List of Unanswered Questions

Answers provided by: Jacob Bor, Sc.D., S.M.
Regression Discontinuity Designs in Public Health Research
Methods: Mind the Gap Webinar
September 27, 2018

Question: Has anyone used RDD for longitudinal data, perhaps within the context of a mixed effect model?

> We do this in an unpublished manuscript (Bor & Barnighausen 2015, "High stakes testing: chronic disease management in low-resource settings", available here: http://sites.bu.edu/jbor/research. We use a mixed effects / "growth curve" approach to model the evolution of CD4 counts for patients presenting just above / below the threshold. The key is that the random slopes (not just intercepts) need to be interacted with the typical RDD model terms. I am not aware of a published paper, though it is certainly possible one exists.

Question: Can we use different concepts of threshold? For example, access to certain public health intervention (access to care).

> If there is a threshold rule on a continuously-measured baseline covariate, you can implement RDD.

Question: It seems like the method is testing the difference at the point of threshold. What is the ability of the method to test beyond that point of threshold? E.g., what if the effect is only local to the threshold but not beyond?

> It is possible to reweight treatment effects beyond the threshold and/or to adjust for factors that account for effect heterogeneity. See https://www.tandfonline.com/doi/abs/10.1080/01621459.2015.1012259

Question: With respect to measurement 'noise' on the threshold variable, is there anything additional to be aware of if using a self-report measure for the variable such as self-reported pain or fatigue? I ask this because treatment targets such as pain and fatigue do not have 'hard' biological measures, and we do have to rely on self-report for treatment decisions for these?

> The key is that the measurement of the assignment variable that you have access to must be the measurement used in care; and that measurement cannot have been precisely manipulated by the patient/provider/anyone else.

Question: How do we choose the regression function (or the shape of regression curve)?

> Local linear regression (1) with a pre-specified bandwidth selector (2,3) is best (4).
> (1) https://onlinelibrary.wiley.com/doi/abs/10.1111/1468-0262.00183;
> (2) https://academic.oup.com/restud/article/79/3/933/1533189;
> (3) https://arxiv.org/pdf/1809.00236.pdf;
> (4) https://www.tandfonline.com/doi/abs/10.1080/07350015.2017.1366909

Question: I would like to know if multivariate RDD is well-developed (multiple dependent variable)?

I have never seen this in a single model (do you mean a seemingly-unrelated regression model?). However, you can certainly run RDD for different outcomes.

Question: In the air quality example, how would RDD be used to change the way the alerts are made? Would they try and get rid of the discontinuity so that AQI below the threshold still trigger the alerts?

If the poor air quality alerts are leading to protective health behaviors (staying indoors, wearing a mask, etc.) and reducing asthma ER visits, then perhaps the threshold for poor air quality should be lowered so that more days trigger alerts. Of course, this would have to be weighed against the costs of people staying indoors and not going to school, e.g., as well as possible attenuation of behavioral responses if more days are labeled "high risk" due to saturation / fatigue.

Question: How likely is that the selection of thresholds may be related to some kind of confounding by indication?

Impossible. If the assignment variable is a continuously-measured random variable that is not precisely manipulated by patients/providers, then random noise in the assignment variable will quasi-randomly assign patients to be above/below. Even if the relationship between the assignment variable and outcomes was a step function (pure confounding by indication), the relationship between the measured assignment variable and outcomes would be smooth, and for a large enough sample, RDD would give you consistent estimates at the threshold.

Question: Can you talk a little more about some ways to address the potential for individuals who receive the treatment to manipulate treatment assignment in RDD? Are there sensitivity tests you might consider?

The key issue is not whether individuals manipulate the actual treatment they receive (of course some will). The key is whether they can manipulate the value of the assignment variable, which is interpreted in RDD as generating quasi-random variation in treatment assignment at the threshold. The key sensitivity test for manipulation of the assignment variable is the density test proposed by McCrary https://www.sciencedirect.com/science/article/pii/S0304407607001133

Question: RDD seems too good to be true – being able to give the intervention to those who need it the most. But there is always a price to pay, so what is the price in RDD? Is it that you need a larger sample size compared to an RCT?

Yes, you need a larger sample size. You also only get a treatment effect at the threshold, rather than a treatment effect across the whole range of assignment variables. For continuous quality improvement, if you cared about effects away from the threshold, you might consider combining RDD with pre-intervention data on the relationship between the assignment variable and outcome, which can give rise to an interrupted time series or diff-in-diff design.

Question: Pennell et al. (2011) published a paper showing how an RDD would work if the unit of assignment was a group or cluster rather than a person. Are you aware of any studies that have implemented a group-RDD?

The Ludwig & Miller "Head Start" paper is a group RDD. As is the Anderson Female HIV Risk paper. I see no reason why causal identification would be any different. However, statistical inference would have to account for the non-independence of error terms within clusters.

As Pennell et al (2011) show, the design effect for a cluster RDD relative to cluster RCT is quite large. Given the expense of cluster-level intervention studies, randomization should be prioritized. When this is not possible, other higher-powered (though less causally rigorous) designs like diff-in-diff and interrupted time series are possible. RDD can certainly be used but may be most applicable in the case where a policy has already been implemented (or is being implemented) at very large scale.