Business and Economics Journal
ISSN: 2151-6219

Sapra, Bus Eco J 2016, 7:2
DOI: 10.4172/2151-6219.1000e106

Editorial Open Access

Causality Inference with Observational Data in Economics

Sapra SK*

Department of Economics and Statistics, California State University, Los Angeles, USA

Editorial

Economists use non-experimental data for causal inference in most empirical research. The causal effect of a treatment for an individual is the difference between potential outcomes with treatment and without treatment. Angrist and Piscke present a simple example from health insurance in which the treatment group consists of individuals with health insurance, the control group consists of individuals without health insurance and the outcome is the health status. Selection bias is a serious problem in measuring causal effect of a treatment with non-experimental data. However, random assignment of treatment to subjects eliminates the selection bias since it makes the treatment variable independent of potential outcomes.

Application of data analysis to answer cause-and-effects questions in economics constitutes the field of applied econometrics. Conclusions derived under ceteris paribus conditions have a causal interpretation. Real-world other things equal comparisons are difficult to accomplish. Applied econometricians use data to achieve other-things-equal in spite of the obstacles, including selection bias or omitted variable bias encountered on the path from raw data to reliable causal knowledge. The path to causal understanding is complicated by selection bias, but applied econometricians employ clever techniques to eliminate or minimize it and link cause and effect. The purpose of this note is to discuss the merits and drawbacks of these techniques, including randomized trials, regression, instrumental variables, regression discontinuity design, and difference in differences. Details can be found in in Angrist and Pischke.

Randomized Trials: The gold standard in cause-effect investigations is a randomized experiment, often called a randomized trial. Experimental random assignment is both a framework for causal questions and a benchmark for comparison for other methods of causal inference. The main challenge for applied econometricians is elimination of the selection bias that arises from unobserved differences between treatment and control groups. In a randomized trial, researchers change the variables of interest, such as the availability of college financial aid for a group selected through randomization using something like a coin toss. Changing circumstances randomly makes it highly likely that the variable of interest is unrelated to the many other factors determining the outcomes we are interested in studying. Thus random assignment has the same effect as holding everything else fixed. Unfortunately, randomized social experiments are expensive to conduct and may be slow to yield results. Often, therefore, applied econometricians turn to less powerful but more accessible research designs. Angrist and Pischke show how wise application of some econometric tools brings us as close as possible to the causality-revealing power of a real experiment. We now turn to a discussion of these tools.

Regression: When random assignment is not practical, researchers look for alternate routes to causal knowledge. There are alternative econometric tools available, which if used skillfully, can have much of the causality-revealing power of a real experiment. The most basic of these tools is regression, which compares treatment and control subjects who have the same observed characteristics. Regression-based causal inference is based on the assumption that when key observed variables have been made equal across treatment and control groups, selection bias from the factors, which we cannot observe is also mostly eliminated.

A regression coefficient approximates the causal effect that might be revealed in an experiment when the conditional expectation function (CEF) is causal. A regression has a causal interpretation when the CEF is causal. The CEF is causal when it describes differences in average potential outcomes for a fixed reference population. Causal relationship between schooling and earning in a controlled environment is studied in Angrist and Pischke. The success of causal inference based on regression depends critically on controlling for observed confounding variables. However, if important confounders are unobserved, researchers try to uncover causal effects using instrumental variables. Due to the difficulty of finding good instruments, other tools are often used, that employ data with a time or cohort dimension to control for unobserved but fixed omitted variables, including fixed effects, difference in differences, and regression discontinuity design.

Instrumental Variables: Statistical control through regression may fail to produce good estimates of causal effects. Fortunately, other techniques are available, which also lead to other-things-equal. As in randomized trials, the forces of nature, including human nature, sometimes manipulate treatment in a manner that eliminates the need for controls. While such forces are rarely the only source of variation in treatment, this is an obstacle easily overcome. The instrumental variable method (IV) exploits partial or incomplete random assignment whether it occurs naturally or is generated by researchers.

Regression Discontinuity Design: Human behavior is controlled by rules. Although many of these rules have little basis in science and experience, they may still be useful. The rules that constrain the role of chance in human affairs often generate interesting experiments. Applied econometricians exploit these experiments with a tool called the regression discontinuity design (RDD). RDD does not work for all causal questions, but when it does, the results have almost the same causal significance as those from a randomized trial.

Difference in Differences: Credible instrumental variables and dramatic policy discontinuities can be hard to find and consequently other econometric tools are needed for causal inference. The difference-in-differences (DD) method recognizes that in the absence of random assignment, treatment and control groups are likely to differ for many reasons. Sometimes, however, treatment and control outcomes move together in the absence of treatment. In such scenarios, divergence of a

***Corresponding author:** Sapra SK, Department of Economics and Statistics, California State University, Los Angeles, USA, Tel: 13233432531; E-mail: ssapra@exchange.calstatela.edu

Received April 06, 2016; **Accepted** April 09, 2016; **Published** April 19, 2016

Citation: Sapra SK (2016) Causality Inference with Observational Data in Economics. Bus Eco J 7: e106. doi:10.4172/2151-6219.1000e106

post-treatment path from the trend established by a comparison group may suggest a treatment effect. Angrist and Pischke demonstrate DD with a study of the effects of monetary policy on bank failures during the Great Depression.

Unlike statisticians, econometricians believe that correlation can occasionally provide evidence of causality even without manipulation of the variable of interest by the researcher. Recent emergence of randomized controlled trials in economics and other social sciences holds promise for measuring the impact of policy interventions in labor economics and development economics among other areas of research.

References

1. Angrist JD, Pischke JS (2008) Mostly Harmless Economics: An Empiricist's Companion. Princeton University Press, Princeton, NJ.
2. Angrist JD, Pischke JS (2015) Mastering metrics: The path from cause to effect. Princeton University Press, Princeton, NJ.