

Dieser Artikel ist eine Übersetzung von Cook, T. D., Steiner, P. M., & Pohl, S. (2010). How bias reduction is affected by covariate choice, unreliability, and mode of data analysis: Results from two types of within-study comparisons. *Multivariate Behavioral Research*, 44(6), 828-847.
doi:10.1080/00273170903333673

Theorie, Methoden und Praxis der Evaluation

Die relative Bedeutung der Kovariatenwahl, Reliabilität und Art der Datenanalyse zur Schätzung kausaler Effekte aus Beobachtungsdaten

Thomas D. Cook

Institute for Policy Research, Northwestern University, Evanston

Steffi Pohl

Otto-Friedrich-Universität Bamberg, Nationales Bildungspanel, Deutschland, Bamberg

Peter M. Steiner

Department of Educational Psychology, University of Wisconsin-Madison

Zusammenfassung

Häufig können in Evaluationsstudien keine randomisierten Experimente zur Schätzung des kausalen Effekts eines Programms durchgeführt werden, so dass auf quasi-experimentelle Designs zurückgegriffen werden muss. Die Güte der Effektschätzung in Quasi-Experimenten hängt jedoch stark von der Qualität des Designs und der statistischen Adjustierung ab. In diesem Artikel wird mithilfe von speziellen Studiendesigns, in denen Effektschätzungen aus (adjustierten) quasi-experimentellen Designs mit denen aus randomisierten, experimentellen Designs verglichen werden, die relative Bedeutung der Kovariatenwahl, der Unreliabilität bei der Messung dieser Kovariaten sowie der Wahl des Analyseverfahrens (Regressionsverfahren oder verschiedene Formen von Propensity Score-Verfahren) bei der Schätzung von kausalen Effekten aus Quasi-Experimenten untersucht. Die Ergebnisse dieser Studienvergleiche zeigen, dass die Kovariatenwahl den größten Einfluss auf die Biasreduktion hat, gefolgt von der reliablen Messung der Kovariaten. Die Art der Datenanalyse spielt hingegen eine untergeordnete Rolle. Die theoretische Überlegenheit von Propensity Score-Verfahren gegenüber Regressionsverfahren konnte in den praktischen Anwendungen bisher nicht bestätigt werden.

Schlagwörter: Kausale Effekte, Wirkungsanalyse, Quasi-experimentelle Designs, Propensity Scores

This article is a translation of Cook, T. D., Steiner, P. M., & Pohl, S. (2010). How bias reduction is affected by covariate choice, unreliability, and mode of data analysis: Results from two types of within-study comparisons. *Multivariate Behavioral Research*, 44(6), 828-847. doi:10.1080/00273170903333673

Theory, Methods and Practice of Evaluation

On the relative importance of covariate choice, reliable measurement and analytical method for estimating causal effects from observational studies

Thomas D. Cook

Institute for Policy Research, Northwestern University, Evanston

Steffi Pohl

University of Bamberg, National Educational Panel Study (NEPS), Germany

Peter M. Steiner

Department of Educational Psychology, University of Wisconsin-Madison

Abstract

In evaluation studies it is often not possible to conduct randomized experiments for estimating causal effects of a program. Instead, quasi-experimental designs are frequently used. The causal interpretation of a statistically adjusted estimate is only warranted if the assumptions of the underlying quasi-experiment are met and the statistical analysis is properly conducted. In this article the relative importance of covariate choice, the covariates' reliable measurement, and the choice of analytic method (regression analysis or different types of propensity score methods) are investigated using a within-study comparison which compares the effect estimate from adjusted quasi-experimental design to the corresponding effect estimate from a randomized experiment. The results show that the covariate choice is the most important factor for removing selection bias. The reliable measurement of covariates is the second most important factor. The choice of an analytic method does not have a strong impact on bias reduction. To date, the superiority of propensity score methods in comparison to regression methods is not yet supported by empirical evidence.

Keywords: Causal Effects, Impact Analysis, Quasi-experimental Designs, Propensity Scores