# Course Title: Quantitative Techniques for Economics
# Course code: ECON6002
# Topic: Impact Evaluation
## *Ph.D. Economics (1st Semester)*


**Dr. Kailash Chandra Pradhan**

**Mahatma Gandhi Central University,**
**Department of Economics**

# Impact Evaluation

- Impact evaluation is the difference between outcomes with the program and without it
- Estimate the *CAUSAL* effect (impact) of program ***P*** on labour market outcome ***Y***
  - For Example: What is the effect of a *job training* ***(P)*** on *employability and labour earnings* ***(Y)***?
- Causal effect:

  Impact = Outcome if in program – Outcome if not in program

# The counterfactuals

- Impact of program (P) as the difference in outcome (Y) for the same individual with and without participation in a program
- The person cannot be observed simultaneously in two different states (in other words, with and without the programs). This is called "the counterfactual problem".
- Hence, we have a problem of a missing counter-factual, a problem of missing data
- The counterfactual is what the outcome (Y) would have been in the absence of a program (P)
- We need a control/comparison group that will allow us to attribute any change in the "treatment" group to the program
- In practice, a key goal of an impact evaluation is to identify a group of program participants (the treatment group) and a group of nonparticipants (the comparison group) that are statistical identically identical in the absence of the program. For example- the average age of in the treatment group should be the same as the average age in the comparison group

# Impact Evaluation Methods

## 1. Randomization

- Sample of Individuals/communities/firms are selected randomly from the population. The sample should be representative of the populations within a certain sampling error
- The targeted population has been defined (say, households below the poverty line, or children under the age of 5, or schools in rural areas). The randomised sample selection procedure allows that every eligible person or unit has the same chance of receiving the program
- Let the treatment, *Ti* , be equal to 1 if subject *i* is treated and 0 if not. Let *Yi*(1) be the outcome under treatment and *Yi*(0) if there is no treatment.
- Where, $Y_i = [T_i.Y_i(1) + (1 - T_i).Y_i(0)]$ .The treatment effect for unit *i* is $Y_i(1) - Y_i(0)$ and the ATE is $ATE = E[Y_i(1) - Y_i(0)]$
- In terms of a regression, This can be expressed as

$$Y_i = \alpha + \beta T_i + \varepsilon_i$$

- where *Ti* is the treatment dummy equal to 1 if unit *i* is randomly treated and 0 otherwise

# Impact Evaluation Methods

## 1. Randomization cntd..

**Advantages**

- Most robust impact evaluation method, quite straight-forward
- Analytically simple (impact = difference in average outcomes)
- fair process of allocation with limited resources

**Disadvantages**

- Requires comparison group to be excluded from the program for duration of impact evaluation
- May be politically more difficult
- Usually run controlled experiment on a pilot, small scale. Difficult to extrapolate the results to a larger population

## 2. Propensity score matching (PSM)

- Each program participant is paired with one or more non-participant that are similar based on observable characteristics
- PSM assumes a conditional independence (namely, that unobserved factors do not affect participation)
- The average treatment effect of the program is calculated as the mean difference in outcomes across two groups.

# Impact Evaluation Methods

## 2. Propensity score matching (PSM) cntd..

**Advantages**

- PSM is a useful approach when only observed characteristics are believed to affect program participation
- Does not require randomization, nor baseline (pre-intervention data)

**Disadvantages**

- Requires very good quality data: need to control for all factors that influence program placement
- Requires significantly large sample size to generate comparison group

# Impact Evaluation Methods

## 3. Double Difference (DD)

- The DD estimator compares the participants and nonparticipants before and after the intervention
- The difference is calculated between the observed mean outcomes for the treatment and control groups before and after the program intervention
- DD methods, compared with PSM, assume that unobserved heterogeneity in participation is present. But such factors are time invariant. This fixed component can be differenced out
- The preceding two-period model can be generalized with multiple time periods, which may be called the *panel fixed-effects model*. This possibility is particularly important for a model that controls not only for the unobserved time-invariant heterogeneity but also for heterogeneity in observed characteristics over a multiple-period setting

# Impact Evaluation Methods

## 3. Double Difference (DD) cntd..

- More specifically, $Y_{it}$ can be regressed on $T_{it}$, a range of time-varying covariates $X_{it}$ and unobserved time-invariant individual heterogeneity $\eta_t$ that may be correlated with both the treatment and other unobserved characteristics $\varepsilon_{it}$. The equation can be written as

$$Y_{it} = \phi T_{it} + \delta X_{it} + \eta_i + \varepsilon_{it}$$

Differencing both the right-and left-hand side of the equation over time, the differenced equation can written as

$$(Y_{it} - Y_{it-1}) = \phi(T_{it} - T_{it-1}) + \delta(X_{it} - X_{it-1}) + (\eta_i - \eta_i) + (\varepsilon_{it} - \varepsilon_{it-1})$$

$$\Rightarrow \Delta Y_{it} = \phi \Delta T_{it} + \delta \Delta X_{it} + \Delta \varepsilon_{it}$$

- Endogeneity (that is, the unobserved individual characteristics $\eta_t$) is dropped from differencing, ordinary least squares (OLS) can be applied to equation to estimate the unbiased effect of the program ($\phi$).

# Impact Evaluation Methods

## 3. Double Difference (DD) cntd..

- With two time periods, ($\phi$) is equivalent to the DD estimate, controlling for the same covariates $X_{it,}$ the standard errors, however, it may needed to be corrected for serial correlation. With more than two time periods, the estimate of the program impact will diverge from DD.
- The double-difference approach described in the previous section yields consistent estimates of project impacts if unobserved community and individual heterogeneity are time invariant
- If comparison areas are not similar to potential participants in terms of observed and unobserved characteristics, then changes in the outcome over time may be a function of this difference. This factor would also bias the DD.
- DD approach might falter if macroeconomic changes during the program affected the two groups differently. In this case, a simple DD might overestimate or underestimate the true effects of a program depending on how the treated and nontreated groups react to the common shock.

# Impact Evaluation Methods

## 4. Instrumental Variable (IV) Estimation

- The IV approach relax the exogeneity assumption of OLS or PSM and that are also robust to time-varying selection bias, unlike DD
- In this approach, it finds a variable (or instrument) that is highly correlated with program participation but that is not correlated with unobserved characteristics affecting outcomes.
- This can be written as

$$Y_i = \alpha X_i + \beta T_i + \varepsilon_i$$

- One needs to find an instrumental variable, denoted *Z*, that satisfies the following conditions:

  1. Correlated with *T*: cov(*Z*,*T* )≠0
  2. Uncorrelated with ε: cov(*Z*, ε) = 0

# Impact Evaluation Methods

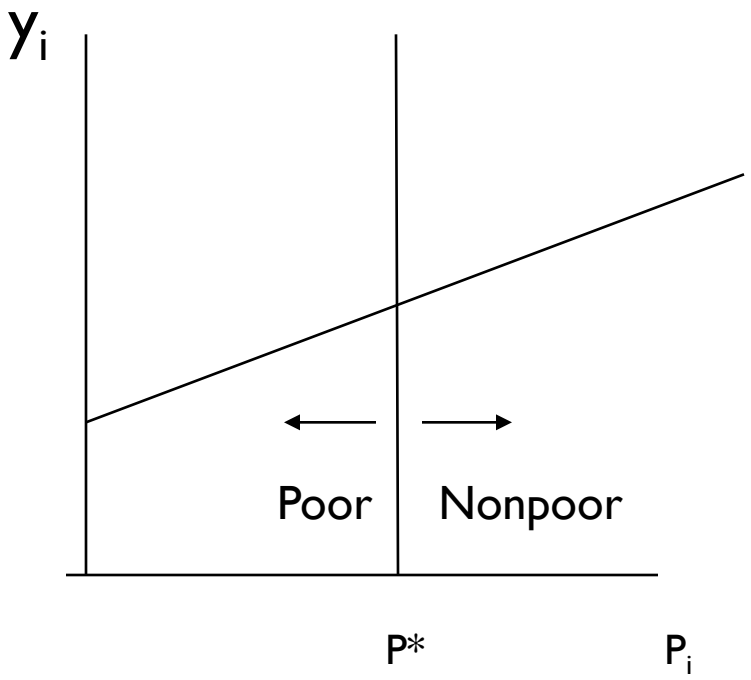
## 5. Regression discontinuity design (RDD)

- The RDD is an impact evaluation method that can be used for the programs that have a continuous eligibility index with a clearly defined cutoff score to determine who is eligible and who is not.
- A continuous eligibility, in other words, a continuous measure on which the population of interest can be ranked, such as poverty index, a test score, or age.
- For example, households with a poverty index score less than 50 out of 100 might be classified as poor; individual age 67 and older might be classified as pensioners; and students with a test score of 90 or more out of 100 might be eligible for a scholarship. The cutoff score in these examples are 50, 67, and 90, respectively
- It does not require randomization of any kind, so can be politically acceptable
- Impact estimates are valid only for the group near the cutoff and cannot be generalized to others

# Impact Evaluation Methods

## 5. Regression discontinuity design (RDD) cntd..

*Outcomes before Program Intervention*



# Impact Evaluation Methods

## 5. Regression discontinuity design (RDD) cntd..

*Outcomes after Program Intervention*