

Evaluating Programs to Increase Women's Economic Empowerment: Introduction to Evaluation (Part 2)

Kenyatta University Workshop #2

April 23, 2021

Dr. Charity Troyer Moore

Yale University

Recap from Session # 1

- Used data to explore *why* female entrepreneurs earn less income than males
- Considered a hypothetical program, EFEADP, aimed at increasing women's acceptance of digital payments in their business
 - Underlying idea: we see a correlation between acceptance of digital payments and income – hope this is causal. We'll test this with EFEADP
- Explored theory of change linking EFEADP to women's improved outcomes
- Groups considered different ways to design an evaluation of EFEADP

Encouraging Female Entrepreneurs' Acceptance of Digital Payments



EFEADP:

Information, training, and incentives to encourage women to accept digital payments

Goal of Your Evaluation

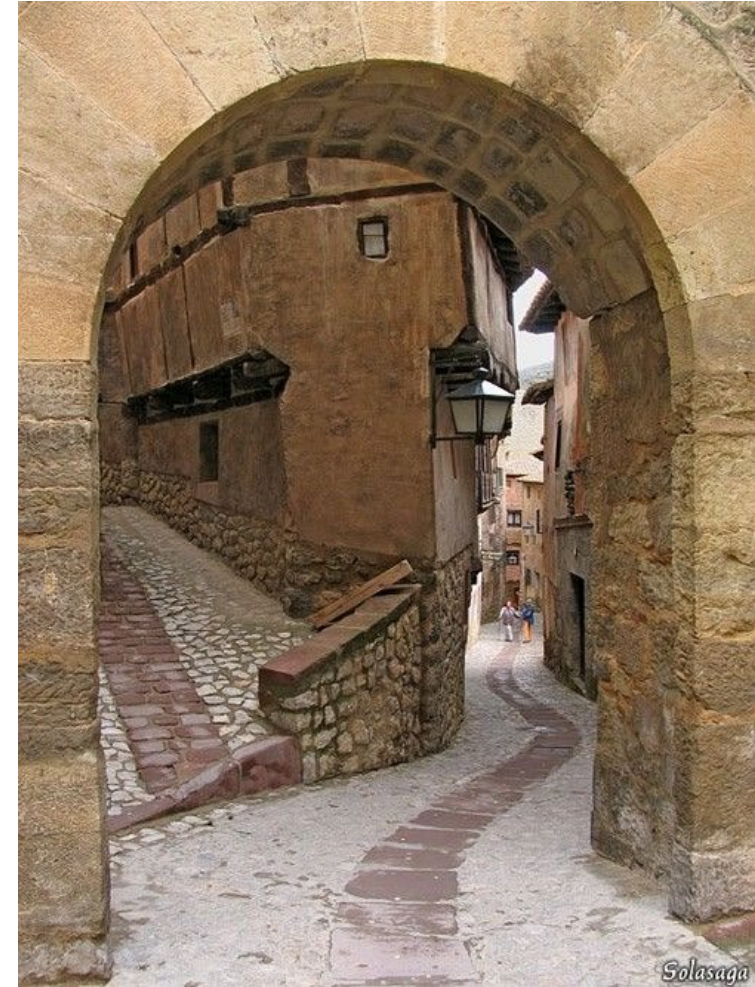
Your goal:

- Understand the *causal impact* of the program – Does the program increase female entrepreneurs' acceptance of digital payments, and does this in turn improve her business outcomes, and broader empowerment?

Counterfactuals

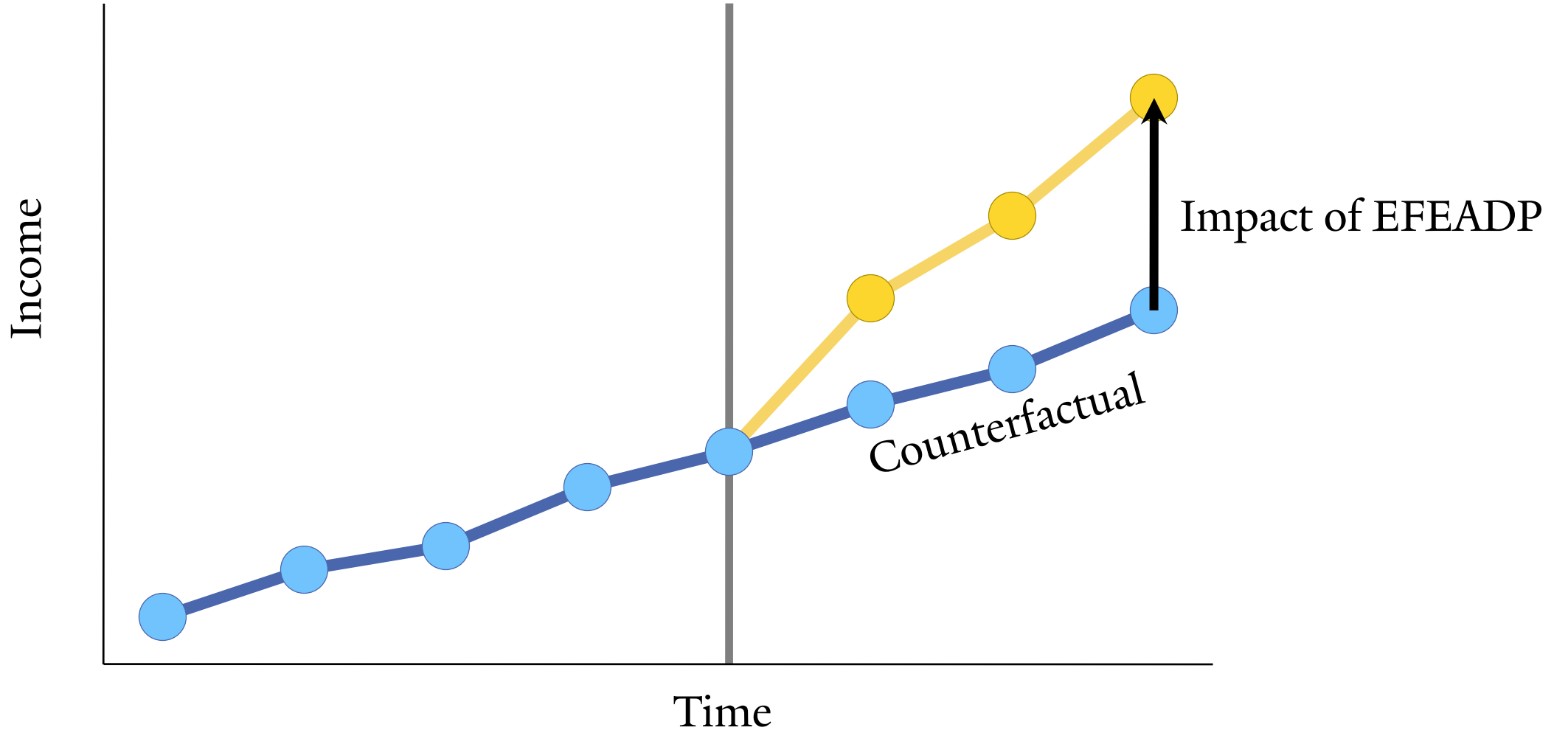
- In research: The outcome that would have occurred for program participants, had they not participated in the program.
- Specifically: The income (outcomes) EFEADP enrollees would have had if they hadn't participated in EFEADP.
- What's the challenge here?

The counterfactual is never observable.



The Key to Causal Identification is
Identifying a Credible Counterfactual
(Comparison) Group

EFEADP Programme



What Makes a Credible Counterfactual?

1. Absent the intervention, the treatment group and the comparison group should, on average, look identical.
 - Women's average characteristics should look the same whether they were *invited* to EFEADP or were not invited – they were in the comparison group.
 - We also assume both groups would react the same to the intervention if they were to receive it.
2. The treatment should not affect the comparison group either directly or indirectly.
 - Women in EFEADP, for example, would not have friends in non-EFEADP group with whom they share all their training.

Identifying a credible counterfactual gives us an *impact evaluation*. Randomized control trials can do this quite robustly.

Questions to ask yourself

- What is my comparison group (“control”)?
 - Does it look the same as the treatment group prior to/absent intervention?
YES
 - Would the comparison group react to treatment the same way the comparison group does?
 - Does treatment affect outcomes of people in the comparison group?
NO

Groupwork: Possible Evaluation Approaches

Approach #1:

Enroll women in one county in EFEADP. 3 months later, collect data on income from women in EFEADP and women in another county where EFEADP wasn't introduced. Compare EFEADP women's income with income of women in this other, non-EFEADP area.

What challenges to identification of a causal impact does this approach present?

Groupwork: Possible Evaluation Approaches

Approach #1:

Enroll women in one county in EFEADP. 3 months later, collect data on income from women in EFEADP and women in another county where EFEADP wasn't introduced. Compare EFEADP women's income with income of women in this other, non-EFEADP area.

What is the comparison group?

Self-employed women in the other county

- Does the comparison group look the same as the treatment group prior to/absent intervention?

Possibly

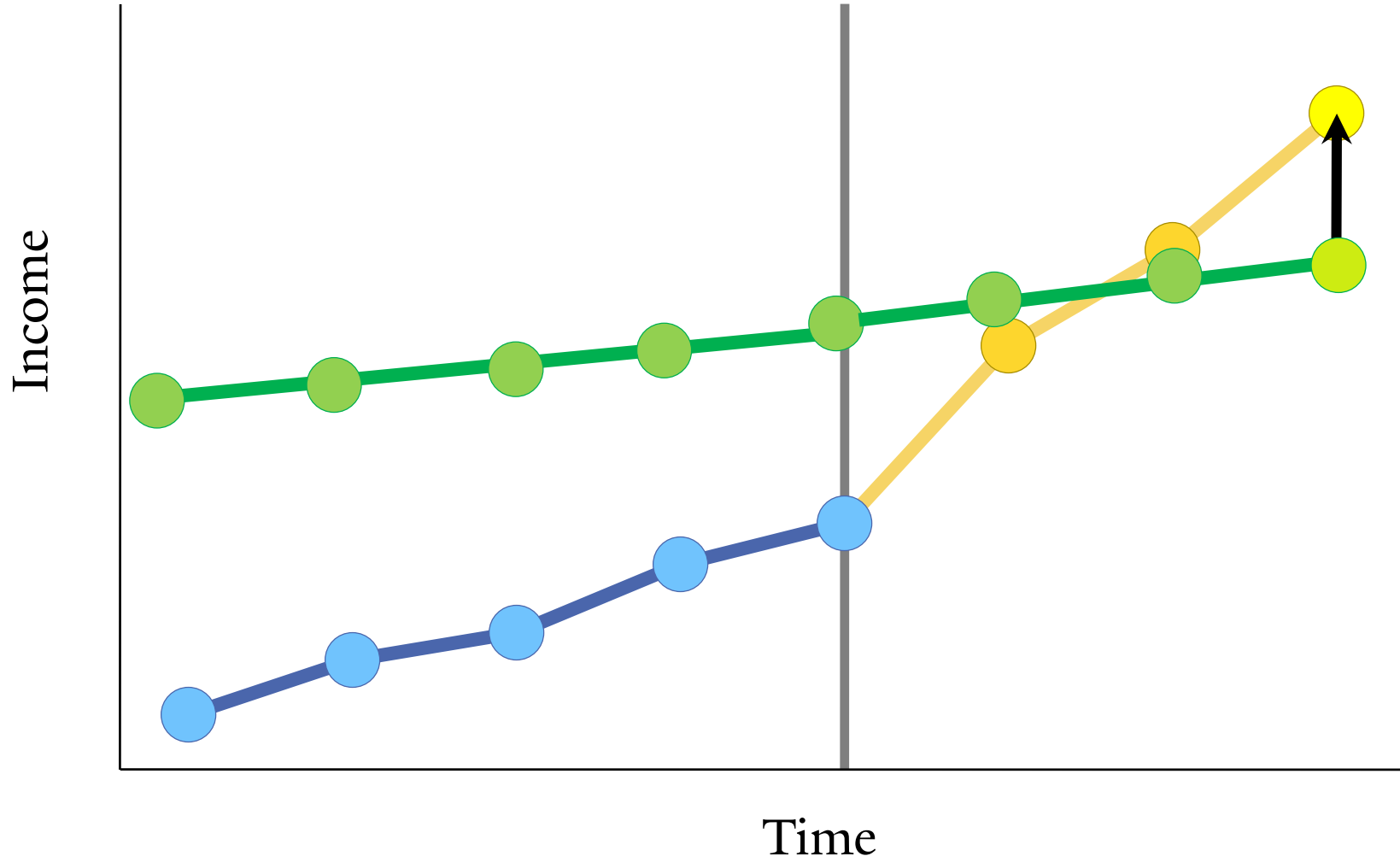
- Does treatment affect outcomes of people in the comparison group?

Probably not

- Would the comparison group react to treatment the same way the treatment group does?

Probably?

EFEADP Programme

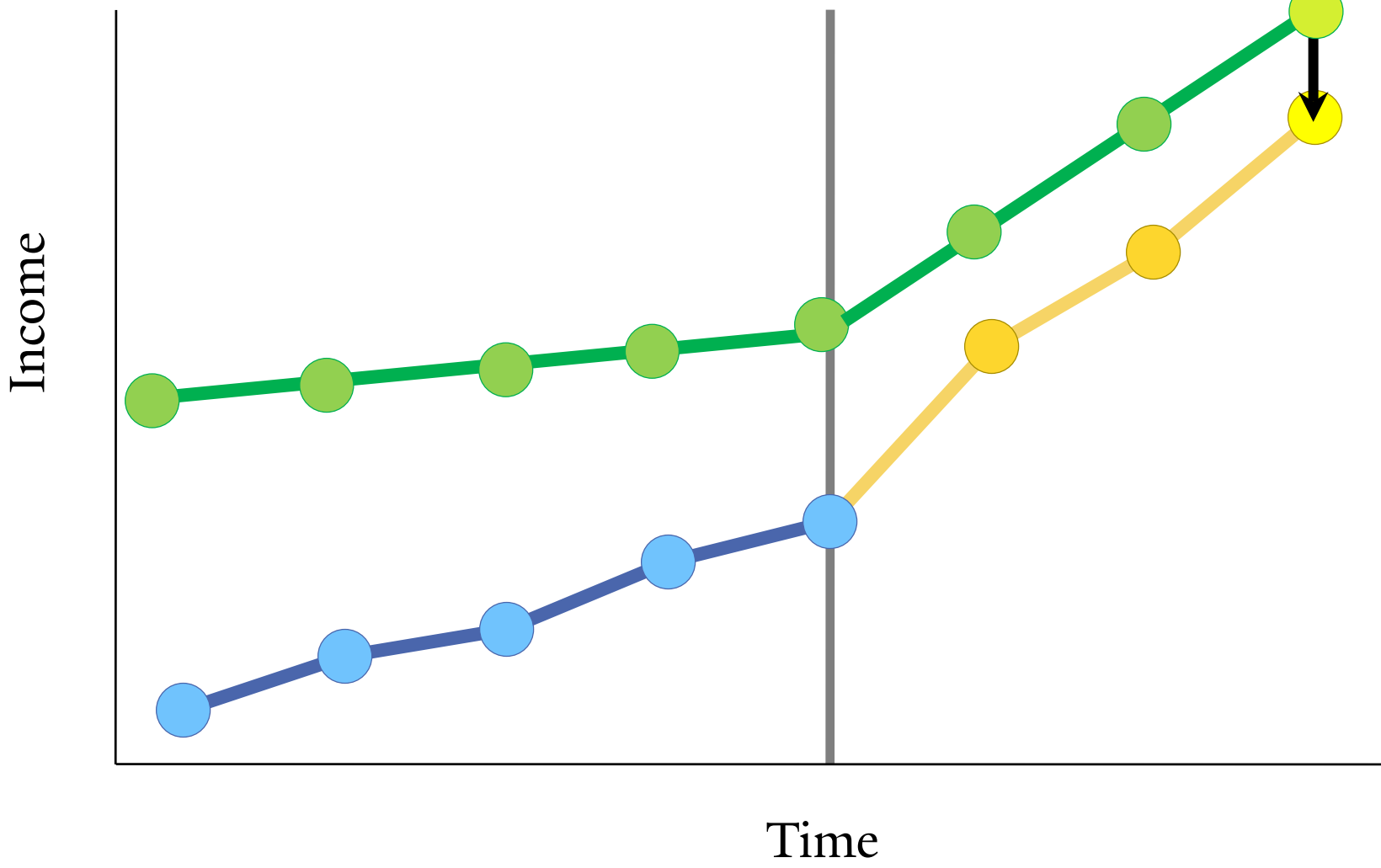


Estimated
Impact of EFEADP =
● - ●

Legend

- Not invited to EFEADP (other county)
- Invited to EFEADP Before training
- Invited to EFEADP After training

EFEADP Programme



Estimated
Impact of EFEADP =



Legend

- Not invited to EFEADP (other county)
- Invited to EFEADP Before training
- Invited to EFEADP After training

Groupwork: Possible Evaluation Approaches

Approach #2:

Enroll women in two counties in EFEADP. Collect data on everyone's income before the training begins and 3 months later. Compare the income over time to see if income is higher among EFEADP women 3 months later.

What challenges to identification of a causal impact does this approach present?

Groupwork: Possible Evaluation Approaches

Approach #2:

Enroll women in two counties in EFEADP. Collect data on everyone's income before the training begins and 3 months later. Compare the income over time to see if income is higher among EFEADP women 3 months later.

What is the comparison group?

Self-employed women, prior to the intervention

- Does the comparison group look the same as the treatment group prior to/absent intervention?

Yes

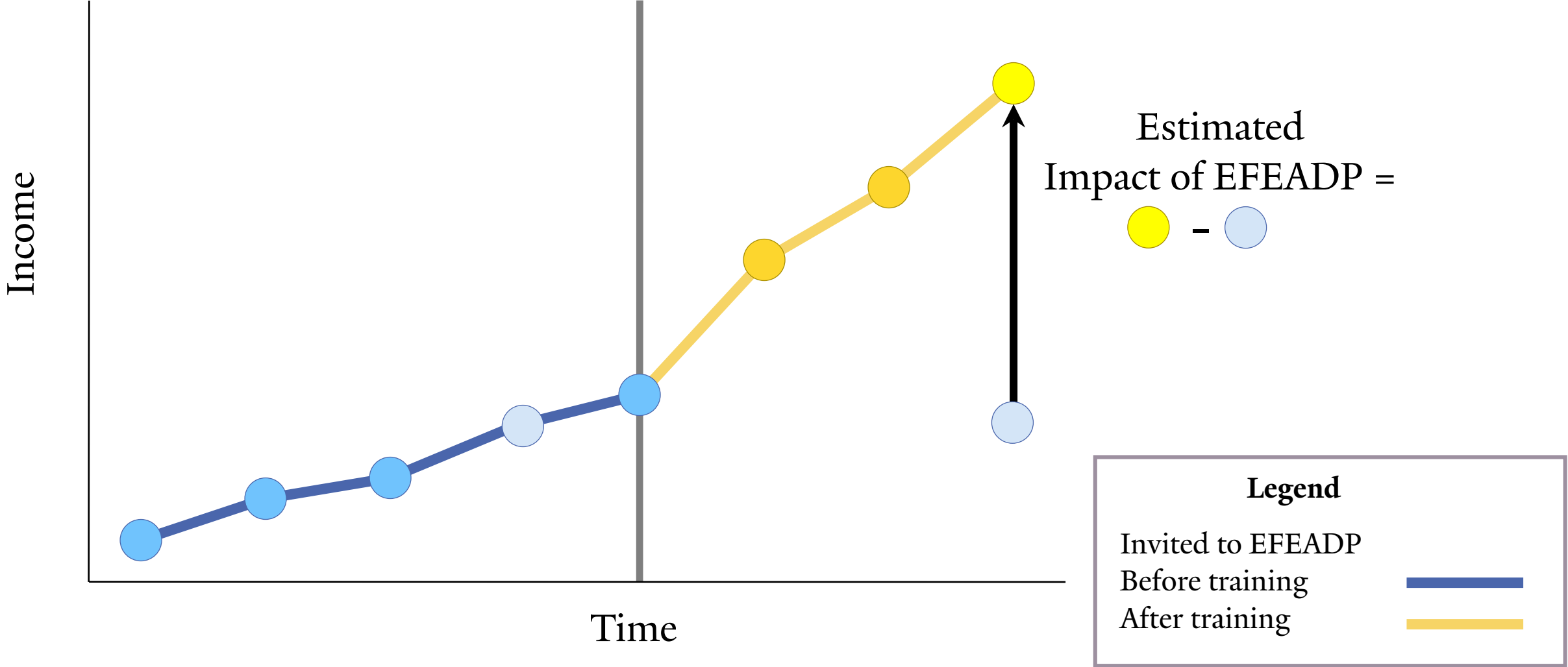
- Does treatment affect outcomes of people in the comparison group?

Yes – but less relevant here

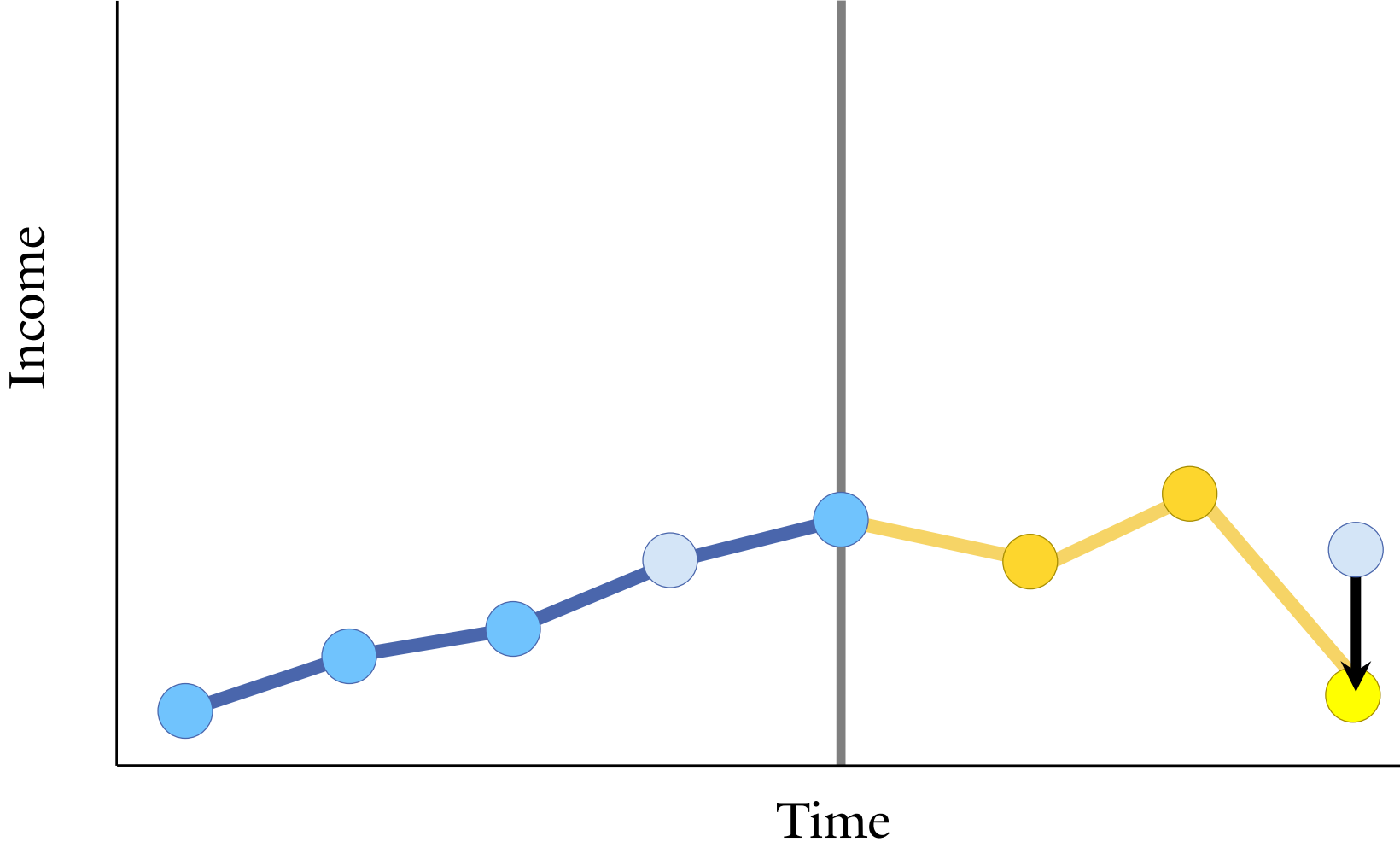
- Would the comparison group react to treatment the same way the treatment group does?

Probably – but don't know about time trends

EFEADP Programme





EFEADP Programme



Legend

Invited to EFEADP
Before training 
After training 

Estimated
Impact of EFEADP =
 - 

Groupwork: Possible Evaluation Approaches

Approach #3:

Enroll any woman who wants to participate in EFEADP across two counties. Collect data on their income **before** the training and 3 months **after** the training; collect the same data from women in the study area who learned about EFEADP but didn't enroll. Compare the *change* in income across the trained and untrained women.

What challenges to identification of a causal impact does this approach present?

Groupwork: Possible Evaluation Approaches

Approach #3:

Enroll any woman who wants to participate in EFEADP across two counties. Collect data on their income **before** the training and 3 months **after** the training; collect the same data from women in the study area who learned about EFEADP but didn't enroll. Compare the *change* in income across the trained and untrained women.

What is the comparison group?

Self-employed women who did not enroll in EFEADP

- Does the comparison group look the same as the treatment group prior to/absent intervention?

Probably not

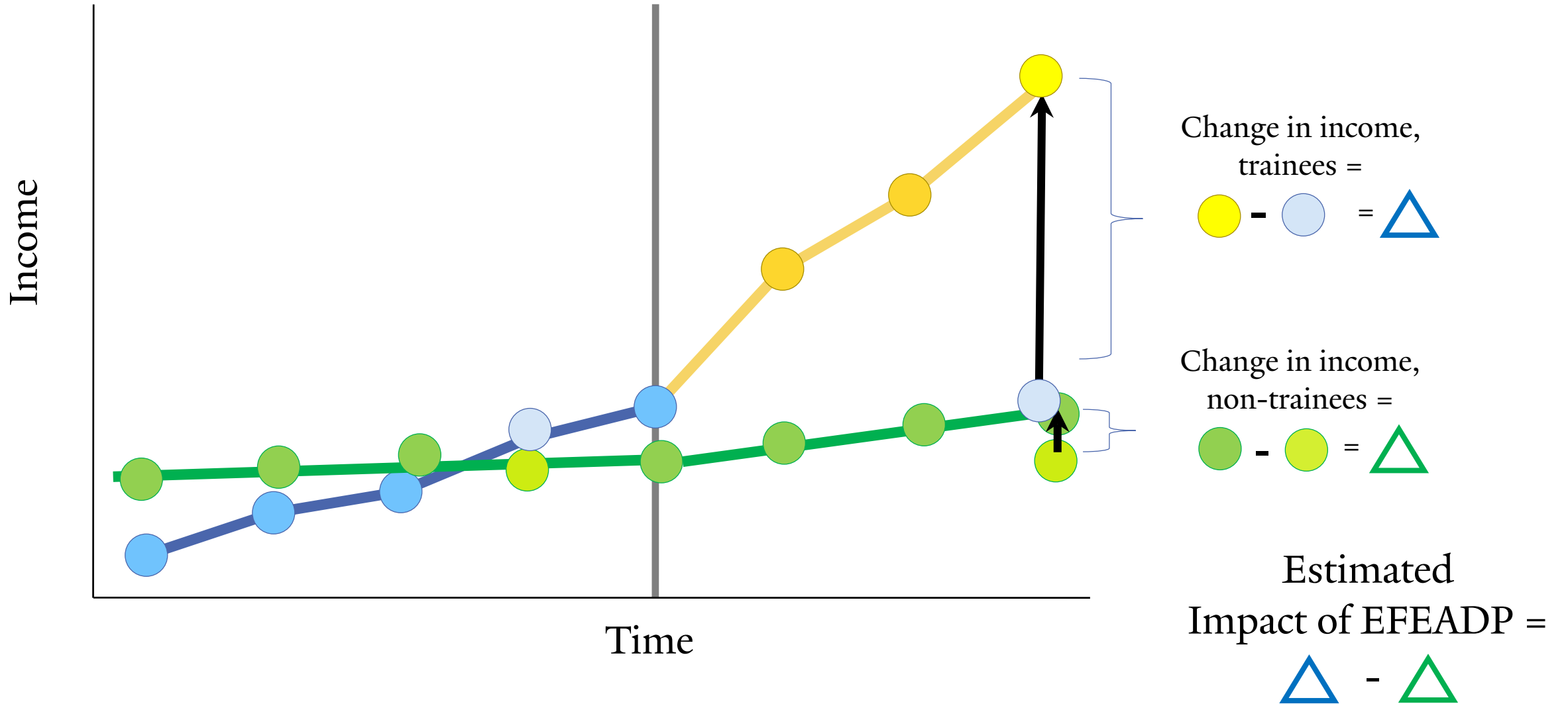
- Does treatment affect outcomes of people in the comparison group?

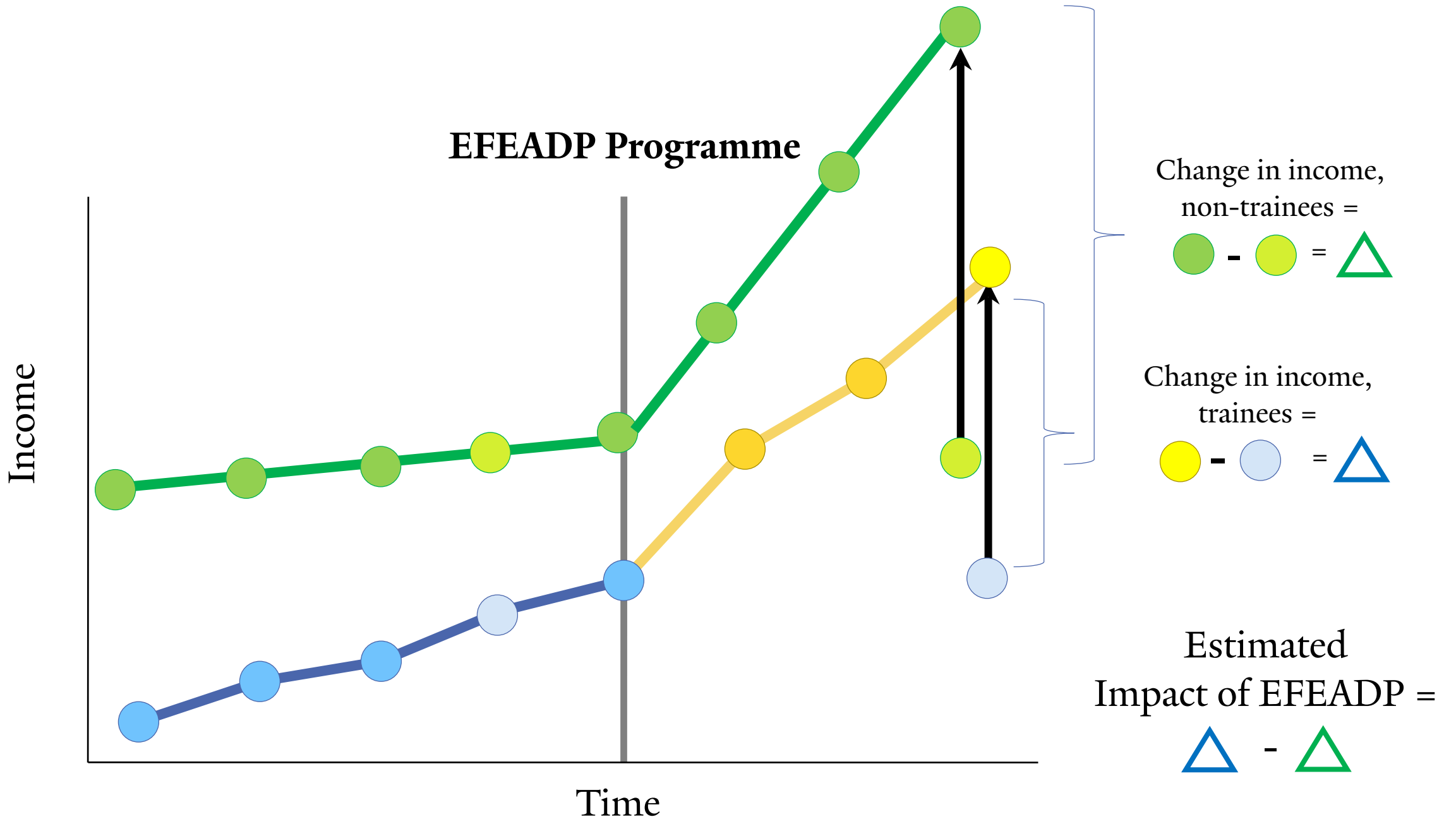
Possibly

- Would the comparison group react to treatment the same way the treatment group does?

Probably not

EFEADP Programme





Groupwork: Possible Evaluation Approaches

Approach #4:

Invite a randomly selected set of women across both counties to enroll in EFEADP. 3 months later, compare income of those who were invited to enroll with those that weren't invited to enroll.

What challenges to identification of a causal impact does this approach present?

Groupwork: Possible Evaluation Approaches

Approach #4:

Invite a randomly selected set of women across both counties to enroll in EFEADP. 3 months later, compare income of those who were invited to enroll with those that weren't invited to enroll.

What is the comparison group?

Women selected by random assignment to be in the control

- Does the comparison group look the same as the treatment group prior to/absent intervention?

Yes

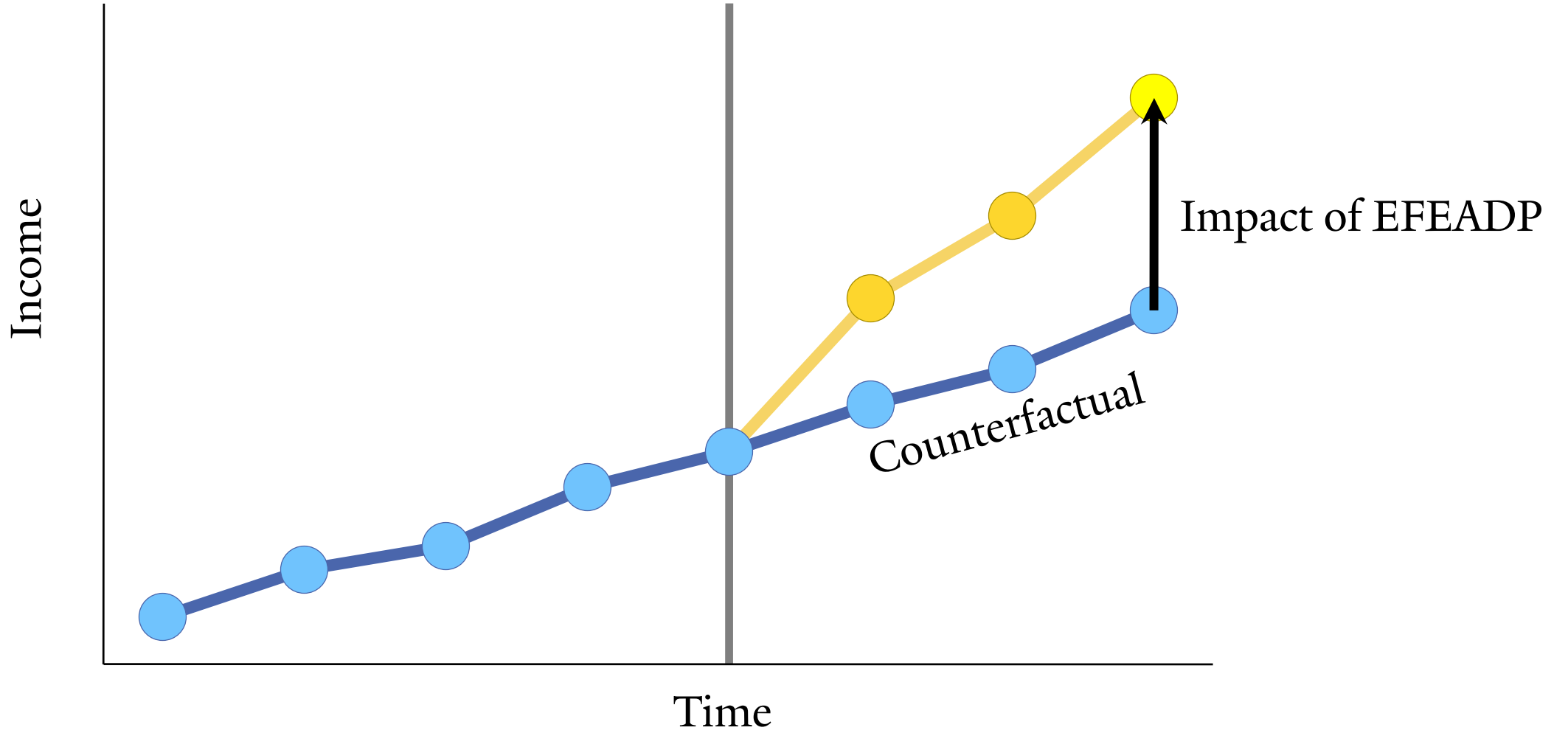
- Does treatment affect outcomes of people in the comparison group?

Yes

- Would the comparison group react to treatment the same way the treatment group does?

Yes

EFEADP Programme



Why is an RCT a good option to test EFEADP's effectiveness?

- We don't know if it will work in this setting
- Limited budget – want to test before scaling broadly
 - Could randomize access to EFEADP if too many women wanted into the program
- May want to test *features* of the intervention
- Could use in context of gradual roll-out

RCTs All Around Us

- *Randomized control trials*: Assignment to the intervention is random. Every eligible unit has the same probability of assignment as a comparable non-treated unit.
- Why does this give us a statistically comparable counterfactual group? In expectation, those assigned to “treatment” (intervention) and “control” (no intervention) are statistically indistinguishable.

Critiques of RCTs

- They can't answer some of the biggest questions in economics
 - Why do countries grow, does aid “work”, etc
- Direct attention (and money) to the little questions – “narrowness of scope”
- Statistical concerns: RCTs tell us about the mean of the treatment effects, but not the distribution of treatment effects (e.g. median), and data mining
- Internal v. external validity
- Excessive focus on what works rather than learning about the world – yet we all have theories of how the world works

These are well-reasoned critiques. What can you do to make sure your evaluations don't fall into these traps?

Anatomy of an RCT Focused on
Women's Economic
Empowerment

The Program: (MG)NREGA



- India's “Employment Guarantee”
- Established ~2007
- Every household → 100 days paid work annually
- Manual labor – ponds, ditches, roads, etc. – low wages
- Significant labor market opportunity for rural women
- 50 million households/year

The Problem

- Payments directly transferred to beneficiary- owned bank accounts
- Wage payments for all household members sent to bank account owned by male household head
- **Women have little-to-no wage control**

The Opportunity

- Central government supportive of female-friendly measures
 - Mandated electronic payments into bank accounts
- At the same time, state initiatives to increase financial inclusion in rural areas

The Partners

- Ministry of Rural Development
- State NREGA Commissioner and Panchayati Raj and Rural Development Department
- District leadership
- Two public sector banks
- Multiple kiosk banking companies



Madhya Pradesh

Panchayat & Rural Development Department

Research Questions (v1)

- Does financial inclusion alone empower women, encourage work under MGNREGA?
- Is additional training required for women to take advantage of access to formal financial services?
- Does giving women a financial use case through direct wage deposits increase further increase her financial engagement, economic activity, and broader empowerment?

Research Questions (v2)

- Can public policy that **strengthens women's income control** help norms-constrained women work more?
 - In NREGA
 - In the private sector
- Can this in turn **reshape social norms**?

Key Design Questions We Had to Address

- **Where** would we run the RCT?
- **Who** would be eligible to participate, and how would we identify potential participants?
- **At what level** would we randomize?
- **How** would we implement the intervention arms?
- **What data** would we collect to monitor implementation?
- **What data** would we collect to assess impact?
- **How** would we pay for all this?

Questions?

Please provide your
feedback!

<https://forms.gle/ZdtC2Xd9nmpAKUcaA>