

Resource Guide 3: Quasi-Experimental Study Methods

This bibliography, designed for users of the Self-Sufficiency Research Clearinghouse, provides several resources that may be consulted to gain a better understanding of evaluation research, and the different methodological approaches that can be taken in conducting such studies. Specifically, this bibliography includes resources that address quasi-experimental methodologies. Researchers may find the reports, briefs, and articles provided in this bibliography useful in helping to plan and conduct studies and analysis of their own. Practitioners and policymakers can utilize these resources to better understand the approaches and data that come out of quasi-experimental evaluations. This annotated bibliography is a tool to support practitioners and policymakers in making decisions about funding and collaborating with such researchers to ensure that high quality quasi-experimental studies are being conducted and account for effective internal performance management practices.

Bloom, H.S. (2009). *Modern regression discontinuity analysis*. New York, NY: MDRC
<http://www.mdrc.org/publications/539/full.pdf>

This paper provides a detailed discussion of the theory and practice of modern regression discontinuity (RD) analysis for estimating the effects of interventions or treatments. Part 1 briefly chronicles the history of regression discontinuity analysis and summarizes its past applications. Part 2 explains how in theory a regression discontinuity analysis can identify an average effect of treatment for a population and how different types of regression discontinuity analyses — “sharp” versus “fuzzy” — can identify average treatment effects for different conceptual subpopulations. Part 3 of the paper introduces graphical methods, parametric statistical methods and nonparametric statistical methods for estimating treatment effects in practice from regression discontinuity data plus validation tests and robustness tests for assessing these estimates. Section 4 considers generalizing regression discontinuity findings and presents several different views on and approaches to the issue. Part 5 notes some important issues to pursue in future research about or applications of regression discontinuity analysis.

Bloom, H.S., Michalopoulos, C., Hill, C.J., & Lei, Y. (2002). *Can nonexperimental comparison group methods match the findings from a random assignment evaluation of mandatory Welfare to Work programs?* New York, NY: MDRC
<http://www.mdrc.org/publications/66/full.pdf>

The present paper addresses two questions: (1) which nonexperimental comparison group methods provide the most accurate estimates of the impacts of mandatory welfare-to-work programs; and (2) do the best methods work well enough to substitute for random assignment experiments?

The authors compare findings for a number of nonexperimental comparison groups and statistical adjustment procedures with those for experimental control groups from a large-sample, six-state random assignment experiment — the National Evaluation of Welfare-to-Work Strategies (NEWWS). The methods examined combine different types of comparison groups (in-state, out-of-state, and multi-state), with different propensity score balancing approaches (sub-classification and one-to-one matching) and different statistical models (ordinary least squares [OLS], fixed-effects models, and random-growth models). These methods are assessed in terms of their ability to estimate program impacts on annual earnings during a short-run follow-up period, comprising the first two years after random assignment, and a medium-run follow-up period, comprising the third through fifth years after random assignment. The tests conducted use data for an unusually rich set of individual background factors, including up to three years of quarterly baseline earnings and employment histories plus detailed socio-economic characteristics.

Khandker, S.R., Koolwal, G.B., & Samad, H.A. (2010). Handbook on impact evaluation: Quantitative methods and practice. Washington, DC: The World Bank.

http://www.esf-agentschap.be/uploadedFiles/Voor_ESF_promotoren/Zelfevaluatie_ESF_project/statistiek%20voor%20quasi%20experimenten.pdf

Identifying the precise effects of a policy is a complex and challenging task. This issue is particularly salient in an uncertain economic climate, where governments are under great pressure to promote programs that can recharge growth and reduce poverty. At the World Bank, our work is centered on aid effectiveness and how to improve the targeting and efficacy of programs that we support. As we are well aware, however, times of crisis as well as a multitude of other factors can inhibit a clear understanding of how interventions work—and how effective programs can be in the long run.

Handbook on Impact Evaluation: Quantitative Methods and Practices makes a valuable contribution in this area by providing, for policy and research audiences, a comprehensive overview of steps in designing and evaluating programs amid uncertain and potentially confounding conditions. It draws from a rapidly expanding and broadbased literature on program evaluation—from monitoring and evaluation approaches to experimental and nonexperimental econometric methods for designing and conducting impact evaluations.

Recent years have ushered in several benefits to policy makers in designing and evaluating programs, including improved data collection and better forums to share data and analysis across countries. Harnessing these benefits, however, depends on understanding local economic environments by using qualitative as well as quantitative approaches. Although this Handbook has a quantitative emphasis, several case studies are also presented of methods that use both approaches in designing and assessing programs.

The vast range of ongoing development initiatives at institutions such as the World Bank, as well as at other research and policy institutions around the world, provide an (albeit wieldy) wealth of information on interpreting and measuring policy effects. This Handbook synthesizes the spectrum of research on program evaluation, as well as the diverse experiences of program officials in the field. It will be of great interest to development practitioners and international donors, and it can be used in training and building local capacity. Students and researchers embarking on work in this area will also find it a useful guide for understanding the progression and latest methods on impact evaluation.

Moore, K.A. (2008). Quasi-experimental evaluations: Part 6 in a series on practical evaluation methods. Washington, DC: Child Trends.

http://www.childtrends.org/Files/Child_Trends_2008_01_16_Evaluation6.pdf

Although experimental studies are described as the “gold standard” for assessing the effectiveness of a program in changing outcomes, in some cases, quasi-experimental studies may be more feasible or appropriate. Many types of quasi-experimental studies are possible. For example, an implementation study can provide valuable information on whether, how, and for whom services are provided, and an outcome monitoring study can provide early information on whether outcomes are changing as expected. In this brief, the author describes varied types of quasi-experimental evaluations and the circumstances under which they are useful.

Seith, D.C., Verma, N., Bloom, H.S., & Galster, G.C. (2003). Exploring the feasibility and quality of matched neighborhood research designs.

New York, NY: MDRC

<http://www.mdrc.org/publications/346/full.pdf>

Many evaluations of neighborhood-level interventions have relied on neighborhood matching strategies, but to the authors’ knowledge the validity of this methodology has not been tested. This paper responds to the demand for rigorous evidence on the use of quasiexperimental neighborhood matches for assessing the effectiveness of community-wide interventions. Using neighborhood-level data in Cleveland, Ohio and Philadelphia, Pennsylvania, to match potential target neighborhoods to virtual comparison neighborhoods, this paper evaluates the number of target neighborhoods that can be matched and how well they stay matched over time. It identifies a range of matching variables and constraints that appear to strike the best balance between matchability and match quality. The results compare favorably to those generated by two less restrictive alternatives. The paper ends with suggestions for replication in other sites, with other outcomes, and in other time periods. It tentatively concludes that the neighborhood-matching algorithm described in this analysis is both operationally feasible and offers respectable accuracy in detecting the magnitude of impacts that might be expected from neighborhood-based employment interventions.

Note: Unless otherwise noted all information is from the publication and/or Web site.