

## Some Practical Guidance for the Implementation of Propensity Score Matching Estimators

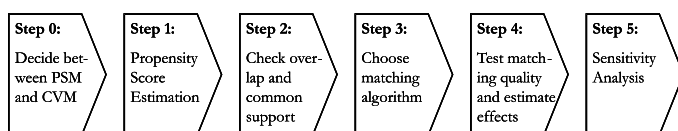
### 3.1 Introduction

In the previous chapter we have presented several different matching estimators. We have also discussed their advantages and disadvantages when used in large and finite samples and we have given some guidance to decide which matching algorithm to choose. It has become clear that exact covariate matching is unfeasible in case of a high dimensional vector of covariates and that propensity score matching (PSM) is a good alternative for such situations.

Once the researcher has decided to use PSM, he is confronted with a lot of questions regarding its implementation. To begin with, a first decision has to be made concerning the estimation of the propensity score. One has not only to decide about the probability model to be used for estimation, but also about the variables which should be included in this model. Following that one has to determine how to check the overlap between treatment and control group and how to implement the common support requirement (see the discussion in section 1.5). Based on these findings and on the trade-offs in terms of bias and efficiency (see section 2.7), one has to decide which matching estimator to choose, e.g. nearest-neighbour or kernel matching. Subsequently, one has to test the matching quality, estimate standard errors and the treatment effects. Additionally, it might be the case that the researcher wants to test the sensitivity of his results with respect to ‘hidden bias’. ‘Hidden bias’ might arise if treatment and control group differ on unobserved variables which simultaneously affect assignment into treatment and the outcome variable. In that case, individuals who look similar in terms of observed covariates may have very different probabilities of receiving treatment. Since it is not possible to estimate the magnitude of ‘hidden bias’ with non-experimental data, we address this problem with a bounding approach suggested by Rosenbaum (2002). Figure 3.1 summarises the necessary steps when implementing PSM. Some of the steps (e.g. the decision between covariate and propensity score matching and the choice of the matching algorithm) have been already discussed in chapter

2. The aim of this chapter is to discuss the remaining implementation issues and give some guidance to researchers who want to use PSM for evaluation purposes, and are confronted with the above mentioned decisions. The chapter is organised as follows. In section 3.2 we will focus on the implementation of PSM estimators. We will start with the estimation of the propensity score (subsection 3.2.1), before we discuss how to assess the matching quality in subsection 3.2.2. Overlap and common support will be the topics of subsection 3.2.3. After that we present the problem of choice-based sampling and discuss the question when to measure programme effects. Section 3.3 will be concerned with the sensitivity of the estimates with respect to ‘hidden bias’.

**Fig. 3.1:** Implementing Propensity Score Matching<sup>1</sup>



<sup>1</sup> CVM = Covariate Matching, PSM = Propensity Score Matching

Section 3.4 deals with more practical issues, like programme heterogeneity (3.4.1), multiple participation of individuals (3.4.2) and the choice of the right control group (3.4.3). Programme heterogeneity may become an issue when evaluating ALMP of countries which do not have only one homogeneous programme, but a variety of different ones, e.g. job creation schemes, vocational training and rehabilitation programmes. It should also be regarded when evaluating one programme with varying characteristics. This is important for our empirical analysis in chapter 6 where we analyse the effects of job creation schemes, which can be operated in very different sectors like AGRICULTURE, CONSTRUCTION & INDUSTRY or OFFICE & SERVICES. It is very likely that the effects differ depending on these sectoral differences and hence, programme heterogeneity has to be considered. Multiple participation has to be taken into account if it can be assumed that individuals participate more than once in (possibly different) programmes in the time span under consideration. In that case we have to move from the static framework to a dynamic one. Finally, choosing the right control group has become an important topic in the last years. One can either choose to define the control group as persons who never participate in the analysed programme or who do not participate until a certain point in time. Finally, section 3.5 concludes by outlining the arising issues when implementing PSM and highlighting our answers once again.