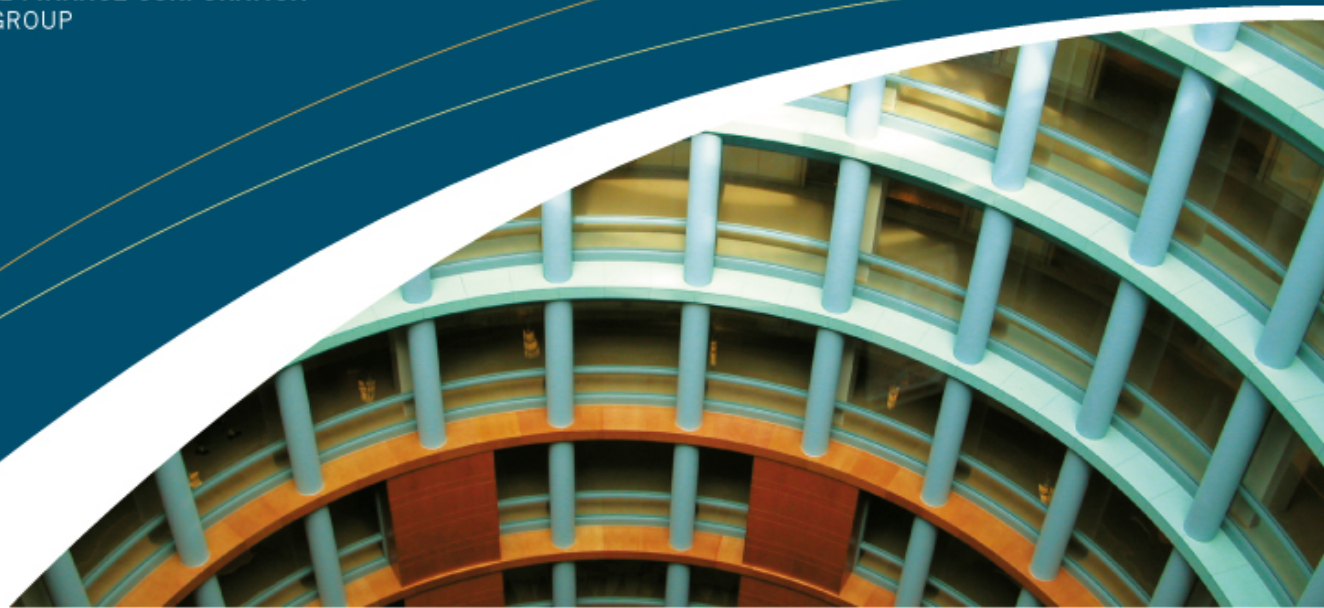




INTERNATIONAL FINANCE CORPORATION
WORLD BANK GROUP



Experimental & Quasi-experimental Design for Program Evaluation

Results Measurement Unit

Monitoring & Evaluation

Monitoring

What were the results?

- Set measurable goals
- Track progress
- Measure results

Evaluation

What's the program effect?

- Define effects in terms of “intervention A” vs. “intervention B”
- Compare results across approaches to determine what works best
- Best when designed before implementation

Ultimately, We Want to Know...

- What is the best mode or approach to TA?
- What was the “additionality” of our efforts?
- What was the attributable or causal effect?
- **So, for example...**
 - eg., which type of training program is best?
 - We have existing training situation (type “E”)
 - Reformed training (type “R”)

What is an experiment?

- A carefully controlled study
 - At least 2 groups
 - Random selection of participants to training/control
 - No systematic differences between the groups
- Goal is to identify program effects

And what is a quasi-experiment?

- When you can't do an experiment and yet want to infer causes and effects (typically ex-post)
 - At least 2 groups (Treatment and Control)
 - No Randomization
 - So what are we giving up over randomization?—
Stronger assumptions on the comparability of the control group
- ...Involves adjustment for selection bias due to lack of randomization

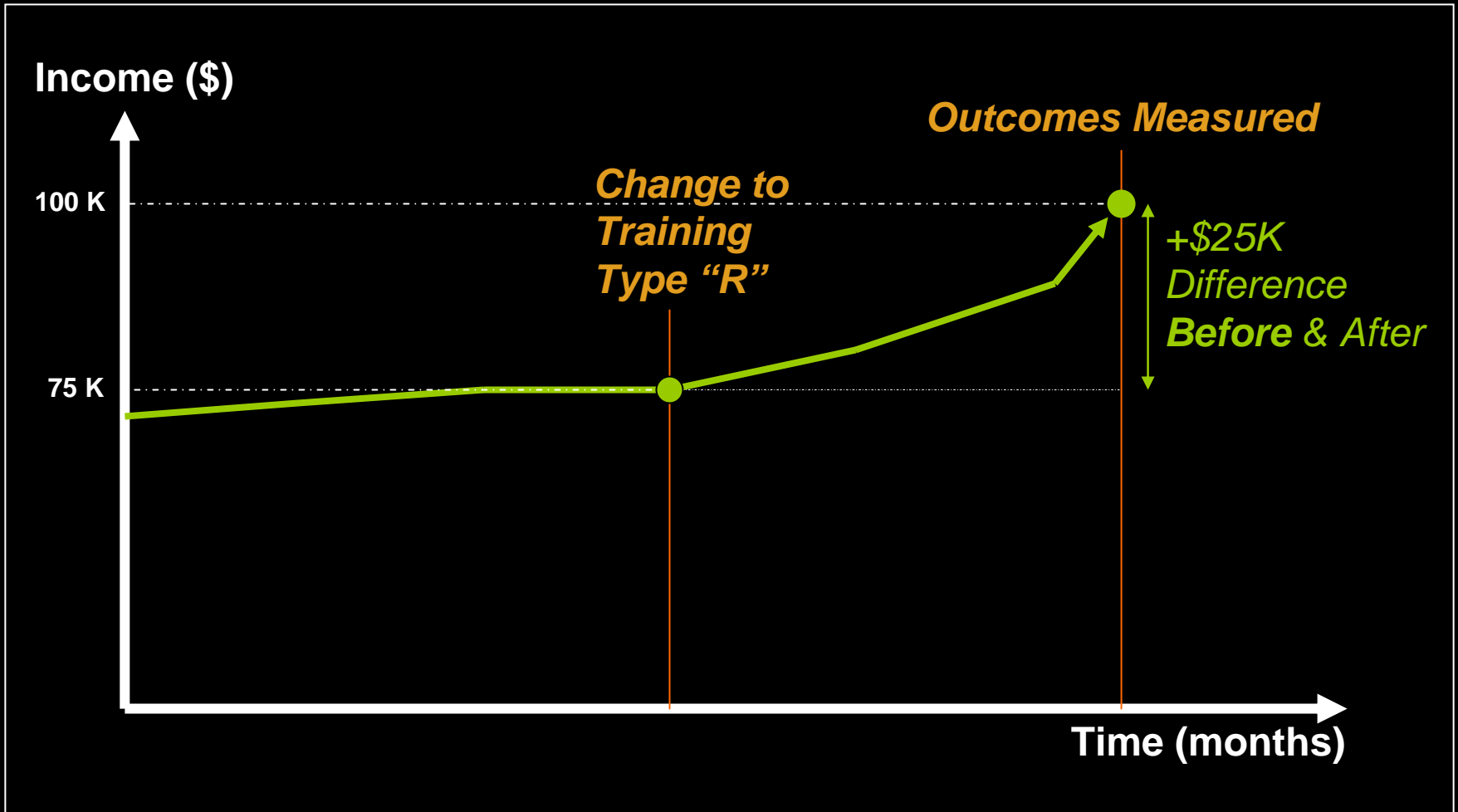
Are These Methods Useful to Evaluate TA?— Our Experience to Date

- Randomize training to seaweed farmers
- Randomize communication strategies to Indian rice growers
- Quasi-experimental design for Serbian firms receiving direct technical assistance
- Quasi-experimental design for Nigerian bank that receives assistance evaluating creditworthiness of female entrepreneurs

Illustrative Example: Which Type of Training is Best?

- You want the effect of training reform (type “R”) vs the pre-existing mode of training (“E”)
- Look at results for those who experienced “R”
- Consider what would have happened if they had continued with existing training “E”?
- Would they be better off, or worse off?

Effect of Training "R" on Income



The Effect of Training *R* vs. *E*

Very Simple Formula

Income observed after training type “R”

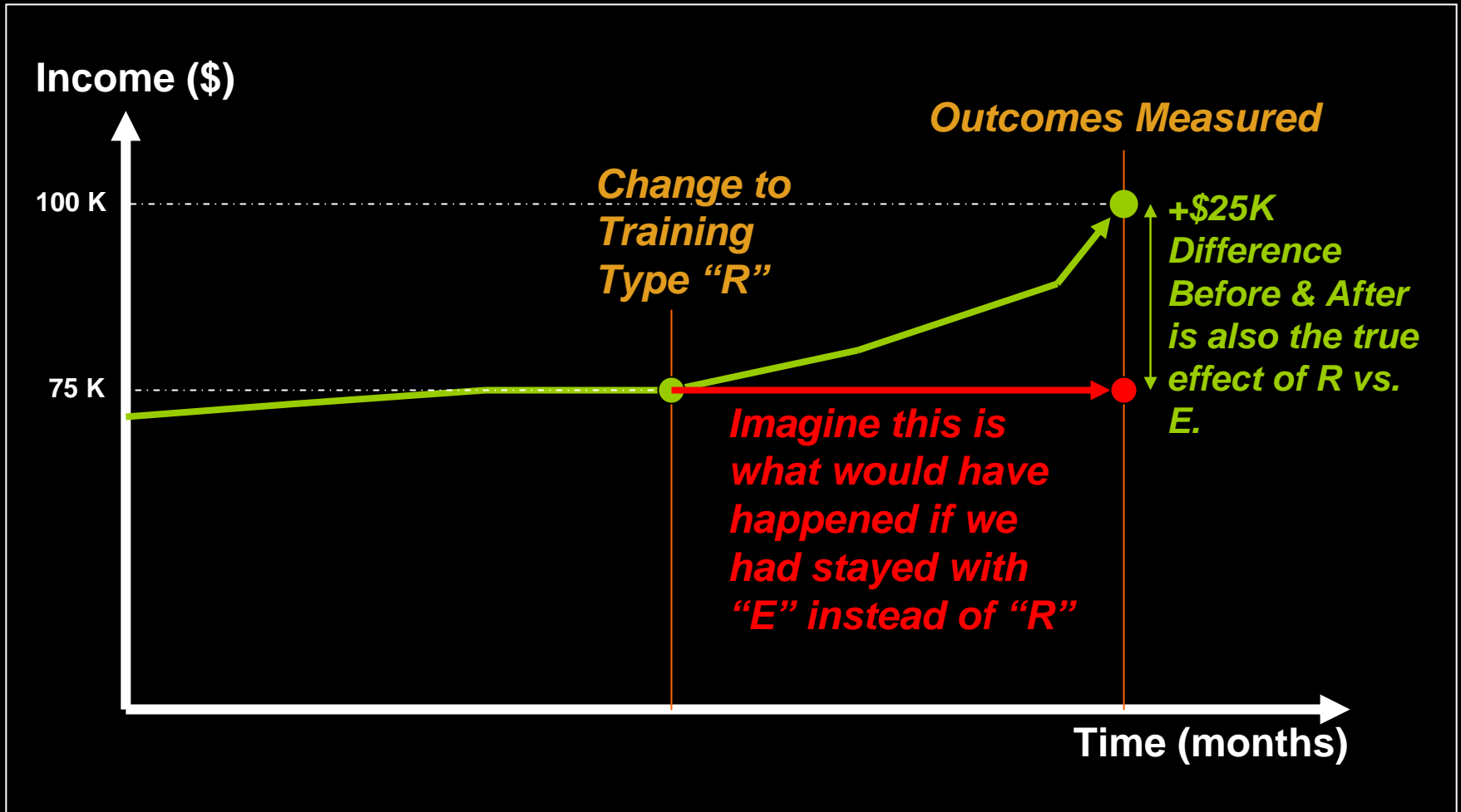
- Income that would have been with type “E”

Effect of “R” versus “E”

We do not observe results in the absence of the reform.

But with the right data and proper design, we can make a sensible inference, backed by evidence. For example...

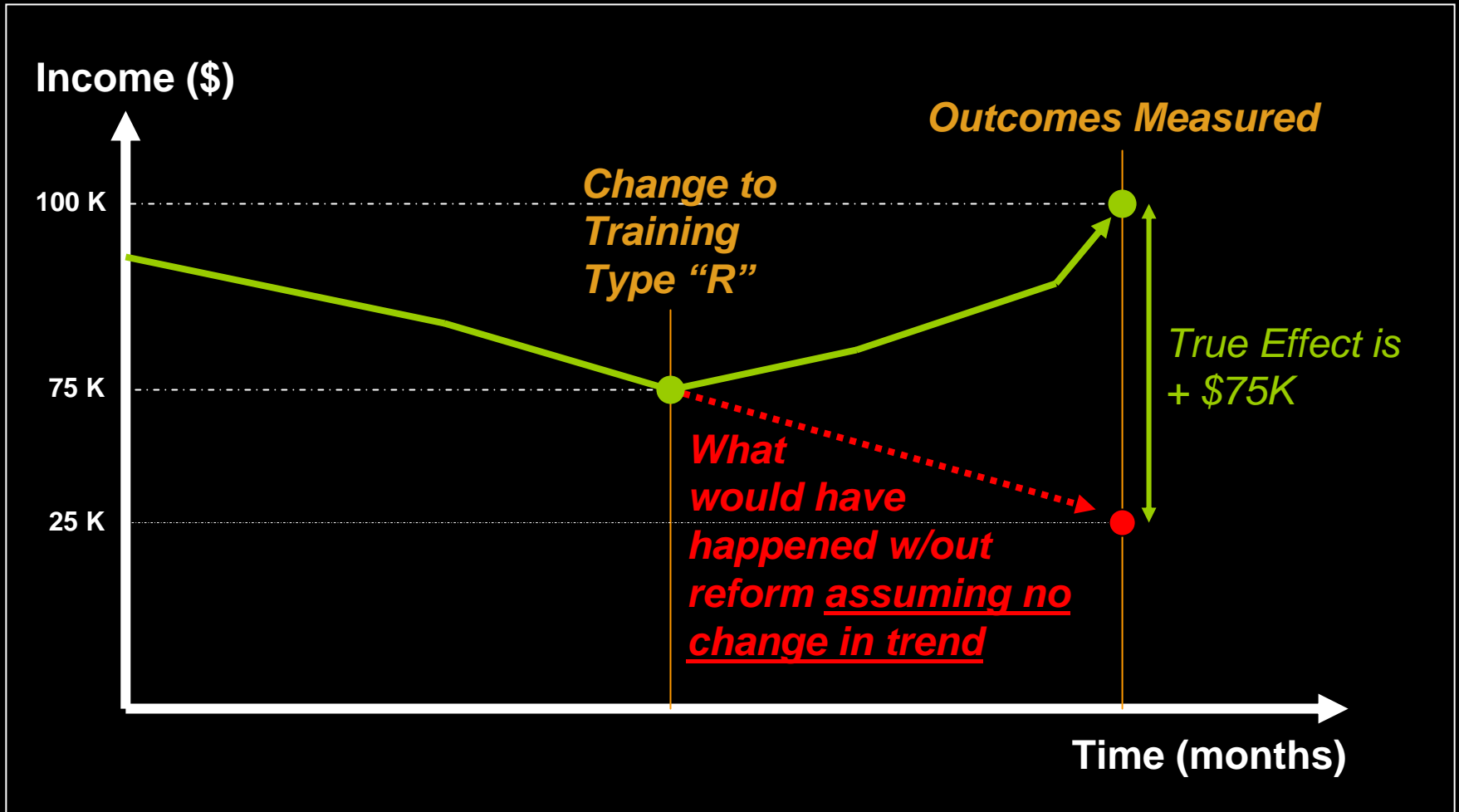
Effect of Training "R" on Income



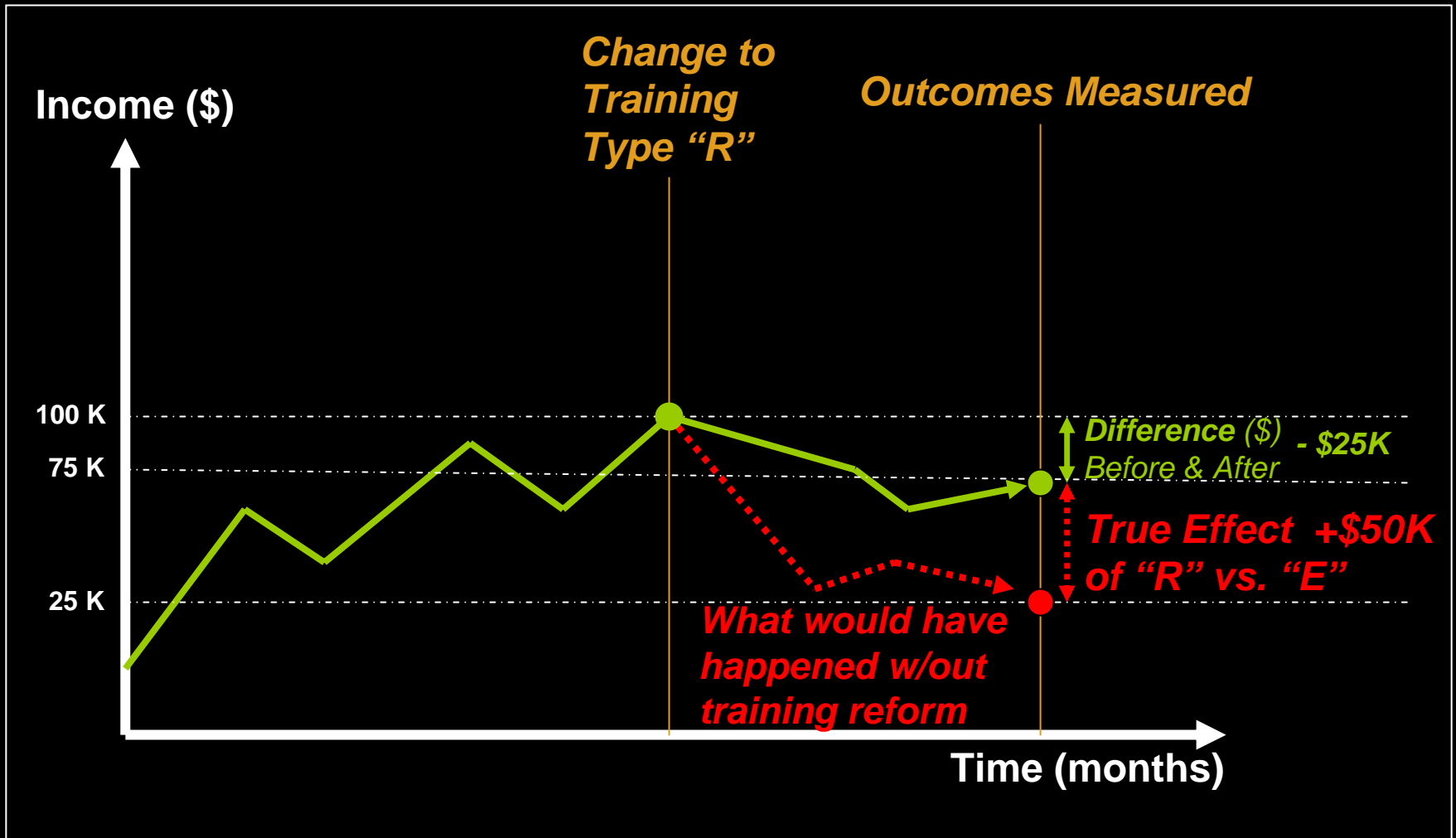
Key Points

- When is “*before vs. after*” the true effect?
 - When there would have been no change w/out the TA
- Can we do better than make assumptions?
 - There may be evidence we can gather to support this claim... typically look at evidence prior to the TA.
- Some assumption is always necessary. Why?

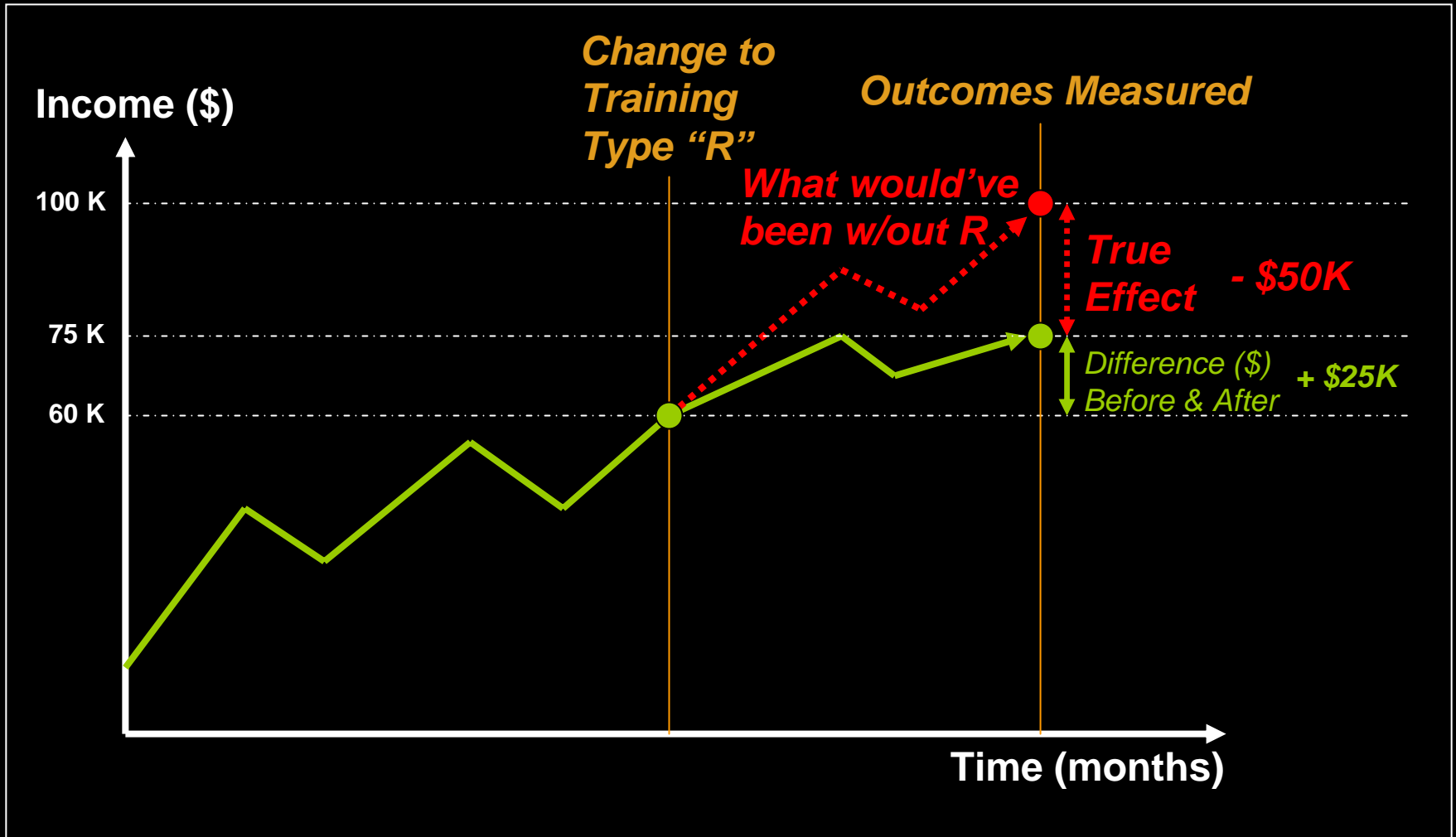
The Effect of Training: Assumptions About Outcomes



Case II: *It would've been worse...*



Case III: *It would've been better...*



Joys of Randomization

- Why do scientists randomize?
 - Eliminates systematic differences between groups
 - Otherwise different outcomes could be due to different underlying characteristics
- Which characteristics?
 - The variables that matter are the ones that relate to both treatment and outcomes
 - Consider a training example
 - Hair color may not matter; education probably would

If Not Experimental, Try Quasi-Experimental

- Even if you can't randomize for ethical or practical reasons...
 - You still want to conceptualize in terms of treatment
 - You still want the different treatment groups to be comparable prior to treatment
- Many alternatives to randomization (quasi)...
 - **Matching**: eg., For every SME with training, we identify a matching SME that did not
 - **Discontinuity** design: eg., credit scoring, look at SMEs just above and below loan approval threshold

Lots of Quasi-Experimental Choices

Consider a TA program that affects SMEs in a certain sector (strawberry farming).

- Maybe you have no control group of SMEs in affected sector?
 - *Differences-in-Differences strategy*
 - Compare against those of SMEs in another sector w/ similar pre-TA growth rates (eg., blackberries)

Lots of Quasi-Experimental Choices

Consider a Business Enabling Environment TA Program which simplifies corporate registration procedures—everyone is eligible to “take treatment” and register.

- How can we randomize?
 - Randomize encouragement via a public information campaign: only some get the info

Summing Up

- These tools can be extremely useful for finding out which modes of implementation work best
- Requires careful planning, but **NOT EXPENSIVE**
 - Consideration of assumptions
 - Collecting data from some other group
 - Sometimes *unaffected group*—or just *differently affected*
- Lots of methodological options to choose from
 - Randomization (fewest assumptions)
 - Quasi-experimental (more flexibility)