

# Introduction in Programme Evaluation

## 1.1 Introduction

The aim of this chapter is to give an introduction into microeconomic evaluation. Empirical microeconomic evaluation is conducted with individual data. The main question is if the outcome variable of interest for an individual is affected by the participation in an ALMP programme or not. Relevant outcome variables can be for example the future employment probability or the future earnings. In any case, we would like to know the difference between the value of the participant's outcome in the actual situation and the value of the outcome if he had not participated in the programme. The fundamental evaluation problem arises because we never observe both states (participation and non-participation) for the same individual at the same time, i.e. one of the states is counterfactual. Therefore finding an adequate control group is necessary to make a comparison possible.<sup>1</sup> This is not an easy task because participants in programmes usually differ in more aspects than just participation from non-participants. Taking simply the difference between their outcomes after treatment will not reveal the true treatment impact, i.e. will lead to a selection bias.

Depending on the data at hand, different evaluation strategies can be thought of. Whenever feasible, experimental evaluation will provide the most compelling evidence in most cases. The basic idea of this approach is to assign individuals randomly to the participant's and the control group. Both groups then differ only with respect to participation and the differences in the outcomes can be taken as treatment effects. Although this approach seems to be very appealing in providing a simple solution to the fundamental evaluation problem, there are also some problems associated with it which we will

---

<sup>1</sup> The terms control group and comparison group will be used interchangeably throughout the book. Either way the group consists of individuals who did not receive treatment.

discuss later on. More importantly, in most European countries experimental data are not available and therefore, the evaluator must choose among non-experimental evaluation estimators.

Whereas in the early stages of treatment evaluation, some analysts like LaLonde (1986) or Ashenfelter and Card (1985) viewed social experiments as the only valid evaluation method, a lot of methodological progress has been made to develop and justify non-experimental evaluation estimators which are based on econometric and statistical methods to solve the fundamental evaluation problem (see e.g. Heckman and Robb (1985b), Heckman and Hotz (1989) or Heckman, LaLonde, and Smith (1999)). They belong now to the standard toolbox of evaluation research. In non-experimental or observational studies, the data are not derived in a process that is completely under the control of the researcher. Instead one has to rely on information how individuals actually performed after the intervention. That is, we observe the outcome with treatment for participants and the outcome without treatment for non-participants. The objective of observational studies is to use this information to restore the comparability of both groups by design. To do so, more or less plausible identification assumptions have to be imposed. We will show that different strategies invoke different identifying assumptions and also require different kinds of data for their implementation.

The different estimators can be classified with respect to two dimensions. The first dimension is the required data for their implementation, where we can distinguish between longitudinal and cross-sectional methods. The second dimension concerns the handling of selection bias, where two categories arise. The first category contains approaches that rely on the so-called unconfoundedness or selection on observables assumption which we will present in section 1.5. The basic idea here is, that based on some observed characteristics the potential outcomes are independent of the assignment to treatment. In that case, the control group with similar characteristics can be used to construct the missing counterfactual outcome. We will present and compare two strategies in this category, namely matching and regression analysis. Clearly, the performance of these estimators depends on the data at hand. For their justification we need a rich dataset that contains all variables that jointly influence the participation decision and the outcome variable.

If one believes that the available data is not rich enough to justify the selection on observables assumption, he has to rely on the second category of estimators which explicitly allows selection on unobservables, too. The methods of instrumental variables and selection models will be presented for that situation. With selection models one tries to model the selection decision completely whereas IV methods focus on finding an instrument which determines participation but does not influence the outcome.

The main part of the conclusions to this chapter will deal with the question on how to choose between different evaluation estimators and what (data) requirements are necessary for their implementation. By doing so, we hope to give some advice for future evaluation analysis. An important topic which has