

Case selection and causal inferences in qualitative comparative research

Thomas Plümper,Vera Troeger,Eric Neumayer

Abstract

Traditionally, social scientists perceived causality as regularity. As a consequence, qualitative comparative case study research was regarded as unsuitable for drawing causal inferences. Apparently, few cases cannot establish regularity. The dominant perception of causality has changed, however. Nowadays, social scientists define and identify causality through the counterfactual effect of a treatment. This brings causal inference in qualitative comparative research back on the agenda since obviously comparative case studies can identify treatment effects. We argue that the validity of causal inferences from the comparative study of cases depends on the employed case-selection algorithm. We employ Monte Carlo techniques to demonstrate that different case-selection rules strongly differ in their ex ante reliability for making valid causal inferences and identify the most and the least reliable case selection rules.

Citation: Thomas Plümper,Vera Troeger,Eric Neumayer Case selection and causal inferences in qualitative comparative research. **protocols.io**

dx.doi.org/10.17504/protocols.io.kjccuiw

Published: 21 Nov 2017

Protocol

MC analysis: setup

Step 1.

We test the performance of 8 simple and 3 more complex case selection algorithms frequently used in comparative political science.

We generate artificial data with known features to understand the conditions under which different case selection algorithms perform better or worse.

MC design

Step 2.

We define various data-generating processes from which we draw a number of random samples, and then select two cases from each sample according to a specific algorithm.

Our analysis explores to what extent a comparison of two cases allows researchers to estimate the effect that one explanatory variable, called x , exerts on a dependent variable, called y . We assume that this dependent variable y is a function of x , a single control variable z , which is observed, and some error term ε :

MC results

Step 3.

We compare the reliability of inference on effect strength. We use the root mean squared error to evaluate the performance of each case selection algorithm for different data generating processes.