Available online at www.sciencedirect.com
ScienceDirect
Journal of Econometrics 142 (2008) 715–730
JOURNAL OF Econometrics
www.elsevier.com/locate/jeconom

Ineligibles and eligible non-participants as a double comparison group in regression-discontinuity designs

Erich Battistin[a,b,*], Enrico Rettore[a]

[a]*Department of Statistics, University of Padova, Italy*
[b]*Institute for Fiscal Studies, London, UK*

Available online 21 May 2007

Abstract

In a sharp regression-discontinuity design (RDD) the participation status deterministically depends on whether a pre-programme characteristic is above or below a specified threshold. The attractiveness of such a design rests on close similarities with a formal experiment. Nevertheless, it is of limited applicability since participation into a programme is seldom determined according to this rule. Besides, in the presence of heterogeneous effects a sharp RDD only allows identification of mean effects for individuals around the threshold for participation. Two results are presented in this paper, and they both partially overcome the two limitations described above. We show that when individuals self-select into participation conditional on some eligibility criteria a sharp RDD provides a natural framework to define a specification test for the non-experimental estimation of programme effects for participants away from the threshold. We also show that, in this set-up, the regularity conditions required for the identification of the mean counterfactual outcome for participants marginally eligible for the programme are essentially the same as in a sharp RDD.

JEL classification: C4; C8

Keywords: Programme evaluation; Second comparison group; Specification tests; Treatment effects

1. Introduction

In this paper we show how discontinuities in the probability of participation induced by the eligibility criteria for being enrolled in a programme can be used to test the performance of alternative non-experimental estimators of the programme effects.

The central issue in the evaluation of the impact of an intervention is to separate its causal effect from the confounding effect of other factors influencing the outcome of interest. Random assignment of individuals to the intervention defines treatment and control groups that are equivalent in all respects, except for their treatment status. Thus, if a randomized experiment is well implemented, any post-intervention difference between treatment and control individuals can safely be attributed to the intervention itself. The main

*Corresponding author. Department of Statistics, University of Padova, Italy. Tel.: +39 049 8274165; fax: +39 049 8274170.
E-mail address: erich.battistin@unipd.it (E. Battistin).

0304-4076/$ - see front matter © 2007 Elsevier B.V. All rights reserved.
doi:10.1016/j.jeconom.2007.05.006

advantage of experiments is that the assumptions they rest upon are generally more plausible than those made in an observational setting (see Heckman and Smith, 1995). In the latter case, the identification of causal relationships rests on assumptions about individuals' behaviour whose plausibility is often controversial.

In the absence of random assignment a fairly favourable situation to the researcher arises when the assignment mechanism leading individuals to participate into the programme is fully specified. This situation applies to those instances in which participation follows a *sharp* regression discontinuity design (RDD; see Thistlethwaite and Campbell, 1960; Rubin, 1977; Trochim, 1984). In this design, assignment to the programme solely depends on whether one or more observable pre-intervention variables satisfy a set of conditions *known* to the analyst. As an example, we can think of situations in which individuals willing to participate are divided into two groups according to whether or not a pre-intervention measure exceeds a known threshold, but only individuals scoring above that threshold are assigned to the programme.

This design features both advantages and disadvantages with respect to its non-experimental competitors. On the one hand, in a neighbourhood of the threshold for selection a sharp RDD presents some features of a pure experiment. In this sense, it is certainly more attractive than other non-experimental designs. Since individuals in the treatment and comparison group solely differ with respect to the variable determining the participation status (and with respect to any other variable correlated to it), one can control for confounding factors by contrasting marginal participants to marginal non-participants. In this context, the term 'marginal' refers to those units *not too far* away from the threshold for selection.

The comparison of mean outcomes for participants and non-participants at the margin identifies the mean impact of the intervention *locally* at the threshold for selection. Intuitively, for identification at the cut-off point to hold it must be the case that any discontinuity in the relationship between the outcome of interest and the variable determining the treatment status is fully attributable to the treatment itself. This requires some regularity conditions at the threshold for selection discussed by Hahn et al. (2001; HTV in the following).

On the other hand, the sharp RDD features two main limitations. First, its feasibility is confined to those instances in which assignment takes place *only* on observable pre-intervention variables; as a matter of fact, this is not often the case. Second, even when such a design applies, it only allows identification of the mean impact of the intervention at the threshold for selection. In the common situation of heterogeneous returns to participation, the local effect might be very different from the effect for individuals away from the threshold for selection. To identify the mean impact on a broader population of participants one has to rely on non-experimental estimators, whose consistency for the intended impact rests on assumptions about the behaviour of individuals.

Throughout this paper we consider the case in which an eligibility rule splits the relevant population into eligible and ineligible individuals, and participation of the former group is determined according to rules potentially *unknown* to the researcher. Examples of such a design are labour market programmes for which participation is voluntary for individuals satisfying a condition on age, or means tested programmes. College enrollment, for which only a subset of applicants is enrolled amongst those passing an achievement test, also fits this design.

The plan of the paper is as follows. First, we show that in the set-up described above the mean impact for participants around the threshold for eligibility is identified essentially under the same regularity conditions required in a sharp RDD, no matter how eligible individuals are selected into the programme (see Section 3). Although references have been made in the literature to the potential of using eligibility rules to identify mean impacts (see Heckman, 1992; Angrist, 1998; Heckman et al., 1999; van der Klaauw, 2002), to the best of our knowledge the regularity conditions required for identification have not been discussed so far.

Second, we show that eligibility rules for participation into a programme can be used to assess the validity of non-experimental estimators for the programme effects (see Section 4). In particular, we show that the selection bias arising from the non-random selection of eligible individuals into the programme is identified at the threshold for eligibility, so that one can formally test whether any of the long array of existing non-experimental estimators can correct for this bias. If this hypothesis is not rejected at the threshold for eligibility, one may feel more confident to use that non-experimental estimator to identify the causal effect on a broader population (typically, the one represented by all participants).

Several links to the literature are established. In particular, we show that our first result is closely related to what discussed by Bloom (1984), Heckman (1990) and Angrist and Imbens (1991). We also stress the