COUNTERFACTUAL EVALUATION

SIMONE BUSETTI
sbusetti@unite.it

CONTENTS OF THIS CLASS

- The problem of causal inference
- Experimental methods
- Quasi-experimental methods

THE PROBLEM OF CAUSAL INFERENCE

THE GOALS OF COUNTERFACTUAL EVALUATION

ENSURING CAUSAL ATTRIBUTION:

- Whether outcomes are to be attributed to program actions and outputs or whether they depend on other variables unrelated to the program

MEASURING A NET EFFECT:

- How much change in the outcomes is produced by the program?

THE NET EFFECT OF THE PROGRAM IS NOT:

THE OUTPUTS

How much was spent, how many scholarships funded, how many meals delivered, how many classes given, etc.

These measures concern what I'm doing with the program, not the change it produces.

THE RAW OUTCOMES

N people with a job, N people food secure, N people with certain skills, etc.

These measures do not tell me whether these results are good or not, nor do they tell me about change: what would have happened without the program.

THE BEFORE AND AFTER DIFFERENCE

N of people who increased their status, N of people who became employed, N of people who now master skills they did not, etc.

The problem is how to disentangle the effects of spontaneous dynamics: things also change without the program.

THE SIMPLE DIFFERENCE BETWEEN TREATED AND UNTREATED

These measures have a problem: people are different. If I enrol in a particular program, I'm different from those who don't.

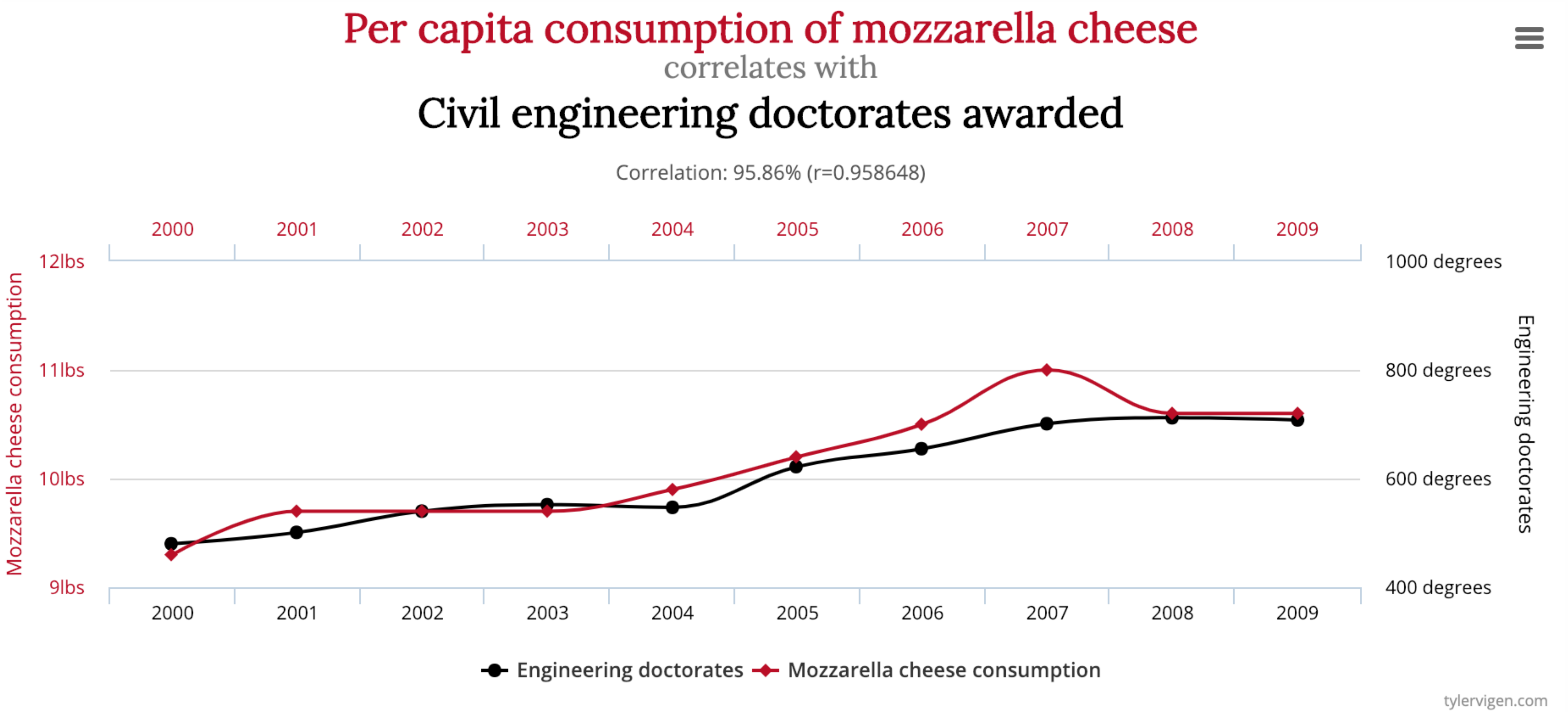
COUNTERFACTUALS AND CAUSAL INFERENCE

COUNTERFACTUAL:

What would have happened without the program?

THE PROBLEM OF CAUSAL INFERENCE

One should measure what is observed after the intervention and what would have happened without the intervention on the same population

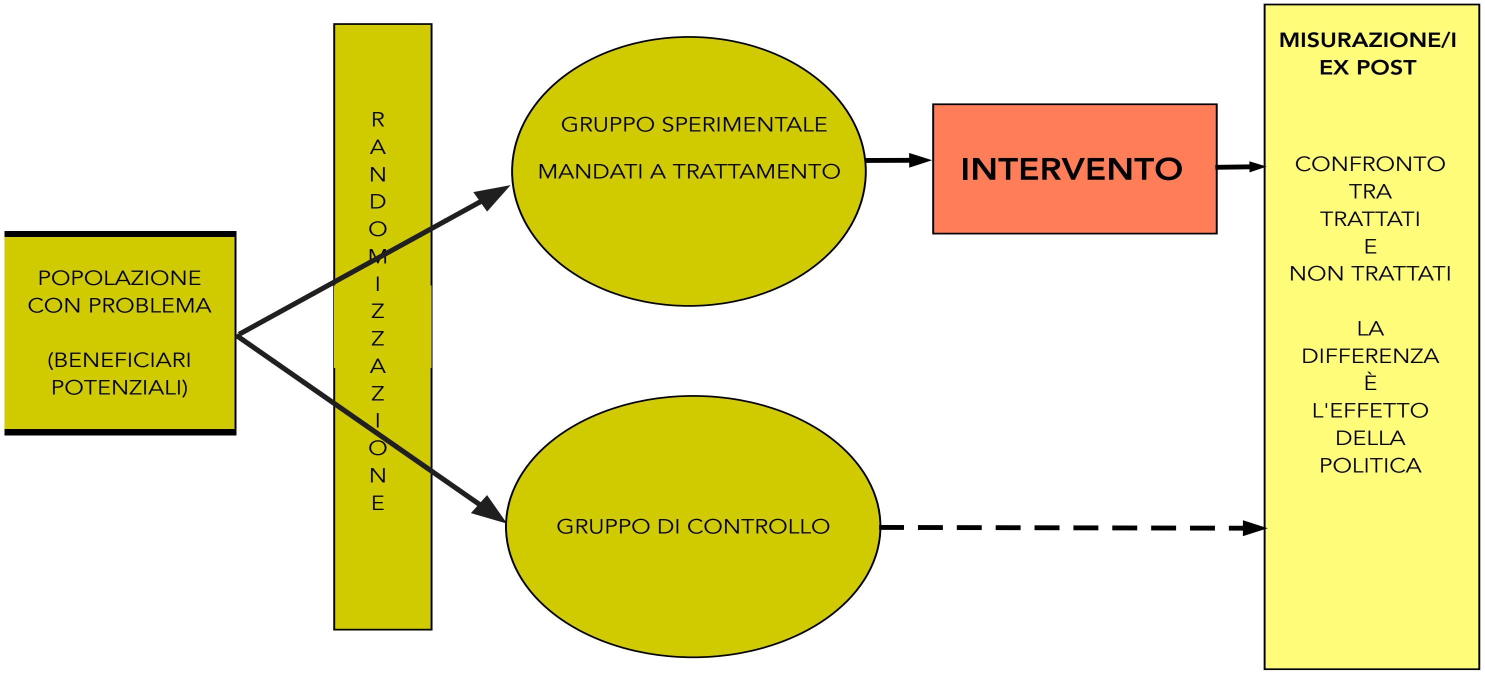


ta sources: U.S. Department of Agriculture and National Science Foundation

EXPERIMENTAL METHODS

EXPERIMENTAL METHODS

A potential target population is randomly assigned to either a treatment group (the ones subject to the program) or control group (without the program)



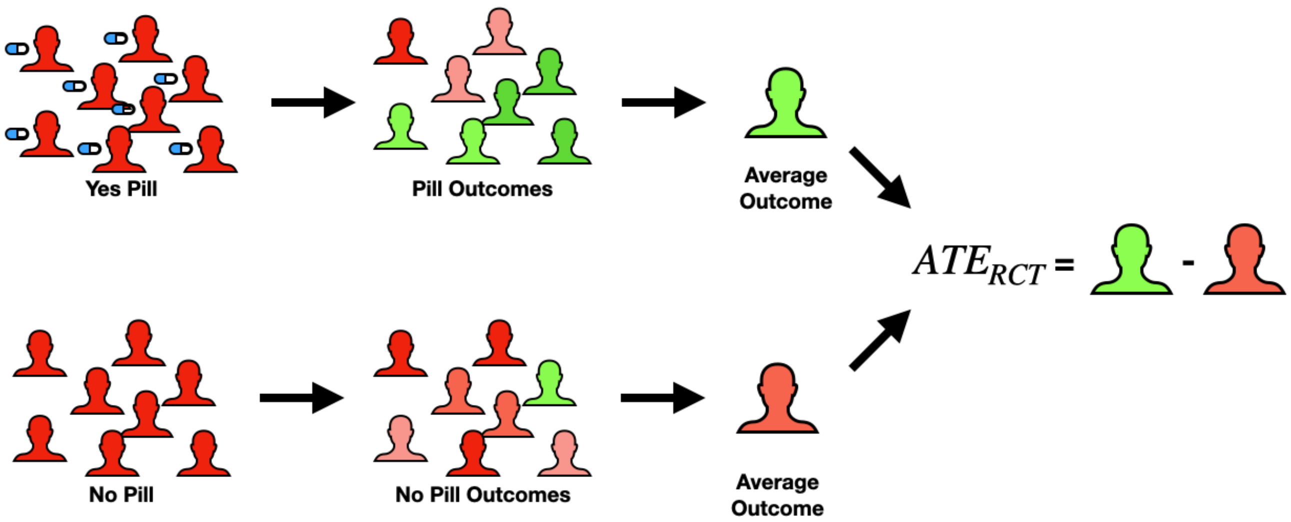
RANDOMISATION

Randomising the people in the treatment and control groups allows for their statistical equivalence:

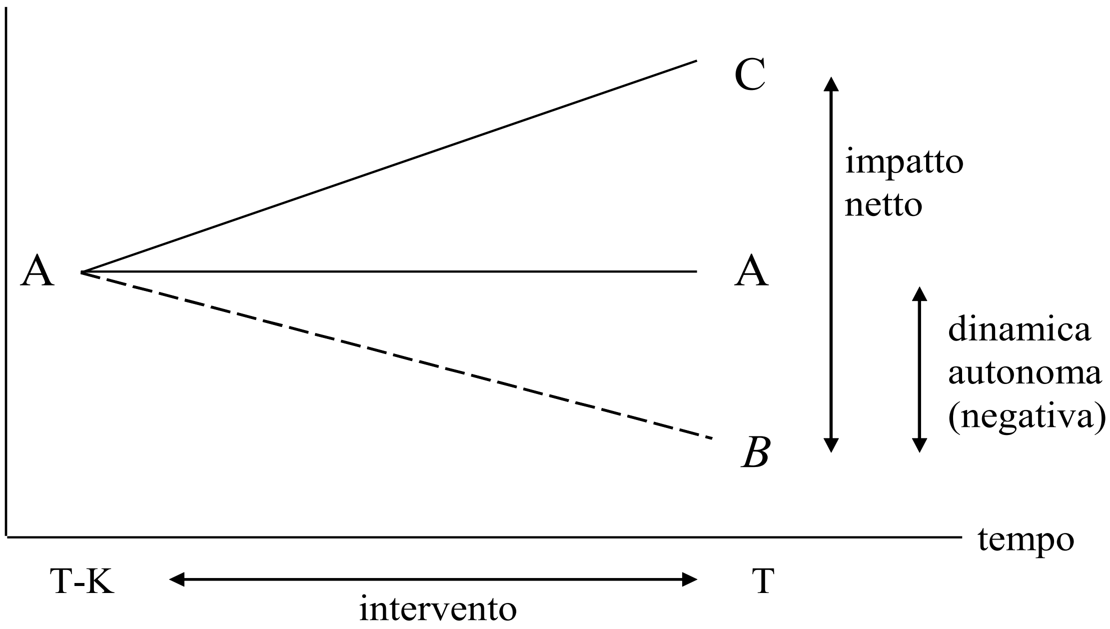
- The two groups will have similar characteristics, both in the observable and unobservable features
- There will be no differences at the start: they will be the same BUT for the treatment
- Every difference observed ex post will be due to the program

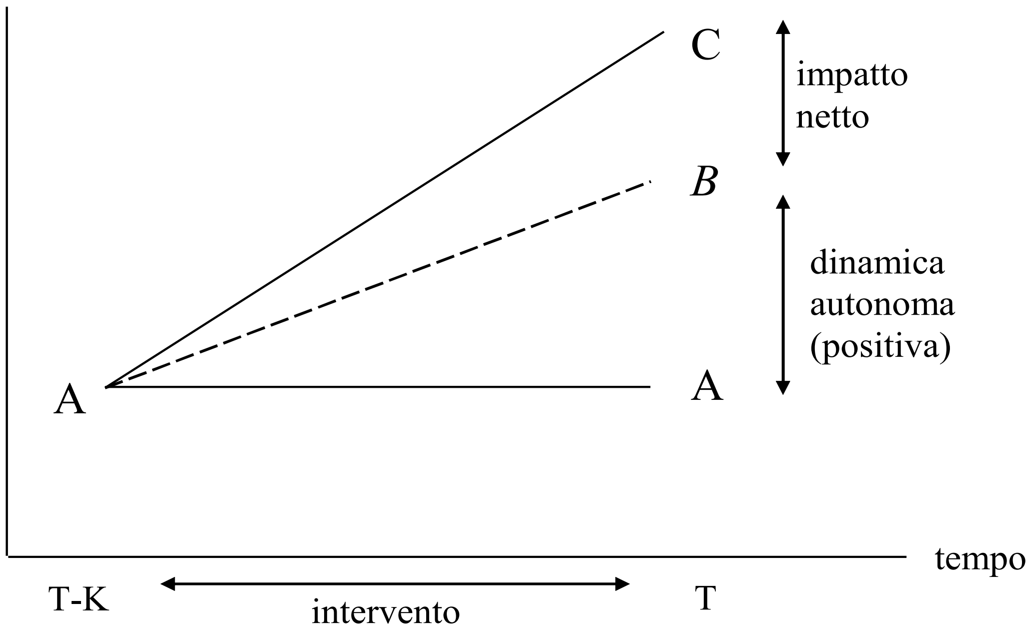
THE AVERAGE TREATMENT EFFECT (ATE)

The average treatment effect (ATE) is a measure in statistics and econometrics that estimates the average causal effect of a treatment or intervention on a population by finding the average difference in outcomes between the treated and untreated groups.



TWO EXAMPLES





USING EXPERIMENTAL EVALUATION

- The evaluation should start BEFORE the program starts
- The numbers should be high to ensure statistical equivalence
- There should be no interference between the two groups
- You need to ensure implementation fidelity so that everyone is treated the same
- In policy programs, you will not have a neat treatment-placebo comparison like in clinical trials
- There can be ethical barriers to treating similar individuals differently
- It is the gold standard in internal validity but it has problems in external validity

For these reasons, experimental evaluation is typically used for pilot interventions

QUASI-EXPERIMENTAL METHODS

QUASI-EXPERIMENTAL METHODS

«An impact evaluation is only as good as the comparison group it uses to mimic the counterfactual»

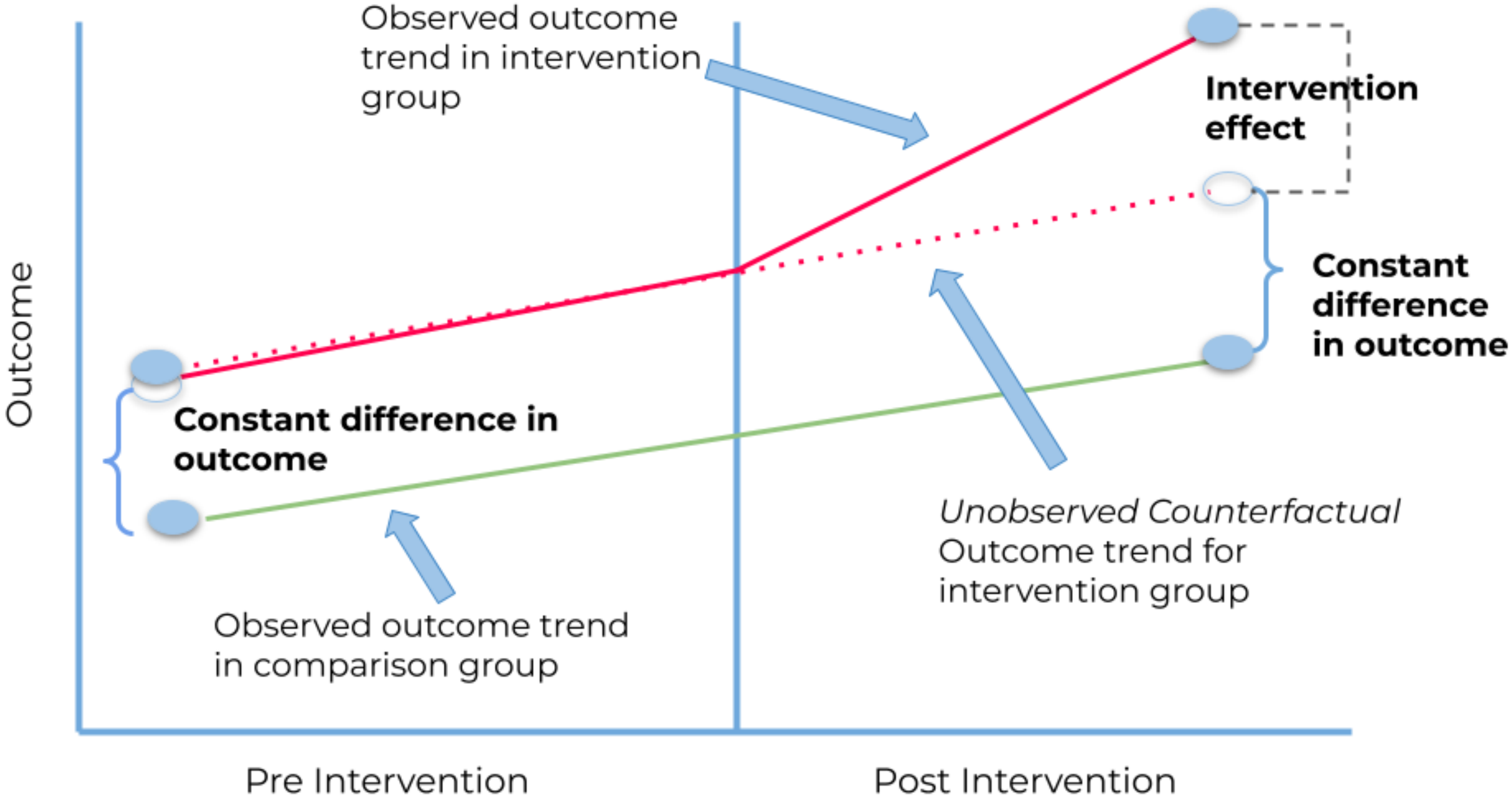
Glennerster & Takavarasha 2013

- The evaluation does not use randomisation but establish a plausible control group.
- It can start after the program and can use 'observational' data.
- It can be used on periodic programs

Several techniques are available:

- Difference in difference
- Matching
- Interrupted time-series
- Etc.

DIFFERENCE IN DIFFERENCE



DIFFERENCE IN DIFFERENCE