

# 合成控制方法的预测区间

研究成果: Prediction Intervals for Synthetic Control Methods

作者: Matias D. Cattaneo, 冯颖杰, Rocío Titiunik

期刊: *Journal of the American Statistical Association*, 2021, 116 (536), 1865-1880.

1990年东西德国统一,这对之后德国的经济增长产生了怎样的影响?1988年美国加利福尼亚州实施控烟法案,这在多大程度上降低了当地的香烟消费?要评估这一类重大历史事件或宏观政策的影响,依照经典因果推断(causal inference)理论,我们须为受政策干预的地区构建“反事实”(counterfactual)结果的预测,即该地区若没有受到干预会怎样。为此,我们通常需要选择一个与受干预地区各方面非常接近的控制组(即未受干预的其他地区),据此预测受干预地区的反事实结果。然而,在类似两德统一这类问题中,政策变动仅发生在某一特定地区(如德国)。研究者通常很难为其选择一个理想的完全可比较的控制组;即使这种选择可行,在实践中也往往具有很大的主观任意性。

为解决这类问题,Abadie和Gardeazabal(2003)提出了“合成控制法”(synthetic control method)。该方法得到了非常丰富的发展,近年来已成为社会政策评估领域的重要分析工具。其核心思想是,尽管我们很难为受干预地区找到理想的直接可比的控制组,但我们仍可以利用其他未受干预的地区构建一个适当的线性组合,即所谓“合成控制地区”(synthetic control region),作为对受干预地区的反事实预测。通过比较实际观察到的受干预地区与构造的合成控制地区,我们可获得对政策效应的估计。合成控制法对反事实预测的构建具有透明性、可解释性以及相对客观性。然而,评估政策效应的统计显著性依赖于我们对合成控制的不确定性的刻画(uncertainty quantification),而此前文献对这方面的研究较少。传统方法是假定政策干预在不同地区间的分配是随机性的唯一来源,利用统计上的“置换检验”(permutation test)来评估所估计的政策效应是否只是由偶然因素驱动。然而,现实中我们很难想象例如两德统一一类的事件会“随机”地发生在其他的国家或

地区;同时,类似人均国民生产总值(per capita GDP)这样的变量作为常见的时间序列,其本身的随机性在统计推断中也不应被忽略。

为解决这些问题,清华大学经济管理学院经济系冯颖杰教授及其合作者Matias Cattaneo和Rocío Titiunik提出了一种新的针对合成控制的统计推断方法,研究成果“Prediction Intervals for Synthetic Control Methods”发表在《*Journal of the American Statistical Association*》(2021)。他们指出,不同于传统因果推断问题,合成控制的目标通常是估计针对某一特定地区的政策效应,这一效应本身应当看作一个随机变量(random variable)而非一个固定的待估参数(fixed parameter),而传统的用于刻画参数估计不确定性的置信区间(confidence intervals)并不适用于合成控制问题。研究提出使用预测区间(prediction intervals)来刻画合成控制下政策效应估计的不确定性,这一概念可以反映出

一个随机变量(例如1992年未经历两德统一的西德人均GDP)在多大程度上会出现在某一区间。具体地,在合成控制法中存在两类随机性。首先,合成控制地区作为其他未受干预地区的线性组合,其构建依赖于各个未受干预地区的权重。这一权重通常利用政策干预发生前的样本来进行估计,因此本身具有(样本内)随机性。其次,合成控制是对政策干预发生后某一时期受干预地区的反事实结果的预测,而在该时期还可能存在一些不可观测的随机冲击影响政策效应的评估,因而统计推断中还须考虑这部分(样本外)随机性。该研究提出的预测区间同时考虑了这两类随机性,并且在实践中这些预测区间可通过统计模拟(simulation)方法很容易地构建。为方便实证研究者使用,三位作者还开发了实现这一方法的统计软件包scpi,可在Python、R和Stata三个软件中使用。

供稿:科研事务办公室 编辑:高晨卉 责编:吴淑媛 赵霞