Multidimensional Regression Discontinuity and Regression Kink Designs with Difference-in-Differences

Rafael P. Ribas

University of Amsterdam

Stata Conference

Chicago, July 28, 2016

Motivation

- Regression Discontinuity (RD) designs have been broadly applied.
- However, non-parametric estimation is restricted to simple specifications.
    - I.e., cross-sectional data with one running variable.
- Thus some papers still use parametric polynomial forms and/or arbitrary bandwidths. For instance,
    - Dell (2010, *Econometrica*) estimates a two-dimensional RD.
    - Grembi et al. (2016, *AEJ:AE*) estimates Difference-in-Discontinuities.
- The goal is to create a program (such as `rdrobust`) that accommodates more flexible specifications.

Motivation

- Regression Discontinuity (RD) designs have been broadly applied.
- However, non-parametric estimation is restricted to simple specifications.
    - I.e., cross-sectional data with one running variable.
- Thus some papers still use parametric polynomial forms and/or arbitrary bandwidths. For instance,
    - Dell (2010, *Econometrica*) estimates a two-dimensional RD.
    - Grembi et al. (2016, *AEJ:AE*) estimates Difference-in-Discontinuities.
- The goal is to create a program (such as `rdrobust`) that accommodates more flexible specifications.

Overview

- The `ddrd` package is built upon `rdrobust` package, including the following options:
  1. Difference-in-Discontinuities (DiD) and Difference-in-Kinks (DiK)
  2. Multiple running variables
  3. Analytic weights (`aweight`)
  4. Control variables
  5. Heterogeneous effect through linear interaction (in progress).
- All options are taken into account when computing the optimal bandwidth, using `ddbwsel`.
  - The estimator changes, so does the procedure.

Difference-in-Discontinuity/Kink, Notation

- Let $\mu_t(x) = \mathbb{E}[Y|X = x, t]$ and $\mu_t^{(v)}(x) = \frac{\partial^v E[Y|X=x,t]}{(\partial x)^v}$.
- Then the conventional sharp RD/RK estimand is:

$$\tau_{v,t} = \lim_{x \to 0^+} \mu_t^{(v)}(x) - \lim_{x \to 0^-} \mu_t^{(v)}(x) = \mu_{t+}^{(v)} - \mu_{t-}^{(v)}$$

- The DiD/DiK estimand is:

$$\Delta\tau_v = \mu_{1+}^{(v)} - \mu_{1-}^{(v)} - \left[\mu_{0+}^{(v)} - \mu_{0-}^{(v)}\right]$$

Optimal Bandwidth, h^*

- Two methods based on the mean square error (MSE):

$$h^*_{MSE} = \left[C(K) \frac{\text{Var}(\hat{\tau}_v)}{\text{Bias}(\hat{\tau}_v)^2} \right]^{\frac{1}{5}} n^{-\frac{1}{5}}$$

  - Imbens and Kalyanaraman (2012), IK.
  - Calonico, Cattaneo and Titiunik (2014), CCT.
- They differ in the way $\text{Var}(\hat{\tau}_v)$ and $\text{Bias}(\hat{\tau}_v)$ are estimated.
- For DiD/DiK, the trick is to replace $\hat{\tau}_v$ by $\Delta\hat{\tau}_v$.
  - That's what `ddbwsel` does.
- While `ddrd` calculates the robust, bias-corrected confidence intervals for $\Delta\hat{\tau}_v$, as proposed by CCT.

Optimal Bandwidth, h^*

- Two methods based on the mean square error (MSE):

$$h^*_{MSE} = \left[C(K) \frac{\text{Var}(\hat{\tau}_v)}{\text{Bias}(\hat{\tau}_v)^2} \right]^{\frac{1}{5}} n^{-\frac{1}{5}}$$

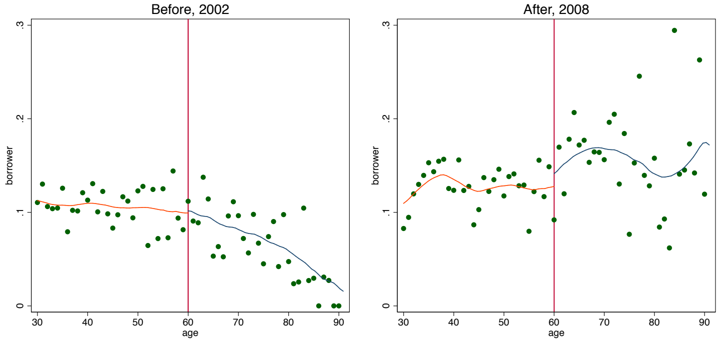
  - Imbens and Kalyanaraman (2012), IK.
  - Calonico, Cattaneo and Titiunik (2014), CCT.
- They differ in the way $\text{Var}(\hat{\tau}_v)$ and $\text{Bias}(\hat{\tau}_v)$ are estimated.
- For DiD/DiK, the trick is to replace $\hat{\tau}_v$ by $\Delta\hat{\tau}_v$.
  - That's what `ddbwsel` does.
- While `ddrd` calculates the robust, bias-corrected confidence intervals for $\Delta\hat{\tau}_v$, as proposed by CCT.

Application: Retirement and Payroll Credit in Brazil

- In 2003, Brazil passed a legislation regulating payroll lending.
  - Loans for which interests are deducted from payroll check (Coelho et al., 2012).
  - It represented a "kink" in loans to pensioners.

Application: Retirement and Payroll Credit in Brazil

- In 2003, Brazil passed a legislation regulating payroll lending.
  - Loans for which interests are deducted from payroll check (Coelho et al., 2012).
  - It represented a "kink" in loans to pensioners.



Application: Retirement and Payroll Credit in Brazil

- Optimal bandwidth for Difference-in-Kink at age 60:

```
. ddbwsel borrower aged [aw=weight], time(time) c(60) deriv(1) all
Computing CCT bandwidth selector.
Computing IK bandwidth selector.

Bandwidth estimators for local polynomial regression


        Cutoff c = 60 |  Left of c  Right of c            Number of obs   =      53757
----------------------+----------------------             NN matches      =          3
 Number of obs, t = 0 |     20836        4484             Kernel type     = Triangular
 Number of obs, t = 1 |     22609        5828
 Order loc. poly. (p) |         2           2
       Order bias (q) |         3           3
 Range of aged, t = 0 |    29.996      29.999
 Range of aged, t = 1 |    29.996      29.996

----------------------------------------------
    Method |        h            b          rho
----------+-----------------------------------
       CCT |  12.45718    18.73484     .6649206
        IK |  14.46675    11.01818     1.312989
----------------------------------------------
```

Application: Retirement and Payroll Credit in Brazil

- ddrd output:

```
. ddrd borrower aged [aw=weight], time(time) c(60) deriv(1) b(‘b’) h(‘h’)
Preparing data.
Calculating predicted outcome per sample.
Estimation completed.

Estimates using local polynomial regression. Derivative of order 1.

        Cutoff c = 60 |  Left of c  Right of c            Number of obs =      27093
----------------------+----------------------            NN matches    =          3
 Number of obs, t = 0 |       6117        3081            BW type       =     Manual
 Number of obs, t = 1 |       7319        4001            Kernel type   = Triangular
 Order loc. poly. (p) |          2           2
       Order bias (q) |          3           3
    BW loc. poly. (h) |     12.457      12.457
          BW bias (b) |     18.735      18.735
            rho (h/b) |      0.665       0.665

Outcome: borrower. Running Variable: aged.
--------------------------------------------------------------------------------------
               Method |     Coef.    Std. Err.      z      P>|z|      [95% Conf. Interval]
----------------------+---------------------------------------------------------------
         Conventional |     .0229      .0221     1.0362    0.300      -.020417    .066218
               Robust |     .0271     .03123     0.8680    0.385      -.034098    .088303
--------------------------------------------------------------------------------------
```

Difference-in-Kink

- What if there is no cutoff and `aged` is a continuous treatment?
- Shift in level represents the *first* difference, while change in the slope represents the *second* difference.
  - Difference-in-Difference with continuous treatment.

Difference-in-Kink

Estimating changes in the first derivative at any part of the function:

```
. ddrd borrower aged [aw=weight], time(time) c(60) deriv(1) nocut
Preparing data.
Computing bandwidth selectors.
Calculating predicted outcome per sample.
Estimation completed.

Estimates using local polynomial regression. Derivative of order 1.

     Reference c = 60 |     Time 0     Time 1              Number of obs =      53757
----------------------+----------------------              NN matches    =          3
        Number of obs |      8433      10395               BW type       =        CCT
 Order loc. poly. (p) |         2          2               Kernel type   = Triangular
       Order bias (q) |         3          3
    BW loc. poly. (h) |    11.489     11.489
          BW bias (b) |    16.813     16.813
            rho (h/b) |     0.683      0.683

Outcome: borrower. Running Variable: aged.
--------------------------------------------------------------------------------------
               Method |     Coef.   Std. Err.      z      P>|z|      [95% Conf. Interval]
----------------------+---------------------------------------------------------------
         Conventional |    .00473     .00161    2.9473    0.003       .001585    .007879
               Robust |    .00528      .0022    2.3988    0.016       .000966    .009598
--------------------------------------------------------------------------------------
```

Multidimensional RD, Notation

- Suppose X has k dimensions, i.e. $X = \{x_1, \cdots, x_k\}$.
- Cutoff doesn't have to be unique.
- Let $\mathbf{c} = \{(c_{11}, \cdots, c_{n1}), \cdots, (c_{1L}, \cdots, c_{nL})\}$ be the cutoff hyperplane.
  - It separates treated and control.
- z_i indicates whether i is "intended for treatment" (in the treated set) or not (in the control set).
- **Trick**: pick one point in $\mathbf{c}$, say $\mathbf{c}_l = (c_{1l}, \cdots, c_{nl})$, and reduce X to one dimension by calculating the distance $d(\mathbf{x}_i, \mathbf{c}_l)$ for every i.
- The new running variable is:

$$r_i = (2 \cdot z_i - 1) \cdot d(\mathbf{x}_i, \mathbf{c}_l).$$

Multidimensional RD, Notation

- Suppose X has k dimensions, i.e. $X = \{x_1, \cdots, x_k\}$.
- Cutoff doesn't have to be unique.
- Let $\mathbf{c} = \{(c_{11}, \cdots, c_{n1}), \cdots, (c_{1L}, \cdots, c_{nL})\}$ be the cutoff hyperplane.
  - It separates treated and control.
- z_i indicates whether i is "intended for treatment" (in the treated set) or not (in the control set).
- **Trick**: pick one point in $\mathbf{c}$, say $\mathbf{c}_l = (c_{1l}, \cdots, c_{nl})$, and reduce X to one dimension by calculating the distance $d(\mathbf{x}_i, \mathbf{c}_l)$ for every i.
- The new running variable is:

$$r_i = (2 \cdot z_i - 1) \cdot d(\mathbf{x}_i, \mathbf{c}_l).$$

Multidimensional RD, Notation

- Suppose X has k dimensions, i.e. $X = \{x_1, \cdots, x_k\}$.
- Cutoff doesn't have to be unique.
- Let $\mathbf{c} = \{(c_{11}, \cdots, c_{n1}), \cdots, (c_{1L}, \cdots, c_{nL})\}$ be the cutoff hyperplane.
  - It separates treated and control.
- z_i indicates whether i is "intended for treatment" (in the treated set) or not (in the control set).
- **Trick**: pick one point in $\mathbf{c}$, say $\mathbf{c}_l = (c_{1l}, \cdots, c_{nl})$, and reduce X to one dimension by calculating the distance $d(\mathbf{x}_i, \mathbf{c}_l)$ for every i.
- The new running variable is:

$$r_i = (2 \cdot z_i - 1) \cdot d(\mathbf{x}_i, \mathbf{c}_l).$$

Multidimensional RD, Notation

- Suppose X has k dimensions, i.e. $X = \{x_1, \cdots, x_k\}$.
- Cutoff doesn't have to be unique.
- Let $\mathbf{c} = \{(c_{11}, \cdots, c_{n1}), \cdots, (c_{1L}, \cdots, c_{nL})\}$ be the cutoff hyperplane.
  - It separates treated and control.
- z_i indicates whether i is "intended for treatment" (in the treated set) or not (in the control set).
- **Trick**: pick one point in $\mathbf{c}$, say $\mathbf{c}_l = (c_{1l}, \cdots, c_{nl})$, and reduce X to one dimension by calculating the distance $d(\mathbf{x}_i, \mathbf{c}_l)$ for every i.
- The new running variable is:

$$r_i = (2 \cdot z_i - 1) \cdot d(\mathbf{x}_i, \mathbf{c}_l).$$

Multidimensional RD

- With one running variable, r, I can apply the previous methods.
- `ddrd` includes the following distance functions:
  - Manhattan (L1)
  - Euclidean (L2)
  - Minkowski (Lp)
  - Mahalanobis
  - Latitude-Longitude
- **Caveat**: If cutoff isn't unique, $\hat{\tau}_v$, $\Delta\hat{\tau}_v$, and h^* depend on the chosen cutoff point.
  - The effect can be heterogeneous.
- **Solution**: Average effect from several different cutoffs.
  - Correlation between cutoffs should be taken into account (in progress).

Multidimensional RD

- With one running variable, r, I can apply the previous methods.
- `ddrd` includes the following distance functions:
  - Manhattan (L1)
  - Euclidean (L2)
  - Minkowski (Lp)
  - Mahalanobis
  - Latitude-Longitude
- **Caveat**: If cutoff isn't unique, $\hat{\tau}_v$, $\Delta\hat{\tau}_v$, and h^* depend on the chosen cutoff point.
  - The effect can be heterogeneous.
- **Solution**: Average effect from several different cutoffs.
  - Correlation between cutoffs should be taken into account (in progress).

Multidimensional RD

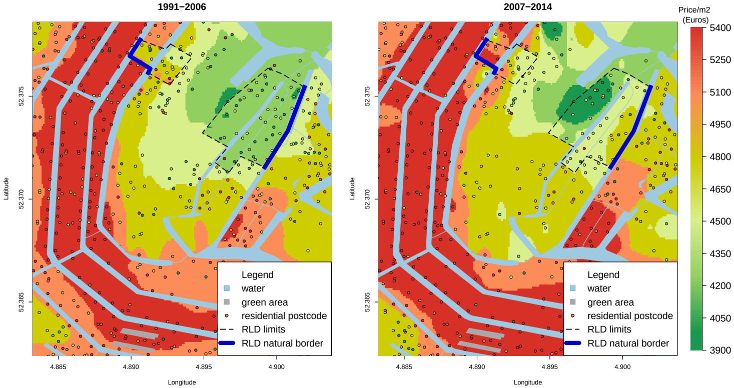
- With one running variable, r, I can apply the previous methods.
- `ddrd` includes the following distance functions:
  - Manhattan (L1)
  - Euclidean (L2)
  - Minkowski (Lp)
  - Mahalanobis
  - Latitude-Longitude
- **Caveat**: If cutoff isn't unique, $\hat{\tau}_v$, $\Delta\hat{\tau}_v$, and h^* depend on the chosen cutoff point.
  - The effect can be heterogeneous.
- **Solution**: Average effect from several different cutoffs.
  - Correlation between cutoffs should be taken into account (in progress).

Multidimensional RD

- With one running variable, r, I can apply the previous methods.
- `ddrd` includes the following distance functions:
  - Manhattan (L1)
  - Euclidean (L2)
  - Minkowski (Lp)
  - Mahalanobis
  - Latitude-Longitude
- **Caveat**: If cutoff isn't unique, $\hat{\tau}_v$, $\Delta\hat{\tau}_v$, and h^* depend on the chosen cutoff point.
  - The effect can be heterogeneous.
- **Solution**: Average effect from several different cutoffs.
  - Correlation between cutoffs should be taken into account (in progress).

Application: The Effect of Prostitution on House Prices

- In Amsterdam, the canals are like natural borders of the red light district (RLD).



Application: The Effect of Prostitution on House Prices

- ddrd output:

```
. ddrd lprice Lat Lon if time==0, itt(rldA) c(52.374611 4.901397) dfunction(Latlong)
Computing Latlong distance
Preparing data.
Computing bandwidth selectors.
Calculating predicted outcome per sample.
Estimation completed.

Estimates using local polynomial regression.

        Cutoff c = 0 |  Left of c  Right of c            Number of obs =      53174
---------------------+----------------------             NN matches    =          3
       Number of obs |        99         124             BW type       =        CCT
 Order loc. poly. (p) |         1           1             Kernel type   = Triangular
      Order bias (q) |         2           2
   BW loc. poly. (h) |     7.445       7.445
         BW bias (b) |    11.258      11.258
           rho (h/b) |     0.661       0.661

Outcome: lprice. Running Variable: Lat Lon.
--------------------------------------------------------------------------------------
              Method |    Coef.   Std. Err.      z      P>|z|      [95% Conf. Interval]
---------------------+----------------------------------------------------------------
        Conventional |  -.27857     .06379    -4.3669   0.000      -.403605   -.153544
              Robust |  -.30377     .09626    -3.1557   0.002      -.492442   -.115104
--------------------------------------------------------------------------------------
```

Application: The Effect of Prostitution on House Prices

- `ddrd` output, with DiD:

```
. ddrd lprice Lat Lon, itt(rldA) time(time) c(52.374611 4.901397) dfunction(Latlong)
Computing Latlong distance
Preparing data.
Computing bandwidth selectors.
Calculating predicted outcome per sample.
Estimation completed.

Estimates using local polynomial regression.

        Cutoff c = 0 |  Left of c  Right of c            Number of obs =      49055
----------------------+----------------------            NN matches    =          3
 Number of obs, t = 0 |        86          90            BW type       =        CCT
 Number of obs, t = 1 |        60          47            Kernel type   = Triangular
 Order loc. poly. (p) |         1           1
       Order bias (q) |         2           2
   BW loc. poly. (h) |     6.937       6.937
         BW bias (b) |    11.963      11.963
           rho (h/b) |     0.580       0.580

Outcome: lprice. Running Variable: Lat Lon.
--------------------------------------------------------------------------------------
              Method |    Coef.   Std. Err.      z      P>|z|      [95% Conf. Interval]
----------------------+---------------------------------------------------------------
        Conventional |    .3801      .1498    2.5374    0.011       .086495    .673705
              Robust |   .51914     .21802    2.3811    0.017       .091824    .946453
--------------------------------------------------------------------------------------
```

Control Variables

- In the previous example, we are interested in residents' willingness to pay for the location.
- However, house prices comprise both quality and location.
  - And house quality is also affected by amenities.
- Solution is to control for house characteristics.
- How?
- I apply the Frisch-Waugh theorem in 3 steps (McMillen and Redfearn, 2010):
  1. Regress variables (x) and y on the running variable (r).
  2. Estimate the coefficient vector β by regressing residuals of y on residuals of x.
  3. Regress $(y - \hat{\beta}'x)$ on the running variable (r).

Control Variables

- In the previous example, we are interested in residents' willingness to pay for the location.
- However, house prices comprise both quality and location.
  - And house quality is also affected by amenities.
- Solution is to control for house characteristics.
- How?
- I apply the Frisch-Waugh theorem in 3 steps (McMillen and Redfearn, 2010):
  1. Regress variables (x) and y on the running variable (r).
  2. Estimate the coefficient vector β by regressing residuals of y on residuals of x.
  3. Regress $(y - \hat{\beta}'x)$ on the running variable (r).

Control Variables

- In the previous example, we are interested in residents' willingness to pay for the location.
- However, house prices comprise both quality and location.
  - And house quality is also affected by amenities.
- Solution is to control for house characteristics.
- How?
- I apply the Frisch-Waugh theorem in 3 steps (McMillen and Redfearn, 2010):
  1. Regress variables (x) and y on the running variable (r).
  2. Estimate the coefficient vector β by regressing residuals of y on residuals of x.
  3. Regress $(y - \hat{\beta}'x)$ on the running variable (r).

Control Variables

- In the previous example, we are interested in residents' willingness to pay for the location.
- However, house prices comprise both quality and location.
  - And house quality is also affected by amenities.
- Solution is to control for house characteristics.
- How?
- I apply the Frisch-Waugh theorem in 3 steps (McMillen and Redfearn, 2010):
  1. Regress variables (x) and y on the running variable (r).
  2. Estimate the coefficient vector β by regressing residuals of y on residuals of x.
  3. Regress $(y - \hat{\beta}'x)$ on the running variable (r).

Control Variables

- In the previous example, we are interested in residents' willingness to pay for the location.
- However, house prices comprise both quality and location.
  - And house quality is also affected by amenities.
- Solution is to control for house characteristics.
- How?
- I apply the Frisch-Waugh theorem in 3 steps (McMillen and Redfearn, 2010):
  1. Regress variables (x) and y on the running variable (r).
  2. Estimate the coefficient vector β by regressing residuals of y on residuals of x.
  3. Regress $(y - \hat{\beta}'x)$ on the running variable (r).

Application: The Effect of Prostitution on House Prices

- ddrd output, with control variables:

```
. ddrd lprice Lat Lon if time==0, itt(rldA) c(52.374611 4.901397) dfunction(Latlong) control(siz
> e date1-date4 monumnt poorcnd luxury rooms floors kitchen bath centhet balcony attic terrace l
> ift garage garden)

(...)

Estimates using local polynomial regression.

         Cutoff c = 0 |  Left of c  Right of c            Number of obs =      72434
----------------------+----------------------             NN matches    =          3
        Number of obs |        117         135            BW type       =     Manual
 Order loc. poly. (p) |          1           1            Kernel type   = Triangular
       Order bias (q) |          2           2
    BW loc. poly. (h) |      7.445       7.445
          BW bias (b) |     11.258      11.258
            rho (h/b) |      0.661       0.661

Outcome: lprice. Running Variable: Lat Lon.
--------------------------------------------------------------------------------------
               Method |    Coef.   Std. Err.      z      P>|z|      [95% Conf. Interval]
----------------------+---------------------------------------------------------------
         Conventional |  -.50715     .22619    -2.2422   0.025      -.950466   -.063836
               Robust |  -.61673     .36225    -1.7025   0.089      -1.32674    .093267
--------------------------------------------------------------------------------------

Control variables: size date1 date2 date3 date4 monumnt poorcnd luxury rooms floors kitchen bath
>  centhet balcony attic terrace lift garage garden.
```