A short note on regression discontinuity design and its use in estimating the effect of month of birth
Claire Crawford and Ellen Greaves

Identifying the causal impact of a particular programme or “treatment” on some outcome of interest has been the subject of a vast literature that has developed across many disciplines over the last 50 years or so. Following standard notation we can denote treatment status by the binary variable $T$, where $T=1$ denotes treatment and $T=0$ denotes no treatment, and potential outcomes when individuals do or do not receive treatment as $Y_1$ and $Y_0$ respectively. The evaluation problem arises because it is impossible to observe both $Y_1$ and $Y_0$ for any given individual: they either receive the treatment or they do not. One must therefore find some way to identify a set of counterfactual outcomes: what would have happened to the outcomes of treated individuals had they not received the treatment?

Many evaluation techniques have been developed to address this problem, which usually involve the construction of an appropriate control group who are expected to be very similar to the treatment group in many ways, and whose outcomes can therefore be used to proxy the counterfactual outcomes for those who receive the treatment. Ideally one would identify a control group who are as similar as possible to the treatment group in ways that are both observed and unobserved to the researcher: for example, one might want individuals in the control group to have the same underlying motivations and preferences as those in the treatment group (characteristics which researchers typically do not observe), as well as being from the same gender, ethnicity and education group, and so on.

The “gold standard” of evaluation designs is often thought to be the randomised control trial (RCT), in which individuals are randomly allocated to either a treatment or control group. With sufficiently large samples, such a design should mean that there are no differences in the distribution of characteristics (observed or unobserved) amongst the treatment and control groups, except by chance. The outcomes of the control group should therefore be a very good proxy for the outcomes of the treatment group in the absence of the treatment.

Regression discontinuity design (RDD) is often regarded as the quasi-experimental technique that comes closest to an RCT in appropriate applications (Lee and Lemieux, 2010). In an RDD setting, the probability of receiving treatment changes discontinuously as a function of one or more underlying variables: for example, pupils scoring below a certain threshold on a test might be assigned to receive additional support (e.g. extra classes), the effect of which one is interested in evaluating.

Under certain conditions (summarised in more detail below), the allocation of treatment on the basis of this underlying variable (the so-called “running variable” – in our example, the test score) is analogous to assignment in a randomised experiment. This means that a comparison of outcomes amongst individuals just below and just above the discontinuity – who, in our example, respectively do and do not receive treatment – recovers the causal effect of treatment at the discontinuity.

As such, RDD provides a way of identifying average treatment effects for a subgroup of the population – those very close to the discontinuity – under minimal assumptions (Hahn et al, 2001). Its main advantage is that it is highly likely that the treatment and control groups will be similar in all respects other than receipt of treatment and thus that we can identify the causal effect of receiving treatment on some outcome; a potential limitation is that the treatment effect for this subgroup may not always constitute a policy relevant parameter (although in many cases it will).
Following Hahn et al. (2001), to identify the causal effect of treatment using RDD, we require the following conditions to be met:

a) **The probability of treatment must vary discontinuously at some point with respect to the running variable** $Z$. Formally: $P(T_i = 1|Z_{i-}) \neq P(T_i = 1|Z_{i+})$, where $Z_{i-}$ refers to the region immediately below the discontinuity and $Z_{i+}$ immediately above.

b) **All other characteristics** ($X$) **must be continuous at the point of discontinuity**. Formally: $E(X_i|Z_i = Z)$ is continuous in $Z$ at the discontinuity, where $X_i$ represents all characteristics of pupils or their families that affect the outcome of interest. This assumption is similarly described by McCall and Bielby (2012) as requiring that the error term – containing the effect of all (unobserved) variables other than the treatment indicator on the outcome of interest – varies continuously with respect to the running variable. This assumption ensures that differences in other factors are not responsible for any differences in outcomes observed between the treatment and control groups around the discontinuity.

c) **Individuals do not select into (or are not selected into) the treatment on the basis of anticipated gains from treatment.** In our example, this assumption might mean that children try their best on the test (i.e. they do not try to perform poorly in order to ensure that they receive additional support); equally, it could mean that teachers do not allocate children into or out of the treatment group on the basis of information other than the test score (e.g. behaviour).

When the discontinuity is “sharp” – i.e. receipt of the treatment is determined solely on the basis of the running variable, and all individuals adhere to their treatment allocation – the average effect of treatment around the discontinuity at $\bar{Z}$ (under the assumptions above) is given by:

$$E(\beta_i|Z_i = Z) = \tau = E(Y_i|Z_i = Z_{i+}) - E(Y_i|Z_i = Z_{i-})$$

McCall and Bielby (2012) equivalently write the estimated treatment effect as follows:

$$\tau = \lim_{d \to 0} E(Y|c + d > Z \geq c) - \lim_{d \to 0} E(Y|c - d < Z < c)$$

where $c$ is the cut-off and $d$ is a small window either side of the cut-off. This notation emphasises that RDD is most likely to produce an unbiased estimate of the treatment effect when comparing individuals just above and just below the cut-off (as $d \to 0$), as this is where the assumptions outlined above are most likely to hold. (The trade-off, of course, is that including observations further away from the discontinuity increases the sample size and thus the statistical power of your estimates.)

However, RDD techniques are often applied to “natural experiments”: discontinuities whose original aim was not to inform an evaluation of the treatment involved. As such, it is often the case that the assignment of treatment is “fuzzy” rather than “sharp” (i.e. that not all individuals appear in the group implied on the basis of their value of the running variable). In the case of our remediation example, this might mean that the teacher decides that some pupils who scored above the cut-off on the relevant test are still in need of additional support and thus are switched – for non-random reasons – from the control group to the treatment group.

This kind of fuzzy allocation does not preclude the use of RDD techniques; it simply changes the implementation and interpretation of the estimated treatment effect. Specifically, one tends to classify individuals according to their “intended” treatment assignment on the basis of the running variable.
Comparing the outcomes of individuals who were supposed to be allocated to the treatment group with those who were supposed to be allocated to the control groups produces an estimate of the “intention to treat” effect: the effect of the treatment on those who were supposed to receive it rather than those who actually received it. The advantage of such an approach is that it avoids the potential bias that might arise if individuals switch groups for reasons that are related to their potential gain from the treatment (which is likely to be the case in many circumstances).

The application to month of birth (encompassing effects on school starting age, age at test, etc)

Several papers – including Cahen and Cohen (1989), Fredriksson and Ockert (2013), Crawford, Dearden and Greaves (2014), and many others – have applied a regression discontinuity design to the issue of estimating the impact of month of birth (or one of the effects associated with month of birth, such as school starting age or age at test) on a series of outcomes of interest.

In many school systems – including England’s – the age at which children start school (and sit public examinations) is determined on the basis of their birthday. For example, in England, children typically start school in the September after they turn four, so a child born on 31st August will start school when they are four years and one day old, while a child born just one day later, on 1st September, will start school when they are five years old, creating a clear discontinuity in terms of school starting age on the basis of date of birth.

In England, it is rare for children to start school early or late, or to be held back or accelerated a year once they have started school. We can therefore think of this discontinuity as being very close to a sharp RDD: if we regard the “treatment” in this case as having started school older, then the vast majority of September-borns will receive the treatment and the vast majority of August-borns will not. Assuming that there are no other differences in characteristics between these two groups in the vicinity of the discontinuity, comparing the outcomes of children born just before and just after the discontinuity should provide us with a causal estimate of the effect of starting school older.

In other countries, such as the US, however, starting school early or late (or being held back or accelerated) is much more common. In such countries, therefore, some children receive the treatment when they should not have done, while others who should have received the treatment do not. In such cases, the discontinuity is fuzzy, because the running variable does not perfectly determine the allocation of treatment. As described above, this is of particular concern when the reasons for deviation are not random, which is highly unlikely in this case.

The papers cited above differ in terms of how they deal with this potential endogeneity of the treatment indicator (i.e. how they deal with the fact that the discontinuity is fuzzy rather than sharp). These differences may be symptomatic of the way in which different disciplines apply RDD techniques.

For example, most papers by economists (e.g. Fredriksson and Ockert, 2013; Crawford et al, 2014) tend to follow the “intention to treat” approach outlined above; that is to say, they use month of birth or expected age at school entry (or similar) to allocate treatment, and compare the outcomes of individuals assigned to the treatment and control groups in this way in order to calculate unbiased intention to treat effects.

The approach taken by Cahan and Cohen (1989) is very different. They, instead, omit children who do not conform to the discontinuity rule, as well as those born immediately either side of the discontinuity (on the grounds that these children were the most likely not to conform). This means that the
treatment effect they estimate applies only to those who conformed to the treatment allocation, and is estimated just on those who were born reasonably far from the discontinuity. In some ways, this runs contrary to the principles underlying the regression discontinuity design, as it omits those for whom the assumptions underlying this approach are most likely to hold. It is also not clear that an estimate based entirely on those who “comply” with the treatment allocation will produce the most policy relevant estimate in this case (although similar assumptions underlie other commonly used evaluation techniques, such as instrumental variables).

RDD techniques will, no doubt, continue to be used by researchers interested in estimating the effects of month of birth, school starting age or age at test on various outcomes. For readers interested in understanding the evidence from a range of disciplines, it will be important to bear in mind these fundamentally different approaches to the application of RDD techniques and interpret the evidence accordingly.

References


