更多资料，请关注微信公众号【草莓科研服务网】

双重差分法(DID) | 空间DID

**一、从聚类标准误到空间相关性**

当使用面板数据进行固定效应模型估计时，考虑到组间异方差和组内自相关，我们必然需要将标准误进行聚类调整。一般来说，聚类调整后的标准误大于异方差稳健标准误，而异方差稳健标准误大于普通标准误，因此，根据聚类标准误做出来的结果是相对最稳健的，这里所说的稳健，指的是系数显著性稳健，因为标准误影响t值，而对系数的本身影响不大。

然而，将标准误聚类调整到什么层次是一个问题。当将个体（id）作为聚类依据时，即假定每个个体不同年份的干扰项存在相关性（组内自相关），而不同个体的干扰项不存在相关性（组间不相关）。但是，这样的假定可能不符合现实情况，比如，对于微观企业来说，同一行业的企业之间必然存在竞合关系，此时同一行业不同企业之间就存在相关性，在这种情况下，将标准误聚类调整到行业层面可能更合理。因此，为了得出更稳健的结果，也为了说服苛刻的读者（或审稿人）接受我们根据实证结果得出的结论，将标准误聚类到更高层级是一种更安全的做法，当然，聚类层级越高系数越不显著。

然而，有时候我们没有甚至不能将标准误聚类到更高层级。除了显著性与稳健性之间的权衡，更多的原因在于聚类层级越高聚类数目越少，而大样本理论要求聚类数目足够大，这样才能保证所估计的标准误收敛到真实值，根据拇指法则，聚类数少于30可能就不太合适了。

退而求其次，为了同时兼顾聚类层级与聚类数目，有些文献将标准误聚类到行业-年份层面（在Stata中可以利用分组函数group生成聚类变量再在回归中进行聚类调整，即：先egen ind\_year = group(industry year)，然后reghdfe y xlist, absorb(id year) cluster(ind\_year)），如李青原和章尹赛楠（2021）、邵朝对等（2021），即假定同一年同一行业之间存在自相关，而不同年或不同行业之间不存在自相关。

根据常识与理论将标准误聚类调整到行业-年份层面，即先验式地承认同一年份同一行业的企业存在相关性，同理，当聚类到省份-年份层面或区域-年份层面时，也可以认为同一年同一地区（省份或区域）不同个体之间存在相关性，这与构建空间计量模型（Spatial Econometric Model）的思想不谋而合。

空间计量模型最大的特征是放松了空间单位之间相互独立的假设，即认为本地区与邻近地区存在相互关系，本地区变量的变动不仅直接影响本地区，也会影响邻近地区，而且通过影响邻近地区最后反作用于本地区。本地区对本地区的影响称为直接效应（Direct Effect），本地区对邻近地区的平均影响称作间接影响或空间溢出效应（Indirect Effect or Spatial Spillover Effect）。对具体空间相关形式的判断，是在OLS模型的基础上对其残差进行拉格朗日乘子检验（LM test）和稳健的拉格朗日乘子检验（robust LM test），以检验OLS回归残差中是否存在某种形式的空间关联，即检验OLS回归残差中是否包含被解释变量的空间滞后项（）抑或扰动项的空间滞后项（）

因此，从这种意义上来说，当标准误聚类调整到地区-年份层面时，此时假定中的同一年份同一地区不同个体之间存在的相关性（组内自相关）本质上是一种空间相关性。也就是说，当我们将标准误聚类调整到地区-年份层面时，实质上就已经变相说明模型中存在一定的空间相关性。

**二、个体处理效应稳定性假设**

在DID中，一个经典的假设是**个体处理效应稳定性假设**（Stable Unit Treatment Value Assumption，SUTVA）。SUTVA最重要的一点是“处理组个体不会影响控制组个体”（Rubin，1974）。换言之，在SUTVA框架下，总体中的任何个体并不会受到其他个体接受处理与否的影响（王金杰和盛玉雪，2020）。然而，这个假设在考虑到空间相关性时被打破了，或者说，当不同空间单元之间存在相关性即存在空间溢出效应时，SUTVA不再成立（Kolak & Anselin，2019）。

事实上，SUTVA在大多数情况下可能都不成立，而现有的DID类实证文章很少会考虑到这一点，并且Ferman（2020）指出忽略空间相关性将导致标准误被低估，从而夸大系数的显著性。

一个简单的道理，处理组在期初实施了某项政策，控制组没有实施，但不是说政策实施之后只会对处理组产生影响，通过示范、学习、要素流动等渠道也会作用于控制组，因此，政策的施行在长期将会在一个较广的区域范围内产生普遍（但不一致）的影响。这里有两个问题，一个是政策实施对控制组作用的力度，另一个是作用的方向。

首先，在处理组实施的某项政策不会对全域内所有地区产生相同的影响，可能两地间经济联系越密切，政策的空间溢出效应越强；可能两地共享同一地理边界（邻接），空间溢出效应越明显；可能两地核心城市之间的地理距离或路程距离越短，空间溢出效应就越强。总之，政策的施行不会对其他地区产生一致的影响，而会随着某种空间相互关系的趋弱存在衰减效应，而我们论文中的SDID就是围绕这种空间相互关系进行研究设计的（空间权重矩阵的设置）。

其次，处理组实施的某项政策对邻近地区的影响有正有负。一方面，政策施行导致资源向先行区倾斜集聚，这不仅意味着先行地区内部资源的重新配置，而且也将导致邻近地区的资源（如劳动力、资金等）向该地区流动富集，这样的结果是抑制了邻近控制组的经济发展，形成了“虹吸效应”。另一方面，先行发展的示范区一般都是经济发展较好的地区（这同时说明DID可能存在严重的样本自选择问题！），政策实施后这些地区将已有的、相对落后的产能向周围转移，连带着将相关人才、技术与资金转移，通过这样的一种方式促进了邻近地区的发展，形成了所谓的“涓滴效应”。总之，政策的施行对邻近地区产生正反两方面的影响，而我们的工作就是要精准稳健地测度出这种影响，并进一步分析出这种异质性影响背后的形成机制或者说机理。

**三、政策实施的空间溢出效应**

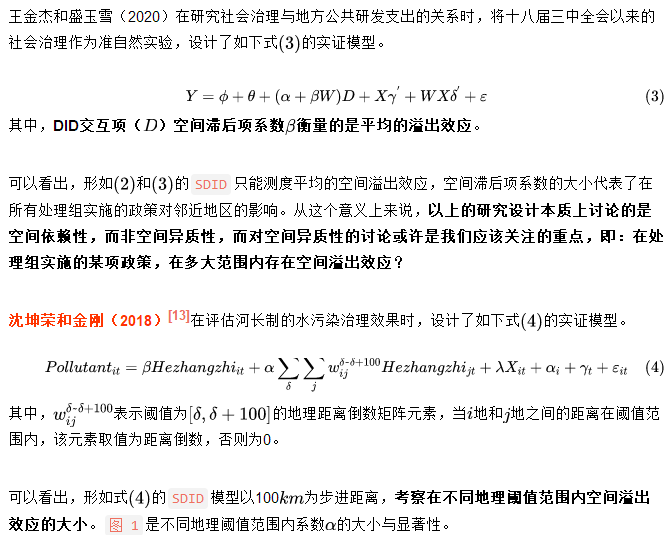
下面结合几个实例给出SDID模型设计的参考。空间计量模型中和DID结合最多的是自变量空间滞后模型（The Spatial Lag of X Model，SLX），而SLX是在普通线性模型等式右端引入（所有）自变量的空间滞后项，形如式下：



式中的即为所有自变量的空间滞后项，由于等式右端不包括被解释变量的空间滞后项，该模型可以视为普通的线性模型，并使用OLS方法进行估计。因此，式中的衡量的是直接效应的大小，测度的是间接效应的大小（Vega & Elhorst，2015）。

SDID等式右边只有DID交互项的空间滞后项，其余自变量均以其初始形式作为控制变量，形如式下：





更多资料，请关注微信公众号【草莓科研服务网】

