Chapter 2
Regression Discontinuity Design: When Series Interrupt

Abstract This chapter introduces the identification and estimation of policy effects when outcome variables can be ordered according to a given variable and when the treatment occurs at a given point. The structural change occurring in the outcome variable is in such a case assumed to be the effect of the policy. The assumptions behind Regression Discontinuity Design are hence discussed alongside with extensions for heterogeneous effect.

Keywords Regression discontinuity design · LATE · Heterogeneous LATE · Geographical discontinuities

2.1 Introduction

The evaluation of policies is driven by the identification of the causal effect of interventions by considering an exogenous variation in the assignment to the treatment.

When observed units can be ordered according to a running variable and then the treatment is assigned above a given threshold, a Regression Discontinuity Design (RDD) can be used. A policy introduced at a given point in time has time as a running variable and the day/month of the introduction as the threshold. European Union Cohesion Policy is particularly generous toward Objective 1/convergence regions, i.e. those regions with a GDP per capita lower than 75 % of EU average. In this case GDP per capita is the running variable (i.e. the variable used to order regions according to their GDP), whereas 75 % of the EU average is the threshold.

The discontinuity design relates to situations where the probability of enrolment into treatment changes discontinuously with some continuous variable. In particular, in sharp design, the probability of receiving the treatment is 1 below a given threshold and 0 above (or vice versa), that is $P(T = 1|x < x*) = 1$ and $P(T = 1|x \geq x*) = 0$ in the case the forcing variable is x and the threshold is x*. In fuzzy design, the probability of receiving the treatment increases (or decreases) with x and shows a discontinuity at the given point x*, that is $1 > P(T = 1|x < x*) > P(T = 1|x \geq x*)$. In this case, the selection variable influences but does not completely determine participation in the treatment. In other words, the jump at x* is smaller than 1.

M. Percoco, *Regional Perspectives on Policy Evaluation*, SpringerBriefs in Regional Science, DOI 10.1007/978-3-319-09519-6_2

In sharp design, the implicit assumption is that the assignment process is driven only by the observable x, so that the discontinuity creates a randomized experiment around the threshold x*. Units on the two sides of the threshold but close to it are expected to be very similar. A difference in their outcomes can within reason be attributed to the treatment:

$$ATT = E(Y|x \in [x*, x * +\delta]) - E(Y|x \in [x*, x * -\delta])$$

However, the discontinuity only identifies the effect at x* and only for "small values" of δ (i.e., for $\delta \to 0$), RDD can reasonably estimate the effect of a policy, that is, RDD estimates a local ATT.

2.2 The Basic Framework

RDD allows for taking into abbount of observed as well as unobserved heterogeneity in the estimation of the treatment effect when there is an eligibility rule for the treatment based on an observable variable x. Indeed, the principle underlying this strategy is that observations just below and above the threshold are likely to be very similar to each other with respect to observed and unobserved characteristics, except for the outcome. Therefore, the mean difference in the outcomes can be attributed to the treatment effect. This average treatment effect (ATE) sacrifices external validity by focusing only on observations close to the cut-off point.

Regression discontinuity may be sharp if the eligibility rule is strictly adhered so that given the threshold level $x*$, the probability of treatment T is $P(T = 1|x < x*) = 1$ and $P(T = 1|x > x*)$ =0. Whenever the rules are not applied sharply, the RDD is said to be fuzzy.

More formally, let y_0 and y_1 denote the counterfactual outcomes without and with treatment T, let x be the forcing variable and consider the following assumptions:

A1. $E(y_g|T, x) = E(y_g|x)$, $g = 0, 1$
A2. $E(y_g|x)$, $g = 0$, 1is continuous at $x = x*$
A3. $P(T = 1|x) \equiv F(x)$ is discontinuous at $x = x*$, i.e. the propensity score of the treatment has a discrete jump at $x = x*$.

In the fuzzy RDD the discontinuity is used as an instrumental variable for treatment status. Following Imbens and Lemieux (2008) the goal is to estimate the parameter ρ on treatment with the following form form:

$$y_{i,T} = \theta + \rho T_i + f(\tilde{x}_{i,T}) + \eta_i \tag{2.1}$$

where $y_{i,T}$ is in our case the outcome of region i whose treatment status is T, θ is a constant, $\tilde{x}_{i,T}$ is the forcing variable properly normalized. Consequently, ρ expresses the impact of the treatment at $x_{i,T} = x_0$. The $f(\tilde{x}_{i,T})$ term is a p-th order parametric polynomial the parameters of which are allowed to differ on the left and the right of the cut-off point (Angrist and Pischke 2009) in order to account for non-linearity in the outcome variable. Lastly η_i is an error term.

Applying OLS estimation to Eq. (2.1) will lead to a biased estimate of the treatment effect in the case of fuzziness of the treatment variable (Imbens and Lemieux 2008; Lee and Lemieux 2010). The treatment dummy T can be instrumented by a first stage regression which takes the form of :

$$T_i = \alpha + \beta R_i + f(\tilde{x}_i) + \varepsilon_i \tag{2.2}$$

where α , β, δ are unknown parameters and ε_i and ν_i are disturbances. The variable R_i denotes the treatment that the unit would have been assigned had the eligibility rule been strictly followed.

In order to have a causal interpretation of the 2SLS the instrument R_i must affect the treatment ($\mathrm{Cov}(R_i T_i) \neq 0$), and it must fulfil the exclusion restriction $\mathrm{Cov}(R_i \ \eta_i) = 0$. The last assumption is that the instrument R_i is independent of the vector of potential outcomes and potential treatment assignments, formally $[y_i(T, R) \forall T, R, T_{i,0}, T_{i,1}] \perp R_i$.

2.2.1 *Example: The Effect of the Point-Record Driving License on Car Accidents*

The penalty points system is a mechanism introduced in Italy on Tuesday 1 July 2003. Each driver was initially awarded 20 points: in the case of infringement of the rules of the road the driver will lose some points and will have to pass a theory test and a driving test, should they lose all their points (the loss of all points cause the automatic termination of the driver's license). The number of points deducted from the license was established by law and varied depending on the severity of the infringement.

Let us assume evaluation of the introduction of the policy in terms of accidents. In this case, our running variable is time, so that we can order observations according to their temporal occurrence. Our approach for estimating the effect of the introduction of the point-record driving license consists of estimating an eventual break in the trend of road accidents in correspondence of the policy adoption in July 2003. In particular, we make use of monthly data; as such our estimation is complicated by seasonality. To deal with this issue, we can estimate the following general model:

$$accidents_{mt} = \sum_{m=1}^{11} d_m + \alpha trend_t + \beta I_{mt} + I_{mt}\left(\sum_{m=1}^{11} d_m + \alpha trend_t\right)$$

where the dependent variable $accidents_{mt}$ indicates the number of accidents occurring in month m in year t, $\sum_{m=1}^{11} d_m$ is a full set of month-specific dummy variables used to take into account seasonality in the data, $trend_t$ is a time trend and I_{mt} is a a dummy variable taking the value of 1 after July 2003 and 0 otherwise, hence it indicates an eventual departure from the trend occurring after the point-record driving license was implemented. Element $I_{mt}\left(\sum_{m=1}^{11} d_m + \alpha trend_t\right)$ indicate that parameters to be estimated are allowed to change on the left and on the right of the threshold represented by July 2003. Finally, the equation is also estimated by considering quadratic trend, in order to deal with possible confoundedness with the parameter of interest β due to non-linearity. Concerning our dependent variables, we will make use of either the monthly number of deaths, of injuries or the total number of accidents over the period 1991–2009.

Table 2.1 reports monthly averages for our three outcome variables. No effect of the policy is reported for the total number of accidents and injuries, whereas a drop of about 20 % is observed in the case of deaths.

Table 2.2 hence reports OLS estimates for our outcome variables across different specifications and time period.[1] The upper panel, in particular, contains policy impact estimates when the dependent variable is the total number of accidents. By considering the pre-treatment average reported in Table 2.1, it emerges that the introduction of the points-record driving license has reduced the number of accidents by 58–66 per month, a contraction of about 0.22–0.25 % with mild significance in terms of

Table 2.1 Summary statistics

	Before the treatment	After the treatment
Accidents	25,698.52	26,727.1
Injuries	17,871.38	18,969.54
Deaths	578.6533	437.8205

Table 2.2 Regression estimates (OLS)

	Whole sample	1997–2009	With Eurocoin as a control
	Accidents		
Treatment	−58.368**	−58.368**	−66.776**
	(25.060)	(25.060)	(30.211)
Observations	120	120	120
	Injuries		
Treatment	−49.817**	−49.817**	−55.385**
	(21.590)	(21.590)	(25.474)
Observations	120	120	120
	Deaths		
Treatment	−40.412***	−41.660***	−51.094***
	(4.426)	(7.849)	(12.306)
Observations	228	156	156

Note Standard errors in parentheses are clustered by month. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

[1] For simplicity, we admittedly omit the issue that data are counts and not continuous.

the parameters. Similar estimates were obtained in the case of injuries. Interestingly, estimates of the policy impact was more significant and hence more reliable in the case of fatalities, with a contraction of 41–51 in the number of death cases per month, corresponding to a decrease of 7.1–8.8 % with respect to the pre-treatment period.

Finally, the third column in Table 2.2 reports estimates when the EUROCOIN coincident indicator of economic activity was used as a control for the business cycle. In this case as well, results were almost robust.

Taken together, our estimates point to a significant effect of the policy, i.e. dramatically reducing the severity of road accidents, with a small effect in the case of the number of accidents and injuries.

In particular, our preferred specifications, i.e. those for the period 1997–2009 with control variables, indicate that the introduction of the points-record driving license has reduced the number of accidents by about 801 annually, the number of injuries by 664, and finally the number of fatalities by 612. We can use these estimates to compute the reduction in the external costs of accident due to the introduction of the policy. In particular, the "Handbook on estimation of external costs in the transport sector" proposes, for Italy, a value of life equal to 1,43 million Euros and a cost per injury in the interval 14,100–183,700 Euros, depending on the severity.

As for injury, as we are not able to detect the severity of the avoided cases, we will use a rough benchmark value of 50,000 Euros. Given these average costs, the introduction of the point-record driving license has had a social benefit equal to 875.16 million Euros for avoided deaths and 33.2 million Euros for avoided injuries, with a total social benefit equal to 908.36 million Euros and a present value, over 20 years with a 3.5 % social discount rate, of 12.9 billion Euros. To be noted is the fact that this estimate is a lower bound estimate of social benefits as it does not contain the cost for physical damage to cars.

2.2.2 Example: The Impact of European Cohesion Policy on Regional Growth

To give a practical example of RDD, in what follows, we will focus on the effect of Objective 1 transfers in the European Union, so that the eligibility rule is that all regions with a per capita GDP lower than 75 % of EU average are eligible. In this example, we will use as an outcome variable regional growth of NUTS3 regions, although our identification assumption will rely on the fact that the treatment is assigned at NUTS2 level. In other words, we will instrument the treatment by using a "theoretical rule of assignment to treatment" to indicate whether or not a NUTS3 region has a GDP per capita lower than 75 % of EU average.

To start our RDD exercise, it is convenient to perform a graphical analysis. Following Imbens and Lemieux (2008) the forcing variable has been divided into equally-sized bins of 1.5 % points width to the left and the right of the threshold level. Figure 2.1 plots the outcome variable (i.e. average growth rate) against the forcing variable (i.e. per capita GDP in PPP). Furthermore, a 5-th order polynomial is added. The jump of the outcome variable at the threshold is evident and amounts to about

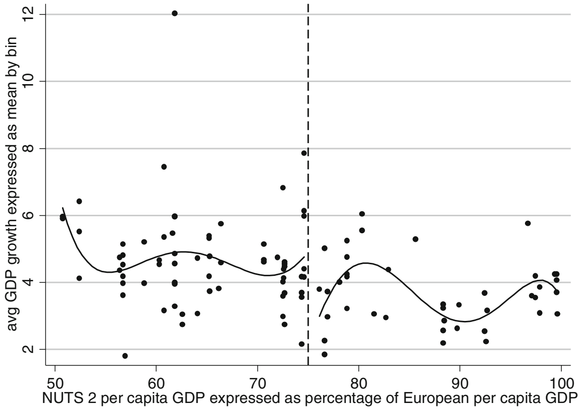

Fig. 2.1 Discontinuity of the outcome at the threshold

Table 2.3 Estimates of the effect of cohesion policy (IV)

	(1)	(2)	(3)
	Baseline	Country FE	With controls
Objective 1	0.758***	1.034***	0.923***
	(0.277)	(0.323)	(0.317)
Observations	1233	1233	998
R. sq	0.338	0.581	0.575

Note Dependent variable is cumulative growth over the period 1999–2008. In model 1 all specifications include a 5th order polynomial in the running variable, that is GDP per capita in 1999. In model 2 we include country dummies, and in model 3 we include population density, employment rate, and the shares of population with secondary and tertiary education, respectively. Robust standard errors in parentheses. Significance levels: *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$

0.8 %. This result suggests that the RDD is a sound approach toward unravelling the effectiveness of the structural funds in promoting the growth of GDP in relatively poorer regions of the EU.

Table 2.3 reports our 2SLS estimation for the whole sample and different groups of regions and different specifications. Model 1 presents the baseline regression where the dependent variable—the average growth rate of GDP over the period under analysis, is related to the treatment status. The Objective 1 status seems to a be significant determinant of the economic performance of NUTS3 regions when the whole sample is taken into account. To further support the reliability of our results, model 2 includes controls for country dummies that should account for any distinctive economic pattern at a national level. Results generally remained consistent with our previous findings. Finally model 3 accounts for a number of additional controls that can be considered standard within the existing literature: population density, employment rate and the share of population with secondary and tertiary education as a proxy for human capital. Additionally, in this case, results remain unchanged.

2.3 Regression Discontinuity with Heterogeneous Effects

RDD, as do most of the evaluation tools, estimates an average effect of the policy, which may hide some interesting heterogeneity driven by certain characteristics of observed units. In other words (and following previous example), the impact of cohesion policy may differ across regions according to local characteristics (e.g. the quality of institutions, economic structure). To deal with this issue, an Heterogeneous LATE (HLATE) can be used. More formally, the heterogeneous local average treatment effect is defined as (Becker et al. 2013):

$$HLATE(x_i = x*, z_i) = HLATE(x*, z_i) = E[y_{i1}|x*, z_i] - E[y_{i0}|x*, z_i] \quad (2.3)$$

where the notation is the same as in Sect. 2.2, whilst z_i is our interaction variable, i.e. the one we hypothesize to drive heterogeneity. The identification of the HLATE in (2.3) needs two further assumptions:

A4. the interaction variable z_i is continuous at $x*$, the threshold;
A5. the interaction variables z_i is uncorrelated with the error term in the outcome equation, conditional on $x*$.

Assuming that the conditional expectation function E$[y_i|x_i, z_i]$ follows an additive process, we can express the two potential outcomes as follows:

$$E[y_{i0}|x_i, z_i] = \alpha + f_0(\tilde{x_i}) + h_0(\bar{z_i}) \quad (2.4)$$

$$E[y_{i1}|x_i, z_i] = E[y_{i0}|x_i, z_i] + \beta + f_1^*(\tilde{x_i}) + h_1^*(\bar{z_i}) \quad (2.5)$$

where α is a constant, β is the coefficient of the treatment dummy, $\tilde{x_i}$, as before, is the deviation from the threshold GDP of region i's GDP while $\bar{z_i}$ is the deviation from the sample mean of region i's interaction variable. The functions $f_0(\tilde{x_i})$, $h_0(\bar{z_i})$, $f_1^*(\tilde{x_i})$ and $h_1^*(\bar{z_i})$ are sufficiently smooth polynomials. They define $f_1(\tilde{x_i})$ and $h_1(\bar{z_i})$ analogously to $f_0(\tilde{x_i})$ and $h_0(\bar{z_i})$ but with the treatment switched on. In addition, it evident that $f_1^*(\tilde{x_i}) = f_1(\tilde{x_i}) - f_0(\tilde{x_i})$ and $h_1^*(\bar{z_i}) = h_1(\bar{z_i}) - h_0(\bar{z_i})$. The equation for generic treatment status can be written as:

$$E[y_i|x_i, z_i] = E[y_{i0}|x_i, z_i] + T_i[\beta + f_1^*(\tilde{x_i}) + h_1^*(\bar{z_i})]. \quad (2.6)$$

With this specification, the LATE is given by β whereas the HLATE is given by $\beta + h_1^*(\bar{z_i})$.

If the RDD is sharp then simple OLS can indeed estimate the parameters without bias using the following specification:

$$y_i = \alpha + f_0(\tilde{x_i}) + h_0(\bar{z_i}) + T_i[\beta + f_1^*(\tilde{x_i}) + h_1^*(\bar{z_i})] + \varepsilon_i \quad (2.7)$$

If the RDD is fuzzy the treatment dummy must be instrumented, for the reasons already mentioned, against the *rule* dummy indicating whether region i satisfies the eligibility criteria or not, and the exogenous variables of the model, therefore the first stage of the 2SLS is given by:

$$T_i = g_0(\tilde{x_i}) + l_0(\bar{z_i}) + R_i[\delta + g_1^*(\tilde{x_i}) + l_1^*(\bar{z_i})] + \nu_i \qquad (2.8)$$

where all the variables have the same notation and the polynomial functions are defined as above. Substituting (2.8) in (2.7) we obtained the reduced form for the fuzzy RDD.

2.3.1 Example: EU Cohesion Policy, Economic Structure and Regional Growth

In the example in Sect. 2.2.1, RDD was used to estimate a LATE of cohesion policy. However, as pointed out by Percoco (2013), the effect of regional development policies is likely to be heterogeneous, depending on local economic structure. In what follows we will consider the case of the share of the service sector (in terms of the gross value added, GVA) as an interaction variable.

Table 2.4 shows the summary statistics of the interaction variable. The first three rows present similar mean and standard deviation, although one sub-sample is double the size of the other, whereas the two other sub-samples have similar sizes but different means highlighting the heterogeneity of the Objective 1 transfers treatment based on the extent of the service sector. Indeed our example will make use of these two sub-samples to estimate the impact of the policy.

Results are showed in Table 2.5, the columns of which refer to the degree of the polynomial in the forcing variable, initial per capita GDP in PPP, while in the horizontal dimension shows the three different specifications of the polynomial in the interaction variable, the regional GVA coming from the tertiary sector as a share of the total regional GVA (SERV). Recall that both variables have previously been centred, the first at the threshold level while the latter at the sample mean. Estimates of the parameters of the forcing variable polynomials, i.e. ρ's, have been omitted for the sake of clarity and simplicity.

Table 2.4 Summary Statistics of service GVA share at time of commission decision

Sample	Observations	Mean	St. Dv.	Min	Max
Whole	1,080	0.653	0.096	0.226	0.935
Below threshold	365	0.644	0.093	0.405	0.883
Above threshold	715	0.658	0.097	0.226	0.935
Below sample mean	556	0.579	0.059	0.226	0.653
Above sample mean	524	0.732	0.057	0.653	0.935

Table 2.5 Objective 1 and tertiary sector: HLATE

	3rd order polyn	4th order polyn	5th order polyn
linear SERV	(1)	(2)	(3)
Object1	0.371	0.407	0.534*
	(0.226)	(0.271)	(0.301)
Object1 × SERV	−2.429***	−2.392***	−2.396***
	(0.812)	(0.812)	(0.803)
SERV	−0.948*	1.302**	−1.561**
	(0.573)	(0.590)	(0.609)
Const.	3.627***	3.639***	3.848***
	(0.197)	(0.250)	(0.290)
Obs.	1080	1080	1080
R^2	0.349	0.354	0.368
quadratic SERV	(1)	(2)	(3)
Object1	0.476**	0.436	0.355
	(0.228)	(0.285)	(0.318)
Object1 × SERV	2.466***	−2.366***	−2.241***
	(0.803)	(0.806)	(0.805)
Object1 × SERV^2	−7.134	−7.696	−7.719
	(5.747)	(5.764)	(5.774)
SERV	−0.960	−1.242*	−1.781***
	(0.624)	(0.635)	(0.663)
SERV^2	−0.844	−5.333	−17.69**
	(6.240)	(6.453)	(7.609)
Const.	3.606***	3.637***	3.824***
	(0.201)	(0.264)	(0.308)
Obs.	1080	1080	1080
R^2	0.350	0.354	0.358
cubic SERV	(1)	(2)	(3)
Object1	0.426*	0.417	0.332
	(0.227)	(0.286)	(0.318)
Object1 × SERV	−5.457***	−5.194***	−5.316***
	(1.383)	(1.390)	(1.388)
Object1 × SERV^2	−1.305	−2.178	−1.771
	(6.221)	(6.241)	(6.228)
Object1 × SERV^3	118.4**	110.4**	122.4***
	(46.44)	(46.60)	(46.67)
SERV	−0.787	−0.949	−0.550
	(0.864)	(0.858)	(0.856)
SERV^2	−1.573	−3.722	−21.38**
	(7.033)	(7.181)	(8.663)

(continued)

Table 2.5 (continued)

	3rd order polyn	4th order polyn	5th order polyn
linear SERV	(1)	(2)	(3)
$SERV^3$	−15.03	−12.57	−52.51*
	(28.16)	(27.66)	(28.84)
Const.	3.624***	3.615***	3.801***
	(0.201)	(0.263)	(0.306)
Obs.	1080	1080	1080
R^2	0.354	0.357	0.362

Note Standard errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

The first striking result was that the treatment, Object1, was not significant per se, instead its interactions with SERV and $SERV^3$ were very significant, meaning that there was indeed an heterogeneity of the treatment according to the level of the SERV variable. Results for Object1 and its interactions remained similar across the columns, which highlights the different specifications of the polynomial in the forcing variable, meaning that a higher order of initial GDP cannot explain the impact of the service share of GVA on the economic growth of the treated regions. The other striking result was that the interactions had negative sign, meaning that a tertiary GVA above the mean, centred at 0, reduced the impact of the transfers on the per capita GDP growth, whereas a service GVA below the mean, i.e. a negative value of SERV, made the transfers more effective.

As an example let us consider the case in column (3) with a cubic SERV polynomial for a treated region whose level of SERV is 0.1, roughly the same as the sample standard deviation. This means that the share of GVA is 10 % points higher than the sample mean causing a disadvantage given by $-5.316 * 0.1 - 1.771 * 0.1^2 + 124.4 * 0.1^3 = -0.42491$ which is not offset by the positive effect of the treatment alone resulting in a negative growth of $0.332 - 0.425 = -0.93$ % points which represents the HLATE. This case might appear a bit extreme as the sample average is 0.653 which becomes 0.753 with the additional 10 % points leaving roughly only 25 % of regional GVA to the other two sectors but confronting the data 119 NUTS 3 region out of 474 treated regions are above such level and among these 119 regions 98 comply with the 75 % rule. Nevertheless even with a smaller but positive amount of SERV the effect is still negative but might be offset by the treatment itself.

2.4 Sensitivity Analysis

In a wide range of scientific matters, sensitivity analysis (SA) plays a crucial role in evaluating the quality of estimated or calibrated models. In particular, SA estimates variation in the output of a given model following the perturbation of given parameters. The literature has so far proposed two distinctive approaches to SA:

(a) Global Sensitivity Analysis (GSA) evaluates the variation in a model's output after imposing probability distributions on the model's parameters and running simulations;
(b) Local Sensitivity Analysis is similar in spirit to comparative statics as it relies on the global perturbations of parameters, which are often based on derivatives.

GSA is very close in spirit to Bayesian Model Averaging; it is likely outside of the scope of this book, but is of interest in verifying the robustness of estimates for the impact of a policy when relaxing some assumptions.

In this section, we will not provide a complete overview of SA in the context of policy evaluation. Rather, we will present some interesting features of the LSA in the context of RDD. Let us consider a sharp RDD:

$$y_i = \alpha + \beta T_i + f(x_i) + \varepsilon_i$$

Suppose for simplicity that $f(x_i) = \gamma x_i$, so that the OLS estimate of β is:

$$\widetilde{\beta} = \frac{cov(T, y)}{var(T)} - \frac{T}{var(T)}\widetilde{\gamma}x = \widehat{\beta} - \frac{T}{var(T)}\widetilde{\gamma}x$$

where $\widetilde{\gamma}$ is an OLS estimate of γ and $\widehat{\beta}$ is a partial OLS estimate of β. Rearranging previous expression, we have:

$$\widetilde{\beta} - \widehat{\beta} = -\frac{T}{var(T)}\widetilde{\gamma}x$$

which can be approximated through a Taylor expansion of $\widehat{\beta}$ perturbated in $\gamma = 0$, i.e.:

$$\widetilde{\beta} - \widehat{\beta} = \widehat{\beta}(\gamma) - \widehat{\beta}(0) = \left.\frac{\partial\widehat{\beta}(\gamma)}{\partial\gamma}\right|_{\gamma=0} + \mathscr{O}_{\mathscr{P}}\left(\frac{1}{n}\right)$$

so that $\frac{\partial\widehat{\beta}(\gamma)}{\partial\gamma} = \frac{cov(T,x)}{var(T)}$. The expression for the first derivative of $\widehat{\beta}$ gives two main pieces of information:

(a) it provides a quantification of the reaction of $\widehat{\beta}$ to local perturbations in γ and that this variation is larger the larger the correlation between T and x;
(b) suppose that we have several variables as potential candidates for perturbation. The first derivative may provide a ranking of the most important variables on the basis of the magnitude of the derivative itself.

It should also be noted that previous derivation of the sentivity of $\widehat{\beta}$ to γ simply provided an LSA representation of the omitted variable bias in OLS. However, it should also be noted that the Taylor expansion representation is more general as, since although it has been proposed as centred at $\gamma = 0$, it might still provide similar information concerning other values of the parameter (e.g. in cases of measurement error in x).

References

Angrist JD, Pischke JS (2009) Mostly harmless econometrics: an empiricist's companion. Princeton University Press, Princeton

Becker SO, Egger PH, von Ehrlich M (2013) Absorptive capacity and the growth effects of regional transfers: a regression discontinuity design with heterogeneous treatment effects. Am Econ J Econ Policy 5(4):29–77

Imbens GW, Lemieux T (2008) Regression discontinuity designs: a guide to practice. J Econom 142(2):615–635

Lee DS, Lemieux T (2010) Regression discontinuity designs in economics. J Econ Lit 48(2):281–355

Percoco M (2013) The impact of European cohesion policy on regional growth: how much do strategy and local economic structure matter?. Università Bocconi, Mimeo

http://www.springer.com/978-3-319-09518-9

Regional Perspectives on Policy Evaluation
Percoco, M.
2014, XII, 52 p. 3 illus., Softcover
ISBN: 978-3-319-09518-9