

Abstract Title Page
Not included in page count.

Title: Understanding the What Works Clearinghouse Standards for Regression Discontinuity Designs

Author(s): John Deke and Jill Constantine

Abstract Body

Limit 5 pages single spaced.

Background / Context:

Description of prior research and its intellectual context.

Regression discontinuity designs (RDDs) are considered to be one of the strongest nonexperimental designs available (Shadish, Cook, & Campbell, 2002) for the purpose of identifying the effects of an intervention. RDD can be used in situations in which assignment to a treatment group is based on a cutoff value on a continuous assignment variable (for example, when all students who score below the 20th percentile on a reading test are offered after-school tutoring.) The impact of the intervention is calculated as the regression-adjusted difference in the predicted outcomes for treatment and control group members at the cutoff value of the assignment variable.

Since Goldberger (1972a and 1972b) showed the theoretical appeal of the approach, numerous researchers have contributed to our understanding of RDD (Cook [2007] reviews this literature). In the past decade, researchers have made a renewed effort to bolster the theoretical underpinnings of RDD and advance the state of the art in estimating impacts and standard errors (Hahn, Todd, & Van der Klauuw, 2001; Imbens & Kalyanaraman, 2009; Lee & Card, 2008).

Meanwhile, the creation of the Institute for Education Sciences (IES) by the Education Sciences Reform Act of 2002 highlights the need for more rigorous education research that can identify causal, not just correlational, relationships between educational practices and student outcomes. The What Works Clearinghouse (WWC) was created to help advance IES' mission "to provide rigorous and relevant evidence on which to ground education practice and policy and share this information broadly" (<http://www.ies.ed.gov/aboutus/>). A key function of the WWC is to determine evidence that is "rigorous." Toward that end, the WWC has posted a set of standards that it uses to assess the rigor of research evidence from studies based on either experimental or matching designs

(http://www.ies.ed.gov/ncee/wwc/pdf/wwc_procedures_v2_standards_handbook.pdf). Recently the WWC has added a new standard for RDD studies

(http://www.ies.ed.gov/ncee/wwc/pdf/wwc_rd.pdf).

Purpose / Objective / Research Question / Focus of Study:

Description of the focus of the research.

The purpose of this paper is to explain the RDD standards and some of their implications for researchers who are designing new RDD studies. Based on our involvement in the development of these standards, we hope to help researchers better understand the standards, what they can do to design and conduct studies that are more likely to meet these standards, and how the standards might evolve in the future.

Setting:

Description of the research location.

(May not be applicable for Methods submissions)

Not applicable.

Population / Participants / Subjects:

Description of the participants in the study: who, how many, key features or characteristics.
(May not be applicable for Methods submissions)

Not applicable.

Intervention / Program / Practice:

Description of the intervention, program or practice, including details of administration and duration.
(May not be applicable for Methods submissions)

Not applicable.

Significance / Novelty of study:

Description of what is missing in previous work and the contribution the study makes.

Unlike other quasi-experimental designs (such as propensity score matching), RDD studies have the potential to be placed in the same WWC evidence category as randomized experiments. However it is one thing for a study to have the *potential* to reach the highest evidence category and quite another thing for a study actually to *achieve* that potential. The WWC RDD standards set a high bar, one that many published RDD studies do not reach.

Given our close involvement in the development of the WWC RDD standards (Jill Constantine is the project director for the WWC and John Deke drafted the standards), we are in a unique position to help other researchers better understand what they need to do in order to produce studies that will meet these evidence standards.

Statistical, Measurement, or Econometric Model:

Description of the proposed new methods or novel applications of existing methods.

These standards apply to studies that follow a regression discontinuity design (RDD).

Usefulness / Applicability of Method:

Demonstration of the usefulness of the proposed methods using hypothetical or real data.

RDD is widely used to estimate impacts of interventions. It is regarded as one of the most rigorous nonexperimental methods available.

Research Design:

Description of research design (e.g., qualitative case study, quasi-experimental design, secondary analysis, analytic essay, randomized field trial).
(May not be applicable for Methods submissions)

Not applicable.

Data Collection and Analysis:

Description of the methods for collecting and analyzing data.

(May not be applicable for Methods submissions)

Not applicable.

Findings / Results:

Description of the main findings with specific details.

(May not be applicable for Methods submissions)

Not applicable.

Conclusions:

Description of conclusions, recommendations, and limitations based on findings.

The WWC RDD standards set a high, but attainable, bar. With a clear understanding of the methodologies and reporting requirements needed to meet the standards, many more researchers should be able to conduct RDD studies that meet WWC standards. Because RDD can be employed in many settings in which experiments are not feasible, the proliferation of RDD studies that meet WWC standards will raise the level of rigor in the education literature beyond what could have been obtained with experiments alone and will therefore facilitate better decision making on the part of education policymakers and practitioners.

Appendices

Not included in page count.

Appendix A. References

- Cook, T. (2008). 'Waiting for life to arrive': A history of the regression-discontinuity design in psychology, statistics and economics. *Journal of Econometrics*, 142, 636-654.
- Goldberger, A. S. (1972a). Selection bias in evaluating treatment effects: Some formal illustrations. Madison, WI: Institute for Research on Poverty.
- Goldberger, A. S. (1972b). Selection bias in evaluating treatment effects: The case of interaction. Madison, WI: Institute for Research on Poverty.
- Hahn, J., Todd, P., & van der Klauuw, W. (2001). Regression discontinuity. *Econometrica*, 69(1), 201-209.
- Imbens, G., & Kalyanaraman, K. (2009). Optimal bandwidth choice for the regression discontinuity estimator. NBER Technical Working Paper Series: Technical Working Paper 14726. Cambridge, MA: National Bureau of Economic Research.
- Imbens, G., & Lemieux, T. (2008). Regression discontinuity designs: A guide to practice. *Journal of Econometrics* 142(2), 615-635.
- Lee, D., & Card, D. (2008). Regression discontinuity inference with specification error. *Journal of Econometrics* 142(2), 655-674.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.

Appendix B. Tables and Figures

Not included in page count.