Regression Discontinuity Design

MARC MEREDITH and EVAN PERKOSKI

Abstract

Social scientists search for interventions in the real world that approximate the conditions of an experiment. One form of such natural experiments that is increasingly used in social science research is regression discontinuity (RD). RD designs are possible when there are thresholds that cause large changes in the assignment of treatments on the basis of small differences in a variable. For example, a high school junior in the state of Pennsylvania who scored 214 out of 240 on the 2012 PSAT test received the treatment of being a National Merit Semi-Finalist, whereas a comparable student who scored 213 did not. The intuition behind a RD design is that we often can learn something about the effects of a treatment by comparing observations that barely receive a treatment (e.g., individuals with scores of 214 and just above on the PSAT) to observations that barely miss receiving a treatment (e.g., individuals who score 213 and just below on the PSAT). We discuss the assumptions under which the effects of treatment that are assigned based on a discontinuous threshold can be estimated using a RD design. We then illustrate how graphical analysis can be used to illustrate whether these assumptions are likely to hold. We conclude by discussing two examples of cutting-edge research that employs RD designs and discussing areas of future research.

INTRODUCTION

Social scientists often seek to understand the effect that different events and policies have on the world: economists study the relationship between the availability of unemployment insurance and the duration of unemployment, criminologist study whether drug and alcohol rehabilitation in prisons reduces recidivism, and political scientists study how media exposure affects voter turnout. We can think of these sorts of events as treatments affecting subsets of the population; the prisoners who receive rehabilitation are considered *treated* while those that do not are *untreated*. The impact that these treatments have on the treated population is referred to as a *treatment effect*. In other words, a treatment effect is a measure of how some intervention, event, or exposure affects an outcome of interest. A variety of approaches are used to estimate treatment effects. One approach

Emerging Trends in the Social and Behavioral Sciences. Edited by Robert Scott and Stephen Kosslyn.
 ISBN 978-1-118-90077-2.

that has been increasingly employed in social science research is regression discontinuity (RD). The increase in the use of RD reflects the ability of RD designs to maintain many of the desirable properties associated with experimentation in situations where experimentation is not ethical, feasible, or practical.

The goal of this essay is to provide a basic overview of RD. To do so, we first define and discuss the appeal of natural experiments more broadly. We then highlight conditions that must exist for an RD design to be feasible. Next, we discuss the assumptions that underlie RD designs, reasons why these assumptions may be violated, and methods people use to judge whether these assumptions are reasonable. The essay ends by illustrating many of these points through a discussion of some well-known applications of RD in the social sciences.

Our goal is to familiarize readers with RD designs and to provide a solid foundation for future learning and research. We do not include much technical discussion of RD designs and the relevant estimation procedures. Readers interested in learning more about the technical details behind RD should refer to work of Guido Imbens and Thomas Lemieux (2008) and David Lee (2009), among many other excellent sources listed at the conclusion of this essay.

FOUNDATIONAL RESEARCH

WHY NATURAL EXPERIMENTS?

Experimentation has long been an essential method of inquiry and discovery in the physical sciences. Students in high school chemistry classes are taught the value of experimentation in laboratories where they compare the properties of a baseline solution against the same solution with a known quantity of another ingredient added. By doing his or her best to make sure there are no other differences between the control solution and the solution treated with that additional ingredient, the student can easily identify the effect of adding the extra ingredient. Laboratory settings are ideal to maintaining experimental control: precise equipment and sterile environments mean that it is relatively easy to apply a treatment to two nearly identical solutions. While the hard sciences have a clear advantage in this regard, social scientists have increasingly come to recognize the benefits of experimentation to their own research, which has lead to a tremendous growth of social science experiments in recent years.

This increased use of experimentation highlights the desire for more *internal validity* in social science research. *Internal validity* refers to one's ability to ensure that the observed differences between the control solution and that which receives the additional ingredient reflect only the effect of that extra ingredient. If the beaker containing the treated solution was not properly washed, for instance, this would reduce internal validity because the residual contents might produce some difference between the treated and control solutions.

In social science research, the greatest threat to internal validity is often the ability for peoples' actions and characteristics to affect whether they receive a treatment, which is called *selection*. Suppose we want to study how watching the presidential debates affects voter turnout. Because people who watch the debates vote at higher rates, we might be tempted to conclude that watching the debates affects whether people vote. However, this difference could also reflect that those who choose to watch the debates already are more likely to vote. The difference in the likelihood that a debate watcher and a nondebate watcher will vote prior to watching the debate is an example of *selection bias*. *Selection bias* refers to the differences that selection causes between the treatment and controls groups before any treatment is administered. We can be more confident that the differences in turnout that we ultimately observe reflect the effect of watching debates, rather than selection bias, if people are brought into a laboratory and randomly assigned to either watch or not watch the debate.

Unfortunately, achieving high internal validity often reduces *external validity*. *External validity* refers to the ability to extrapolate the results of a study to the broader world. Do we expect to find similar results if we did the same experiment on another group of people at another point in time, or are these results only applicable to the current test conditions? The findings from a study with high external validity are relevant to the world beyond the experimental population. Returning to our hypothetical presidential debate experiment, our findings would be externally valid if the effect of watching a debate in the laboratory setting were similar to the effect of watching a debate at home. We might be concerned, for example, that people pay more attention to the debate when watching in a laboratory than they would if they were watching at home, which may cause the laboratory study to overestimate the effect that watching the debate will have on most people.

This tension between internal and external validity has lead social scientists to seek out *natural experiments*. Natural experiments are situations in the real world that approximate experimental conditions. For example, the draft lottery in the United States during the Vietnam War caused men born on March 2nd, 1951 to be more likely to serve in the army than men born on March 3rd, 1951. From this we can learn something about how serving in the military affects political ideology by comparing the political beliefs of those who were born on March 2nd, 1951 and March 3rd, 1951. The advantage of such natural experiments is that they overcome many of the internal validity concerns that might result from simply comparing those who, by their own choice, select into and do not select into the army. In addition, natural experiments are useful in situations when reproducing the condition is simply not feasible, either for ethical or practical reasons. Military service and a draft lottery, for obvious reasons, would be impossible to reproduce in a lab.

When Is An RD Design Feasible?

The use of natural experiments in the social sciences is limited only by their existence. How many things naturally occur in the world to cause two otherwise similar groups of individuals to receive different treatments? It turns out that there are more than you might expect. In particular, discontinuous thresholds, which are a required part of any RD design feasible, frequently occur.

A *discontinuous threshold* refers to a situation where a treatment is assigned on the basis of whether the value of some variable, often called a *forcing variable*, is above or below a certain value. It is called *discontinuous* because there is a jump in the probability of treatment at this threshold. To illustrate this point, consider the example of the National Merit Scholarship Program—a prestigious scholarship that many high school juniors compete to receive by taking a standardized test. To be named a National Merit Semi-Finalist in 2012, a high school junior in the state of Pennsylvania needed to score at least 214 out of 240 on the PSAT test. In this example, the treatment of being a National Merit Semi-Finalist varies depending on whether a forcing variable, the test score, is above or below the 214-point threshold. Those who score above 214 are treated, while those scoring under 214 are untreated. As a result, there is a 100% increase in the probability of being a National Merit Semi-Finalist at the 214-point threshold.

There are myriad examples of discontinuous thresholds that determine treatment. US Citizens can vote when they turn 18, so whether one's age is above a threshold of 18 years determines whether he or she receive the treatment of being eligible to vote. The Earned Income Tax Credit, a tax credit that is designed to incentivize people to work in low-income jobs, is only available to a single individual who earned less than $13,980. Thus, whether one's income is below a threshold of $13,980 determines eligibility for the credit. Finally, a 33-year-old male must run a marathon in 3 h and 5 min to qualify to compete in the 2014 Boston Marathon. Whether or not such an individual's previous marathon time is less than 185 min determines if he is eligible to run in Boston.

The intuition behind an RD design is that we can compare people who happen to fall just above or just below one of these discontinuous thresholds to estimate a treatment effect. Returning to the case of the National Merit Scholarship Program, we may be interested in knowing whether receiving this scholarship increases college attendance. Selection bias makes it so we cannot assess the impact of the scholarship simply by comparing the rates of college attendance among those who do and do not receive the scholarship; there are too many other differences besides Semi-Finalist status between those students who score, for example, 235 and students who score 150 to attribute differences in college attendance solely to the scholarship. However, we do expect that students who score 213 and who score 214 on the PSAT would be very similar. Thus, observing that those who scored 214 are substantially more likely to attend college than those who score 213 would be suggestive that the National Merit Scholarship Program increases college attendance.

Assumptions of RD Designs

While a discontinuous threshold is necessary for an RD design to be feasible, its presence alone is insufficient to guarantee that one can be used. First, it is essential that the discontinuous threshold affect the assignment of treatment. If people above the threshold are no more likely to be treated than people below the threshold, then it cannot be used. However, this does not mean that everyone above the threshold has to receive a different treatment than everyone below the threshold. When the probability of treatment goes from 0% to 100% around the threshold, it results in a *sharp discontinuity*. The National Merit Scholarship Program is an example of a sharp discontinuity because everyone who scores above the threshold is a semi-finalist, while no one who scores below the threshold is a semi-finalist.

Following are two graphs that use simulated data from our National Merit Scholarship Program example to illustrate what sort of patterns appear in the presence of sharp discontinuities. Figure 1 demonstrates a sharp discontinuity because the probability of becoming a National Merit Semi-Finalist jumps to one for those scoring at least 214 points on the PSAT, whereas everyone scoring less than 214 on the PSAT has probability zero of becoming a semi-finalist. Figure 2 plots the college attendance rate against PSAT score. It is clear from the graph that there is a jump upward in college attendance among those who score above 214 on the PSAT, which is consistent with the idea that National Merit Semi-Finalist status increases college attendance rates.

Discontinuous thresholds do not always generate sharp discontinuities. Many times those assigned to treatment never actually receive the treatment. For example, a person eligible for an Earned Income Tax Credit might file his or her taxes without knowing that the credit is available. These situations are referred to as *fuzzy discontinuities* wherein the probability of treatment

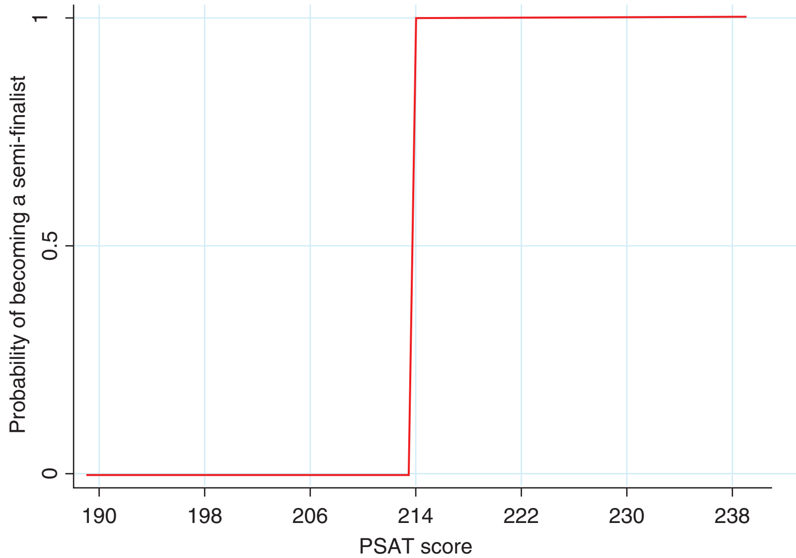

Figure 1 Sharp RD design—probability of becoming a semi-finalist by PSAT score.

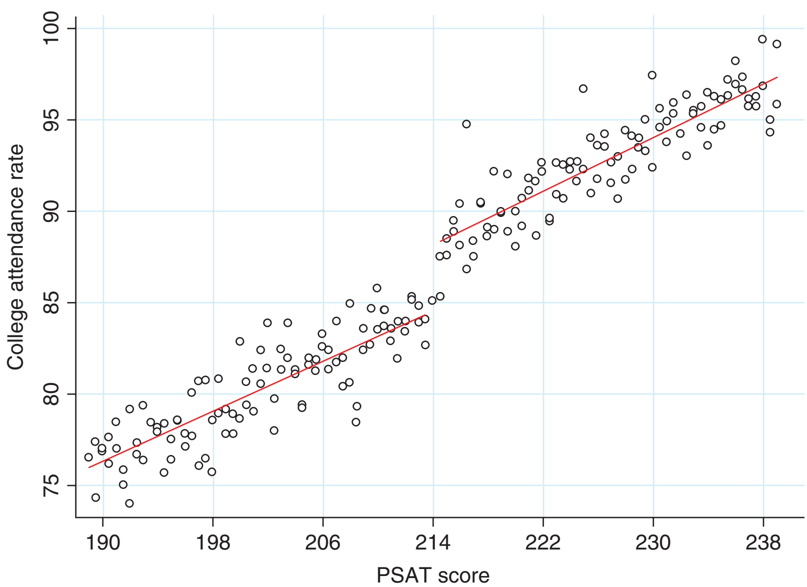

Figure 2 Sharp RD design—college attendance rate by PSAT score.

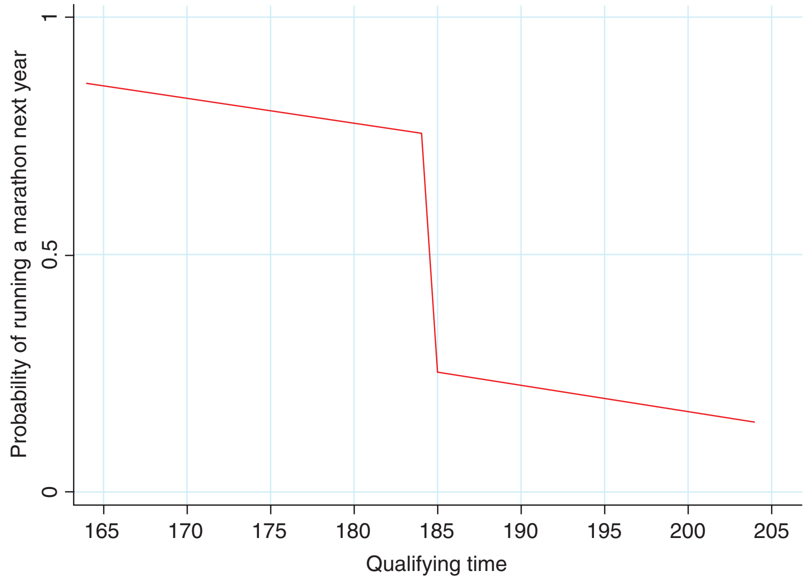

Figure 3 Fuzzy RD design—probability of running in a marathon by qualifying time.

changes at the threshold but not by 100%. In other words, not everyone above the threshold necessarily gets treated while some people below the threshold might get treated.

In the following, we simulate some data to illustrate a fuzzy discontinuity. Figure 3, the first graph, plots a runner's qualifying time against a measure of whether he or she runs another marathon in the next year. While some people who qualify for the Boston Marathon do not run it and many people who do not qualify run some other marathon, we observe a discontinuous decrease in the probability of running a marathon in the next year for those who just missed qualifying. We might be interested in using this discontinuous threshold to explore whether running marathons reduces blood pressure. This is evident from Figure 4; those who ran the marathon in just under 185 min have lower levels of diastolic pressure at the end of the next year. We would generally expect that people who run marathons in similar times be in similar health. Thus observing that blood pressure discontinuously changes at the same point that there is a discontinuous change in the probability of running another marathon is consistent with marathon running being the cause of this discontinuous change in blood pressure.

Another assumption of RD designs is that the characteristics of people with values of the forcing variable just below the discontinuous threshold are similar to the characteristics of people with values of the forcing variable just above the discontinuous threshold. That is, there cannot be systematic

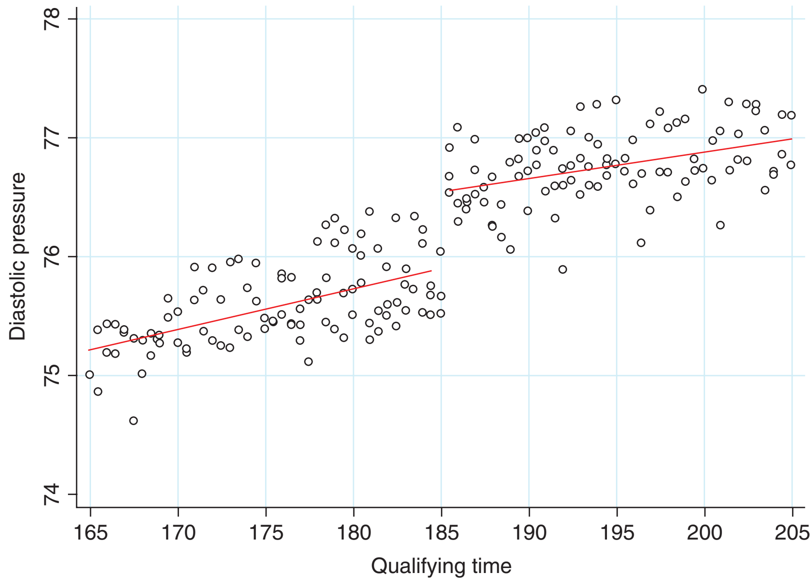

Figure 4 Fuzzy RD design—diastolic pressure by qualifying time.

differences between those who are just above and just below the discontinuous threshold except for the receipt of the treatment. This assumption is most problematic when individual agents manipulate the value of the forcing variable that determines treatment. For example, runners who are good at pacing themselves may be more likely to finish a marathon in just under 185 min than in just over 185 min. As we discussed earlier, this phenomena is known as *selection* and when it affects our findings we call the effects selection bias. Selection is a serious concern for RD because it can invalidate the assumption that people with values of the forcing variable just below the discontinuous threshold are similar to people with values of the forcing variable just above the discontinuous threshold. In such situation, RD produces biased estimates of the treatment effect.

Sorting around the threshold can even be problematic in cases where people do not manipulate the value of the forcing variable to affect their treatment. Suppose a city allows people to vote by mail, while a neighboring city does not. We do not expect that the availability of vote by mail to affect where someone lives, so we might be tempted to estimate the mobilizing effect of vote by mail by comparing the turnout rates of people who live near the border of the two cities. However, parents are likely to consider schools when deciding where to move. If one city's schools are known to be better, then there may be sorting around the boundary so that children can attend a certain school. Because those parents who intentionally move into the better district may be more politically involved, this sorting is likely to cause selection bias when comparing the turnout rates of people in the two cities.

A variety of statistical approaches can be used to estimate treatment effects using an RD design. The goal of the estimation procedure is twofold. First, control for any direct effect of the forcing variable on outcomes. Returning to our scholarship example, we would expect there to be some small difference in college attendance between those who score 213 and 214 on the PSAT absent any differences in scholarships. A statistical approach is likely to use additional information, like the change in college attendance between those who score 212 and 213 on the PSAT, to control for these differences. Second, a statistical approach is going to estimate the certainty that the differences in outcomes above and below a discontinuous threshold are caused by the treatment and not some other unmodeled factors. In other words, the model will tell us not only how much the threshold affects college attendance rates but also how confident we can be that the scholarship has its own significant effect.

Even when a natural experiment adheres to all of these assumptions and the necessary conditions, the estimation procedure could potentially produce misleading findings. While it is not the goal of this essay to provide a technical discussion of RD estimation, it is important to be able to recognize some of these pitfalls. One basic concern is whether the relationship between the forcing variable and the outcomes is modeled correctly. Modeling this relationship incorrectly can lead to either underestimating or overestimating treatment effects. Problems can also arise when the researchers uses too much or too little data; while observations right around the discontinuous threshold are thought to be most comparable, using too few observations makes it difficult to fit a model with confidence. A number of techniques have been developed recently to help researchers select models and data in a systematic way to help avoid these issues.

Finally, with any RD design it is worth considering the external validity of the findings. RD designs can be used to estimate a treatment effect for observations with a value of the forcing variable just around the discontinuous threshold. In our PSAT example, the RD design would be unable to estimate the effect of a National Merit Scholarship for individuals who instead of scoring around 214 on the PSAT scores about 114 instead. The same effect may not generalize to the general population, for example, if students who score lower on the tests may be more likely to attend college because they receive a scholarship.

The Importance of Graphing

Using graphs to better understand the data being studied in an RD design is an important aspect of the overall process. Here we will discuss two graphs that are essential to any RD: first, a plot of how the outcome varies as a function of the forcing variable, and second, a plot of how other variables that cannot plausibly be affected by a treatment vary as a function of the forcing variable.

Plotting the outcome variable against the forcing variable is extremely useful. It is primarily helpful for detecting whether a discontinuity actually exists. If there is no visible jump in the outcome variable around the discontinuous threshold, then it is unlikely that the treatment has a significant effect. In addition, it is useful to check if similar jumps exist elsewhere in the data. For example, suppose we observe a jump in college attendance around PSAT scores of 150 although there is no discontinuous threshold that affects treatment. If so, we might be less certain that the difference in outcomes near the 214-point threshold is caused by the treatment and not something else.

Plotting other variables against the forcing variable is useful for detecting the presence of selection. Take the previous example of the Boston Marathon qualifying time. Suppose we are concerned that experienced runners will pace themselves better, and thus will be more likely to finish a marathon in just under 185 min. To investigate this possibility, we can plot the age, previous marathon experience, and other observable characteristics of runners as a function of their finishing time. Figure 5 uses simulated data to show what such a plot might look like. The figure shows that while more experience is associated with a faster time, there are no systematic differences in experience of runners who finish in just over and just under 185 min. Showing that runners who finish in just under 185 min have similar observable characteristics to those who finish in just over 185 min helps to reassure us that the only difference between those who finish in just under and just over 185 min is the probability of running a marathon in the next year.

CUTTING-EDGE RESEARCH

Here we discuss two exemplary uses of RD design in recent literature. We first demonstrate how an RD is used to study how a municipality's revenue and funding affect levels of corruption and the quality of political candidates. In other words, does more funding result in more corrupt behavior? We then discuss how an RD is used to examine the political advantage that results from being the incumbent in the US House of Representatives?

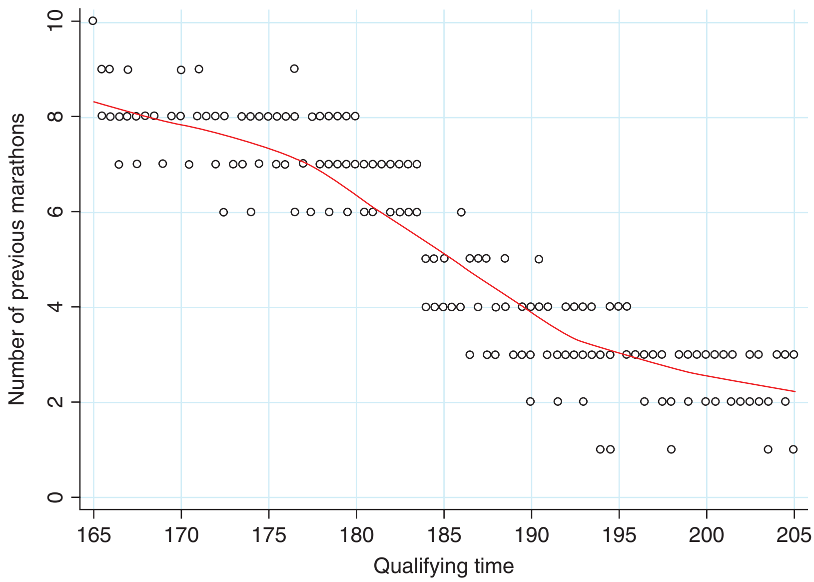

Figure 5 Previous marathon experience against qualifying time.

How Does Governmental Revenue Affect Corruption?

What is the relationship between political corruption, government revenue, and the quality of political candidates? On the one hand, greater revenue may make government jobs more attractive and as a result increase the quality of candidates seeking the job. However, on the other hand, greater revenue might also increase opportunities for rent seeking and other corrupt behavior, and thus increase the number of corrupt political candidates. Understanding this relationship is complicated by several facts. First, the state's willingness to provide local governments with money may depend on their perceptions how corrupt it already is. Second, other variables, such as income, could both affect the amount of government revenue and also the quality of political candidates. It is therefore nearly impossible to study this question without a research design that can untangle these highly correlated and seemingly interdependent factors.

Fernanda Brollo, Roberto Perotti, Tommaso Nannicini, and Guido Tabellini circumvent these problems using an RD design that is made possible by some unique features of Brazilian law. Brazilian municipalities receive federal funding based on their size and which state they are in. There are city population thresholds that increase a city's federal funding discontinuously. For example, a city with 34,999 citizens might receive substantially less money than a nearby city with 35,000 citizens. When states cross these discontinuous thresholds, they automatically receive additional funds. The authors examine how corruption levels and political candidate characteristics differ in town just below and just above the population thresholds. Because places with a similar population size should, on average, be similar in terms of corruption and the quality of political candidates, any observed differences can be attributed to the additional funding.

The authors find support for the hypothesis that additional revenues increase the level of corruption. They show that politicians in cities just above the population thresholds engage in more corruption than politicians in cities just below the population thresholds. Candidates for municipal office in cities just above population thresholds are also less likely to have a college degree than those in cities just below the population thresholds. In other words, where there is more money there is more political corruption and less qualified candidates to run the municipality.

As we discussed in the previous section, the authors' design hinges on the assumption that towns just below and just above the population thresholds are similar. In a number of graphical and empirical tests, they find no systematic differences between cities on either side of these thresholds. Thus, we are more confident that the increase in corruption above the discontinuous threshold is a result of additional revenues and not other factors that might differ between cities with more and less money.

DOES HOLDING OFFICE HELP YOU WIN OFFICE?

It is often said that political incumbents have a much higher chance of being reelected as their incumbency status affords them a number of advantages. For instance, while in office they can enact policies that will benefit constituents thereby increasing their favorability among them. Yet assessing the degree to which incumbents receive more support because they are incumbents is a much trickier question than initially meets the eye. How can we separate the effect of the variables that caused a candidate to win in the first place from the effect of incumbency? Both the importance and complexity of answering this question have generated a substantial amount of academic attention in recent years.

David Lee attempts to overcome these issues by employing RD to estimate the incumbency advantage a party receives from holding a seat in the US House of Representatives. Rather than looking at all winning and losing candidates, he focuses on those candidates that barely won and barely lost. Candidates that won and that lost by very small percentages should be extremely similar in terms of past experience, ability to fundraise, charisma, and other features that help candidates win elections. However, only those that win are treated with incumbency.

In this context, the forcing variable is the two-party vote share (i.e., percent of the votes cast for one of the two major parties) a candidate receives. Because a candidate wins a US House seat when he or she receives a plurality of the votes, there is a sharp discontinuity when the vote shares cross the 50% threshold. When the Democrat's candidate receives just under 50% of the vote, the Democrats have a 0% chance of being the incumbent party, as compared to when the Democrat's candidate receives just over 50% of the vote and the Democrat's have a 100% chance of being the incumbent party. This is a clear case of a sharp RD design.

Overall, Lee finds that incumbency has a significant and a positive impact on the chance of running again and subsequently, the chance of winning in future elections. The party that barely wins the election receives about an 8% increase in their vote share in the next election. As a result, this party is about 40% more likely to win the seat again in the next election. Candidates who barely win are also about 40% more likely to run again in the next election. These findings are consistent with the presence of large electoral benefits to incumbents that deter strong challenging candidates.

The validity of Lee's RD design hinges on the traditional consideration of whether candidates who barely win differ systematically from those who barely lose. Lee argues that it is arbitrary which candidate wins a close US House election. He presents a series of graphs that demonstrate the similarity of candidates across several dimensions, but overall his argument rests on the assumption that in these very close elections, some part of the vote is essentially random. For example, the composition of the electorate who votes depends on weather conditions on Election Day. This random component makes it almost equally likely that a candidate will win or lose an election by a small number of votes.

But is this assumption believable? Devin Caughey and Jasjeet Sekhon argue that it is not. Building on Lee's original data and adding in a number of new covariates, they find candidates who barely win House elections are actually quite different than candidates who barely lose. For example, they show that winners of close elections were more likely to be favored in Congressional Quarterly's October predictions of House Races. Candidates are also more likely to win close elections when their party controls the part of the state government that is in charge of counting votes. Such findings generate concern that selection bias may cause Lee's RD design to overstate the incumbency advantage.

CONCLUSIONS

As social scientists continue to look outward in search of natural experiments, we are likely to see more and more instances of RD designs. Compared to a study that analyzes observational data, the benefits of natural experiments are clear: while we primarily focused here on the tradeoff between internal and external validity, natural experiments also offer a better chance of understanding causality and a lower likelihood of biased inferences.

However, social scientists must take precautions. Before we begin analyzing natural experiments, we have to be confident that basic experimental conditions hold. With RD, we must be certain that the treated and untreated groups are similar and that selection is not occurring around the threshold for treatment. A violation of these basic assumptions could lead researchers to produce incorrect findings.

Moving forward, we expect further research to make RD estimation procedures more straightforward to implement. Currently there are many choices that researchers must make like how to specify the model and which data to include in their study. While we did not delve into these issues here, these choices can have important consequences of the inferences that readers draw from a study. We expect more research will be done, like recent work by Guido Imbens and Karthik Kalyanaraman, to generate theoretically motivated protocols on how these decisions can automatically be implemented.

We also expect more work on how to deal with violations of the assumptions that we laid out for RD designs. Almost anyone who has implemented an RD design has been forced to deal with something in their data that violates one the theoretical assumptions of RD designs. For example, sometimes treatments are assigned on the basis of multiple forcing variables rather than a single forcing variable. In other cases, the forcing variable by which treatment is assigned may be observed with some measurement error. Future work will help us understand how we can best deal with these violations, while preserving the benefits of something that approximates an experiment.

This future work is important because natural experiments and RD design will surely feature prominently in modern scholarship. New and unexpected natural experiments provide social scientists with unparalleled opportunities for learning. With natural experiments occurring around us every day, there is no limit to the types of questions that it can be used answered.

FURTHER READING

Brollo, F., Perotti, R., Nannicini, T., & Tabellini, G. (2010). The political resources curse. NBER Working Paper #15705.

Caughey, D., & Sekhon, J. S. (2011). Elections and the regression discontinuity design: Lessons from close U.S. house races, 1942–2008. *Political Analysis*, *19*(4), 385–408.

Hahn, J., Todd, P., & Van der Klaauw, W. (2001). Identification and estimation of treatment effects with a regression discontinuity design. *Econometrica*, *69*(1), 201–209.

Imbens, G., & Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, *79*(3), 933–959.

Imbens, G., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics*, *142*(2), 615–635.
Lee, D. S. (2008). Randomized experiments from non-random selection in U.S. house elections. *Journal of Econometrics*, *142*(2), 675–697.
Lee, D. S., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics*, *142*(2), 655–674.
Lee, D. S., & Lemieux, T. (2009). Regression discontinuity designs in economics. *Journal of Economic Literature*, *48*(2), 281–355.

MARC MEREDITH SHORT BIOGRAPHY

Marc Meredith is an Assistant Professor of political science at the University of Pennsylvania. His research examines the political economy of American elections, with a particular focus on the application of causal inference methods. Professor Meredith's substantive research interests include election administration, local political institutions, political campaigns, and voter decision-making, particularly as it relatives to economic conditions. His work can be found at www.sas.upenn.edu/~marcmere/.

EVAN PERKOSKI SHORT BIOGRAPHY

Evan Perkoski is a PhD candidate in political science at the University of Pennsylvania and a research fellow at the Belfer Center for Science and International Affairs at the Harvard Kennedy School of Government. Evan's research focuses on important issues in subnational conflict and political violence. In particular, his work seeks to better understand the dynamics and decision-making of violent nonstate actors like terrorist, insurgent, and rebel organizations. His work can be found at www.evanperkoski.com.

RELATED ESSAYS

Social Epigenetics: Incorporating Epigenetic Effects as Social Cause and Consequence *(Sociology)*, Douglas L. Anderton and Kathleen F. Arcaro
To Flop Is Human: Inventing Better Scientific Approaches to Anticipating Failure *(Methods)*, Robert Boruch and Alan Ruby
Repeated Cross-Sections in Survey Data *(Methods)*, Henry E. Brady and Richard Johnston
Ambulatory Assessment: Methods for Studying Everyday Life *(Methods)*, Tamlin S. Conner and Matthias R. Mehl
The Evidence-Based Practice Movement *(Sociology)*, Edward W. Gondolf
Meta-Analysis *(Methods)*, Larry V. Hedges and Martyna Citkowicz
The Use of Geophysical Survey in Archaeology *(Methods)*, Timothy J. Horsley