Dinamiche e persistenze nel mercato del lavoro italiano ed effetti di politiche (basi di dati, misura, analisi)

Progetto di ricerca cofinanziato dal MIUR
(Ministero dell'Istruzione, dell'Università e della Ricerca) – Assegnazione: 2001
Coordinatore: Ugo Trivellato

Another look at the Regression Discontinuity Design

E. Battistin*, E. Rettore**

* *Istitute for Fiscal Studies, London*
** *Dip. di Scienze Statistiche, Univ. di Padova*

Working Paper n. 44 novembre 2002

Unità locali del progetto:

Dip. di Economia "S. Cognetti De Martiis", Univ. di Torino	(coord. Bruno Contini)
Dip. di Scienze Economiche, Univ. "Ca' Foscari" di Venezia	(coord. Giuseppe Tattara)
Dip. di Metodi Quantitativi, Univ. di Siena	(coord. Achille Lemmi)
Dip. di Scienze Statistiche, Univ. di Padova	(coord. Ugo Trivellato)
Dip. di Politiche Pubbliche e Scelte Collettive, Univ. del Piemonte Orientale	(coord. Alberto Martini)

Dip. di Scienze Statistiche
via C. Battisti 241-243, 35121 Padova

1 Introduction[1]

The central issue in the evaluation of public policies is to separate their causal effect from the confounding effect of other factors influencing the outcome of interest. Random assignment of units to the intervention produces treatment and control groups that are equivalent in all respects, except for their treatment status. Thus, in a completely randomized experiment any post-intervention difference between outcomes for the two groups doesn't reflect pre-intervention differences by construction. As a result, differences between exposed and control units are entirely due to the intervention itself.

However, in most instances randomization is unfeasible either for ethical reasons or simply because assignment to the treatment can't be controlled by the analyst. Besides, even in those instances in which the analyst can randomize the assignment, units may not comply with the assigned status and either drop out of the intervention or seek an alternative program (see Heckman and Smith, 1995). A well-known and widely used example of randomized assignment is the JTPA program in the United States, which currently serves close to one million economically disadvantaged people every year (see Friedlander *et al.*, 1997). Random assignment occurs prior to the actual enrolment in the program, but a consistent fraction of those randomized into the treatment group don't participate. For certain components of the JTPA, such a non-complying behavior seems to be non-negligible (see, for example, Heckman *et al.*, 1998b). In this situation, the ideal experiment is not fully realized since participation turns out (at least partly) voluntary: training is provided only to those individuals who meet certain criteria of need and comply with the result of randomization. It follows that participation depends on observable and unobservable characteristics of individuals that might be correlated with the outcome of interest. Accordingly, differences in outcomes for treated and control groups might be the result of units' self-selection into the intervention.

The assessment of whether observed changes in the outcome of interest could be attributed to the intervention itself and not to other possible causes turns out to be even more complicated in a non-experimental setting. In this situation estimating cause-effect relationships that one might think to hold between the program and its outcomes typically depends on not testable assumptions about individuals' behavior. Given that the ideal situation for policy evaluators is the complete knowledge of the mechanism leading individuals to participate into the treatment, and given that in most instances such a mechanism is unknown (either because of non-compliance of individuals in an experimental setting or because of the lack of knowledge arising from observational studies), the question then arises of how to make the most out of each design to obtain reasonable estimates of program effects.

There are instances in which the so called Regression Discontinuity Design (RDD) arises (see Thistlethwaite and Campbell, 1960, Rubin, 1977, Trochim, 1984). According to this design, assignment is solely based on pre-intervention variables observable to the

[1]First draft 13^{th} February 2002. This paper benefited from helpful discussion with David Card, Hide Ichimura, Andrea Ichino and comments by audiences at ESEM 2002, CEPR/IZA Conference "Improving Labor Market Performance: The Need for Evaluation" (Bonn, October 2002), Statistics Canada Symposium 2002 and LABORatorio Conference "New perspectives in public policy evaluation" (Turin, November 2002). Financial support from MIUR to the project "Dynamics and inertia in the Italian labour market and policies evaluation (data-bases, measurement issues, substantive analyses)" is gratefully acknowledged.

analyst and the probability of participation changes discontinuously as a function of these variables. To fix ideas, consider the case in which a pool of units willing to participate into a program is split into two groups according to whether a pre-intervention measure is above or below a known threshold. Those who score above the threshold are exposed to the intervention while those who score below are denied it.

This design features both advantages and disadvantages relative to its competitors. On the one hand, in a neighborhood of the threshold for selection a RDD presents some features of a pure experiment. In this sense, it is certainly more attractive than a non-experimental design. Since subjects in the treatment and control group solely differ with respect to the variable on which the assignment to the intervention is established (and with respect to any other variable correlated to it), one can control for the confounding factors contrasting marginal participants to marginal non-participants. In this context, the term marginal refers to those units *not too far* from the threshold for selection. The comparison of mean outcomes for marginally treated and marginally control units identifies the mean impact of the intervention locally with respect to 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 outcomes 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 RDD features two main limitations. Firstly, its feasibility is by definition confined to those instances in which selection takes place on observable pre-intervention variables; as a matter of fact, this is not often the case. Secondly, even when such a design applies, in the presence of heterogeneous impacts it only permits to identify the mean impact of the intervention at the threshold for selection. In the realistic situation of heterogeneous impacts across units, this local effect might be very different from the effect for units away from the threshold for selection.

To identify the mean impact on a broader population one can only resort to a non-experimental estimator whose consistency for the intended impact intrinsically depends on behavioral (and not testable) assumptions. In this paper we show that the range of applicability of the RDD includes all those instances in which the relevant population is split into two subpopulations, eligibles and non-eligibles, provided that (i) the eligibility status is established with respect to a continuous variable and (ii) information on both ineligibles and eligible non-participants is available. Then, the mean impact on participants in a neighborhood of the threshold for eligibility is identified, no matter how eligible units self-select into the program.

We make an explicit connection to the literature on RDDs (only implicit references so far; Angrist, 1998) and we derive the regularity conditions necessary for identification to hold (van der Klaauw, 2002, and Heckman *et al.*, 1999, point out that a RDD arises from eligibility rules but they do not discuss conditions for identification). We also show that the discontinuity in the eligibility rule leads to regularity conditions for identification weaker than those arising in the so called fuzzy RDD design (Hahn *et al.*, 2001), to which the set-up we discuss *prima facie* belongs.

Secondly, as a straightforward corollary of the previous result, the selection bias at the threshold for eligibility turns out identifiable. Then, one can formally test whether any of the long arrays of existing non-experimental estimators can correct for this bias. If a non-experimental estimator can correct for the selection bias even if only for a particular subpopulation - namely, the units in a neighborhood of the threshold for eligibility - one might feel more confident to use it to estimate the causal effects on a broader population.

Links to related literature are established. In particular, we show that our first result is closely related to Bloom (1984) and to Angrist and Imbens (1991). We also stress that our result is closely related to the idea stated by Rosenbaum (1987) of using two alternative comparison groups for the identification of causal effects. Finally, we point out the similarities between our specification test of a non-experimental estimator and the specification tests derived by Heckman and Hotz (1989) as well as the link to the characterization of the selection bias provided by Heckman et al. (1998a).

The remaining of this paper is organized as follows. Section 2 discusses similarities between a fully randomized experiment and a RDD. Section 3 generalizes the use of a RDD when participation into the treatment group is determined by self-selection. Section 4 shows how to validate the use of non-experimental estimators for the treatment effect using a RDD. Threats to the validity of the design are discussed in section 5. Section 6 presents some concluding remarks.

2 The Regression Discontinuity Design

This section presents the basic features of the RDD and highlights its similarities with a randomized experiment. The discussion of identification issues arising in the RDD is based on HTV, to which the interested reader is referred for further details.

Following the notation of the potential outcome approach to causal inference (see Rubin, 1974), let (Y_1, Y_0) be the potential outcomes that would be experienced by participating and not participating into the program, respectively.

The causal effect of the treatment is then defined as the difference between these two potential outcomes, $\beta = Y_1 - Y_0$, which is not observable since being exposed to (denied) the program reveals Y_1 (Y_0) but conceals the other potential outcome. In what follows we will discuss the case in which the program impact β varies across individuals, which in most instances is the relevant case.

Let I be the binary variable for the treatment status, with $I = 1$ for participants and $I = 0$ for non-participants. If the assignment is determined by randomization, the following condition holds true by construction

$$(Y_1, Y_0) \perp I. \tag{1}$$

The attractiveness of randomization is that the difference between mean outcomes for participants and non-participants identifies the mean impact of the program

$$E(\beta) = E(Y_1|I = 1) - E(Y_0|I = 0), \tag{2}$$

since conditioning on I in the right-hand side of (2) is irrelevant by construction. In other words, randomization allows using information on non-participants to identify the mean counterfactual outcome for participants, namely what participants would have experienced had they not participated into the program.

Suppose that the participation status depends on an observable individual characteristic S which we assume to be continuous on the real line. A RDD arises when there exists a *known* point in the support of S where the probability of participation changes discontinuously. Formally, if $\bar{s}$ is the discontinuity point, a RDD arises when

$$Pr\{I(\bar{s}^+) = 1\} \neq Pr\{I(\bar{s}^-) = 1\}. \tag{3}$$

Here and in the following $\bar{s}^+$ and $\bar{s}^-$ refer to those individuals *marginally* above and below $\bar{s}$, respectively. Moreover, to ease the exposition and without any loss of generality, we will deal with the case in which the assignment probability increases as S crosses the threshold $\bar{s}$, so that

$$Pr\{I(\bar{s}^+) = 1\} - Pr\{I(\bar{s}^-) = 1\} > 0.$$

Following Trochim (1984), it is usual to distinguish between *sharp* RDD and *fuzzy* RDD depending on the size of this jump. The former design occurs when the probability of selection into the treatment conditional on S steps from zero to one as S crosses the threshold $\bar{s}$. That is, the treatment status deterministically depends on whether individuals' values of S are above $\bar{s}$, so that

$$I = \mathbb{1}(S \geq \bar{s}). \tag{4}$$

A fuzzy RDD occurs when the size of the discontinuity at $\bar{s}$ is smaller than one, implying that the assignment to the treatment is no longer a deterministic function of S.

Although the RDD lacks random assignment of units to the treatment group, it shares an attractive feature with an experimental design. We will address this issue by considering a sharp RDD; we will discuss the fuzzy case further below in this section. Let

$$Y(s) = Y_1(s)I(s) + Y_0(s)(1 - I(s))$$

be the outcome observed for an individual scoring $S = s$. The mean outcome difference for individuals in a neighborhood of $\bar{s}$

$$E\{Y(\bar{s}^+)\} - E\{Y(\bar{s}^-)\} \tag{5}$$

can be written as

$$E\{Y_0(\bar{s}^+) - Y_0(\bar{s}^-)\} + E\{I(\bar{s}^+)\beta(\bar{s}^+)\} - E\{I(\bar{s}^-)\beta(\bar{s}^-)\}, \tag{6}$$

which simplifies to

$$E\{Y_0(\bar{s}^+) - Y_0(\bar{s}^-)\} + E\{\beta(\bar{s}^+)\}$$

because of (4). The following condition is then sufficient for the mean impact of the treatment at $\bar{s}^+$ to be identified in a sharp RDD.

Condition 1. The mean value of $Y_0(s)$ conditional on $S = s$ is a continuous function of S at $\bar{s}$.

Accordingly, Condition 1 requires that the only discontinuity taking place at the threshold for selection is in the assignment probability.

The attractiveness of the RDD is apparent here. By controlling for S one can identify the average impact of the program on subjects belonging to a right-neighborhood of $\bar{s}$, thus a local version of the parameter in (2)

$$E\{\beta(\bar{s}^{+})\} = E\{Y(\bar{s}^{+})\} - E\{Y(\bar{s}^{-})\}.$$

In most instances however, the policy relevant parameter is likely to be $E\{\beta(\bar{s}^{-})\}$, that is the expected mean impact from expanding the program to the marginally excluded individuals. To identify such a parameter the following additional condition needs to hold.

Condition 2. The mean value of $Y_1(s)$ conditional on $S = s$ is a continuous function of S at $\bar{s}$.

Note that the sharp RDD represents a special case of selection on observables and, in a neighborhood of $\bar{s}$, presents the same features of a pure randomized experiment (see Rubin, 1977). By exploiting the relationship between S and I in (3), the following condition holds true

$$(Y_1, Y_0) \perp I | S = \bar{s}. \tag{7}$$

Because of this property the RDD is often referred to as quasi-experimental design (Cook and Campbell, 1979). It allows inferring causal relationships by exploiting the fact that in a neighborhood of $\bar{s}$ the treatment status is determined nearly randomly.

It is worth stressing again that to meaningfully define marginal units (with respect to $\bar{s}$) S needs to be a continuous variable. Moreover, note that the estimation of the right-hand side (left-hand side) of the mean outcome difference in (5) makes use of data only in a neighborhood on the right (left) side of the discontinuity point. Unless one is willing to make some parametric assumptions about the regression curve away from $\bar{s}$, only data local to the discontinuity help to estimate the jump. Asymptotically the neighborhood needs to shrink as with usual non-parametric estimation, implying a non-standard asymptotic theory for the resulting estimator of the mean impact (see HTV and Porter, 2002).

In some cases units do not comply with the mandated status as it results from the *sharp* RDD, dropping out of the program if assigned to it or seeking alternative treatments if denied it (see, for example, Battistin and Rettore, 2002). The *fuzzy* RDD arises in these instances. Any violation of the original assignment invalids the orthogonality condition in (7). Accordingly, the mean impact at $\bar{s}$ cannot be identified by simply comparing mean outcomes for marginally participants and non-participants as in (5).

In fact, if Conditions 1 and 2 are satisfied, the expression in (6) becomes

$$E\{\beta(\bar{s})|I(\bar{s}^{+}) \neq I(\bar{s}^{-})\}Pr\{I(\bar{s}^{+}) \neq I(\bar{s}^{-})\}.$$

Then, under the additional condition

Condition 3. Participation into the program is *monotone* around $\bar{s}$, that is it is either the case that $I(\bar{s}^{+}) \geq I(\bar{s}^{-})$ for all subjects or the case that $I(\bar{s}^{+}) \leq I(\bar{s}^{-})$ for all subjects.

the mean impact on *compliers* (LATE) in a neighborhood of $\bar{s}$ is identified by

$$E\{\beta(\bar{s})|I(\bar{s}^+) \neq I(\bar{s}^-)\} = \frac{E\{Y(\bar{s}^+)\} - E\{Y(\bar{s}^-)\}}{E\{I(\bar{s}^+)\} - E\{I(\bar{s}^-)\}}, \tag{8}$$

the compliers being those individuals who would switch from non participation to participation if their score S crossed the threshold (see Imbens and Angrist, 1994). Note that Condition 3 is an assumption on individuals' behavior which is not testable. Moreover, even if the size of the group of complying individuals is observable, it is not observable which individuals it consists of.[2] Whether the resulting mean impact is a policy relevant parameter it is case-specific (see Heckman, 1997, for a discussion).[3]

Apparently, identification of the mean impact in a fuzzy RDD is more demanding as well as restricted to a narrower group of individuals than in a sharp RDD. Heckman *et al.* (1999) emphasize it by saying that much of the simplicity of the design is lost switching from the sharp RDD to the fuzzy one.

Two major drawbacks hamper the applicability of RDDs. Firstly, in an observational study it is very often the case that units self-select into the treatment rather than being exogenously selected on a pre-program measure. If this is the case, the RDD set-up as introduced so far no longer applies. Secondly, even in those instances in which the RDD applies, if the impact is heterogeneous across individuals such a design is not informative about the impact on individuals away from $\bar{s}$. These are the two issues we will look at in the next sections.

3 A generalization of the 'sharp' Regression Discontinuity Design

It is often the case that social programs are targeted to specific sub-groups of individuals meeting a fully specified set of conditions for eligibility. For example, the New Deal for Young People in the United Kingdom offers job-search assistance followed by a package of subsidized options to all individuals aged between eighteen and twenty-four who have been claiming unemployment insurance for six months (see Blundell *et al.*, 2002). To prevent long-term unemployment, the Swedish UVG program guarantees the assignment to some labor market programs to the unemployed younger than 25 (see Carling and Larsson, 2002).

To fix ideas, let S be a continuous pre-program characteristic and let the *eligibility status* I be established according to the rule in (4). Throughout this paper we will assume that S is observable for all individuals. If all eligibles participated into the program a sharp RDD then would arise. For example, if participation into the program were mandatory for all eligible individuals the effect of the policy at the threshold for eligibility would be identified by (5).

As a matter of fact, it is often the case that some eligible individuals self-select into the program while some others do not. Across individuals heterogeneity about information

[2]The proportion of compliers at $\bar{s}$ equals the denominator in the right-hand side of (8), that is $Pr\{I(\bar{s}^+) = 1\} - Pr\{I(\bar{s}^-) = 1\}$.

[3]As an alternative to Condition 3 HTV also consider the local orthogonality condition $(Y_1, Y_0) \perp I | S = \bar{s}$. Apparently, this condition rules out non-random selection based either on (Y_1, Y_0) or on any variable correlated to (Y_1, Y_0).

on program's availability, about preferences and about opportunity costs are factors likely to influence the participation decision in several instances. Accordingly, the population turns out split into three subgroups: *ineligibles*, *eligible non-participants* and *participants*. Our analysis develops with reference to the general case in which the analyst does not know anything about the rule leading eligible individuals to self-select into the program.

As a result of the eligibility rule and of self-selection the probability of participation into the program steps at $\bar{s}$ from zero to less than one originating a fuzzy RDD. Van der Klaauw (2002, p.1284) explicitly mentions the potential for using the RDD arising from the eligibility criteria for a social program. Heckman *et al.* (1999, pp.1971-1972) recognize the fuzzy RDD nature of this set-up and point out that in this case the LATE estimand (8) identifies the mean impact on participants. In what follows we show that in this set-up, despite its *prima facie* fuzzy RDD nature, the regularity conditions required to identify the mean impact of the program are the same as those conditions required in a sharp RDD. In this sense the set-up described in this section provides a generalization of the sharp RDD.

We address this issue by stressing once more the analogy with an experimental setting. Bloom (1984) notes that even if some individuals randomly assigned to the treatment group do not eventually show-up, the identification of the mean impact on participants is still secured by the experiment (see also Little and Yau, 1998, and Heckman *et al.*, 1998b). The key relationship on which the result rests is the following equality implied by (1)

$$E(Y_0|I=1) = E(Y_0|I=0), \tag{9}$$

where I is the mandated status as it results from randomization. The left hand side of (9) can be written as the weighted average of the mean outcome for randomized-in participants and for randomized-in non-participants, respectively

$$E(Y_0|I=1, D=1)\phi + E(Y_0|I=1, D=0)(1-\phi).$$

Here D represents the *actual* status of subjects randomly assigned to the treatment group as it results from self-selection and $\phi = Pr(D=1|1=1)$ the probability of self-selection into the program conditional on assignment to the treatment group. Substituting the last expression in (9) we obtain

$$E(Y_0|I=1, D=1) = \frac{E(Y_0|I=0)}{\phi} - E(Y_0|I=1, D=0)\frac{1-\phi}{\phi}. \tag{10}$$

Namely, the *counterfactual* mean outcome for participants is identified as a linear combination of the *factual* mean outcome for randomized-out individuals and of the *factual* mean outcome for randomized-in non-participants. The coefficients of the linear combination add up to one and are function of the probability ϕ which in turn is identifiable. Hence, equation (10) implies that the mean impact on participants is identified.

By analogy, exploiting the fact that the eligibility rule (4) defines a randomized experiment in a neighborhood of $\bar{s}$ with eligible non-participants playing the role of Bloom's (1984) no-shows, the intuition suggests that the mean impact on participants in a neighborhood of $\bar{s}$ is also identified. To recover the regularity conditions needed for identification consider again the difference in (5). Since the probability of participation for marginally ineligibles is zero (since $I(\bar{s}^-) = 0$), the expression in (6) becomes

$$E\{Y_0(\bar{s}^+) - Y_0(\bar{s}^-)\} + E\{I(\bar{s}^+)\beta(\bar{s}^+)\}.$$

The second term of the previous expression equals

$$E\{\beta(\bar{s}^+)|I(\bar{s}^+) = 1\}Pr\{I(\bar{s}^+) = 1\},$$

so that the ratio

$$E\{\beta(\bar{s}^+)|I(\bar{s}^+) = 1\} = \frac{E\{Y(\bar{s}^+)\} - E\{Y(\bar{s}^-)\}}{E\{I(\bar{s}^+)\}} \tag{11}$$

identifies the mean impact on participants in a right-neighborhood of $\bar{s}$ provided that $Y_0(s)$ is continuous at $S = \bar{s}$. As before, Condition 2 is needed to extend the identification result to the left-neighborhood of $\bar{s}$. Otherwise stated, if in the absence of self-selection among the eligibles the conditions needed for the identification of the mean impact at $S = \bar{s}$ are met, then the mean impact on participants at $S = \bar{s}$ is also identifiable in the presence of self-selection. Results by HTV and Porter (2002) on non-parametric inference in a RDD straightforwardly apply to the estimation of (11).

Note that, although the RDD arising from the eligibility rule is fuzzy, the conditions needed to identify the impact of the program on complying individuals are those required in a sharp RDD, and therefore they are *weaker*. This key result rests on the fact that the probability of participation at the left-hand side of $\bar{s}$ is zero by design ($I(\bar{s}^-) = 0$), and this simplifies the expression in (6) without further assumptions on individuals' behavior. Moreover, note that this feature of the design implies that the the pool of compliers in a neighborhood of the threshold for selection corresponds to the pool of participants, so that in this case the identifiable LATE refers to a larger and well defined sub-population
.

Few comments are in order. Firstly, neither ineligibles nor eligible non-participants alone would allow the identification of any interesting parameter. As it is apparent from Bloom's key result in (10) adapted to the RDD context, it is the joint use of information on either group to secure the identification of the mean counterfactual outcome on participants in a neighborhood of $\bar{s}$. In this sense, the availability of sampling information on three different groups of subjects (participants, eligible non-participants and ineligibles) is crucial for identification.

Secondly, note that to derive the result we don't need to specify how eligible units self-select into the treatment. Thus, the identifiability of the mean impact doesn't require any behavioral assumption on the process itself.

Finally, our result (as well as Bloom's one) can also be derived as a special case of Angrist and Imbens (1991). The authors prove that even if participation into the program takes place as a result of self-selection, the mean impact on participants is identifiable provided that (i) there exists a random variable Z affecting the participation into the program and orthogonal to the potential outcomes and (ii) the probability of participation conditional on Z is zero in at least one point of the support of Z. Condition (i) qualifies Z as an Instrumental Variable for the problem. In our case, since I is orthogonal to the potential outcomes in a neighborhood of $\bar{s}$ and $Pr(I(\bar{s}^-) = 1) = 0$, I meets the conditions stated by Angrist and Imbens (1991) in a neighborhood of $\bar{s}$.[4] The identification of the

[4] See the discussion in HTV on the property of the IV estimator in this instance.

mean impact on participants at $\bar{s}$ follows. In this sense, the RDD can be interpreted as a combination of regression control and instrumental variables identification strategy in a neighborhood of the threshold for selection.[5]

4 Validating non-experimental estimators of the mean impact on participants

4.1 Specification tests

The previous section has discussed a set of sufficient conditions for the identification of the mean impact of the program on participants *marginally* eligible for it, even if they might represent a self-selected group from the pool of eligible individuals. This result relies on the existence of an eligibility rule depending on continuous characteristics observable to the analyst.

If the gain of the treatment is heterogeneous with respect to S, the mean impact for individuals in a neighborhood of the threshold for eligibility is not informative on the impact of the intervention for individuals away from this point. Nor ineligible units and eligible non-participants can be used as valid comparison groups, since they differ systematically from participants (the former with respect to S and the latter with respect to the variables driving self-selection).

In order to identify the mean impact on the overall population of participants

$$E\{Y_1(s)|I(s) = 1\} - E\{Y_0(s)|I(s) = 1\}, \quad s \geq \bar{s}$$

one has to resort to one of the long array of non-experimental estimators available in the literature which adjust for selection bias under different assumptions (see Heckman *et al.*, 1999, and Blundell and Costa Dias, 2000, for a review). The main problem with non-experimental identification strategies is that they rely on assumptions intrinsically not testable.

Over the years the literature has taken two main routes to deal with this problem. The first route amounts to seek whether any over-identifying restrictions on the data generating process arise from a theory on the phenomenon under investigation. If this is the case, one might exploit such restrictions to test the assumptions on which the non-experimental estimator rests (see Rosenbaum, 1984, and Heckman and Hotz, 1989).

The second route is feasible only if an experimental estimate of the impact is available. This solution is quite demanding, since it requires data from an experimental design. If this is the case, one can exploit the experimental set-up to characterize the selection bias and to assess whether non-experimental estimators are able to reproduce the experimental estimate (see LaLonde, 1986 and Heckman *et al.*, 1998a). Accordingly, one can learn which non-experimental strategy to follow when experimental data are not available.

[5] This idea is easily implemented in a fully parametric set-up. From a sample with information for participants, eligible non-participants and ineligibles, estimate the regression

$$Y = \alpha_0 + \alpha_1 S + \alpha_2 I + \varepsilon$$

using the eligibility status $Z = \mathbb{1}(S \geq \bar{s})$ as an instrumental variable for I, where I indexes actual participants and non-participants. Note that S enters the equation to control for the selection induced by the eligibility rule.

This section shows that if information is available on the three groups of individuals resulting from the set-up of Section 3, then one can test the validity of any non-experimental estimator on a *specific* subpopulation of participants. To fix the ideas, we will focus on the widely used *matching estimator* (see Rosenbaum and Rubin, 1983), but the same line of reasoning applies to other non-experimental estimators.

The key assumption on which the matching estimator rests is that all the variables driving the self-selection process *and* correlated to the potential outcome Y_0 are observable to the analyst. Formally, the assignment to the treatment is told to be *strongly ignorable* given a set of characteristics x if

$$(Y_1, Y_0) \perp I | x, \quad 0 < Pr(I = 1|x) < 1. \tag{12}$$

If strong ignorability holds, then it is as if units were randomly assigned to the treatment with a probability depending on x, provided that such probability is non-degenerate at each value x. The counterfactual outcome for participants presenting characteristics x can be approximated by the actual outcome of non-participants presenting the same characteristics. Since units presenting x have a common probability to enter the program, then an operational rule to obtain an *ex post* experimental-like data set is to match participants to non-participants on such probability (the so called *propensity score*), whose dimension is invariant with respect to the dimension of x (see Rosenbaum and Rubin, 1983).

The critical identifying assumption is that the available set of variables x is rich enough to guarantee the orthogonality condition in (12). In other words, the ignorability condition requires that the selection bias affecting the raw comparison of outcomes for participants and non-participants only depends on observables. In principle, this imposes strong requirements on data collection. Moreover, the violation of the second condition in (12) causes the so called common support problem (see for example Heckman *et al.*, 1998a, and Lechner, 2001).[6]

Let

$$sb(s) = E\{Y_0(s)|I(s) = 1\} - E\{Y_0(s)|I(s) = 0\}, \quad s \geq \bar{s} \tag{13}$$

be the *selection bias* affecting the raw comparison of participants to eligible non-participants scoring $S = s$, with $S \geq \bar{s}$. The first term on the right hand side is the mean outcome that participants would have experienced had they not taken part into the program. The second term is the actual mean outcome experienced by eligible non-participants. This quantity captures pre-intervention differences between eligible subjects self-selected in and out of the program, respectively, at each level of S, with $S \geq \bar{s}$.

Using the results of the previous section, the mean counterfactual outcome for participants on the right hand side of (13) is identifiable in a neighborhood of $\bar{s}$ provided that Conditions 1 and 2 hold. It follows that the selection bias for marginally eligible subjects, $sb(\bar{s})$, is identifiable as the difference between

$$E(Y_0(\bar{s}^+)|I(\bar{s}^+) = 1) = \frac{E(Y_0(\bar{s}^-)|I(\bar{s}^-))}{\phi(\bar{s}^+)} - E(Y_0(\bar{s}^+)|I(\bar{s}^+) = 1)\frac{1 - \phi(\bar{s}^+)}{\phi(\bar{s}^+)} \tag{14}$$

[6] It is worth noting that the RDD represents a particular case of selection on observables, but locally with respect to the threshold for eligibility. In this sense, the RDD estimator is a limit form of matching at the threshold for eligibility.

and the factual mean outcome experienced by marginally eligible non-participants, $E(Y_0(\bar{s}^+)|I(\bar{s}^+) = 0)$.

Note that the identification of the counterfactual term on the right-hand side of (13) at $\bar{s}$ crucially requires information on the subgroup of non-eligibles closest to the group of eligibles, thus in a neighborhood of the threshold for eligibility. Apparently, identification is precluded as S moves away from $\bar{s}$.

Consider now the bias term for a specific subpopulation of eligibles indexed by x and s

$$sb(s, x) = E\{Y_0(s)|I(s) = 1, x\} - E\{Y_0(s)|I(s) = 0, x\}, s \geq \bar{s}$$

where x are the variables claimed to properly account for the selection bias in a matching estimation of the program effect. If the orthogonality condition in (12) holds, then

$$sb(s, x) = 0$$

uniformly with respect to x and s. In particular, a necessary condition for the matching estimator to work is that $sb(\bar{s}, x) = 0$, which is directly testable since in a neighborhood of $\bar{s}$ $sb(s, x)$ is identifiable.

In a neighborhood of $\bar{s}$ any test of the equality of the mean outcomes experienced by non-eligibles and eligible non-participants conditional on x is a test of the ignorability of the self-selection into the program, thus a test on the validity of the matching estimator.[7] Clearly, the rejection of the null hypothesis is sufficient to conclude that condition (12) does not hold. On the other hand, on accepting the null hypothesis one might feel more confident in using the matching estimator. However, it can't be said that the validity of the estimator has been proved: in fact, the test is not informative on whether the ignorability condition holds away from $\bar{s}$.[8]

4.2 Related results

Since the RDD can been seen as a formal experiment at $S = \bar{s}$, the specification test developed above displays an apparent similarity to what Heckman *et al.* (1998a) develop in an experimental set-up. In both cases there is a benchmark estimate of the mean impact - the RDD estimate in the former, the experimental estimate in the latter - to which the analyst is ready to attach credibility. Then, the analyst compares non-experimental estimates to the benchmark value interpreting any discrepancies as a violation of the maintained assumptions.

The similarity between the two approaches stops here. The availability of an experimental set-up as in Heckman *et al.* (1998a) allows to fully characterize the selection bias

[7] In fact, one can test simultaneously the equality of the two mean outcomes as well as the equality of their derivatives at $\bar{s}$, thus gaining power.

[8] To give a flavor of the specification test discussed in this section, we rely again on the parametric set-up introduced in the previous footnote. Restricting the sample to ineligible and eligible non-participant individuals, estimate the regression

$$Y = \alpha_0 + \alpha_1 S + \alpha_2 X + \alpha_3 Z + \varepsilon,$$

where $Z = \mathbb{1}(S \geq \bar{s})$ denotes the eligibility status, and then test for $\alpha_3 = 0$.

and to test non-experimental estimators with reference to the pool of participants. If a RDD is available, this is feasible only for participants in a neighborhood of $S = \bar{s}$.

However, the availability of experimental data is rarely encountered in the evaluation of social policies, especially in EU countries. On the other hand, it is very often the case that a policy is targeted to a population of eligible individuals whose participation into the program is on a voluntary basis. In this situation, information on the three groups of individuals needed to implement the results in this paper is in principle available.[9] This opens the door to a routinely application of the specification test based on the RDD as a tool to validate non-experimental estimators of the mean impact on participants.

Rosenbaum (1987) in his discussion on the role of a second control group in an observational study gives an example which resembles, albeit loosely, the set-up we refer to (example 2, p.294). The Advanced Placement (AP) Program provides high school students with the opportunity to earn college credits for work done in high school. Not all high schools offer the AP program, and in those that do, only a small minority of students participate. Two comparison groups naturally arise in this context: students enrolled in high school not offering the program and students enrolled in high schools offering the program who did not participate.

Then, Rosenbaum (1987) goes on discussing how the availability of two comparison groups can be exploited to test the strong ignorability condition needed to believe the results of a matching estimator. Apparently, his first comparison group resembles our pool of ineligibles while the second group resembles our pool of eligible non-participants. The crucial difference between Rosenbaum's example and our set-up is that in the former case the rule according to which high schools decide whether to offer the AP program is unknown to the analyst. In our set-up the eligibility rule is fully specified. It is exactly this feature of our set-up to allow identifying the mean impact on participants at $S = \bar{s}$ as well as to test the validity of any other non-experimental estimator even if only locally at $S = \bar{s}$.

5 Threats to validity

5.1 Substitution effects

The estimation of any causal implications of the treatment on the outcome of interest rests on the *Stable Unit Treatment Value Assumption* (SUTVA) for all members in the population (Rubin, 1977). According to this assumption, the outcome experienced by each individual is not affected by assignment and receipt of treatment by other individuals. It follows that SUTVA rules out general equilibrium effects among potential participants that could occur because of their participation decision.

Consider the case in which the participants' gain from the treatment hampers the outcome of those individuals not participating into the program. For instance, if the program is targeted to improve job opportunities for participants, it might happen that non-participants experience a deterioration of their job opportunities because as a consequence of the program employers find comparatively more attractive to hire participants. Following the literature (see Heckman *et al.*, 1999), we will refer to such a problem as 'substitution effect' and we will discuss its implications for the RDD set-up.

[9] Whether or not this is the case depends on the design and on data collection problems.

To deal with this problem, the two potential outcomes discussed in Section 2 are redefined and a third new potential outcome is introduced. Let Y_0 be the outcome experienced by individuals if the program is *not rolled out*, Y_1 be the outcome experienced by participating into the program and Y_2 be the outcome experienced by not participating if the program is actually rolled out and someone else take part into it. Consequently, the effect of the program on participants is $Y_1 - Y_0$, while the effect on non-participants is $Y_2 - Y_0$. A substitution effect arises when in the presence of the intended effect on participants the latter effect has sign opposite to the former one.

If the participation status is randomly assigned, the comparison of mean effects for participants and non-participants identifies

$$E(Y_1 - Y_2) = E(Y_1 - Y_0) - E(Y_2 - Y_0),$$

which is larger (in absolute value) than the parameter of interest $E(Y_1 - Y_0)$. That is, although data from a fully randomized experiment are exploited, the raw comparison of outcomes experienced by participants and non-participants yields an upward biased estimate of the true effect (substitution bias). Note that if the impact on non-participants is zero, no substitution effects arise and standard results for an experimental design straightforwardly apply.

In a RDD context, things are likely to be even worse. To fix ideas, consider the case of a *sharp* RDD and suppose that the participation status depends on individuals' age S. Then, the analogue of the previous expression corresponds to

$$E(Y_1 - Y_2|\bar{s}) = E(Y_1 - Y_0|\bar{s}) - E(Y_2 - Y_0|\bar{s}),$$

where $\bar{s}$ is the threshold age for being enrolled into the program. Marginally non-participants are those individuals closest to participants with respect to S. If age enhances the attractiveness of potential employees to employers, then a marginal non-participant individual is more at risk of being substituted out by participants than a randomly drawn non-participant. Formally,

$$|E(Y_2 - Y_0|\bar{s})| \geq |E(Y_2 - Y_0)|$$

implying that the substitution bias might be larger in a RDD than in a randomized experiment setting.

However it is worth noting that, while the occurrence of substitution bias might preclude the identification of program effects, it does not preclude testing for a non-zero program impact. In fact, for the mean impact on non-participants $E(Y_2 - Y_0|\bar{s})$ to be different from zero it must be that the mean impact on participants $E(Y_1 - Y_0|\bar{s})$ is not zero. Hence, the mean impact on participants is different from zero if and only if $E(Y_1 - Y_2|\bar{s}) \neq 0$, which is testable.

5.2 Entry effects

A further threat to the validity of the RDD occurs when subjects ineligible for the program on the basis of their characteristic S purposively behave to modify their S to become eligibles. For example, mean-tested policies targeted to improve family income might induce some subjects to deliberately reduce their labor income to become eligible.

If this is the case, eligibility for the policy is no longer a deterministic function of S as in (4), requiring stronger conditions for identification than those discussed in Section 3. In what follows we will refer to this problem as 'entry effect'.

Two necessary conditions must be met for such an effect to take place. Firstly, the characteristic S with respect to which eligibility is established has to be under individuals' control. Income and unemployment duration are two notable examples of variables potentially affected by this problem, while age is not. Secondly, eligibility for the program must not come 'as a surprise', otherwise individuals could not act to modify their own S so that it meets eligibility conditions.

If an entry effect takes place, the nearly randomized experiment set-up in a neighborhood of $\bar{s}$ as in (7) no longer applies. This is because the pool of *actual* ineligibles turns out made up of those subjects who *chose* to self-select out of the program by not acting on their S while the pool of *actual* eligibles turns out to include some subjects who *chose* to become eligible by acting on S.

Two simple tests are available to check whether any entry effects took place. The first test is presented in Lee (2001). It exploits the intuition that if the local orthogonality condition (7) does not hold, then the treatment and the control groups are no longer equivalent with respect to any conceivable variable. To simultaneously test whether an entry effect took place *and* whether it is causing a selection bias, it suffices to find a variable stochastically related to Y_0 and *logically* unaffected by the program. Then, check whether the two groups are balanced with respect to such variable.

The second test is based on the intuition that, if some ineligible subjects alter their S to switch from values below $\bar{s}$ to values above it, a discontinuity in the cumulative distribution of the *observed* S shows up at $\bar{s}$. To formalize this intuition, let S_{obs} be the observed value of the variable on which the eligibility is established and let the following condition be satisfied.

Condition 4. The distribution of S, F_S, is smooth in a neighborhood of $\bar{s}$ and the entry effect takes place as an alteration of the value of S for a fraction of ineligible subjects such that the corresponding observed value, S_{obs}, belongs to a right neighborhood of $\bar{s}$.

Let p be the proportion of individuals in the population switching from anywhere below $\bar{s}$ to just above it to become eligible. As a result, the distribution of S_{obs}, F_{obs}, in a *left*-neighborhood of $\bar{s}$ is

$$F_{obs}(\bar{s}^-) = F_S(\bar{s}^-) - p,$$

while in a *right*-neighborhood of $\bar{s}$ it is

$$F_{obs}(\bar{s}^+) = F_S(\bar{s}^+) + p.$$

As a result, F_{obs} turns out discontinuous at $\bar{s}$

$$F_{obs}(\bar{s}^+) - F_{obs}(\bar{s}^-) = 2p,$$

and the discontinuity jump equals two times the proportion of switchers.

6 Conclusions

The main message from this paper is that every time an intervention is targeted to a population of eligible individuals but is actually administered to a sub-set of self-selected eligibles, it is worth collecting information separately on *three* groups of subjects: ineligibles, eligible non-participants and participants. Also, the variables with respect to which eligibility is established have to be recorded to allow the identification of marginally eligibles and ineligibles, respectively.

The relevance of distinguishing between non-eligibles and eligible non-participants to improve the comparability between the treated and the comparison groups has already been pointed out in the literature (see, amongst others, Heckman *et al.*, 1998a). Here we have derived the regularity conditions required to identify the mean impact on marginally eligible participants by *jointly* exploiting these two comparison groups. Despite the *prima facie* fuzzy RDD nature of the resulting set-up, we have shown that to secure identification of the mean impact it only requires the same continuity condition as a sharp RDD.

Then, we have shown that as a straightforward consequence of the previous result also the selection bias for subjects on the margin between eligibility and non-eligibility is identifiable. This opens up the door to a specification test in a neighborhood of the threshold for eligibility so that the properties of non-experimental estimators can be assessed. By design, such a test is informative only for a particular subgroup of individuals, thus results cannot be generalized to the whole population (unless one is willing to impose further identifying restrictions). The value of the specification test is that if it rejects the non-experimental estimator locally then this is enough to reject it altogether.

References

[1] Angrist, J.D (1998), *Estimating the Labor Market Impact of Voluntary Military Service Using Social Security Data on Military Applicants*, Econometrica, 66, 2, pp. 249-288

[2] Angrist, J.D. and Imbens, G.W. (1991), *Sources of Identifying Information in Evaluation Models*, NBER Technical Working Paper 117

[3] Battistin, E. and Rettore, E. (2002), *Testing for programme effects in a regression discontinuity design with imperfect compliance*, Journal of the Royal Statistical Society A, Vol. 165, No. 1, 1-19

[4] Bloom, H.S. (1984), *Accounting for No-Shows in Experimental Evaluation Designs*, Evaluation Review, Vol. 8, 225-246

[5] Blundell, R. and Costa Dias, M. (2000), *Evaluation methods for non-experimental data*, Fiscal Studies, Vol. 21, No. 4, 427-468

[6] Blundell, R. Costa Dias, M. Meghir, C. and Van Reenen, J. (2002), *Evaluating the Employment Impact of a Mandatory Job Search Assistance Program: The New Deal for Young People in the UK*, unpublished manuscript, Institute for Fiscal Studies, London

[7] Carling K. and L. Larsson (2002), *Does early intervention help unemployed youth?*, IFAU Working Paper 2002:10

[8] Cook, T.D. and Campbell, D.T. (1979), *Quasi-Experimentation. Design and Analisys Issues for Field Settigs*, Boston: Houghton Mifflin Company

[9] Friedlander, D. Greenberg, D.H. and Robins, P.K. (1997), *Evaluating Government Training Programs for the Economically Disadvantaged*, Journal of Economic Literature, Vol. 35, No. 4, 1809-1855

[10] Hahn, J. Todd, P. and van der Klaauw, W. (2001), *Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design*, Econometrica, Vol. 69, No. 3, 201-209

[11] Heckman, J.J. (1997), *Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations*, Journal of Human Resources, XXXII, 441-462.

[12] Heckman, J.J. and Smith, J. (1995), *Assessing the case for social experiments*, Journal of Economic Perspectives, 9, 2, 85-110

[13] Heckman, J.J. and Hotz, V.J. (1989), *Choosing Among Alternative Nonexperimental Methods for Estimating the Impact of Social Programs: The Case of Manpower Training*, Journal of the American Statistical Association, Vol. 84, No. , 862-874

[14] Heckman, J.J. Ichimura, H. Smith, J. and Todd, P. (1998a), *Characterizing Selection Bias Using Experimental Data*, Econometrica, Vol. 66, No. , 1017-1098

[15] Heckman, J.J. Smith, J. and Taber, C. (1998b), *Accounting for Dropouts in Evaluations of Social Experiments*, The Review of Economics and Statistics, Vol. 80, No. 1, 1-14

[16] Heckman, J.J. Lalonde, R. and Smith, J. (1999), *The Economics and Econometrics of Active Labor Market Programs*, Handbook of Labor Economics, Volume 3, Ashenfelter, A. and Card, D. (eds.), Amsterdam: Elsevier Science

[17] Imbens, G.W. and Angrist, J.D. (1994), *Identification and Estimation of Local Average Treatment Effects*, Econometrica, 62, 467-476.

[18] van der Klaauw, W. (2002), *Estimating the Effect of Financial Aid Offers on College Enrollment: A Regression-Discontinuity Approach*, International Economic Review, 43, 4, 1249-1287

[19] LaLonde, R. (1986), *Evaluating the econometric evaluations of training programs with experimental data*, American Economic Review, Vol. 76, No. , 604-20

[20] Lechner, M. (2001), *A note on the common support problem in applied evaluation studies*, Discussion Paper 2001-01, Department of Economics, University of St. Gallen

[21] Lee, D.S. (2001), *The Electoral Advantage to Incumbency and Voters' Valuation of Politicians' Experience: A Regression Discontinuity Analysis of Elections to the U.S. House*, NBER Working Paper Series 8441

[22] Little, R.J.A. and Yau, L. (1998), *Statistical Techniques for Analyzing Data from Prevention Trials: Treatment of No-Shows Using Rubin's Causal Model*, Psychological Methods, Vol. 3, No. 2, 147-159

[23] Porter, J. (2002), *Asymptotic bias and optimal convergence rates for semiparametric kernel estimators in the regression discontinuity model*, Discussion Paper 1989, Harvard Institute of Economic Research.

[24] Rosenbaum, P.R. (1984), *From Association to Causation in Observational Studies: The Role of Tests of Strongly Ignorable Treatment Assignment*, Journal of the American Statistical Association, Vol. 79, No. 385, 41-48

[25] Rosenbaum, P.R. (1987), *The Role of a Second Control Group in an Observational Study*, Statistical Science, Vol. 2, No. 3, 292-306

[26] Rosenbaum, P.R. and Rubin, D.B. (1983), *The central role of the propensity score in observational studies for causal effects*, Biometrika, Vol. 70, No. , 41-55

[27] Rubin, D.B. (1974), *Estimating causal effects of treatments in randomized and non-randomized studies*, Journal of Educational Psycology, Vol. 66, No. , 688-701

[28] Rubin, D.B. (1977), *Assignment to Treatment Group on the Basis of a Covariate*, Journal of Educational Statistics, Vol. 2, 4-58

[29] Thistlethwaite, D.L. and Campbell, D.T (1960), *Regression discontinuity analysis: an alternative to the ex post facto experiment*, Journal of Educational Psycology, Vol. 51, No. 6, 309-317

[30] Trochim, W. (1984), *Research Design for Program Evaluation: the Regression-Discontinuity Approach*, Beverly Hills: Sage Publications

Another look at the Regression Discontinuity Design

Summary

The attractiveness of the Regression Discontinuity Design (RDD) either in its *sharp* or *fuzzy* formulation rests on close similarities with a formal experimental design. On the other hand, it is of limited applicability since rarely individuals are assigned to the treatment group on the basis of a pre-program measure observable to the analyst. Besides, it only allows to identify the mean impact of the program on a very specific sub-population of individuals. In this paper we show that the *sharp* RDD straightforwardly generalizes to the instances in which the eligibility for the program is established with respect to an observable pre-program measure and eligible individuals can self-select into the treatment group according to an unknown process. This set-up also turns out very convenient to define a specification test on conventional non-experimental estimators of the program effect. Data requirements are made explicit.

Keywords

Program evaluation; Second comparison group; Specification tests.
JEL Classification: C4; C8.

Working Papers già pubblicati

1. E. Battistin, A. Gavosto e E. Rettore, *Why do subsidized firms survive longer? An evaluation of a program promoting youth entrepreneurship in Italy*, Agosto 1998.
2. N. Rosati, E. Rettore e G. Masarotto, *A lower bound on asymptotic variance of repeated cross-sections estimators in fixed-effects models*, Agosto 1998.
3. U. Trivellato, *Il monitoraggio della povertà e della sua dinamica: questioni di misura e evidenze empiriche*, Settembre 1998.
4. F. Bassi*, Un modello per la stima di flussi nel mercato del lavoro affetti da errori di classificazione in rilevazioni retrospettive*, Ottobre 1998.
5. Ginzburg, M. Scaltriti, G. Solinas e R. Zoboli, *Un nuovo autunno caldo nel Mezzogiorno? Note in margine al dibattito sui differenziali salariali territoriali*, Ottobre 1998.
6. M. Forni e S. Paba, *Industrial districts, social environment and local growth. Evidence from Italy*, Novembre 1998.
7. B. Contini*, Wage structures in Europe and in the USA: are they rigid, are they flexible?*, Gennaio 1999.
8. B. Contini, L. Pacelli e C. Villosio, *Short employment spell in Italy, Germany and Great Britain: testing the "Port-of-entry" hypothesis*, Gennaio 1999
9. B. Contini, M. Filippi, L. Pacelli e C. Villosio*, Working careers of skilled vs. unskilled workers*, Gennaio 1999
10. F. Bassi, M. Gambuzza e M. Rasera, *Il sistema informatizzato NETLABOR. Caratteristiche di una nuova fonte sul mercato del lavoro*, Maggio 1999.
11. M. Lalla e F. Pattarin, *Alcuni modelli per l'analisi delle durate complete e incomplete della disoccupazione: il caso Emilia Romagna*, Maggio 1999.
12. A. Paggiaro*, Un modello di mistura per l'analisi della disoccupazione di lunga durata*, Maggio 1999.
13. T. Di Fonzo e P. Gennari*, Le serie storiche delle forze di lavoro per il periodo 1984.1-92.3: prospettive e problemi di ricostruzione*, Giugno 1999.
14. S. Campostrini, A. Giraldo, N. Parise e U. Trivellato*, La misura della partecipazione al lavoro in Italia: presupposti e problemi metodologici di un approccio "time use"*, Ottobre 1999.
15. A. Paggiaro e N. Torelli, *Una procedura per l'abbinamento di record nella rilevazione trimestrale delle forze di lavoro*, Ottobre 1999.
16. A. D'Agostino, G. Ghellini e L. Neri*, A Multiple Imputation Method for School to Work Panel Data*, Ottobre 1999.
17. G. Betti, B. Cheli e A. Lemmi, *Occupazione e condizioni di vita su uno pseudo panel italiano: primi risultati, avanzamenti e proposte metodologiche*, Ottobre 1999.
18. B. Anastasia, M. Gambuzza e M. Rasera*, La durata dei rapporti di lavoro: evidenze da alcuni mercati locali del lavoro veneti*, Marzo 2000.
19. F. Bassi, M. Gambuzza e M. Rasera*, Struttura e qualità delle informazioni del sistema NETLABOR. Una verifica sui dati delle Scica delle province di Belluno e Treviso*, Marzo 2000.
20. N. Rosati, *Permanent and Temporary Inequality in Italy in the 1980s and 1990s*, Marzo 2000.
21. G. Betti, B. Cheli e A. Lemmi, *Analisi delle dinamiche di povertà e disoccupazione su uno pseudo panel italiano*, Marzo 2000.
22. A. D'Agostino, G. Ghellini e L. Neri, *Modelli statistici per l'analisi dei comportamenti di transizione scuola lavoro*, Marzo 2000.

23. A. Paggiaro e U. Trivellato*, Assessing the effects of the "Mobility List" programme in an Italian region: do (slightly) better data and more flexible models matter?,* Marzo 2000.
24. F. Bassi, M. Gambuzza, M. Rasera e E. Rettore, *L'ingresso dei giovani nel mercato del lavoro: prime esplorazioni dall'archivio Netlabor*, Giugno 2000.
25. A. D'Agostino, G. Ghellini e L. Neri, *Percorsi di ingresso dei giovani nel mercato del lavoro*, Giugno 2000.
26. E. Battistin, E. Rettore e U. Trivellato, *Measuring participation at work in the presence of fallible indicators of labour force state,* Giugno 2000.
27. E. Battistin e E. Rettore, *Testing for the presence of a programme effect in a regression discontinuity design with non compliance*, Novembre 2000.
28. A. Ichino, M. Polo e E. Rettore, *Are judges biased by labor market conditions?*, Novembre 2000.
29. N. Rosati, *Further results on inequality in Italy in the 1980s and the 1990s*, Aprile 2001.
30. F. Bassi, M. Gambuzza e M. Rasera, *Imprese e contratti di assunzione: prime analisi da Netlabor*, Novembre 2001.
31. F. Bassi e U. Trivellato, *Gross flows from the French labour force survey: a reanalysis*, Novembre 2001.
32. A. Borgarello e F. Devicienti, *Trend nella distribuzione dei salari italiani 1985-1996*, Novembre 2001.
33. B. Contini, *Earnings mobility and labor market segmentation in Europe and USA: preliminary explorations*, Novembre 2001.
34. B. Contini e C. Villosio, *Job changes and wage dynamics*, Novembre 2001.
35. A. Borgarello, F. Devicienti e C. Villosio, *Mobilità retributiva in Italia 1985-1996*, Novembre 2001.
36. L. Pacelli, *Fixed term contracts, social security rebates and labour demand in Italy*, Novembre 2001.
37. B. Anastasia, M. Gambuzza e M. Rasera, *Le sorti dei flussi: dimensioni della domanda di lavoro, modalità di ingresso e rischio disoccupazione dei lavoratori extracomunitari in Veneto*, Novembre 2001.
38. N. Torelli e A. Paggiaro, *Estimating transition models with misclassification*, Novembre 2001.
39. G. Barbieri, P. Gennari e P. Sestito, *Do public employment services help people in finding a job? An evaluation of the italian case*, Novembre 2001.
40. A. Giraldo, E. Rettore e U. Trivellato, *The persistence of poverty: true state dependence or unobserved heterogeneity? Some evidence form the Italian survey on household income and wealth*, Novembre 2001.
41. A. Giraldo, E. Rettore e U. Trivellato, *Attrition bias in the bank of Italy's survey on household income and wealth*, Novembre 2001.
42. F. Devicienti, *Estimating poverty persistence in Britain*, Novembre 2001.
43. B. Contini, F: Cornaglia, C. Malpede, E. Rettore, *Measuring the impact of the Italian CFL programme on the job opportunities for the youths,* Novembre 2002.
44. E. Battistin, E. Rettore, *Another look at the regression discontinuity design,* Novembre 2002.
