`ted`:
a Stata Command for Testing Stability of Regression Discontinuity Models

Giovanni Cerulli

IRCrES, Research Institute on Sustainable Economic Growth
National Research Council of Italy

2016 Stata Conference
Chicago, Illinois
July 28–29

Introduction

Given a running variable X, a threshold c, a treatment indicator T, and an outcome Y, Regression Discontinuity (RD) models identify a local average treatment effect (LATE) by associating a jump in mean outcome with a jump in the probability of treatment T when X crosses the threshold c.

Example: Jacob and Lefgren (2004): You are likely to be sent to summer school if you fail a final exam. T indicates summer school, $-X$ is test grade, $-c$ is grade needed to pass, Y is later academic performance.

- Dong and Lewbel (2015) construct the Treatment Effect Derivative (TED) of estimated RD. TED is nonparametrically identified and easily estimated.
- They argue TED is useful because, under a local policy invariance assumption, TED $=$ MTTE (Marginal Threshold Treatment Effect). MTTE is the change in the RD treatment effect resulting from a marginal change in c.
- We argue here that even without policy invariance, TED provides a useful measure of stability of RD estimates, in both sharp and fuzzy RD designs.
- We also define a closely related concept, the CPD (Complier Probability Derivative). We show that this is another useful measure of stability in fuzzy designs.

- The RD treatment effect (RD LATE) only applies to a small subpopulation: people having $X = c$.
- In fuzzy RD it's an even smaller group: only people who are both compliers and have $X = c$.
- RD Stability: Would people with $X \neq c$ but X near c experience similar treatment effects, in sign and magnitude to those having $X = c$?
- If a small ceteris paribus change in X greatly changes either the ATE or the set of compliers, that should raise doubts about the generality and hence external validity of the estimates.
- This is what TED and CPD estimate. We therefore recommend calculating TED (and CPD for fuzzy designs) in virtually all RD empirical applications.

Angrist and Rokkanen (2015) recognize the issue. They estimate LATE away from the cutoff, but require a strong running variable conditional exogeneity assumption.

In contrast, the only thing we impose to identify TED, beyond standard RD assumptions is additional smoothness: some differentiability (instead of just continuity) of potential outcome expectations.

Similar additional smoothness is already always imposed in practice - differentiability is included in the regularity assumptions needed for local regressions.

TED and CPD are trivial to estimate. In sharp designs TED equals a coefficient people were already estimating and throwing away, not knowing it was meaningful.

Literature Review

General RD identification and estimation: Thistlethwaite and Campbell (1960), Hahn, Todd, and van der Klaauw (2001), Porter (2003), Imbens and Lemieux (2008), Angrist and Pischke (2008), Imbens and Wooldridge (2009), Battistin, Brugiavini, Rettore, and Weber (2009), Lee and Lemieux (2010), many others.

RD derivatives: Card, Lee, Pei, and Weber (2012) regression kink design models (continuous kinked treatment). Dong (2014) shows standard RD models can be identified from a kink in probability of treatment. Slope changes also used by Calonico, Cattaneo and Titiunik (2014).

Dinardo and Lee (2011) informal Taylor expansion at the threshold for ATT.

Policy invariance (outcome doesn't depend on some features of the treatment assignment mechanism, a form of external validity) Abbring and Heckman (2007), Heckman (2010), Carneiro, Heckman, and Vytlacil (2010).

Literature Review - continued

Sufficient assumptions and tests for RD validity: Hahn, Todd and Van der Klaauw (2001), Lee (2008), Dong (2016).

Almost all tests or analyses of internal or external validity of RD require covariates with certain properties: McCrary (2008), Angrist and Fernandez-Val (2013), Wing and Cook (2013), Bertanha and Imbens (2014), and Angrist and Rokkanen (2015).

TED and CPD do not require any covariates other than those used to estimate RD.

Identification and estimation of TED and CPD requires no additional data or information beyond what is needed for standard RD models.

All that is needed for TED and CPD are slightly stronger smoothness conditions than for standard RD. Similar required differentiability assumptions are already imposed in practice when one uses local linear or quadratic estimators.

Regression Discontinuity: Model Definitions

T is a treatment indicator: $T = 1$ if treated, $T = 0$ if untreated. example: Jacob and Lefgren (2004), T indicates going to summer school.

Y is an outcome, e.g. academic performance in higher grades.

X is a running or forcing variable that affects T and may also affect Y, e.g, $-X$ is a final exam grade.

c is a threshold constant, e.g., $-c$ is the grade needed to pass the exam.
The RD instrument is $Z = I(X \geq c)$, e.g. $Z = 1$ if fail the exam, zero if pass it.

A "complier" is an individual i who has $T_i = 1$ if and only if $Z_i = 1$ (e.g. a complier is one who goes to Summer school if and only if he fails the exam).

Sharp RD design: Everybody is a complier. The probability of treatment at $X = c$ jumps from zero to one.

Fuzzy RD design: Some people are not compliers, e.g., teachers sometimes overrule the exam results.

RD Model Treatment Effects

Average Treatment Effect, ATE: The average difference in outcomes across people randomly assigned treatment (e.g. average increase in academic performance Y if randomly chosen students switched from not attending to attending Summer school T).

RD LATE denoted $\pi(c)$: The ATE at $X = c$ among compliers. (e.g. the ATE just among complier students at the borderline of passing or failing the exam).

The RD LATE is identified under very weak conditions by associating the jump in $E(Y \mid X = c)$ with the jump in $E(T \mid X = c)$.

RD Intuition: π can be identified at c, because for X near c assignment to treatment is almost random. Assumes no manipulation: individuals can't set X precisely.

The Definition of TED - sharp case

For any function h and small $\varepsilon > 0$, define the left and right limits of the function h as

$$h_{+}(x) = \lim_{\varepsilon \to 0} h(x+\varepsilon) \quad \text{and} \quad h_{-}(x) = \lim_{\varepsilon \to 0} h(x-\varepsilon).$$

Let $g(x) = E(Y \mid X = x)$.

Sharp RD LATE is defined by $\pi(c) = g_{+}(c) - g_{-}(c)$.

Define the left and right derivatives of the function h as

$$h'_{+}(x) = \lim_{\varepsilon \to 0} \frac{h(x+\varepsilon) - h(x)}{\varepsilon} \quad \text{and} \quad h'_{-}(x) = \lim_{\varepsilon \to 0} \frac{h(x) - h(x-\varepsilon)}{\varepsilon}.$$

Sharp RD TED is $\pi'(c) = g'_{+}(c) - g'_{-}(c)$.

The intuition behind TED - sharp case

Let $Y = g_0(X) + \pi(X)\,T + e$.

e is an error term that embodies all heterogeneity across individuals.
Endogeneity: X, T, and e may all be correlated.

$\pi(x)$ is a LATE. Its the ATE among compliers having $X = x$.
The treatment effect estimated by RD designs is $\pi(c)$.

Let $\pi'(x) = \partial\pi(x)/\partial x$. TED is just $\pi'(c)$.

How can we identify and estimate TED, which is $\pi'(c)$?

Consider sharp design first, so $Y = g_0(X) + \pi(X) Z + e$ where $Z = I(X \geq c)$.

Looking at individuals in a small neighborhood of c, approximate $g_0(X)$ and $\pi(X)$ with linear functions making

$$Y \approx \beta_1 + Z\beta_2 + (X - c)\beta_3 + (X - c) Z\beta_4 + e$$

This is local linear estimation yielding $\widehat{\beta}_1$, $\widehat{\beta}_2$, $\widehat{\beta}_3$ and $\widehat{\beta}_4$.

(Local quadratic just adds $(X - c)^2 \beta_5 + (X - c)^2 Z\beta_6$ to the right).

Under the standard RD and local linear estimation assumptions we get $\widehat{\beta}_2 \to^p \pi(c)$ and $\widehat{\beta}_4 \to^p \pi'(c)$. So $\widehat{\beta}_2$ is the usual RD LATE estimate, and $\widehat{\beta}_4$ is the estimated TED.

Fuzzy Design TED and CPD

For fuzzy design have two local linear (or local polynomial) regressions:

$$
\begin{aligned}
T &\approx \alpha_1 + Z\alpha_2 + (X - c)\,\alpha_3 + (X - c)\,Z\alpha_4 + u \\
Y &\approx \beta_1 + Z\beta_2 + (X - c)\,\beta_3 + (X - c)\,Z\beta_4 + e
\end{aligned}
$$

First is the instrument equation, second is the reduced form outcome equation.

First is local linear approximation of $f(x) = E(T \mid X = x)$, second is local approximation of $g(x) = E(Y \mid X = x)$, recalling that $Z = I(X \geq c)$.

Recall a complier is one having T and Z be the same random variable.

$$
\begin{aligned}
T &\approx \alpha_1 + Z\alpha_2 + (X-c)\,\alpha_3 + (X-c)\,Z\alpha_4 + u \\
Y &\approx \beta_1 + Z\beta_2 + (X-c)\,\beta_3 + (X-c)\,Z\beta_4 + e
\end{aligned}
$$

Let $p\,(x)$ denote the conditional probability that someone is a complier, conditioning on that person having $X = x$. Let $p'\,(x) = \partial p\,(x)\,/\partial x$.

By the same logic as in sharp design (replacing Y with T), we have:

$p\,(c) = f_+(c) - f_-(c)$ and $p'\,(c) = f'_+(c) - f'_-(c)$,

$\widehat{\alpha}_2 \to^p p\,(c)$ and $\widehat{\alpha}_4 \to^p p'\,(c)$.

$p'\,(c)$ is what we call the CPD (Complier Probability Derivative), consistently estimated by $\widehat{\alpha}_4$.

$$
\begin{aligned}
T &\approx \alpha_1 + Z\alpha_2 + (X-c)\,\alpha_3 + (X-c)\,Z\alpha_4 + u \\
Y &\approx \beta_1 + Z\beta_2 + (X-c)\,\beta_3 + (X-c)\,Z\beta_4 + e
\end{aligned}
$$

Let $q(x) = E(Y(1) \mid X=x) - E(Y(0) \mid X=x)$, so $q(c) = g_+(c) - g_-(c)$.

The fuzzy RD Late is $\pi_f(c) = q(c)/p(c)$, $\widehat{\pi}_f(c) = \widehat{\beta}_2/\widehat{\alpha}_2$. Applying the formula for the derivative of a ratio,

$$
\pi_f'(x) = \frac{\partial \pi_f(x)}{\partial x} = \frac{\partial \frac{q(x)}{p(x)}}{\partial x} = \frac{q'(x)}{p(x)} - \frac{q(x)\,p'(x)}{p(x)^2} = \frac{q'(x) - \pi_f(x)p'(x)}{p(x)},
$$

so the fuzzy design TED $\pi_f'(c)$ is consistently estimated by

$$
\widehat{\pi}_f'(c) = \frac{\widehat{\beta}_4 - \widehat{\pi}_f(c)\widehat{\alpha}_4}{\widehat{\alpha}_2} = \frac{\widehat{\beta}_4 - (\widehat{\beta}_2/\widehat{\alpha}_2)\widehat{\alpha}_4}{\widehat{\alpha}_2}
$$

Stability

TED $\pi'(c)$ measures stability of the RD LATE, since $\pi(c+\varepsilon) \approx \pi(c) + \varepsilon\pi'(c)$ for small ε.

Zero TED means $\pi(c+\varepsilon) \cong \pi(c)$, so individuals with x near c have almost the same LATE as those with $x = c$.

Large TED means a small change in x away from c yields large changes in LATE, i.e., instability.

Same stability argument holds for fuzzy designs, with $\pi_f(c+\varepsilon) \approx \pi_f(c) + \varepsilon\pi_f'(c)$.

$$\pi_f'(c) = \frac{q'(c)}{p(c)} - \frac{q(c)p'(c)}{p(c)^2} = \frac{q'(c)}{p(c)} - \frac{p'(c)\pi_f(c)}{p(c)}$$

Fuzzy has two potential sources of instability. Fuzzy can be unstable because $q'(c)$ is far from zero or because $p'(c)$ is far from zero.

$q'(c)$ term large means the treatment effect for the average compiler changes a lot as x moves away from c.

$p'(c)$ term large means that population of compliers changes a lot as x moves away from c.

TED combines both effects.

CPD is just $p'(c)$.

MTTE (Marginal Threshold Treatment Effect)

Define:
$S(x, c) = E[Y(1) - Y(0) \mid X = x, \text{ being a complier, threshold is } c]$

The level of cutoff c is the policy.

$S(x, c)$ is the average treatment effect for individuals having running variable equal to X when the threshold is c.

$S(c, c)$ is the RD LATE

When $x \neq c$, the function $S(x, c)$ is a counterfactual. It defines what the expected treatment effect would be for a complier who is not actually at the cutoff c.

The TED and the MTTE - continued

$$S(x,c) = E[Y(1) - Y(0) \mid X = x, \text{ being a complier, threshold is } c]$$

Let $\tau(c) = S(c,c)$. The TED vs. the MTTE are defined by

$$\begin{aligned} TED &= \frac{\partial S(x,c)}{\partial x}\Big|_{x=c} \\ MTTE &= \frac{\partial \tau(c)}{\partial c} = \frac{\partial S(c,c)}{\partial c} = \frac{\partial S(x,c)}{\partial x}\Big|_{x=c} + \frac{\partial S(x,c)}{\partial c}\Big|_{x=c} \end{aligned}$$

Define local policy invariance as $\frac{\partial S(x,c)}{\partial c}|_{x=c} = 0$: The expected effect of treatment on any particular individual having x near c would not change if the policy cutoff c were marginally changed.

If local policy invariance, then TED = MTTE. Given MTTE, we can evaluate how the treatment effect would change if c marginally changed.

If local policy invariance holds, then we estimate that the LATE would change if the cutoff were changed.

Why might local policy invariance may fail to hold? General equilibrium effects.

Example: in Jacob and Lefgren (2004) treatment is Summer school, the cutoff is an exam grade. Changing the cutoff grade would change the size and composition of the Summer school student body possibly affecting outcomes.

Many policy debates center on whether to change thresholds. Examples:

- Minimum wage levels.
- Legal age for drinking, smoking, voting, medicare or pension eligibility.
- Grade levels for promotions, graduation or scholarships.
- Permitted levels of food additives or environmental pollutants.
- ...

A popular type of experiment is to compare outcomes before and after a threshold change. In contrast, we do not observe a change in the threshold, but MTTE still identifies what the effect would be of a (marginal) change in the threshold.

Even if local policy invariance fails, TED provides useful information for these debates, by comprising a large component of the MTTE.

Stata implementation using `ted`

Cerulli, G., Dong, Y., Lewbel, A., and Paulsen, A. (forthcoming 2016), "Testing Stability of Regression Discontinuity Models", *Advances in Econometrics*, Volume 38. Special issue on "Regression Discontinuity Designs: Theory and Applications", Eds: Matias D. Cattaneo (University of Michigan) and Juan-Carlos Escanciano (Indiana University).

Calonico, Cattaneo and Titiunik (2014): Robust Data-Driven Inference in the Regression-Discontinuity Design, *Stata Journal* 14(4): 909-946.

Overview of RD packages

https://sites.google.com/site/rdpackages

- **rdrobust package**: estimation, inference and graphical presentation using local polynomials, partitioning, and spacings estimators.
  - rdrobust: RD inference (point estimation and CI; classic, bias-corrected, robust).
  - rdbwselect: bandwidth or window selection (IK, CV, CCT).
  - rdplot: plots data (with “optimal” block length).
- **rddensity package**: discontinuity in density test at cutoff (a.k.a. manipulation testing) using novel local polynomial density estimator.
  - rddensity: manipulation testing using local polynomial density estimation.
  - rdbwdensity: bandwidth or window selection.
- **rdlocrand package**: covariate balance, binomial tests, randomization inference methods (window selection & inference).
  - rdrandinf: inference using randomization inference methods.
  - rdwinselect: falsification testing and window selection.
  - rdsensitivity: treatment effect models over grid of windows, CI inversion.
  - rdrbounds: Rosenbaum bounds.

Stata implementation using `ted`

Title

ted - Testing Stability of Regression Discontinuity Models

Description

ted estimates the "local average treatment effect" (LATE), the "compliers' probabilty discontinuity" (CPD), and "treatment effect derivative" (TED) for either sharp or fuzzy Regression Discontinuity (RD) models. Estimation and inference for TED are especially useful for testing the stability of LATE estimates in RD models when infinitesimal changes of the threshold value are allowed.

According to Dong and Lewbel (2015) and Cerulli, et al. (2016), a TED which is significantly different from zero signals that LATE estimate is instable any time very small changes of the threshold value are considered, thus questioning the validity of RD results. In the fuzzy case, standard errors for LATE, CPD, and TED are estimated using the delta method. ted provides also a graphical representation of LATE and TED, by jointly plotting the potential outcome functions and their tangents at threshold.

Syntax of `ted`

Syntax

ted *outcome run_var treat_var* [*if*] [*in*] [*weight*], **model**(*modeltype*) **m**(*number*) **h**(*number*) **k**(*kerneltype*) [**l**(*number*) ***graph*** **vce**(*robust*)]

fweights, *iweights*, and *pweights* are allowed; see weight.

Options

Options

model(*modeltype*) specifies the RD model to be estimated, where modeltype must be one of the following two models: "sharp", "fuzzy". it is always required to specify one model.

m(*number*) sets the polynomial degree of the left and right "conditional expectation of the outcome given the running variable" equal to the number specified in parenthesis.

h(*number*) sets a specific value of the bandwidth for the local RD estimation. For identifying optimal bandwidth, please refer to the user-written command rdbwselect provided by Calonico, Cattaneo, and Vazquez-Bare (2014).

c(*number*) sets the threshold (or cut-off).

l(*number*) sets the interval of the running variable to consider in the graphical representation.

k(*kerneltype*) sets the type of kernel function to consider in the local polynomial estimation of the potential outcomes at threshold.

graph allows for a graphical representation of both sharp and fuzzy RD.

vce(robust) allows for robust regression standard errors. It is optional for all models.

Options - continued

```
 modeltype_options            Description
--------------------------------------------------------------------------
Model
  sharp                        Sharp RD design
  fuzzy                        Fuzzy RD design
  ------------------------------------------------------------------------
  kerneltype_options           Description
  ------------------------------------------------------------------------
k
  epan                         Epanechnikov weighting scheme
  normal                       Normal weighting scheme
  biweight                     Biweight (or Quartic) scheme
  uniform                      Uniform weighting scheme
  triangular                   Triangular weighting scheme
  tricube                      Tricube weighting scheme
  ------------------------------------------------------------------------
```

Example 1: RDD-sharp

Ludwig and Miller (2007) assess the impact of the Head Start program.

Head Start was established in the United States in the year 1965. Its objective is to provide preschool, health, and other social services for poor children ages three to five, as well as their families.

The 300 counties with the highest poverty rates received aid writing grants, thus creating a large, persistent discontinuity in Head Start funding.

Their main result focuses on Head Start funding's effect on mortality due to causes Head Start should have an effect on, using poverty rates as their running variable.

Example 1: code for RDD-sharp

```
*****************************************************************************************
* KEY VARIABLES
*****************************************************************************************
gen y=age5_9_sum2
gen s=povrate60 // running variable
global s_star=59.1984
gen w=(s > $s_star)
global L=6
global M=2
global kernel triangular
*****************************************************************************************
* GENERATE THE OPTIMAL BANDWIDTH USING "rdbwselect"
*****************************************************************************************
rdbwselect y s, c($s_star) p($M) q(3) kernel($kernel) all
global bw_CCT=e(h_CCT) // bandwidth proposed by Calonico, Cattaneo, and Titiunik (2014).
global bw_IK=e(h_IK)   // bandwidth proposed by Imbens and Kalyanaraman (2012).
global bw_CV=e(h_CV)   // bandwidth using cross-validation proposed by Ludwig and Miller (2007).
*****************************************************************************************
* CHOOSE ONE OF THE THREE OPTIMAL BANDWIDTHS
*****************************************************************************************
global band=$bw_CCT
*****************************************************************************************
* ESTIMATE TED USING "ted"
*****************************************************************************************
ted y s w , model(sharp) h($band) c($s_star) m($M) l($L) k($kernel) graph vce(robust)
*****************************************************************************************
```

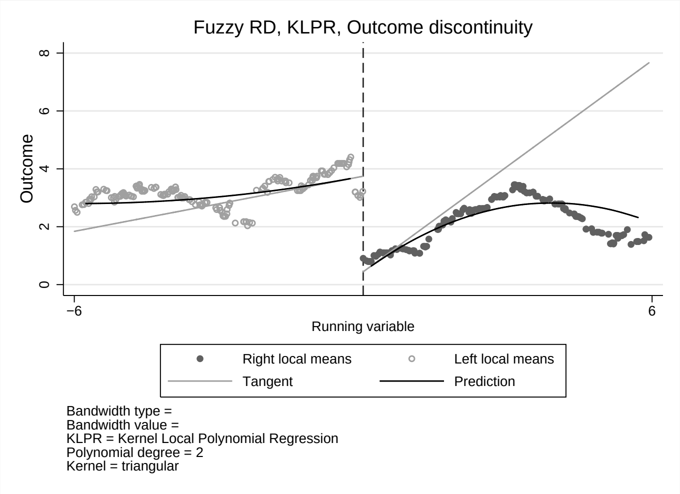
Example 1: `ted` output for RDD-sharp - 1

```
*********************************************************************
*********************** DISCONTINUITY IN THE OUTCOME ***********************
*********************************************************************
(sum of wgt is   9.9503e-01)

Linear regression                               Number of obs =      482
                                                F(  5,   476) =     3.62
                                                Prob > F      =   0.0032
                                                R-squared     =   0.0192
                                                Root MSE      =   5.6933

------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        _x_1 |   .3175989    .604472     0.53   0.600    -.8701644    1.505362
        _x_2 |   .0266717   .0766494     0.35   0.728    -.1239413    .1772847
          _T |  -3.302928   1.290168    -2.56   0.011    -5.838057   -.7677993
      _T_x_1 |   .8990958   .7675621     1.17   0.242    -.6091331    2.407325
      _T_x_2 |  -.1822127   .1038724    -1.75   0.080    -.3863179    .0218924
       _cons |   3.745666   1.187967     3.15   0.002     1.411358    6.079974
------------------------------------------------------------------------------
```

Example 1: `ted` output for RDD-sharp - 2

```
*********************** Test of significance for LATE ***********************

       LATE:  _b[_T]

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        LATE |  -3.302928   1.290168    -2.56   0.010    -5.831611   -.7742453
------------------------------------------------------------------------------

*********************** Test of significance for TED ***********************

        TED:  _b[_T_x_1]

------------------------------------------------------------------------------
           y |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         TED |   .8990958   .7675621     1.17   0.241    -.6052981     2.40349
------------------------------------------------------------------------------
```

Example 1: `ted` output for RDD-sharp - 3



Example 2: RDD-fuzzy

We considers the fuzzy RD model in Clark and Martorell (2010, 2014), which evaluates the signaling value of a high school diploma.

In about half of US states, high school students are required to pass an exit exam to obtain a diploma. The random chance that leads to students falling on either side of threshold passing score generates a credible RD design.

Clark and Martorell takes advantage of the **exit exam rule** to evaluate the **impact on earnings of having a high school diploma**.

The outcome Y is the present discounted value (PDV) of earnings through year 7 after one takes the last round of exit exams. The treatment T is whether a student receives a high school diploma or not. The running variable X is the exit exam score (centered at the threshold passing score).

Example 2: code for RDD-fuzzy

```
*************************************************************************************
* KEY VARIABLES
*************************************************************************************
gen y=pdvrwage_y
gen s=test_lcs_min // running variable is the hs exit exam score-from the last try
gen w=awards_hsd4
global s_star=0
global L=25
global M=2
global kernel triangular
*************************************************************************************
* GENERATE THE OPTIMAL BANDWIDTH USING "rdbwselect"
*************************************************************************************
rdbwselect y s, c($s_star) p($M) q(3) kernel($kernel) all
global bw_CCT=e(h_CCT)
global bw_IK=e(h_IK)
global bw_CV=e(h_CV)
*************************************************************************************
* CHOOSE ONE OF THE THREE OPTIMAL BANDWIDTHS
*************************************************************************************
global band=$bw_CCT
*************************************************************************************
* ESTIMATE TED USING "ted"
*************************************************************************************
ted y s w , model(fuzzy) h($band) c($s_star) m($M) l($L) k($kernel) graph vce(robust)
*************************************************************************************
```

Example 2: `ted` output for RDD-fuzzy - 1

```
*******************************************************************************
************************ DISCONTINUITY IN THE OUTCOME *************************
*******************************************************************************
(sum of wgt is   1.0000e+00)

Linear regression                               Number of obs =      3321
                                                F(  5,   3315) =      0.56
                                                Prob > F      =    0.7320
                                                R-squared     =    0.0006
                                                Root MSE      =     27878

------------------------------------------------------------------------------
             |               Robust
           y |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        _x_1 |   426.5014   813.6623     0.52   0.600     -1168.83    2021.833
        _x_2 |   28.42907   58.76563     0.48   0.629     -86.7915    143.6497
          _T |  -599.5591     2976.5    -0.20   0.840    -6435.522    5236.404
      _T_x_1 |  -1065.823   1103.396    -0.97   0.334    -3229.229    1097.583
      _T_x_2 |   38.93901   83.25466     0.47   0.640    -124.2967    202.1748
       _cons |   27977.25   2413.719    11.59   0.000     23244.72    32709.78
------------------------------------------------------------------------------
```

Example 2: `ted` output for RDD-fuzzy - 2

```
******************************************************************
********************* DISCONTINUITY IN THE PROBABILITY *********************
******************************************************************
(sum of wgt is   1.0000e+00)

Linear regression                               Number of obs =     3321
                                                F(  5,  3315) =   324.72
                                                Prob > F      =   0.0000
                                                R-squared     =   0.3241
                                                Root MSE      =   .39923

------------------------------------------------------------------------------
             |               Robust
           w |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        _x_1 |   .0187589   .0139671     1.34   0.179    -.0086261    .0461438
        _x_2 |   .0011624   .0009774     1.19   0.234     -.000754    .0030788
          _T |   .5130246   .0452378    11.34   0.000     .4243277    .6017215
      _T_x_1 |  -.0270201   .0155274    -1.74   0.082    -.0574643    .0034241
      _T_x_2 |  -.0004856   .0011129    -0.44   0.663    -.0026676    .0016963
       _cons |   .4046514   .0418516     9.67   0.000     .3225938    .4867091
------------------------------------------------------------------------------
```

Example 2: `ted` output for RDD-fuzzy - 3

```
************************ Test of significance for CPD ************************

        CPD:  _b[_T_x_1]

------------------------------------------------------------------------------
           w |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         CPD |  -.0270201   .0155274    -1.74   0.082    -.0574531     .003413
------------------------------------------------------------------------------

************************ Test of significance for LATE ************************

       LATE:  [y]_b[_T]/[w]_b[_T]

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        LATE |  -1168.675   4609.017    -0.25   0.800    -10202.18    7864.832
------------------------------------------------------------------------------

************************ Test of significance for TED ************************

        TED:  ([y]_b[_T_x_1]-([w]_b[_T_x_1]*[y]_b[_T]/[w]_b[_T]))/[w]_b[_T]

------------------------------------------------------------------------------
             |      Coef.   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         TED |  -2139.079   1984.695    -1.08   0.281     -6029.01    1750.851
------------------------------------------------------------------------------
```

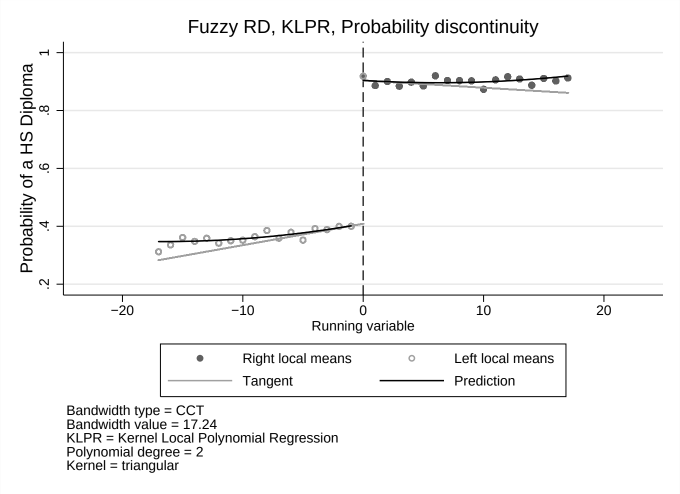


Figure: Fuzzy RD discontinuity in the probability and tangents lines at threshold. Dataset: Clark and Martorell (2010).

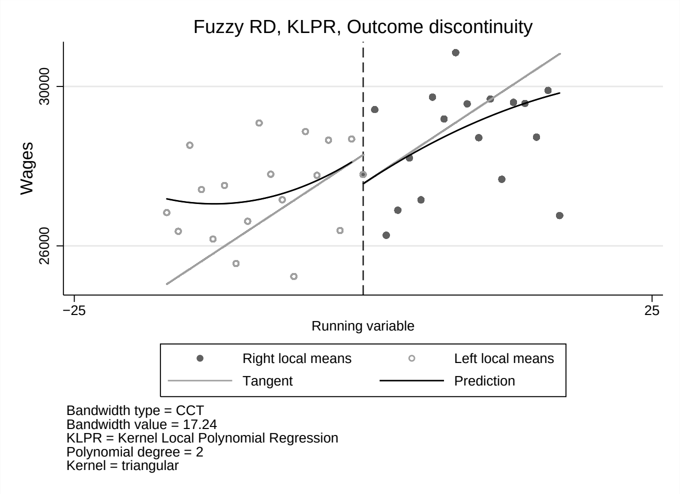


Figure: Fuzzy RD discontinuity in the outcome and tangents lines at threshold. Dataset: Clark and Martorell (2010).

Conclusions

1. Dong and Lewbel (2015) define CPD along with TED, and show they are almost always useful as tests of RD LATE stability.
2. TED and CPD are numerically simple to estimate, and require no more data than needed for RD estimation itself.
3. `ted` is a Stata module to estimate LATE, TED and CPD. It easily provides correct inference for these parameters.
4. We recommend calculating TED (and CPD for fuzzy designs) in virtually all RD empirical applications.