	(0.37)	(0.42)	(0.01)
Senior Secondary	0.48	0.50	0.45	0.06**
	(0.50)	(0.50)	(0.50)	(0.03)
Tertiary/vocational training	0.25	0.26	0.23	0.03
	(0.43)	(0.44)	(0.42)	(0.19)
Higher level	0.02	0.01	0.03	-0.02***
	(0.14)	(0.10)	(0.17)	(0.01)
Region of birth and residence				
Born in Greater Banjul Area	0.56	0.54	0.58	-0.04
	(0.50)	(0.50)	(0.49)	(0.14)
Residence in Greater Banjul Area	0.69	0.64	0.74	-0.10***
	(0.46)	(0.48)	(0.44)	(0.00)
Other characteristics				
Parents' income	4,767	4,627	4,923	-296
	(5,603)	(5,810)	(5,370)	(0.49)
# of adults in HH	5.1	5.3	5.0	0.3
	(3.5)	(3.9)	(3.1)	(0.11)
# of children in HH	5.0	4.9	5.2	-0.2
	(4.7)	(5.1)	(4.2)	(0.35)
Interview score				
Average interview score	15.5	16.4	14.3	2.1***
	(2.7)	(2.1)	(2.8)	(0.00)
Observations	1,402	760	642	

Center for Evaluation and Development

European Commission

- What about the table as a whole
  - Do these two groups look similar?

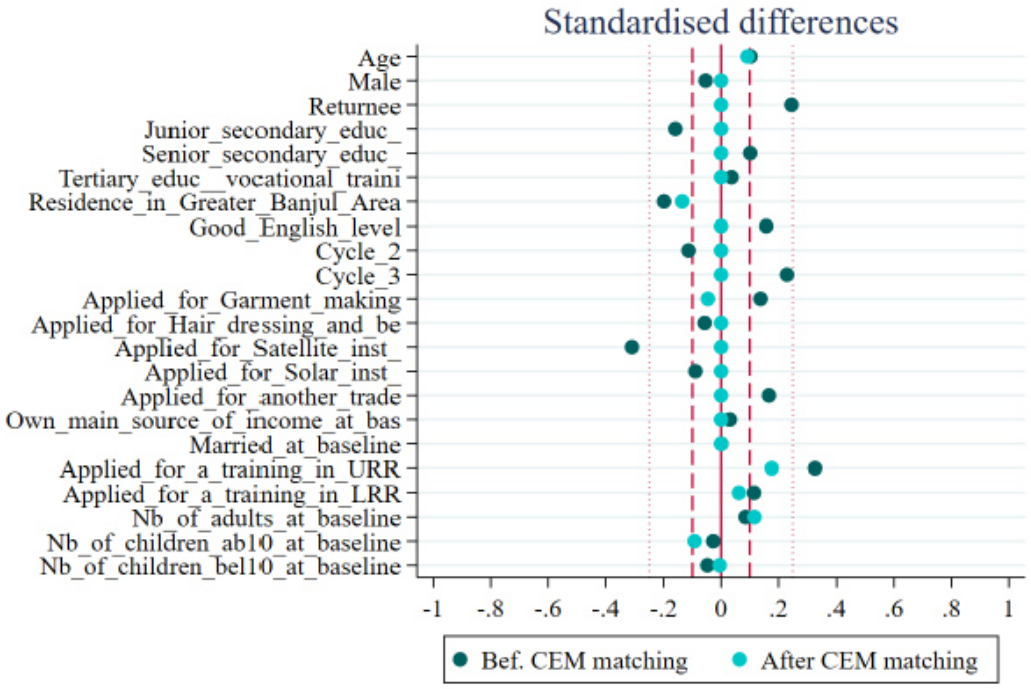
	(1) Full sample	(2) Treatment	(3) Comparison	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age	25.0	25.2	24.8	0.4*
	(4.3)	(4.3)	(4.4)	(0.07)
Female	0.50	0.52	0.47	0.05*
	(0.50)	(0.50)	(0.50)	(0.05)
Returnee	0.09	0.12	0.05	0.07***
	(0.29)	(0.33)	(0.22)	(0.00)
Not married	0.85	0.85	0.85	0.00
	(0.36)	(0.36)	(0.36)	(0.88)
Education				
Primary or lower	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.25)	(0.48)
Junior Secondary	0.20	0.17	0.23	-0.06***
	(0.40)	(0.37)	(0.42)	(0.01)
Senior Secondary	0.48	0.50	0.45	0.06**
	(0.50)	(0.50)	(0.50)	(0.03)
Tertiary/vocational training	0.25	0.26	0.23	0.03
	(0.43)	(0.44)	(0.42)	(0.19)
Higher level	0.02	0.01	0.03	-0.02***
	(0.14)	(0.10)	(0.17)	(0.01)
Region of birth and residence				
Born in Greater Banjul Area	0.56	0.54	0.58	-0.04
	(0.50)	(0.50)	(0.49)	(0.14)
Residence in Greater Banjul Area	0.69	0.64	0.74	-0.10***
	(0.46)	(0.48)	(0.44)	(0.00)
Other characteristics				
Parents' income	4,767	4,627	4,923	-296
	(5,603)	(5,810)	(5,370)	(0.49)
# of adults in HH	5.1	5.3	5.0	0.3
	(3.5)	(3.9)	(3.1)	(0.11)
# of children in HH	5.0	4.9	5.2	-0.2
	(4.7)	(5.1)	(4.2)	(0.35)
Interview score				
Average interview score	15.5	16.4	14.3	2.1***
	(2.7)	(2.1)	(2.8)	(0.00)
Observations	1,402	760	642	

Center for Evaluation and Development

European Commission

- This intervention did **not** use random assignment
- Whether someone took part in the intervention depended on an interview score
  - Made up of various criteria
- Used propensity score matching to reduce the differences between groups

	(1) Full sample	(2) Treatment	(3) Comparison	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age	25.0	25.2	24.8	0.4*
	(4.3)	(4.3)	(4.4)	(0.07)
Female	0.50	0.52	0.47	0.05*
	(0.50)	(0.50)	(0.50)	(0.05)
Returnee	0.09	0.12	0.05	0.07***
	(0.29)	(0.33)	(0.22)	(0.00)
Not married	0.85	0.85	0.85	0.00
	(0.36)	(0.36)	(0.36)	(0.88)
Education				
Primary or lower	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.25)	(0.48)
Junior Secondary	0.20	0.17	0.23	-0.06***
	(0.40)	(0.37)	(0.42)	(0.01)
Senior Secondary	0.48	0.50	0.45	0.06**
	(0.50)	(0.50)	(0.50)	(0.03)
Tertiary/vocational training	0.25	0.26	0.23	0.03
	(0.43)	(0.44)	(0.42)	(0.19)
Higher level	0.02	0.01	0.03	-0.02***
	(0.14)	(0.10)	(0.17)	(0.01)
Region of birth and residence				
Born in Greater Banjul Area	0.56	0.54	0.58	-0.04
	(0.50)	(0.50)	(0.49)	(0.14)
Residence in Greater Banjul Area	0.69	0.64	0.74	-0.10***
	(0.46)	(0.48)	(0.44)	(0.00)
Other characteristics				
Parents' income	4,767	4,627	4,923	-296
	(5,603)	(5,810)	(5,370)	(0.49)
# of adults in HH	5.1	5.3	5.0	0.3
	(3.5)	(3.9)	(3.1)	(0.11)
# of children in HH	5.0	4.9	5.2	-0.2
	(4.7)	(5.1)	(4.2)	(0.35)
Interview score				
Average interview score	15.5	16.4	14.3	2.1***
	(2.7)	(2.1)	(2.8)	(0.00)
Observations	1,402	760	642	


Standardised differences
Age
Male
Returnee
Junior_secondary_educ_
Senior_secondary_educ_
Tertiary_educ__vocational_traini
Residence_in_Greater_Banjul_Area
Good_English_level
Cycle_2
Cycle_3
Applied_for_Garment_making
Applied_for_Hair_dressing_and_be
Applied_for_Satellite_inst_
Applied_for_Solar_inst_
Applied_for_another_trade
Own_main_source_of_income_at_bas
Married_at_baseline
Applied_for_a_training_in_URR
Applied_for_a_training_in_LRR
Nb_of_adults_at_baseline
Nb_of_children_ab10_at_baseline
Nb_of_children_bel10_at_baseline
-1
-.8
-.6
-.4
-.2
0
.2
.4
.6
.8
1
Bef. CEM matching
After CEM matching

Center for Evaluation and Development

European Commission

- If we use random assignment we should expect that participants and non-participants are similar

- We can check this again using a difference in means test
  - Referred to as a balance table

	(1) Full sample	(2) Treatment (T1 & T2)	(3) Control	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age (in April 2022)	23.4	23.3	23.5	-0.1
	(4.5)	(4.6)	(4.5)	(0.44)
Female	0.61	0.63	0.56	0.07***
	(0.49)	(0.48)	(0.50)	(0.00)
Refugee	0.43	0.44	0.43	0.01
	(0.50)	(0.50)	(0.49)	(0.65)
Married	0.43	0.42	0.44	-0.02
	(0.49)	(0.49)	(0.50)	(0.23)
Education				
No formal education	0.06	0.06	0.05	0.01
	(0.24)	(0.24)	(0.22)	(0.28)
Primary	0.71	0.72	0.71	0.01
	(0.45)	(0.45)	(0.46)	(0.53)
Secondary	0.22	0.21	0.23	-0.02
	(0.41)	(0.41)	(0.42)	(0.15)
Tertiary	0.01	0.01	0.01	0.00
	(0.09)	(0.10)	(0.08)	(0.22)
Observations	3,330	2,243	1,087	

Note: Columns (1), (2) and (3) present the sample means (proportions when % is shown in the variable name or in the table) of selected variables for the full sample, the treatment group and the control group, respectively. Standard deviations in parentheses. Column (4) presents the mean difference between the treatment and control groups. P-value of the corresponding t-test in parentheses. Significance level: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.
Source: C4ED elaboration

Center for Evaluation and Development

European Commission

- Does it appear that the randomization was successful?

	(1) Full sample	(2) Treatment (T1 & T2)	(3) Control	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age (in April 2022)	23.4	23.3	23.5	-0.1
	(4.5)	(4.6)	(4.5)	(0.44)
Female	0.61	0.63	0.56	0.07***
	(0.49)	(0.48)	(0.50)	(0.00)
Refugee	0.43	0.44	0.43	0.01
	(0.50)	(0.50)	(0.49)	(0.65)
Married	0.43	0.42	0.44	-0.02
	(0.49)	(0.49)	(0.50)	(0.23)
Education				
No formal education	0.06	0.06	0.05	0.01
	(0.24)	(0.24)	(0.22)	(0.28)
Primary	0.71	0.72	0.71	0.01
	(0.45)	(0.45)	(0.46)	(0.53)
Secondary	0.22	0.21	0.23	-0.02
	(0.41)	(0.41)	(0.42)	(0.15)
Tertiary	0.01	0.01	0.01	0.00
	(0.09)	(0.10)	(0.08)	(0.22)
Observations	3,330	2,243	1,087	

Note: Columns (1), (2) and (3) present the sample means (proportions when % is shown in the variable name or in the table) of selected variables for the full sample, the treatment group and the control group, respectively. Standard deviations in parentheses. Column (4) presents the mean difference between the treatment and control groups. P-value of the corresponding t-test in parentheses. Significance level: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.
Source: C4ED elaboration

	(1) Full sample	(2) Treatment (T1 & T2)	(3) Control	(4) (2)-(3) (p-value)
Has a stable job	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.26)	(0.34)
Employee	0.07	0.06	0.09	-0.03
	(0.26)	(0.25)	(0.29)	(0.44)
Family worker	0.17	0.19	0.12	0.07*
	(0.37)	(0.39)	(0.33)	(0.08)
Self employed	0.51	0.51	0.50	0.01
	(0.50)	(0.50)	(0.50)	(0.89)
Apprentice	0.04	0.04	0.03	0.01
	(0.19)	(0.20)	(0.16)	(0.49)
Casual worker	0.17	0.16	0.19	-0.03
	(0.38)	(0.37)	(0.40)	(0.53)
Other	0.04	0.03	0.06	-0.04
	(0.20)	(0.16)	(0.25)	(0.24)
Observations	3,168	2,131	1,037	

Note: Columns (1), (2) and (3) present the sample means (proportions when % is shown in the variable name or in the table) of selected variables for the full sample, the treatment group and the control group, respectively. Standard deviations in parentheses. Column (4) presents the mean difference between the treatment and control groups. P-value of the corresponding t-test in parentheses. Significance level: $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.
Source: C4ED elaboration

Center for Evaluation and Development

European Commission

- Does it appear that the randomization was successful?

	(1) Full sample	(2) Treatment (T1 & T2)	(3) Control	(4) (2)-(3) (p-value)
Principal sociodemographic characteristics				
Age (in April 2022)	23.4	23.3	23.5	-0.1
	(4.5)	(4.6)	(4.5)	(0.44)
Female	0.61	0.63	0.56	0.07***
	(0.49)	(0.48)	(0.50)	(0.00)
Refugee	0.43	0.44	0.43	0.01
	(0.50)	(0.50)	(0.49)	(0.65)
Married	0.43	0.42	0.44	-0.02
	(0.49)	(0.49)	(0.50)	(0.23)
Education				
No formal education	0.06	0.06	0.05	0.01
	(0.24)	(0.24)	(0.22)	(0.28)
Primary	0.71	0.72	0.71	0.01
	(0.45)	(0.45)	(0.46)	(0.53)
Secondary	0.22	0.21	0.23	-0.02
	(0.41)	(0.41)	(0.42)	(0.15)
Tertiary	0.01	0.01	0.01	0.00
	(0.09)	(0.10)	(0.08)	(0.22)
Observations	3,330	2,243	1,087	

Note: Columns (1), (2) and (3) present the sample means (proportions when % is shown in the variable name or in the table) of selected variables for the full sample, the treatment group and the control group, respectively. Standard deviations in parentheses. Column (4) presents the mean difference between the treatment and control groups. P-value of the corresponding t-test in parentheses. Significance level: * $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.
Source: C4ED elaboration

	(1) Full sample	(2) Treatment (T1 & T2)	(3) Control	(4) (2)-(3) (p-value)
Has a stable job	0.06	0.06	0.07	-0.01
	(0.24)	(0.24)	(0.26)	(0.34)
Employee	0.07	0.06	0.09	-0.03
	(0.26)	(0.25)	(0.29)	(0.44)
Family worker	0.17	0.19	0.12	0.07*
	(0.37)	(0.39)	(0.33)	(0.08)
Self employed	0.51	0.51	0.50	0.01
	(0.50)	(0.50)	(0.50)	(0.89)
Apprentice	0.04	0.04	0.03	0.01
	(0.19)	(0.20)	(0.16)	(0.49)
Casual worker	0.17	0.16	0.19	-0.03
	(0.38)	(0.37)	(0.40)	(0.53)
Other	0.04	0.03	0.06	-0.04
	(0.20)	(0.16)	(0.25)	(0.24)
Observations	3,168	2,131	1,037	

Note: Columns (1), (2) and (3) present the sample means (proportions when % is shown in the variable name or in the table) of selected variables for the full sample, the treatment group and the control group, respectively. Standard deviations in parentheses. Column (4) presents the mean difference between the treatment and control groups. P-value of the corresponding t-test in parentheses. Significance level: $p \leq 0.1$, ** $p \leq 0.05$, *** $p \leq 0.01$.
Source: C4ED elaboration

Center for Evaluation and Development

European Commission

- After all the steps setting up an evaluation you finally get to the end…

- The impact estimate…

Center for Evaluation and Development

European Commission

Table 5.5 Impact estimates for education and employment outcomes

	DiD	Comp. group: Mean (st. dev.)	(1) ATT	(2) ITT	(3) LATE
(i) Education					
Enrolled	Yes	**0.513**	**0.124*****	**0.016**	**0.041**
		0.500	0.046	0.036	0.094
(ii) Labour market					
Employed	Yes	**0.286**	**-0.081***	**-0.063***	**-0.165***
		0.452	0.040	0.034	0.088
Unemployed	Yes	**0.180**	**0.007**	**-0.007**	**-0.018**
		0.385	0.042	0.031	0.082

	ATT		POM	Observations Total	T	C
Employment						
Has a job (last 7 days)	0.15	***	0.56	1229	657	572
	(0.03)					

	Reduced form coefficient estimate (s.e.)		IV-2SLS coefficient estimate (s.e.)		Control group mean (s.d.)	Number of Observations
	Vocational education	Vocational Education *W2	Vocational Education	Vocational Education*W2		
Panel A: Total Earnings[1]						
Ln(Total earnings, past month)	-0.260**	0.405**	-0.103	0.128	7.644	1,549
	(0.123)	(0.167)	(0.142)	(0.202)	(1.272)	
Ln(Total earnings, past month), including zeros	-0.146	0.096	-0.094	0.040`	6.264	1,892
	(0.282)	(0.387)	(0.340)	(0.465)	(3.159)	

Center for Evaluation and Development

European Commission

Table 5.5 Impact estimates for education and employment outcomes

	DiD	Comp. group: Mean (st. dev.)	(1) ATT	(2) ITT	(3) LATE
(i) Education					
Enrolled	Yes	**0.513**	**0.124*** **	**0.016**	**0.041**
		0.500	0.046	0.036	0.094
(ii) Labour market					
Employed	Yes	**0.286**	**-0.081***	**-0.063***	**-0.165***
		0.452	0.040	0.034	0.088
Unemployed	Yes	**0.180**	**0.007**	**-0.007**	**-0.018**
		0.385	0.042	0.031	0.082

	ATT		POM	Observations Total	T	C
Employment						
Has a job (last 7 days)	0.15	***	0.56	1229	657	572
	(0.03)					

	Reduced form coefficient estimate (s.e.)		IV-2SLS coefficient estimate (s.e.)		Control group mean (s.d.)	Number of Observations
	Vocational education	Vocational Education *W2	Vocational Education	Vocational Education*W2		
Panel A: Total Earnings[1]						
Ln(Total earnings, past month)	-0.260**	0.405**	-0.103	0.128	7.644	1,549
	(0.123)	(0.167)	(0.142)	(0.202)	(1.272)	
Ln(Total earnings, past month), including zeros	-0.146	0.096	-0.094	0.040`	6.264	1,892
	(0.282)	(0.387)	(0.340)	(0.465)	(3.159)	

Center for Evaluation and Development

European Commission

- Common ingredients of a impact estimate table..

Center for Evaluation and Development

European Commission

- Estimate of impact based on model(s)
  - Measure of central tendency
    - Absolute – i.e. 500 dollars per month
    - Relative
      - % increase
      - Percentage point increase
  - Measure of dispersion
    - Standard deviation
    - Standard errors
- Reference point of comparison
- Statistical testing (e.g. p-value)
- Treatment effect type

	ATT		POM	Observations		
				Total	T	C
Employment						
Has a job (last 7 days)	0.15	***	0.56	1229	657	572
	(0.03)					
Months in employment in the last 6 months	0.02		4.75	795	462	333
	(0.15)					
Has a stable job	0.09	***	0.61	1229	657	572
	(0.03)					
Self-employed in stable job	0.11	***	0.27	1229	657	572
	(0.03)					

Center for Evaluation and Development

European Commission

- Estimate of impact based on model(s)
  - Measure of central tendency
    - Absolute – i.e. 500 dollars per month
    - Relative
      - % increase
      - Percentage point increase
  - Measure of dispersion
    - Standard deviation
    - Standard errors
- Reference point of comparison
- Statistical testing (e.g. p-value)
- Treatment effect type

	ATT		POM	Observations		
				Total	T	C
Employment						
Has a job (last 7 days)	0.15	***	0.56	1229	657	572
	(0.03)					
Months in employment in the last 6 months	0.02		4.75	795	462	333
	(0.15)					
Has a stable job	0.09	***	0.61	1229	657	572
	(0.03)					
Self-employed in stable job	0.11	***	0.27	1229	657	572
	(0.03)					

Center for Evaluation and Development

European Commission

- Estimate of impact based on model(s)
  - Measure of central tendency
    - Absolute – i.e. 500 dollars per month
    - Relative
      - % increase
      - Percentage point increase
  - Measure of dispersion
    - Standard deviation
    - Standard errors
- Reference point of comparison
- Statistical testing (e.g. p-value)
- Treatment effect type

	ATT		POM	Observations		
				Total	T	C
Employment						
Has a job (last 7 days)	0.15	***	0.56	1229	657	572
	(0.03)					
Months in employment in the last 6 months	0.02		4.75	795	462	333
	(0.15)					
Has a stable job	0.09	***	0.61	1229	657	572
	(0.03)					
Self-employed in stable job	0.11	***	0.27	1229	657	572
	(0.03)					

Center for Evaluation and Development

European Commission

- Estimate of impact based on model(s)
  - Measure of central tendency
    - Absolute – i.e. 500 dollars per month
    - Relative
      - % increase
      - Percentage point increase
  - Measure of dispersion
    - Standard deviation
    - Standard errors
- Reference point of comparison
- Statistical testing (e.g. p-value)
- Treatment effect type

	ATT		POM	Observations		
				Total	T	C
Employment						
Has a job (last 7 days)	0.15	***	0.56	1229	657	572
	(0.03)					
Months in employment in the last 6 months	0.02		4.75	795	462	333
	(0.15)					
Has a stable job	0.09	***	0.61	1229	657	572
	(0.03)					
Self-employed in stable job	0.11	***	0.27	1229	657	572
	(0.03)					

Center for Evaluation and Development

European Commission

- Estimate of impact based on model(s)
  - Measure of central tendency
    - Absolute – i.e. 500 dollars per month
    - Relative
      - % increase
      - Percentage point increase
  - Measure of dispersion
    - Standard deviation
    - Standard errors
- Reference point of comparison
- Statistical testing (e.g. p-value)
- Treatment effect type

	ATT		POM	Observations		
				Total	T	C
Employment						
Has a job (last 7 days)	0.15	***	0.56	1229	657	572
	(0.03)					
Months in employment in the last 6 months	0.02		4.75	795	462	333
	(0.15)					
Has a stable job	0.09	***	0.61	1229	657	572
	(0.03)					
Self-employed in stable job	0.11	***	0.27	1229	657	572
	(0.03)					

- Treatment effect types
  - Outside the scope of this workshop

- Examples of type of effect
  - ATT: Average treatment effect on the treated
    - Considers those that received the treatment
  - ITT: Effect of Intention to treat
    - Considers only original assignment to programme or not
    - Particularly useful for policy considerations

Center for Evaluation and Development

European Commission

- What initial conclusions can we draw based on this output table for an impact estimate?

	ATT		POM	Observations		
				Total	T	C
Income						
Average monthly income over last 6 months	472.83	**	1475	1189	636	553
	(238.79)					

Notes: *, **, & *** represent statistical significance at the 10%, 5%, & 1% level respectively.

Results from IPWRA regressions. Regressions include covariates.

POM is expressed in the outcome's original unit.

Center for Evaluation and Development

European Commission

- Important thing to consider when interpreting estimates of impact

- Effect Sizes
- Internal Validity
- External Validity

Center for Evaluation and Development

European Commission

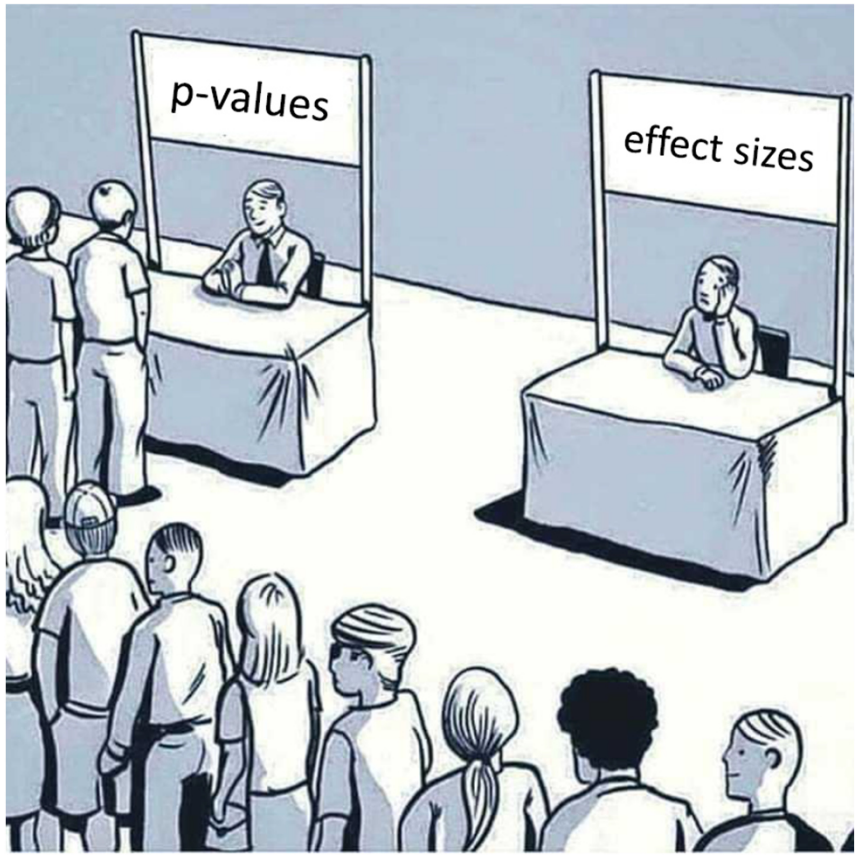
- Important thing to consider when interpreting estimates of impact

- **Effect Sizes**
- Internal Validity
- External Validity

Center for Evaluation and Development

European Commission

$p<0.05$



Center for Evaluation and Development

European Commission

- Statistical testing and p-values are extremely important for the reasons we have discussed
- We need to have confidence that our results are credible and not simply down to random chance

- However a low p-value for an impact estimate on an outcome we're interested in does not mean an intervention is great and should be expanded!

Center for Evaluation and Development

European Commission

- We can have a low p-value but a very small effect
- What is the % increase in income due to the programme?
- Would you consider the programme to be effective?

	Treatment	Control	ATT	p-value
Income	1,002	1,000	2 **	(0.048)
	(20)	(25)	(1.01)	
N	1,000	1,000		

Center for Evaluation and Development

European Commission

- Important thing to consider when interpreting estimates of impact

- Effect Sizes
- **Internal Validity**
- External Validity

- **Internal validity**
  - Confidence that the measured impact is an unbiased estimate of the true impact
- Just as with statistical testing – impact estimates are simply an output of a mathematical formula!
- Internal validity depends on all of the factors that we've discussed
  - CIE design
  - Sampling
  - Data Quality

Center for Evaluation and Development

European Commission

- Important thing to consider when interpreting estimates of impact

- Effect Sizes
- Internal Validity
- **External Validity**

Center for Evaluation and Development

European Commission

- External validity
  - Degree to which the estimated effect is generalizable

- Can we assume that our impact estimates will be the same when scaled up?

- There can often a trade-off between internal and external validity
- Small and highly controlled randomized trial may be very effective in limiting bias in the impact
- But may focus on a specific geographical location, sub-section of the population
- “Lab setting” vs “Real world”

Center for Evaluation and Development

European Commission

END OF SESSION 6