Regression-Discontinuity Analysis

Wilbert van der Klaauw
Federal Reserve Bank of New York[1]

The regression discontinuity (RD) data design is a quasi-experimental evaluation approach first introduced by Thistlethwaite and Campbell (1960) as an alternative method for evaluating social programs. The design is characterized by a treatment assignment or selection rule which involves the use of a known *cutoff* point with respect to a continuous variable, generating a discontinuity in the probability of treatment receipt at that point. Under certain comparability conditions, a comparison of average outcomes for observations just left and right of the cutoff can be used to estimate a meaningful causal impact. While interest in the design had previously been mainly limited to evaluation research methodologists (Cook and Campbell 1979; Trochim, 1984), the design is currently experiencing a renaissance among econometricians and empirical economists (Hahn et al, 1999, 2001; Angrist and Krueger, 1999; Porter 2003). Among the main econometric contributions have been the formal derivation of identification conditions for causal inference and the introduction of semi-parametric estimation procedures for the design. At the same time, a large and rapidly growing number of empirical applications are providing new insights into the applicability of the design, which have led to the development of several sensitivity and validity tests.

The popularity of the RD design in applied economic research can be linked to several of its features. First, the assignment rules in many existing programs and procedures for allocating social resources, frequently lend themselves to RD evaluations. In many cases, program resources are allocated based on some type of formula that has a cutoff structure. One area of economic research where the design has proven especially fruitful in recent years has been the evaluations of educational interventions. Education programs are frequently assigned to schools or students who score below a cutoff on some scale (student performance, poverty), and school and program funding decisions are often based on allocation formulas containing discontinuities. Similarly, the design has proven useful in evaluating the socioeconomic impacts of a diverse set of government programs and laws, many of which use eligibility cutoffs or funding formulas involving thresholds in allocating scarce resources to those potential recipients who need or deserve them most (see van der Klaauw 2006). A second attractive feature of the design is that it is intuitive and its results can be easily communicated, often with a visual portrayal of sharp changes in both treatment assignment and average outcomes around the cutoff value of the assignment variable (Bloom, 2005). Third, a researcher can choose from among several different estimation methods to estimate effects that have credible causal interpretations (Hahn et al, 2001).

Consider the general problem of evaluating the impact of a binary treatment on an outcome variable, using a random sample of individuals where for each individual i we

[1] The views and opinions offered in this article do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve System as a whole.

observe an outcome measure y_i and a binary treatment indicator t_i, equal to one if treatment was received and zero otherwise. The evaluation problem that arises in determining the effect of t on y, is due to the fact that each individual either receives or does not receive treatment and is never observed in both states. Let $y_i(1)$ be the outcome given treatment, and $y_i(0)$ the outcome in absence of treatment. Then the actual outcome we observe equals $y_i = t_i y_i(1) + (1 - t_i) y_i(0)$. A common regression model representation for the observed outcome can then be written as

$$y_i = \beta + \alpha_i t_i + u_i \tag{1}$$

where $\alpha_i = y_i(1) - y_i(0)$ and $y_i(0) = E[y_i(0)] + u_i = \beta + u_i$. Non-random assignment or selection into treatment implies that a comparison of average outcomes of treatment recipients and non-recipients would generally not provide us with a valid treatment effect estimate.

Hahn, Todd and Van der Klaauw (2001, HTV in what follows) analyzed the conditions under which a discontinuity in the treatment assignment or selection rule can be exploited to solve the selection bias problem and to identify a meaningful causal effect. Following Trochim (1984) they distinguish between two different forms of the design, depending on whether the treatment assignment is related to the assignment variable by a deterministic function (**sharp design**) or a stochastic one (**fuzzy design**). In case of a sharp RD design, individuals are assigned to or selected for treatment solely on the basis of a cutoff score on an observed continuous variable x. This variable, alternatively called the assignment, selection, running or ratings variable, could represent a single characteristic or a composite variable constructed using several characteristics. Those who fall below some distinct cutoff point $\bar{x}$ are placed in the control group ($t_i = 0$), while those on or above that point are placed in the treatment group ($t_i = 1$) (or vice versa). Thus, assignment occurs through a known and measured deterministic decision rule: $t_i = t(x_i) = 1\{x_i \geq \bar{x}\}$ where $1\{.\}$ is the indicator function. As the assignment variable itself may be correlated with the outcome variable, the assignment mechanism is clearly not random.

However, if we have reason to believe that persons close to the threshold with very similar x values are comparable, then we may view the design as almost experimental near $\bar{x}$, suggesting that we could evaluate the causal impact of treatment by comparing the average outcome for those with ratings just above to those with ratings just below the cutoff. More formally, consider the following *Local Continuity (LC) assumption*:

$E[u_i|x]$ and $E[\alpha_i|x]$ are continuous in x at $\bar{x}$, or equivalently, $E[y(1)|x]$ and $E[y(0)|x]$ are continuous at $\bar{x}$,

then assuming that the density of x is positive in a neighborhood containing $\bar{x}$,

$$\begin{aligned} \lim_{x \downarrow \bar{x}} E[y_i|x] - \lim_{x \uparrow \bar{x}} E[y_i|x] &= \lim_{x \downarrow \bar{x}} E[\alpha_i t_i|x] - \lim_{x \uparrow \bar{x}} E[\alpha_i t_i|x] + \lim_{x \downarrow \bar{x}} E[u_i|x] - \lim_{x \uparrow \bar{x}} E[u_i|x] \\ &= E[\alpha_i|x = \bar{x}]. \end{aligned} \tag{2}$$

The RD approach therefore identifies the average treatment effect for individuals close to the discontinuity point. Note that the continuity assumption formalizes the idea that individuals just above and below the cutoff need to be 'comparable', requiring them to have similar average potential outcomes when receiving and when not receiving treatment. While in absence of additional assumptions (such as a common effect assumption) one could only learn about treatment effects for a subpopulation of persons near the discontinuity point, as pointed out by HTV this local effect is highly relevant to policymakers who are contemplating less restrictive eligibility rules and marginal expansions of programs via a change in the cutoff.

The continuity assumption required for identification is not innocuous. Even if treatment receipt is determined solely on the basis of a cutoff score on the assignment variable, this is not a sufficient condition for the identification of a meaningful causal effect. The continuity assumption rules out coincidental functional discontinuities in the x-y relationship such as those caused by other programs employing assignment mechanisms based on the exact same assignment variable and cutoff. In addition, the continuity restriction generally rules out certain types of behavior both on the part of potential treatment recipients who exercise control over their value of x, as well as program administrators in choosing the assignment variable and cutoff point. Lee (2005) analyzes the conditions under which an ability to manipulate the assignment variable may invalidate the RD identification assumptions. He shows in the context of a sharp RD design that as long as individuals do not have *perfect* control over the position of the assignment variable relative to the cutoff score, the continuity assumption will be satisfied.

While in the sharp RD design treatment assignment is known to depend on the selection variable x in a deterministic way, in case of a fuzzy design (Campbell 1969), treatment assignment depends on x in a stochastic manner but in such a way that the propensity score function $Pr(t = 1|x)$ is again known to have a discontinuity at $\bar{x}$. Instead of a 0-1 step function, the selection probability as a function of x would now contain a jump smaller than 1 at $\bar{x}$. The fuzzy design can occur in case of mis-assignment relative to the cutoff value in a sharp design, with values of x near the cutoff appearing in both treatment and control groups. This situation is analogous to having no-shows (treatment group members who do not receive treatment) and/or crossovers (control group member who do receive the treatment) in a randomized experiment. This could occur if in addition to the position of the individual's score relative to the cutoff value, assignment is based on additional variables observed by the administrator, but unobserved by the evaluator.

A comparison of average outcomes of recipients and non-recipients, even if near the cutoff, would not generally lead to correct inferences regarding an average treatment effect. However, as shown by HTV, one can again exploit the discontinuity in the selection rule to identify a causal impact of interest by noting that under the LC assumption and with a locally constant treatment effect ($\alpha_i = \alpha$ in a neighborhood around $\bar{x}$), the treatment effect α is identified by

$$\frac{\lim_{x\downarrow\bar{x}} E[y_i|x] - \lim_{x\uparrow\bar{x}} E[y_i|x]}{\lim_{x\downarrow\bar{x}} E[t_i|x] - \lim_{x\uparrow\bar{x}} E[t_i|x]}, \qquad (3)$$

where the denominator is always nonzero because of the known discontinuity of $E[t|x]$ at $\bar{x}$.

In case of varying treatment effects, HTV show that under the local continuity assumption, and a local conditional independence assumption requiring t_i to be independent of α_i conditional on x near $\bar{x}$, the ratio above identifies $E[\alpha_i|x = \bar{x}]$, the average treatment effect for cases with values of x close to $\bar{x}$. The conditional independence assumption is a strong assumption which may be violated if individuals self-select into or are selected for treatment on the basis of expected gains from treatment. HTV show that under a weaker local monotonicity assumption, similar to that assumed by Imbens and Angrist (1994), the ratio (3) will instead identify a local average treatment effect (LATE) at the cutoff point, which represents the average treatment effect of the 'compliers', that is the subgroup of individuals whose treatment status would switch from non-recipient to recipient if their score x crossed the cutoff. More recently Battistin and Rettore (2003) considered the special case where an eligibility rule divides the population into eligibles and non-eligibles according to a sharp RD design, and with eligible individuals self-selecting into treatment. In this case the LC assumption alone is sufficient for the ratio to identify $E[\alpha_i|t_i = 1, x = \bar{x}]$, the average treatment effect of the treated, for those near the cutoff.

As indicated by these identification results, estimation of treatment effects in an RD design involves estimating boundary points of conditional expectation functions. The most common empirical strategy in the literature has been to adopt parametric specifications for the conditional expectations functions. Consider the following alternative representation of outcome equation (1) in case of a sharp RD design:

$$y_i = m(x_i) + \delta t_i + e_i, \tag{4}$$

where $e_i = y_i - E[y_i|t_i, x_i]$, $t_i = 1\{x_i \geq \bar{x}\}$, $m(x) = E[u_i|x] + (E[\alpha_i|x] - E[\alpha_i|\bar{x}])1\{x \geq \bar{x}\}$. Then under the local continuity assumption $m(x)$ will be a continuous function of x at $\bar{x}$, and $\delta = E[\alpha_i|\bar{x}]$ (the average treatment effect at $\bar{x}$) will measure the discontinuity in the average outcome at the cutoff. This suggests that if the correct specification of $m(x)$ were known, and was included in the regression, we could consistently estimate the treatment effect for the sharp RD design. This idea of including a specification of $m(x)$ in the regression of y on t in order to correct for selection bias caused by selection on observables, is in the econometrics literature known as the *control function approach* (Heckman and Robb, 1985). A popular choice among empirical researchers has been to use global polynomials or to use splines (piecewise polynomials) which, even though globally continuous, have a knot at the cutoff (Trochim 1984, van der Klaauw 2002, McCrary 2005).

In case of a fuzzy RD design, when assuming local independence of t_i and α_i conditional on x, then in a neighborhood of $\bar{x}$,

$$y_i = m(x_i) + \delta E[t_i|x_i] + w_i, \tag{5}$$

where $w_i = y_i - E[y_i|x_i]$ and $m(x) = E[u_i|x] + (E[\alpha_i|x] - E[\alpha_i|\bar{x}])E[t|x]$. With the local continuity assumption again implying that $m(x)$ will be continuous at the cutoff, and with $E[t_i|x_i]$ being discontinuous at $\bar{x}$, δ in this regression will measure the ratio in (3), which in this case equals the average local treatment effect $E[\alpha_i|\bar{x}]$. Similarly, δ can be interpreted as a local average treatment effect if we replaced the local independence assumption with the local monotonicity condition of Imbens and Angrist (1994).

This naturally leads to the two-stage procedure adopted by van der Klaauw (2002), where in the first stage we estimate the propensity score function specified as $t_i = E[t_i|x_i] + v_i = f(x_i) + \gamma 1\{x_i \geq \bar{x}\} + v_i$ where $f(\cdot)$ is continuous in x at $\bar{x}$ and γ measures the discontinuity in the propensity score function at $\bar{x}$. In the second stage the control function-augmented outcome equation is then estimated with t_i replaced by the first-stage estimate of $E[t_i|x_i] = Pr[t_i = 1|x_i]$ as in Maddala and Lee (1976). With correctly specified $f(x)$ and $m(x)$ functions, this two-stage procedure yields a consistent estimate of the treatment effect. The approach is similar in spirit to those proposed earlier in the RD evaluation literature by Spiegelman (1979) and Trochim and Spiegelman (1980). Note that in case of a parametric approach, if we assume the same functional form for $m(x)$ and f(x), then the two-stage estimation procedure described here will be equivalent to two-stage least squares (in case of linear-in-parameter specifications) with $1\{x_i \geq \bar{x}\}$ and the terms in $m(x)$ serving as instruments. Because of the popularity of this particular parameterization, the RD approach is often interpreted as being equivalent to an Instrumental Variable approach, as it implicitly imposes an exclusion restriction by excluding $1\{x_i \geq \bar{x}\}$ as a variable in the outcome equation.

Valid parametric inference for the estimation of the treatment effect requires a correct specification of the control function $m(x)$ and of $f(x)$ in the treatment equation. To mitigate the potential for misspecification bias, several semi-parametric estimation procedures have been proposed for estimating $m(x)$ and $f(x)$, or equivalently for estimating the limits $\lim_{x\downarrow\bar{x}} E[z|x]$ and $\lim_{x\uparrow\bar{x}} E[z|x]$ in (3) semi-parametrically. These methods rely on less-restrictive smoothness conditions away from the discontinuity, with estimates based mainly on data in a neighborhood on either side of the cutoff point. Asymptotically this neighborhood needs to shrink as with usual non-parametric estimation, implying that we should expect a slower than parametric rate of convergence in estimating treatment impacts. HTV considered the use of Kernel and local linear regression estimators, while Porter (2003) proposed estimating the limits using local polynomial regression and partially linear model estimation. RD estimators based on local polynomial regression and partially linear model estimation have better boundary behavior than the Kernel-based estimator and as shown by Porter, achieve the optimal rate of convergence. This result is based on a known degree of smoothness of the conditional expectation functions. Sun (2005) proposed an adaptive estimator to first estimate the degree of smoothness in the data prior to implementing either estimator.

The internal validity of the RD approach relies on local continuity of conditional expectations of potential outcomes near the discontinuity point. While this assumption is fundamentally untestable, a number of validity tests have been developed to bolster the credibility of the RD design. First, economic behavior may lead to sorting of individuals around the cutoff point, where those below the cutoff may differ on average from those just above the cutoff. Such precise sorting around the cutoff would generally be accompanied by a discontinuous jump in the density of the assignment variable at the cutoff. Several approaches have been used for assessing this possibility (McCrary 2005; Lee 2005; Chen and van der Klaauw 2005; Lemieux and Milligan, 2005). Second, one can test for evidence that individuals on either side of the cutoff are observationally similar, by directly comparing average characteristics (McEwan and Urquiola, 2005) or by repeating the RD analysis with the covariate as outcome variable (van der Klaauw 2005). Alternatively, one can test for an imbalance of relevant characteristics by assessing the sensitivity of RD estimates to the inclusion of observed characteristics as controls (van der Klaauw 2002, Lee 2005). Third, in some applications data are available from a baseline period in which the program did not yet exist, or for a group of individuals that was not eligible for treatment. In such a case the credibility of the design can be significantly enhanced by repeating the RD analysis with such data. Finding a zero treatment effect in such a falsification test would suggest that a nonzero post-program effect was not an artifact of the specific RD model specification, estimation approach chosen or caused by another program using the same cutoff and assignment variable.

Finally, while this exposition has focussed on the binary treatment case with a selection rule containing a single discontinuity at a known cutoff, the approach can be readily extended to one where there are multiple treatment dose levels and multiple cutoffs or 'cutoff ranges' within which the treatment dose varies continuously (van der Klaauw 2006). Similarly, the approach can be modified to cover cases where the assignment or selection variable is discrete instead of continuous (Lee and Card 2006).

References

[1] Angrist, Joshua D. and Alan B. Krueger, "Empirical Strategies in Labor Economics", in O. Ashenfelter and D. Card (eds.), *Handbook of Labor Economics*, ed 1, vol 3, chapter 23, pages 1277-1366, Elsevier, 1999.

[2] Battistin, Erich and Enrico Rettore, "Another look at the Regression Discontinuity Design", CeMMAP working papers from Centre for Microdata Methods and Practice, Institute for Fiscal Studies, No CWP01/03, 2003.

[3] Bloom, Howard S., James Kemple, Beth Gamse and Robin Jacob, "Using Regression Discontinuity Analysis to Measure the Impacts of Reading First", Paper presented at the annual conference of the American Educational Research Association, Montreal Canada, 2005.

[4] Campbell, Donald T., "Reforms as Experiments", *American Psychologist*, 24, 1969, 409-429.

[5] Chen, Susan and Wilbert van der Klaauw, "The Work Disincentive Effects of the Disability Insurance Program in the 1990s", working paper UNC-Chapel Hill, 2005, forthcoming in *Journal of Econometrics*.

[6] Cook, T.D and Campbell, D.T., *Quasi-Experimentation: Design and Analysis Issues for Field Settings*. Boston: Houghton-Mifflin, 1979.

[7] Hahn, Jinyong, Petra Todd and Wilbert van der Klaauw, "Evaluating the Effect of an Antidiscrimination Law using a Regression-Discontinuity Design", NBER working paper 7131, May 1999.

[8] Hahn, Jinyong, Petra Todd and Wilbert van der Klaauw, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", *Econometrica* 69, 2001, 201-209.

[9] Heckman, James J. and Richard Robb, "Alternative Methods for Evaluating the Impact of Interventions" in J. Heckman and B. Singer, eds., *Longitudinal Analysis of Labor Market Data*, Cambridge University Press, New York, 1985, 156-245.

[10] Imbens, Guido W. and Joshua Angrist, "Identification and Estimation of Local Average Treatment Effects", *Econometrica* 62(2), 1994, 467-476.

[11] Lee, David S., "Randomized Experiments from Non-random Selection in U.S. House Elections", working paper University of California, Berkeley, 2005, forthcoming in *Journal of Econometrics*.

[12] Lee, David S. and David Card, "Regression Discontinuity Inference with Specification Error", NBER Technical Working Paper No. 322, March 2006.

[13] Lemieux, Thomas and Kevin Milligan, "Incentive effects of social assistance: A regression discontinuity approach", NBER Working Paper, No. 10541, 2004.

[14] McCrary, Justin, “Testing for Manipulation of the Running Variable in the Regression Discontinuity Design”, working paper, University of Michigan, 2005, forthcoming in *Journal of Econometrics.*

[15] McEwan, Patrick J. and Miguel Urquiola, “Economic Behavior and the regression-discontinuity design: Evidence from class size reduction”, working paper, Columbia University, 2005.

[16] Porter, J., “Estimation in the Regression Discontinuity Model”, unpublished manuscript, Harvard University, 2003.

[17] Spiegelman, C., “Estimating the effect of a large scale pretest posttest social program”, *Proceedings of the Social Statistics Section*, American Statistical Association, 1979, 370-373.

[18] Sun, Yixiao, “Adaptive Estimation of the Regression Discontinuity Model”, working paper, University of California, San Diego, 2005.

[19] Thistlethwaite, D., and D. Campbell, “Regression-Discontinuity Analysis: An alternative to the ex post facto experiment”, *Journal of Educational Psychology* 51, 1960, 309-317.

[20] Trochim, W.K., *Research Design for Program Evaluation: the Regression-Discontinuity Approach*, Beverly Hills, CA: Sage Publications, 1984.

[21] Trochim, W. and C. Spiegelman, “The Relative Assignment Variable Approach to Selection Bias in Pretest-Posttest Group Designs”, *Proceedings of the Survey Research Section*, American Statistical Association, 1980, 376-380.

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

[23] van der Klaauw, W., “Breaking the Link Between Poverty and Low Student Achievement: An Evaluation of Title I”, working paper UNC-Chapel Hill, 2005, forthcoming in *Journal of Econometrics.*

[24] van der Klaauw, W., “Regression-Discontinuity Analysis: A Review”, unpublished manuscript, Federal Reserve Bank of New York, 2006.