Regression Discontinuity Donuts

Connor Dowd

University of Chicago Booth School of Business

UChicago Econometrics Lunch
May 21, 2019

Question

When we use a donut, how much can we learn using standard RD assumptions?

What assumptions does it make sense to consider adding?

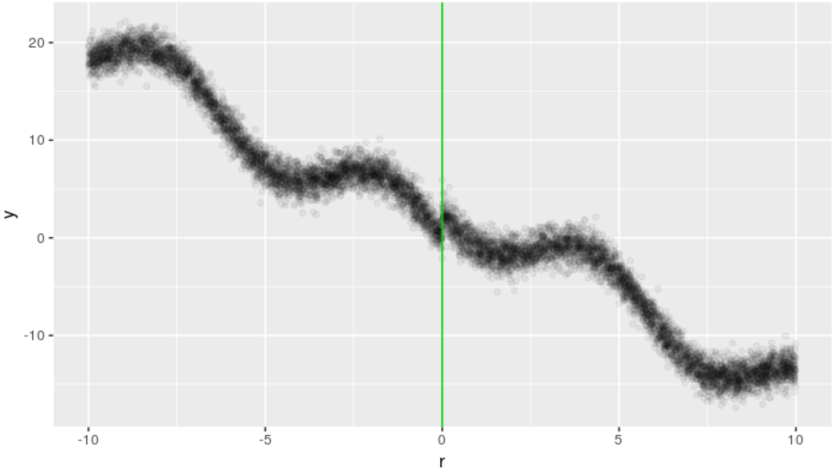
Regression Discontinuity – A reminder

Some Notation:

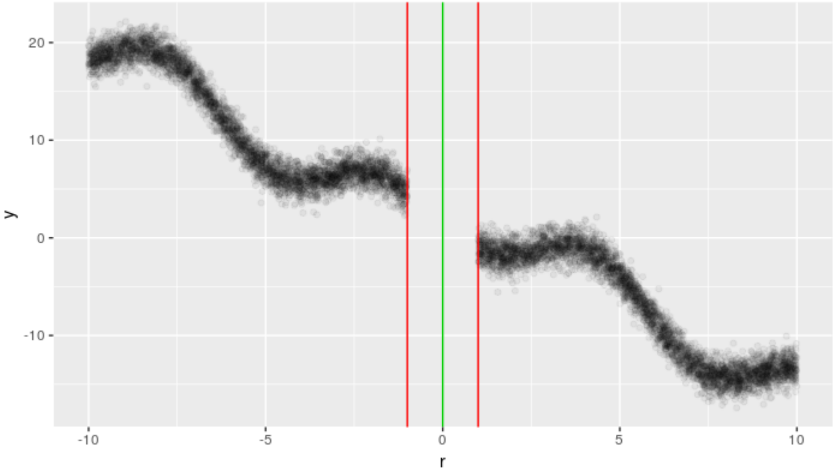
- Running Variable: R
- Outcome: Y
- Treatment: T, induced at some threshold $\bar{R}$
- Potential Outcomes: Y_1, Y_0, with CEFs $Y_1(r), Y_0(r)$

WLOG shift R so that $\bar{R} = 0$.

RD plot



Donut Plot



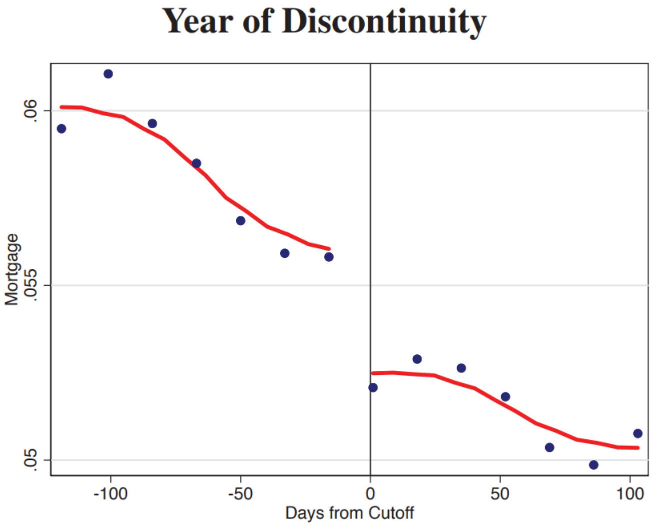
Donut: (-1,1)

Sample Donut

- Policy Threshold for birthday in new calendar year
- Policy affects 24 year olds
- Research indicates some degree of manipulation around end of year births
- Outcome of interest is homeownership – as measured by mortgage possession.

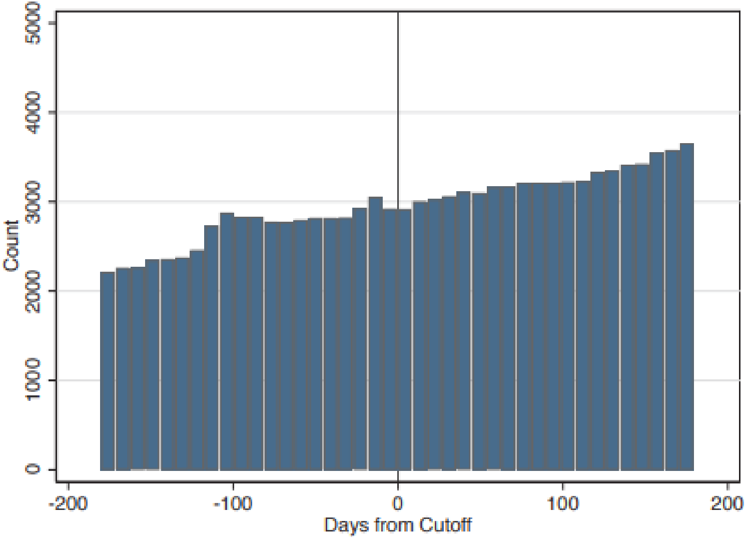
Plots and numbers are straight out of Goodman, Isen, and Yannelis [2018]

RD Donut Example – Mortgages



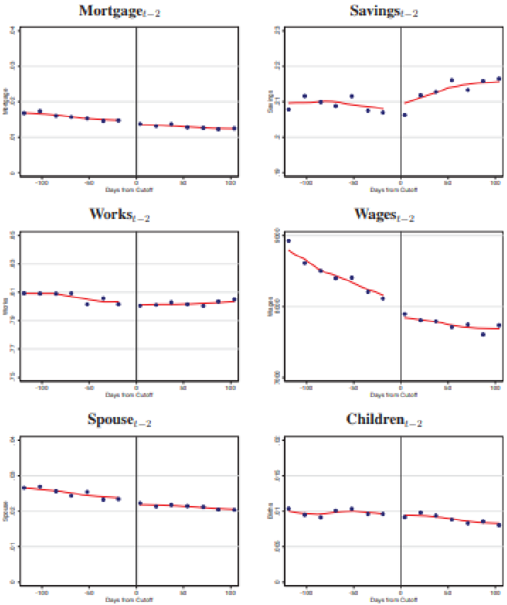
Donut: [-3,3]

RD Donut Example – Density of Running Variable



McCrary Test pval: 0.49

RD Donut Example – Covariates



Simulation Results

- TE: 1, Donut: (-3,3), Var(ϵ) $=$2
- Min price of manipulation: 0.5/unit, correlated with ϵ_i
- 1% choose to manipulate
- 25% of them fail to cross threshold
- Bias of Standard RD: -0.465
- Coverage of Standard RD: 14.9%
- McCrary test power: 56%
- McCrary conditional Bias: -0.205
- McCrary conditional Coverage: 60.4%
- Donut Exclusion: 100%
- Nsims $=$ 3200

RD Donut Example – Effects on Mortgages

	(1) In Year of Discontinuity	(2) Year After Discontinuity
Above Cutoff	.005245**	.0073083**
	(.002384)	(.003398)
Observations	464,008	464,008

How do we get these numbers?

RD Donut Example – Functional Form Assumption

Estimated Model:

$$Y_{it} = \beta_0 + \beta_1 \; \mathbb{1}\left[R \geq \bar{R}\right] + \left(\sum_{j=1}^{2} \delta_j R_{it}^{j} + \varphi_j \mathbb{1}\left[R_{it} \geq \bar{R}\right] R_{it}^{j} \right) + \gamma_t + \epsilon_{it}$$

This only needs to hold within the bandwidth of 50 days, and is estimated outside the donut of 3 days.

Bandwidth is based on Calonico, Cattaneo, and Titiunik [2014] and the polynomial order is based on Gelman and Imbens [2014].

Comments

- The Donut size is not shrinking asymptotically – rather it is a feature of people´s ability to manipulate, and so we can take it as fixed.
- Some assumptions beyond functional form are needed for this to identify the treatment effect.

Moving away from Functional Form

Notation:

- Running Variable: R
- Outcome: Y
- Treatment: T, induced at some threshold $\bar{R}$
- Potential Outcomes: Y_1, Y_0, with CEFs $Y_1(r), Y_0(r)$
- Running Variable after Manipulation: $R' = R + m + e$.
  where m is the manipulation an individual engaged in, and e is the error in their final location.
- Donut: (d_-, d_+)

Continue to focus on Sharp case.

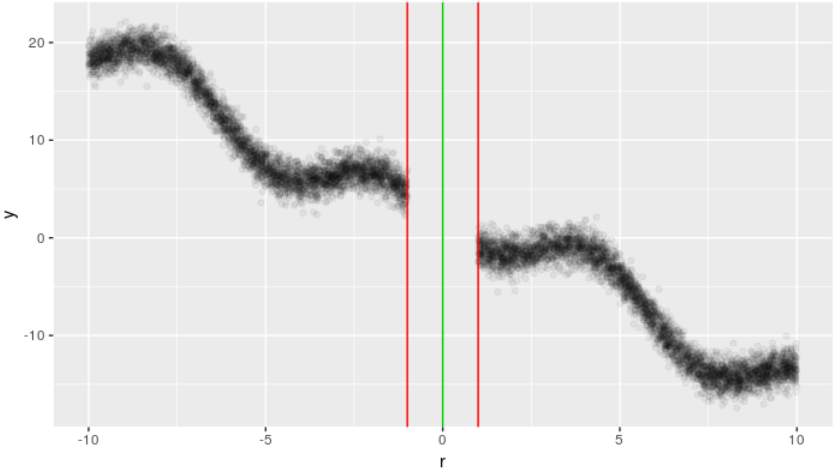
Donut Assumptions: Threshold Exclusion

Threshold exclusion Assumption

There are no other policies with thresholds inside the donut.

- This is about knowing what treatment effect we identify. Standard RD just needs there to not be a co-located policy threshold, but donuts need more.
- The functional form assumptions above imply that the only location such a policy could have a threshold is at our threshold, $\bar{R}$ – i.e. no stronger than standard RD.

Donut Example



Donut: (-1,1)

Donut Assumptions: Exclusion

Donut Exclusion

There is no manipulation outside of the donut.

- Specifically, individuals who manipulate the running variable do not move *to or from* points outside the donut.
- This is necessary to get away from a ´selected´ group.

Why would we believe Donut Exclusion?

- Increasing cost to 'long-distance' manipulation.
  - Must grow larger than TE.
- No real reason for local manipulation outside donut.
  - Outcome as a function of true value of running variable and treatment status. $E[Y|R, R', T] = E[Y|R, T]$
  - Only treatment status changes after manipulation – $T(R')$.

What do these assumptions give us?

With donut exclusion, we can trust all the data outside of the donut (d_-, d_+).
But we still know nothing about the interior.

- $E[Y|r = d_-]$ and $E[Y|r = d_+]$ can both be identified.

Most papers using donuts today will implicitly make a functional form assumption here.

Moving beyond functional forms – Donut Difference

With standard RD assumptions, Donut exclusion, and threshold exclusion, we can identify the donut difference.

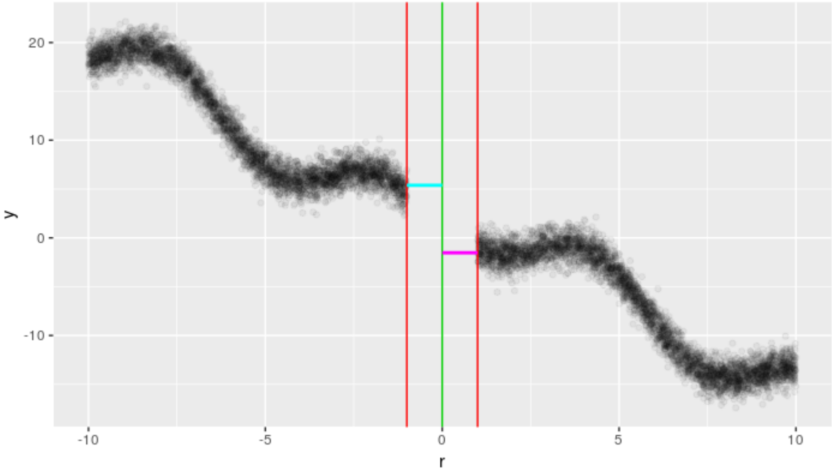
Donut Difference

$$\tau_{dd} = E[Y|r = d_+] - E[Y|r = d_-]$$

Intuition for estimation: Drop the donut, move all points in towards the threshold. Estimate using standard RD.

Calling this the TE would implicitly assume that the CEFs are behaving very similarly inside the donut. (Constant, or identical slopes both would make that true).

Donut Difference Example



Donut: (-1,1)

Partial Identification

Derivative bounds are a standard RD assumption.

For donuts with finite width, derivative bounds identify finite sets for the treatment effect.

Derivative Bounds

Knowing that we can estimate both $E[Y|r = d_-]$ and $E[Y|r = d_+]$, if we know that:

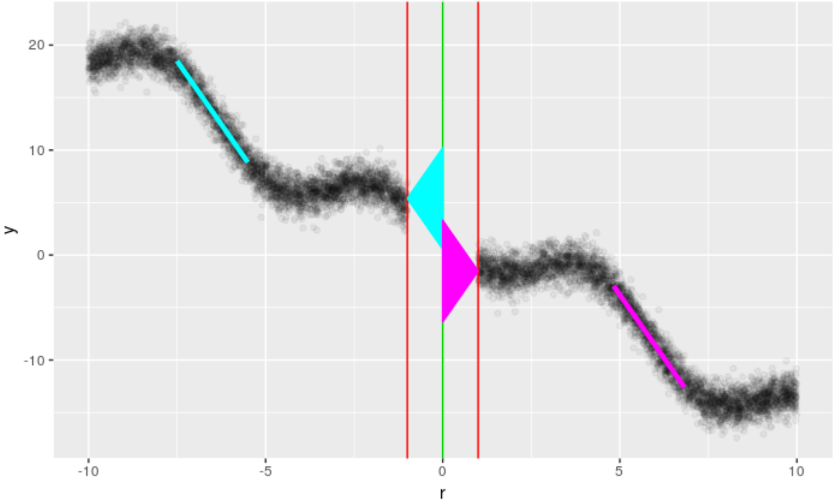
$$\frac{\partial Y_t(r)}{\partial r} \leq F_t$$

Where $t = 0, 1$.
In the oracle world where we know $Y_0(d_-)$ and $Y_1(d_+)$ exactly, we know that $Y_0(0) \in [Y_0(d_-) + F_0 * d_-, Y_0(d_-) - F_0 * d_-]$. Analogously $Y_1(0)$ can be identified, and with both of them, τ.

Conditional on a fixed F_t, the only variation is in the estimation of $Y_0(d_-)$ and $Y_1(d_+)$. Building a confidence region for the identified set of τ then requires only handling that noise on each side.

Higher order derivative bounds require estimating more parameters at the boundary, but are not functionally different.

Derivative Bound Example



What if we do not know F_t? (In progress)

Suppose there is some global bound on derivatives, but we don´t know it. Can we learn it well enough to take advantage of it? Specifically assume:

Lipschitz Continuity

$\forall x_1, x_2 \in \quad |Y_t(x_1) - Y_t(x_2)| \leq F_t(x_1 - x_2)$ for $t = 0, 1$

As well as:

Observed Derivatives Assumption

$\exists x_1, x_2 \in (min(R), d_{-} - h)$, such that the Lipschitz Bound holds with equality.

We also need the symmetric assumption for Y_1 – where the bound holds for $x_1, x_2 \in (d_{+} + h, max(R))$.

Procedure for Unknown F_t

Set a sequence p_n such that $p \to \infty$ and $p/n \to 0$. We can estimate Y_0 at $p_n + 1$ different equally spaced points over the range $(min(R), d_- - h)$.

Each point is asymptotically normal, and there are asymptotically many points.

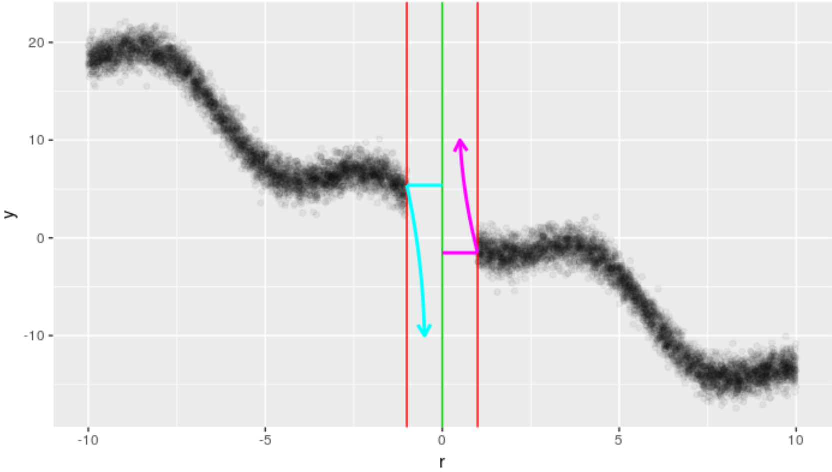
Between each pair of points we can estimate a derivative, which will also be asymptotically normal, with estimable covariance.

The max of the derivatives gives us a point estimate for our identified set.

Extensions – Monotonicity

- Monotonicity is often a much weaker assumption than derivative bounds.
- However, identified set will never be finite.
- Treatment effect can only be bounded away from 0 if it is in the correct direction.

Monotonicity Example



Donut: (-1,1)

Extensions – Monotonicity+

- If we know Y_0 is monotonic, but not which direction, learning the direction can help our set bounds.
- Perhaps more interesting here is an asymmetric case. If Y_0 is monotonic, but not Y_1, there is one point in the donut where the identified set for $Y_1(x) - Y_0(x)$ is not $(-\infty, \infty)$. That point is d_+. Is it meaningful to talk about the treatment effect over there?

Wrap up

We have discussed:

- Conditions for functional form assumptions to identify treatment effects with a donut.
- Standard RD type derivative conditions under which set identification is possible.
- Data-driven conditions under which we can improve on the above.

Future Questions

- If we are allowing τ to be estimated anywhere, does that affect the best functional form procedure?
- Can we give more guidance on the selection of donut sizes?
- When manipulation has an error term which is unbounded, donut sizes grow asymptotically. Is there anything to do about this?