

研究方法前沿

[Methodology]

公共管理研究对控制变量的使用正确吗？*

——基于系统文献综述的发现

韩 啸 马 亮 陈睿菲**

【摘 要】定量研究无疑是公共管理学科重要的研究范式,但一直以来不乏对其批评之声,学界争论的焦点之一是定量研究的有效性正面临来自科学性危机的挑战。在定量研究中,使用统计方法控制额外变量的效应,是提高核心变量之间关系估计精度的主要方式。然而,由于标准不明或操作疏忽,实践中研究人员对控制变量的使用存在诸多错误,严重威胁研究发现的有效性。本研究聚焦统计控制应用中的规范性问题,对2014—2023年发表于8本公共管理学术期刊的760篇论文展开分析,重点关注控制变量的选择、分析和结果报告存在的不足。研究发现,控制变量的使用缺乏理论依据,其选择具有同质化倾向,存在缺少多重共线性检验和比较分析结果、报告完整性缺失等问题。本研究针对如何正确使用控制变量提出建议,以期提高公共管理研究的信度和效度。

【关键词】公共管理;控制变量;同质化;遗漏;偏差;效度

一、问题的提出

在实证研究中使用控制变量(control variable),旨在排除对研究结果的其他解释

* 本文受国家社会科学基金重大项目“数字政府建设成效测度与评价的理论、方法及应用研究”(项目批准号:23&ZD080)和广东省社会科学规划一般项目“广东省政府首席数据官政策效应、作用机制与优化路径研究”(项目批准号:GD25CGG08)的资助。衷心感谢匿名审稿专家对本文提出的宝贵建议,感谢电子科技大学刘宇阳同学对本文的贡献。

** 韩啸(通信作者),电子科技大学公共管理学院副教授,zhengguan_yjs@126.com;马亮,北京大学政府管理学院教授;陈睿菲,电子科技大学公共管理学院本科生。

引文格式:韩啸,马亮,陈睿菲. 2025. 公共管理研究对控制变量的使用正确吗? ——基于系统文献综述的发现[J]. 公共管理评论,7(3):217-235.

Cite this article: Han X, Ma L, Chen R F. 2025. Have we controlled properly in public administration research? Findings based on a systematic literature review[J]. *China Public Administration Review*, 7(3): 217-235. (in Chinese)

因素,并对变量间关系做出更精准的估计,提高研究的内外部效度。正因如此,控制变量与核心变量在研究中具有同等重要的地位(Becker,2005)。目前,实证研究对控制变量的使用方法主要分为实验控制和统计控制两类。实验控制法是在变量产生影响前进行控制,帮助研究者识别更真实的因果关系;统计控制则是在变量影响发生后,通过移除与控制变量有关的变异,修正数据收集过程中的缺陷,得到更有效度的分析结果(曹江雨等,2020)。实践中,受限于研究议题的特殊性,部分研究并不适合实验控制,使用统计控制更为普遍。

控制变量的错误使用会给统计结果带来损害。许多研究者认为在分析中加入控制变量,可以净化核心变量之间的关系,减少 I 类错误的发生^①。然而,不当地添加控制变量不仅无法减少偏差,还会导致 II 类错误的出现(Li,2021)。可见,如果没有明确标准或处理不当,都有可能参数估计无法解释、变量相关性扭曲及分析结果有效性降低等问题,制约研究发现的科学性与准确性(Becker et al.,2016)。为此,政治学(Lenz and Sahn,2021)、社会学(York,2018)、经济学(Newey and Stouli,2021)、心理学(Bernerth and Aguinis,2016)及工商管理(Curado et al.,2024)等学科已就如何正确使用控制变量展开系统反思与讨论。例如,有研究对管理学顶级期刊发表的 580 篇论文进行分析发现,只有 3% 的研究为选择控制变量提供了充分的理论基础(Bernerth et al.,2018)。曹江雨等(2020)对中文期刊组织管理定量研究展开分析,发现绝大部分研究存在过度控制或控制不足的现象。

公共管理曾被视为传统文科专业,而后定量研究的应用日益普遍,然而公共管理学科的定量研究“不规范”问题十分突出(马亮,2017)。尽管近年来国内学者开始关注公共管理定量研究的规范性议题(于文轩和樊博,2020),但针对控制变量的正确选择与使用的讨论尚付阙如。研究者们尚未深刻认识到控制变量的重要性以及应如何正确使用控制变量,“误用”“滥用”的情况十分普遍。因而,强化控制变量使用的规范性、正确性,可提升公共管理定量研究的严谨性、科学性,推动学科知识发现与积累。鉴于此,本文围绕三个问题展开分析:

- (1) 国内公共管理研究对控制变量的基本应用情况如何?
- (2) 国内公共管理研究在使用控制变量方面存在哪些错误?
- (3) 在研究中应如何正确使用控制变量?

为系统回答上述问题,本文对 2014—2023 年发表于国内 8 本公共管理权威期刊的 760 篇论文使用控制变量的情况,进行分析并提出针对性建议。

二、统计控制的基本原理

受实证主义哲学影响,科学研究的目標转向讨论变量间稳定的因果关系,尝试

^① I 类错误(Type I Error)是指拒绝了正确的、成立的假设,也称为假阳性(false positive)。II 类错误(Type II Error)是指接受一个实际上是错误的、不成立的假设,也称为假阴性(false negative)。

解释外部世界的运行规律与变化动因。密尔指出因果关系存在三个必备条件:原因在时间上早于结果;原因必须与结果存在相关关系;除原因外,不存在对结果的替代解释(Cook and Campbell, 1979)。借助文氏图(Venn Diagram)来理解统计控制的基本原理。图1中 X 为解释变量, Y 为被解释变量, Z 为额外变量。使用统计方法对额外变量的影响进行控制后,实际使用的解释变量 X 已不再是原始的 $X(b+c+e+g)$,而是移了解释变量 X 和额外变量 Z 的共同变异($c+g$)后的剩余解释变量 $X(b+e)$ 。

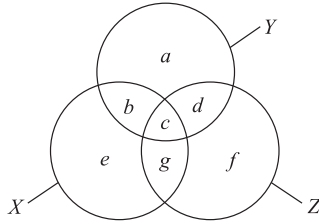


图1 解释变量、被解释变量和额外变量的文氏图

控制变量通过排除其他解释来加强因果论证,实现“净化原则”(purification principle)。如果在定量分析中没有加入相关控制变量,其研究会质疑存在遗漏变量偏差。

本文通过下列等式呈现控制变量如何净化 x 和 y 之间的关系,假设要估计 x_1 对 y 的影响(等式1中的 β_1),控制变量可以消除 x_2 到 x_k 造成的影响。可以使用两阶段回归来估计 β_1 :第一阶段,将 x_1 作为因变量,对 x_2 到 x_k 进行回归,得到残差 u (等式2),剔除了所有其他所有控制变量的影响。根据统计假设,残差与等式2右侧的所有变量都没有相关性。第二阶段,用 y 对残差 u 进行回归(等式3), u (b_1)的回归系数与 β_1 (控制 x_1 的正确模型中等式1的 x_1 回归系数)相同。显然,如果研究者遗漏了一个重要的控制变量,在接下来的分析中会发现 x_1 和误差项之间存在相关性,这违反了等式基本假设,会产生有偏估计。

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k + e \quad (1)$$

$$x_1 = \delta_0 + \delta_2 x_2 + \cdots + \delta_k x_k + u \quad (2)$$

$$y = b_0 + b_1 u + r \quad (3)$$

在统计控制中,研究者很可能会遗漏一些重要变量,这些遗漏变量会导致估计结果出现偏差。遗漏变量偏差(Omitted Variable Bias)是指 b_i (没有控制 x_k 的错误模型中 x_i 的回归系数,等式4)与 β_i (控制 x_k 的正确模型中等式1的 x_i 回归系数)之间的差值。对遗漏变量 x_k (等式5)进行回归来量化遗漏变量偏差,使用等式6计算任何变量 x_i 缺失 x_k 导致的偏差。遗漏变量偏差会导致解释变量的估计系数不准确,为规避遗漏变量带来的内生性问题,研究需要针对性引入控制变量,减轻遗漏变量偏差造成的负面影响。

$$y = b_0 + b_1 x_1 + b_2 x_2 + \cdots + b_{k-1} x_{k-1} + e \quad (4)$$

$$x_k = \delta_0 + \delta_1 x_1 + \cdots + \delta_{k-1} x_{k-1} + u \quad (5)$$

$$b_i = b_i - \beta_i = \beta_k \times \delta_i \quad (6)$$

根据可观测性和时间变化性,变量分为可观测随时间变化、可观测不随时间变化、不可观测不随时间变化三种类型。固定效应主要控制不可观测不随时间变化的变量,在此基础上,固定效应也是一种控制变量,其包括个体固定效应、时间固定效应和双向固定效应等。下文介绍常用的三类固定效应模型。

个体固定效应模型主要用于处理面板数据中的个体差异问题。预设每个个体的截距项是固定的,不随时间变化,但允许不同个体之间存在差异。这意味着不同个体之间的差异主要体现在截距项上,而不是斜率。这类模型特别适用于研究个体特征对因变量的影响,如个人的教育背景、性别等因素,可有效解决多重共线性和多水平变量的问题,减少分析误差,获得更准确的结果。例如,在一项讨论我国城镇养老保险制度对居民消费影响的研究中,研究者通过构建异方差稳健标准误的个体固定效应模型对结果进行估计(李珍和赵青,2015)。

$$y_{it} = \lambda_i + \sum_{k=2}^K \beta_k x_{kit} + u_{it}$$

时间固定效应模型预设每个时间点都有一个固定的效应,这个效应代表在该时间点所有个体共同的特征。通过在模型中引入时间固定效应,可以控制那些随时间变化但不随个体变化的因素,消除遗漏变量的影响,从而更准确地估计自变量和因变量之间的关系。

$$y_{it} = \gamma_t + \sum_{k=2}^K \beta_k x_{kit} + u_{it}$$

双向固定效应模型是在面板数据分析中被使用频率最高的模型之一,通过同时控制个体和时间的固定效应,减少遗漏变量和不可观测因素带来的偏误,从而更为精确地估计变量间的因果关系。例如,在讨论“独生子女”政策的反馈效应分析中,研究者不仅加入了可能影响个体生育观念的个体特征、家庭特征和省份特征作为控制变量,还加入省份固定效应和时间固定效应以提高分析结论的准确性和稳健性(王培杰等,2022)。又如,为了论证政府间协作治理能否有效降低雾霾污染,研究者选择双向固定效应模型对时间和个体产生的差异进行控制,以得到环境协作治理的净效应(徐换歌和王峰,2022)。

$$y_{it} = \lambda_i + \gamma_t + \sum_{k=2}^k \beta_k x_{kit} + u_{it}$$

研究中需要注意,一些变量不需要被控制,被错误控制后会产生偏差,例如后处理偏差(Post-treatment Bias)和碰撞偏差(Collider Bias)。后处理偏差是指误将中介变量当作控制变量纳入模型所导致的偏差(如图2所示)。在实际研究中,对事物的认识是螺旋式推进的,有时会把潜在的中介变量当作控制变量,不仅无法让控制变量发挥净化作用,而且会造成控制变量解释方差过大,影响对核心变量关系的判断(冯帅帅和罗教讲,2021)。

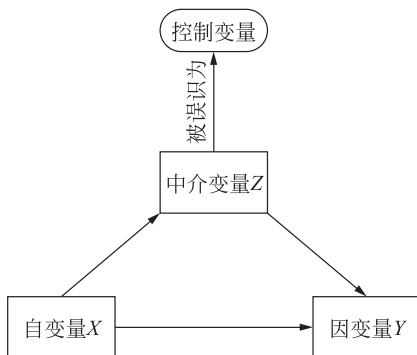


图2 后处理偏差

当自变量和因变量都被第三个变量共同影响,该变量在统计中被控制时,就会发生碰撞偏差。换言之,就是把相关关系误认为因果关系。一个典型的例子就是:冰激凌的销量和溺水事件呈现出来非常明显的正相关关系,如果忽视炎热天气这一因素,得出冰激凌的高销量导致了溺水事件这一伪因果关系,便出现了碰撞偏差(图3)。

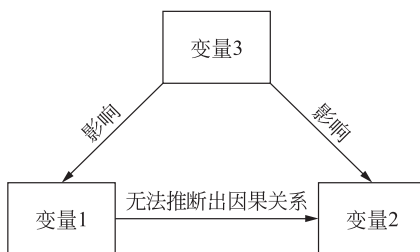


图3 碰撞偏差

综上所述,如果研究者在数据分析时,不了解统计控制原理和其使用规范时,很有可能会出现误用、滥用控制变量的情况,导致数据分析结果出现偏差。下面将对国内8本公共管理领域权威期刊在2014—2023年间发表的文章使用统计控制的情况进行分析总结,对使用控制变量的常见错误进行归纳,找出原因并提出建议,为后续研究提供参考。

三、研究设计

(一) 样本来源与选择

本文选择2014—2023年在《中国行政管理》《公共管理学报》《公共管理评论》《公共行政评论》《公共管理与政策评论》《行政论坛》《电子政务》《治理研究》上发表的实证论文,对定量研究中的控制变量使用情况进行分析。选择上述期刊的理由是,其均为国内公共管理权威期刊,具有较好的代表性。在判断文章是否纳入

时,采用逐篇阅读的方式进行判断,判断依据为:一是该论文是否属于定量研究(含混合研究);二是该论文是否进行统计控制。筛选后共纳入 760 篇定量论文作为分析样本。

(二) 样本分布

如图 4 所示,将纳入分析的 760 篇文章按年份进行分类后发现,总体上定量研究呈现逐年增加趋势。

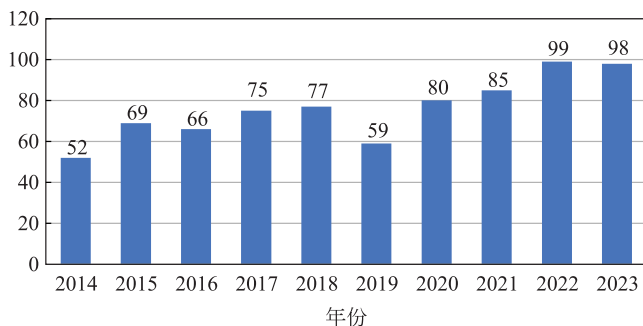


图 4 样本论文的发表时间分布情况(2014—2023 年)

期刊分布如图 5 所示,其中《中国行政管理》的文章数量最多,有 133 篇,占比 17.5%;其次是《公共行政评论》,有 130 篇,占比 17.1%;《公共管理学报》有 122 篇,占比 16.05%;《公共管理与政策评论》有 102 篇,占比 13.4%;《电子政务》有 94 篇,占比 12.4%;《公共管理评论》有 82 篇,占比 10.8%;《治理研究》有 51 篇,占比 6.7%;《行政论坛》有 46 篇,占比 6.05%^①。

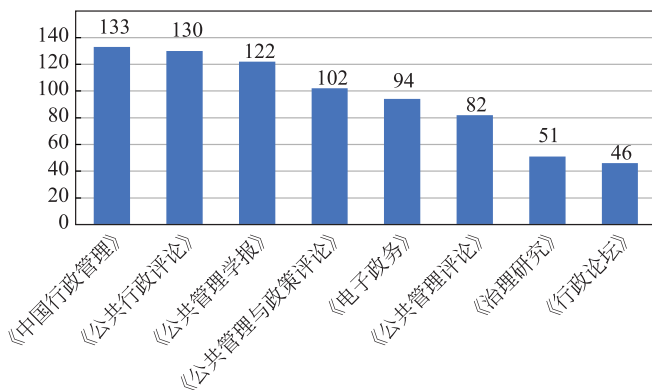


图 5 样本论文的期刊分布情况

^① 这里需要说明的是,笔者采用绝对值进行计算,没有考虑期刊创刊时间、期刊创刊以来的总体发文量以及期刊类型(如月刊、双月刊和季刊等)。因此,这不能说明某一期刊对定量类型文章的倾向。

四、控制变量在公共管理研究中的使用情况

对纳入分析的每篇论文,研究团队检查并记录了以下内容:是否报告控制变量选择的理论依据与参考来源、是否将其包含在假设中、是否进行信度和效度分析、是否报告控制变量的测量方式、是否使用固定效应模型、是否进行多重共线性检验、是否对控制变量进行数据插补等。为确保编码信度,研究参照编码信度公式进行检验(迈尔斯和休伯曼,2008)。从两位研究者的编码数据中随机抽取 50%,让另一位研究者独立对其再次编码,基于该部分数据计算得出编码信度为 95.27%。

(一) 描述性统计

统计结果显示,每篇文章平均使用 6.52 个控制变量,标准差 3.34。其中,21.1% 的研究使用 1~3 个控制变量,40.6% 使用 4~6 个,20.1% 使用 7~9 个,18.2% 使用 10 个及以上。图 6 呈现了分析样本中使用频次最高的前十个控制变量:年龄、性别、受教育程度、经济发展水平、婚姻状况、政治面貌、户籍、人口密度、民族、产业结构。

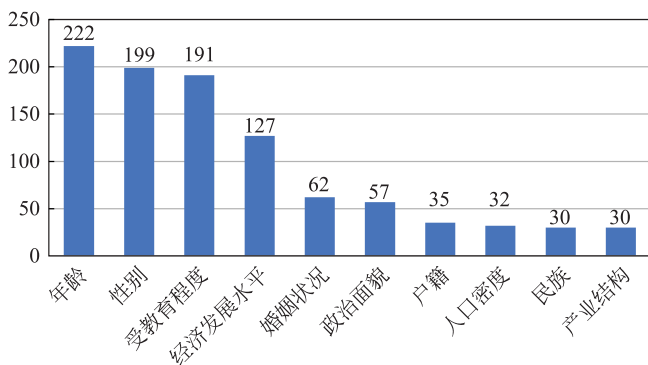


图 6 使用频率最高的 10 个控制变量

(二) 控制变量使用存在的问题

表 1 为国内公共管理研究在控制变量使用上存在的典型问题,例如未报告理论依据、未纳入研究假设、未完整报告统计信息等。透过上述问题可以发现,由于缺乏理论指导,研究者在控制变量使用上,存在选择不当、操作不规范、报告不完整等错误。这些不仅会影响数据分析的准确性,还会降低研究结果的信度与效度。基于此,本部分重点围绕控制变量的选择、分析和报告三个部分,对其存在的典型问题展开剖析。

表 1 公共管理实证研究在控制变量使用上存在的典型问题

存在的问题	2014 年—2018 年 (N=336)	2019 年—2023 年 (N=424)	总计 (N=760)
没有报告理论依据	327(97.3%)	412(97.2%)	739(97.2%)

续表

存在的问题	2014 年—2018 年 (N=336)	2019 年—2023 年 (N=424)	总计 (N=760)
没有说明代理变量依据	301(89.6%)	336(79.2%)	637(83.8%)
没有报告引证证据	272(81.0%)	248(58.5%)	520(68.4%)
没有使用固定效应模型 ^①	85(83.3%)	149(68.7%)	234(73.3%)
没有比较分析结果	308(91.7%)	354(83.5%)	662(87.1%)
没有检验多重共线性	321(95.5%)	374(88.2%)	695(91.4%)
没有完整报告统计信息	173(51.5%)	214(50.5%)	387(50.9%)
没有完整报告信效度	149(44.3%)	225(53.1%)	374(49.2%)
没有进行数据插补 ^②	57(87.7%)	123(93.9%)	180(91.8%)

1. 控制变量的选择问题

(1) 缺乏理论依据。具备坚实理论基础的控​​制变量,能够帮助研究者明晰其在核心变量关系中所能发挥的作用。反之,若缺乏理论依据,研究者可能无法判断应纳入哪些控制变量,以及这些控制变量能否实现预期效果(Bernerth et al., 2018)。尽管选择有理论意义的控制变量十分重要,但样本分析显示,研究者对此重视程度依然不足。

在纳入分析的样本中,超过 97% 的文章没有报告选择控制变量的理论依据。研究者根据经验判断控制变量与核心变量的相关性,却未从理论层面阐释这种关联的内在逻辑。例如,部分研究选择年龄、性别作为控制变量,却未说明选择的理论依据。这种做法会引发两方面负面影响,一是陷入“以讹传讹”的虚假控制,理论缺位会导致研究者借鉴已有文献的做法进行控制变量的选择,若已有研究存在变量选择错误或有遗漏,那么将影响后续研究统计控制的目的。二是流于“东施效颦”的形式主义,研究选择控制变量并非基于相关理论,而是因循文献惯例——若不采用相同变量,一方面在形式上显得与主流研究“格格不入”,另一方面在评审过程中可能会被质疑。

(2) 没有说明代理变量的使用依据。研究中,当某一变量不可观测或无法直接测量时,就需要使用代理变量,且应详细阐明选择依据,例如代理变量是否与研究变量存在相关性,二者是否具备替代性,以及已有研究的使用情况与评价等。

① 2014—2018 年共有 102 篇需要使用固定效应模型的文章,其中 85 篇未使用;2019—2023 年共有 217 篇需要使用固定效应模型的文章,其中 149 篇未使用;2014—2023 年总计有 319 篇需要使用固定效应模型的文章,其中 234 篇未使用。

② 2014—2018 年共有 65 篇涉及数据插补的文章,其中 57 篇未规范报告;2019—2023 年共有 131 篇涉及数据插补的文章,其中 123 篇未规范报告;2014—2023 年总计有 196 篇数据插补文章,其中有 180 篇未规范报告。

在 760 篇分析样本中,31.9%(242 篇)未使用任何代理变量,49.7%(378 篇)对部分变量选择了代理变量进行测量(占比),仅 18.4%(140 篇)全部使用代理变量。在使用代理变量的 311 篇文章中,60.5%(188 篇)没有说明选择代理变量的原因,23.8%(74 篇)说明较为模糊,仅有 15.7%(49 篇)进行了详细阐释。

事实上,大多数代理变量缺乏理论基础,研究者无法基于理论对代理变量与研究变量之间的关系进行解释,也难以通过代理变量阐释核心变量之间的关系,因此在研究中对代理变量的控制可能会误删有效变异,降低研究发现的可信度。以“市场化程度”为例,有研究使用代理变量“国有单位职工在当地人口占比”来衡量地区的市场化程度,该处理方式缺乏理论依据,同时该代理变量对市场化程度的反映效度也值得商榷。可能更合理的做法是:要么有理论支撑,要么直接选择樊纲和王小鲁编制的“中国市场化指数”作为测量数据而非代理变量(李辉和杨肖光,2019)。

(3) 控制不当与选择同质化。纳入分析的 760 篇文章中,控制变量使用数量方面存在较大差异,最多的一篇文章纳入 34 个控制变量,最少的仅有 1 个,平均值为 6.52 个。尽管引入有理论支撑的控制变量有助于非实验研究中的因果推断,但当前研究普遍存在控制变量滥用与选择同质化等问题。通过分析发现,研究者遵循“控制变量越多越好”的理念盲目增加变量数量,而这些缺乏理论依据的控制变量被纳入模型中,既浪费资源又漏洞百出。需要认识到,控制变量的质量远比数量更重要。

笔者发现,控制变量的选择存在较严重的同质化现象。这是一个令人惊讶的现象,一方面是差异化的研究议题,另一方面却是同质化的控制变量。在 760 篇文章中,有 144 篇文章阐述了部分控制变量借鉴自前人研究经验,96 篇则声称研究涉及的全部控制变量均依据先前文献的选择。更值得注意的是,两篇研究议题差异较大的文章,在控制变量的使用上几乎趋同。这表明,大部分文章以“参考以往研究、过去做法”作为选择控制变量的标准,而非来自核心变量相关的理论。这种“随大流式”的选择模式只是模仿前人研究,并未依据自身研究需要识别潜在控制变量,也未充分考虑控制变量在核心变量关系中可能发挥的作用机制。

(4) 未使用固定效应模型。面板数据常常存在个体和时间异质性,会导致普通最小二乘法(OLS)估计参数不一致。固定效应模型通过控制该异质性可增强估计参数一致性,从而更好地处理动态面板数据问题并提高模型解释力。在分析样本中,319 篇研究使用了面板数据,这意味着在数据分析中这些研究需要使用双向固定效应模型对个体和时间的异质性进行控制,但超过 73% 的文章没有明确说明是否采用该模型,或对模型使用情况的说明存在模糊性。

2. 控制变量的分析问题

(1) 没有比较分析结果。为了更好地探究控制变量对核心变量关系的潜在影响,研究应分别汇报包含与不包含控制变量的分析结果,通过比较发现具体差异。在实践中,鲜有研究者会这样做。例如,Carlson and Wu(2012)对 266 篇管理学顶级期刊上的实证研究分析发现,仅 8% 的文章比较了包含与不包含控制变量的结果。

本研究亦发现,87.1%(662 篇)的文章没有对是否纳入控制变量的分析结果进行比较,仅 12.9%(98 篇)进行了结果对照。这种对比缺失引出了一个问题,如果没有对基线模型和控制模型的结果进行比较,就无法判断控制变量是否发挥作用,也让控制变量的使用流于形式,成为一种象征。因此,研究必须规范呈现控制变量纳入前后的结果差异,以验证控制变量是否发挥作用。

(2) 相关度过低。控制变量与核心变量之间存在显著相关性是进行统计控制的前提。然而,由于变量选择缺乏理论依据,实践中控制变量与核心变量的关系并不显著,例如,学者发现管理学最常使用的 5 个控制变量,即性别、任期、年龄、教育水平和组织规模,在超过 2/3 的研究中与核心变量关系不显著(Bernerth and Aguinis, 2016)。有研究运用荟萃分析法对 2003—2014 年发表的实证研究进行分析发现,所有控制变量的平均效应量仅为 0.04(冯帅帅和罗教讲,2021),表明在实际研究中控制变量未能发挥“净化作用”。因此,在研究中汇报控制变量与核心变量之间的相关系数非常有必要。如果控制变量与核心变量的相关度过低甚至不相关,这类变量不仅无法排除对核心变量的其他解释,还会因降低自由度而成为“无效控制变量”。

(3) 没有进行多重共线性检验。在定量研究中,当模型中纳入多个变量时,可能会出现多重共线性,对结论产生干扰,降低研究品质。多重共线性是指两个及以上自变量之间存在高度相关的现象。为避免控制变量解释系数失真,回归分析前需要进行多重共线性检验。在纳入分析的 760 篇文献中,绝大部分文章(91.4%)未进行该检验,忽视了该问题的重要性。当变量之间存在高度相关时,模型的稳定性会受到挑战,即便是微小的数据变化也可能导致回归系数出现较大波动,导致模型预测结果不稳定。此外,多重共线性还会增大回归系数的标准误差,降低统计检验的效力,使得原本显著的变量变得不显著,或者原本不显著的变量错误地呈现显著性。

3. 控制变量的报告问题

(1) 没有报告统计信息。严谨的定量研究应完整汇报包括控制变量在内的所有变量统计信息。其中,控制变量的核心统计指标应涵盖均值、标准差、最大值和最小值。如果没有完整报告与控制变量有关的统计结果,一方面难以判定其是否有效发挥了控制作用,另一方面也会削弱研究的可复制性与延续性,无法为后续研究提供详细参考(Carlson and Wu, 2012)。在分析样本中,仅 49%(373 篇)的文章完整、详细地报告了控制变量统计信息。这表明国内公共管理定量研究的规范性还需要加强,学界对报告控制变量统计信息重要性的认知需要提升,也反映出国内公共管理学对定量研究方法的教学和训练存在欠缺。

(2) 信度和效度报告不完整。信度是指测量工具或评估方法在重复测量同一构念时所得结果的一致性和稳定性,效度则是指其测量的构念的准确性与正确性。二者是研究中评估测量工具质量的两个关键维度。控制变量的测量信度和效度,直接影响着统计控制目的能否顺利实现。若缺乏控制变量信度和效度信息,研究者不仅

无法确定这些变量在重复测量中的表现是否一致,而且无法评估控制变量是否真正测量了构念。因此,如实、准确地报告控制变量的信度和效度十分重要。笔者发现,研究者常常会忽略这一环节,只有不足 51% 的文章报告了控制变量的信度和效度。

(3) 没有报告控制变量的数据插补情况。在使用面板数据分析时,当某时间段、城市或样本数据存在缺失时,常采用插补法将残缺数据集补全,减轻数据缺失导致的估计偏差。作为恢复数据完整性的有效手段,数据插补在定量研究中的应用逐渐增多。控制变量作为非核心变量也不例外,同样需要通过数据插补来补充缺失的数据。然而,现有研究对控制变量重视不足,使用数据插补的规范性较低。本研究发现,仅有 16 篇文章规范报告了数据插补的使用情况。需注意的是,数据插补方法的应用具有严格的前提条件,例如,当数据符合正态分布时,才能使用线性插补方法,在插补前需要考虑数据的分布特征、方差等因素,以选择更合适的插补方法。

4. 分析总结

本文在对 760 篇实证研究分析后,发现样本文献在选择、分析和报告控制变量上存在以下主要问题,具体情况见表 2。

表 2 公共管理研究中控制变量使用存在的问题

问题的产生阶段	存在的问题	具体说明
控制变量选择	缺乏理论依据	所选择的控制变量缺乏理论基础
	没有说明代理变量的使用依据	使用代理变量随意化且关联程度低
	控制不当与同质化选择	加入过多无效的控制变量或忽视重要控制变量,不同研究选择的控制变量趋同
	未使用固定效应	无法减少遗漏变量和不可观测因素带来的偏误
控制变量分析	没有比较分析结果	缺乏比较有/无控制变量的分析结果
	相关度过低	与核心变量的关系不显著,说明控制无效
	没有进行多重共线性检验	变量之间存在高度相关会降低模型的解释力
控制变量报告	没有报告统计信息	缺失报告控制变量的最大值、最小值、均值、标准差等统计信息
	信度和效度报告不完整	对控制变量的信度和效度报告不完整,无法做出准确判断
	没有报告数据插补情况	忽视数据特征,无法正确选择最合适的数据插补方法,制约控制变量的有效性

五、研究建议

本文从控制变量的选择、分析和报告三个方面对 2014—2023 年发表于国内 8 本公共管理权威期刊的 760 篇文章使用统计控制的情况,进行逐篇分析。结果发现,这些文章无论是在选择控制变量,还是分析和报告控制变量上都存在进一步提升空间。为此,提出以下建议。

(一) 关于控制变量选择的建议

1. 根据理论选择控制变量

在选择控制变量前,研究者需明确“为何要在研究中使用控制变量”,这是科学运用控制变量的起点。唯有厘清这一问题,后续对控制变量的使用才不会出现严重错误。事实上,合理且有意义的控制变量选择往往基于理论推导,而非简单参照或模仿。尽管当前国内公共管理研究中理论导向的控制变量选择尚未被广泛采纳,但个别研究已开启这一探索,这无疑是的趋势,而其进一步推广则需要学术共同体持续推动并逐渐形成共识。

为此,笔者从以下三个方面提出建议。一是在选择控制变量时,研究者应系统回溯与核心变量相关的理论体系,深入理解理论涉及的核心概念、命题、预设等。这是正确选择控制变量的基础。例如,若某一理论明确指出需要控制某些变量后命题才成立,那么研究者就应将这些变量纳入控制范畴。二是研究者需要明确所研究的问题是否存在其他理论解释路径。尤其当存在替代性理论时,若该理论所涉及的前因变量与研究者关注的前因变量之间存在显著相关性,则需将其作为控制变量纳入模型。三是若某一控制变量确实缺乏理论支撑,可以基于实证研究的统计证据或逻辑推导,论证控制变量与核心变量之间的关联性,从而为控制变量的选择提供依据。

2. 详细说明代理变量的选择依据

当部分控制变量缺乏可利用的数据或无法进行可信测量时,可策略性地采用代理变量。控制变量测度中使用代理变量常伴随潜在问题,本文提出以下规范建议。一是明确代理变量与控制变量的理论边界,二者不能等同使用,且在使用前要从理论层面对代理变量进行分析、讨论,仔细辨识代理变量与核心变量的关系,明晰代理变量对核心变量的影响机制。二是注重代理变量选择的规范性与数据测度的科学性,避免陷入“错误的输入导致错误的输出”的方法论陷阱(Shiau et al., 2024)。三是注意固定效应模型作为控制变量的统计分析前提,例如采用 Breusch-Pagan 拉格朗日乘数检验判断个体随机效应是否存在,并采用 Hausman 检验比较固定效应模型与随机效应模型的优劣等等。综上所述,笔者绘制决策树模型(见图 7),为后续研究在选择控制变量时提供参考。

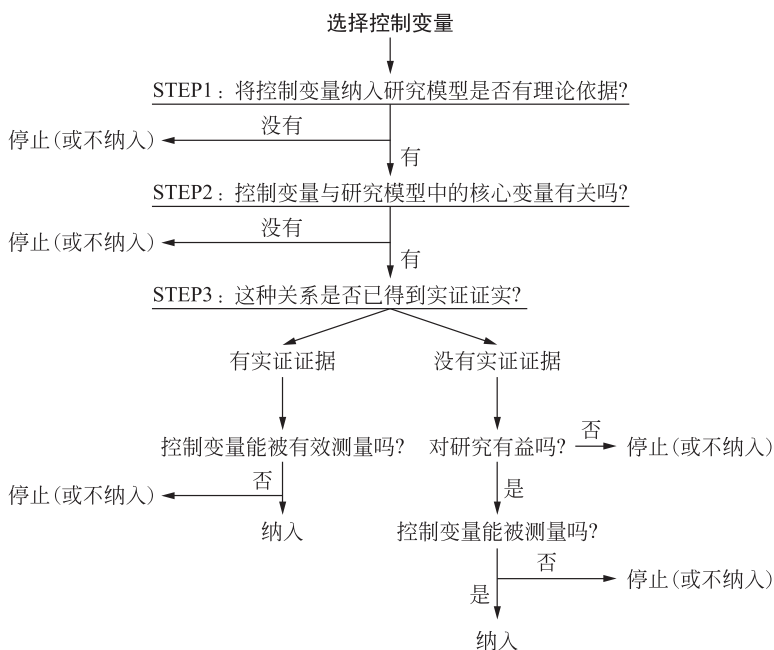


图7 选择控制变量的决策树模型

(二) 关于控制变量分析的建议

在控制变量的选择时,建议“根据理论选择控制变量”,以使数据分析过程具有意义,这也是确保研究发现精度和效度的关键所在。在对控制变量分析的建议中,本文呼吁重视规范性、科学性,让整个过程更加规范、严谨,可参考笔者总结的“选择控制变量的决策树模型”。透过前文分析可知,控制变量问题的核心并非数量过少或过多,而是没有正确控制变量。同时,需对使用的控制变量进行充分比较和讨论,明确加入控制变量前后的具体变化。现有分析表明,国内公共管理实证研究在此方面存在明显不足,超过80%的研究未对加入控制变量前后的结果变化进行比较。缺乏比较,就无法判断控制变量是否发挥了预期作用。因此,本文建议在研究中构建与控制变量有关的基线假设,系统考量控制变量可能产生的影响(包括加入或没有控制变量的差异),通过对比分析实现统计控制的目的。

(三) 关于控制变量报告的建议

大部分研究者容易忽视控制变量的重要性,甚至对其作用存在错误认知,认为相较于核心变量而言,控制变量在研究中仅扮演辅助性、仪式性角色。据统计,国内公共管理研究文献中超过94%的研究未详细报告控制变量的统计信息。这会引发两个严重问题:第一,控制变量效度存疑。例如,在基于问卷调查的数据研究中,若未报告控制变量的构念效度、聚合效度及区分效度等重要统计信息,研究者就无法

判断控制变量测量的有效性。第二,极端值扭曲控制变量与核心变量的关系难以识别。缺少报告控制变量的均值、标准差等描述性统计信息,可能会忽视极端值产生的扭曲影响,进而引发误判。为此,本文建议:一是详细呈现控制变量的描述性统计信息,二是完整报告控制变量的信度、效度以及相关系数。通过上述信息的规范报告,研究可有效控制误差,提高参数估计的准确性。

六、研究展望

定量研究作为国际公共管理学科的支配性研究方法,也是最契合后实证主义常规科学的范式,在认识和改造世界,以及知识的积累和发展上,以及学科体系建设和人才培养等层面具有独特优势。随着定量研究方法在国内公共管理领域的应用日益广泛,其一方面丰富了公共管理学科的研究手段,有力推动了国内公共管理学科研究的科学化、规范化和国际化进程,另一方面,针对定量研究的批评和质疑之声也在持续提醒研究者们要重视暴露出的问题(颜海娜和蔡立辉,2008;王印红和王刚,2013;刘润泽和巩宜萱,2020)。

当下出现的批评和质疑(如代理变量不匹配、数据来源缺乏信度、样本选取不符合科学抽样原理),以及本文对控制变量使用进行分析发现的问题,这些均指向一个核心症结,即国内公共管理定量研究的规范性不足。这与于文轩和樊博(2020)的见地一致——“中国公共管理学对定量研究方法的教学和训练仍显薄弱,与公共管理现实需求及公众期待存在显著差距”。

当今公共管理已步入数智时代,大数据、人工智能等新兴技术蓬勃兴起,推动研究发现精准化,助力政府决策高效化,促进公共服务智能化,这些应用也构成了当前公共管理研究的重要内容(范如国,2018;郁建兴等,2023)。与此同时,“乌卡时代”的公共管理问题日益复杂化,传统定量分析方法正遭遇“统计失灵”的困境,而以机器学习、ABM 等为代表的计算社会科学新方法应运而生,为公共管理研究提供了更为丰富的研究手段和分析方法(陈云松等,2020;刘闻博和王丛虎,2024)。这些变化的趋势也对未来公共管理定量研究提出了更高要求。

在数智时代开展公共管理研究,一方面,需把握前沿研究方法,立足原理基础与基本规范,避免“炫技式”研究,让方法的选择与运用更加贴合研究问题的解决,以便更真切、精准地认识和理解世界(马亮,2024);另一方面,公共管理研究需要“想象力”,需要聚焦国家发展的“大问题”,这与前文强调的科学性和规范性并不冲突,反而从规范、严谨的分析过程出发,走向科学可信、有效回应现实问题,进一步推动学科知识的增长和积累,强化公共管理学科的身份认同和合法性地位。

参考文献

曹江雨,王忠军,唐云,等. 2020. 组织管理研究中的控制变量使用:问题与策略[J].

- 中国人力资源开发, 37(8): 48-65.
- Cao J Y, Wang Z J, Tang Y, et al. 2020. Control variable usage in organizational and management research: Problems and recommendations[J]. *Human Resources Development of China*, 37(8): 48-65. (in Chinese)
- 陈云松, 吴晓刚, 胡安宁, 等. 2020. 社会预测: 基于机器学习的研究新范式[J]. 社会学研究, (3): 94-117.
- Chen Y S, Wu X G, Hu A N, et al. 2020. Social prediction: A new research paradigm based on machine learning[J]. *Sociological Studies*, (3): 94-117. (in Chinese)
- 范如国. 2018. 公共管理研究基于大数据与社会计算的方法论革命[J]. 中国社会科学, (9): 74-91, 205.
- Fan R G. 2018. The big data/social computing methodological revolution in public management studies[J]. *Social Sciences in China*, (9): 74-91, 205. (in Chinese)
- 冯帅帅, 罗教讲. 2021. 社会学量化研究控制变量方法的反思与超越[J]. 深圳社会科学, 4(6): 95-105, 115.
- Feng S S, Luo J J. 2021. Rethinking and transcending the method of controlling variables in sociological quantitative research[J]. *Social Sciences in Shenzhen*, 4(6): 95-105, 115. (in Chinese)
- 李辉, 杨肖光. 2019. 市场化与腐败类型的地区差异——基于职务犯罪起诉书数据的多层分析[J]. 公共行政评论, 12(3): 139-155, 193.
- Li H, Yang X G. 2019. Marketization and structures of corruption: A multilevel analysis based on data of duty-crime indictment[J]. *Journal of Public Administration*, 12(3): 139-155, 193. (in Chinese)
- 李珍, 赵青. 2015. 我国城镇养老保险制度挤进了居民消费吗? ——基于城镇的时间序列和面板数据分析[J]. 公共管理学报, 12(4): 102-110, 158.
- Li Z, Zhao Q. 2015. Has urban old-age insurance crowded-in household consumption in China? —An empirical study based on time-series and panel data analysis[J]. *Journal of Public Management*, 12(4): 102-110, 158. (in Chinese)
- 刘润泽, 巩宜萱. 2020. 回顾与反思: 定量研究在公共管理学科的滥用[J]. 公共管理学报, 17(1): 152-158, 176.
- Liu R Z, Gong Y X. 2020. Review and reflection: The misuse of quantitative research in public administration[J]. *Journal of Public Management*, 17(1): 152-158, 176. (in Chinese)
- 刘闻博, 王丛虎. 2024. 公共管理研究中的机器学习方法——原理、应用及挑战[J]. 公共管理与政策评论, 13(5): 152-168.
- Liu W B, Wang C H. 2024. Machine learning in public administration research: Tutorial, application and challenges[J]. *Public Administration and Policy Review*, 13(5): 152-168. (in Chinese)

- 马亮. 2017. 实证公共管理研究日趋量化: 因应与调适[J]. 学海, (5): 194-201.
- Ma L. 2017. The increasing quantification of empirical public management research: Responses and adaptations[J]. *Academia Bimestris*, (5): 194-201. (in Chinese)
- 马亮. 2024. 智能时代的公共管理研究向何处去? ——基于研究方法的反思与前瞻[J]. 行政论坛, 31(3): 42-50.
- Ma L. 2024. Where is the future of public administration research in the era of artificial intelligence—Reflection and prospection based on methodology [J]. *Administrative Tribune*, 31(3): 42-50. (in Chinese)
- 迈尔斯, 休伯曼. 2008. 质性资料的分析: 方法与实践[M]. 张芬芬, 译. 重庆: 重庆大学出版社.
- Miles M B, Huberman A M. 2008. Analysis of qualitative data: Methods and applications [M]. Zhang F F, trans. Chongqing: Chongqing University Press. (in Chinese)
- 王培杰, 彭雨馨, 张友浪. 2022. 政策设计、政策认同和生育偏好——基于“独生子女”政策的反馈效应分析[J]. 公共行政评论, 15(5): 27-48, 196-197.
- Wang P J, Peng Y X, Zhang Y L. 2022. Policy design, policy support, and fertility preference: The feedback effects of China's "One-child" policy[J]. *Journal of Public Administration*, 15(5): 27-48, 196-197. (in Chinese)
- 王印红, 王刚. 2013. 对公共管理研究方法中定量推崇的批判[J]. 重庆大学学报(社会科学版), 19(1): 64-71.
- Wang Y H, Wang G. 2013. Criticism of preferring quantitative research methods of public administration[J]. *Journal of Chongqing University (Social Science Edition)*, 19(1): 64-71. (in Chinese)
- 徐换歌, 王峰. 2022. 纵向干预下的城市群协作治理何以降低雾霾污染? [J]. 公共管理与政策评论, 11(3): 65-79.
- Xu H G, Wang F. 2022. How does collaborative governance of urban agglomerations under vertical intervention reduce haze pollution? [J]. *Public Administration and Policy Review*, 11(3): 65-79. (in Chinese)
- 颜海娜, 蔡立辉. 2008. 公共行政学研究方法: 问题与反思[J]. 公共管理学报, 5(4): 109-115.
- Yan H N, Cai L H. 2008. The research method of public administration: Question and reflection[J]. *Journal of Public Management*, 5(4): 109-115. (in Chinese)
- 郁建兴, 高翔, 王诗宗, 等. 2023. 数字时代的公共管理研究范式革命[J]. 管理世界, 39(1): 104-115.
- Yu J X, Gao X, Wang S Z, et al. 2023. The research paradigm shift of public management in the digital era[J]. *Journal of Management World*, 39(1): 104-115. (in Chinese)
- 于文轩, 樊博. 2020. 公共管理学科的定量研究被滥用了吗? ——与刘润泽、巩宜萱一

- 文商榷[J]. 公共管理学报, 17(1): 159-166, 176.
- Yu W X, Fan B. 2020. Is quantitative research methods abused in public management research? [J]. *Journal of Public Management*, 17(1): 159-166, 176. (in Chinese)
- Becker T E. 2005. Potential problems in the statistical control of variables in organizational research: A qualitative analysis with recommendations [J]. *Organizational Research Methods*, 8(3): 274-289.
- Becker T E, Atinc G, Breaugh J A, et al. 2016. Statistical control in correlational studies: 10 essential recommendations for organizational researchers [J]. *Journal of Organizational Behavior*, 37(2): 157-167.
- Bernerth J B, Aguinis H. 2016. A critical review and best-practice recommendations for control variable usage [J]. *Personnel Psychology*, 69(1): 229-283.
- Bernerth J B, Cole M S, Taylor E C, et al. 2018. Control variables in leadership research: A qualitative and quantitative review [J]. *Journal of Management*, 44(1): 131-160.
- Carlson K D, Wu J P. 2012. The illusion of statistical control: Control variable practice in management research [J]. *Organizational Research Methods*, 15(3): 413-435.
- Cook T D, Campbell D T. 1979. *Quasi-experimentation: Design & analysis issues for field settings* [M]. Chicago: Rand McNally College Publishing Company.
- Curado C, Oliveira M, Schniederjans D G, et al. 2024. Control variable use and reporting in operations management: A systematic literature review and revisit [J]. *Management Review Quarterly*, 74(3): 1809-1839.
- Lenz G S, Sahn A. 2021. Achieving statistical significance with control variables and without transparency [J]. *Political Analysis*, 29(3): 356-369.
- Li M X. 2021. Uses and abuses of statistical control variables: Ruling out or creating alternative explanations? [J]. *Journal of Business Research*, 126: 472-488.
- Newey W, Stouli S. 2021. Control variables, discrete instruments, and identification of structural functions [J]. *Journal of Econometrics*, 222(1): 73-88.
- Shiau W L, Chau P Y K, Thatcher J B, et al. 2024. Have we controlled properly? Problems with and recommendations for the use of control variables in information systems research [J]. *International Journal of Information Management*, 74: 102702.
- York R. 2018. Control variables and causal inference: A question of balance [J]. *International Journal of Social Research Methodology*, 21(6): 675-684.

Have We Controlled Properly in Public Administration Research? Findings Based on a Systematic Literature Review

HAN Xiao¹ MA Liang² CHEN Ruifei¹

(1. School of Public Administration, University of Electronic Science and Technology of China;

2. School of Government, Peking University)

Abstract: Quantitative research methodologies have gained significant traction in the field of public administration research, garnering attention for their systematic and objective approach to understanding complex phenomena. Despite its growing popularity, quantitative research has not been without its share of criticism, with a central point of contention being the challenges associated with ensuring its reliability and validity. Scholars argue that while quantitative methods offer a structured framework for analysis, they also present unique obstacles that may compromise the integrity of the research outcomes.

A key strategy employed in quantitative research to strengthen the precision of estimating the relationships among core variables is the utilization of statistical techniques to account for extraneous variables. By controlling for these extraneous factors, researchers seek to isolate the effects of any variables of interest, thereby enhancing the accuracy of their findings. However, the effectiveness of this approach hinges on the proper selection and application of the control variables. In practice, the use of control variables is fraught with challenges stemming from ambiguous methodological guidelines and potential oversights by the researcher. These challenges often lead to errors in the implementation of the control variables, which may cast a shadow over the validity of the entire study.

This study presents an extensive analysis of 760 articles published across eight well-regarded public administration journals over the decade from 2014 to 2023. The focus of the analysis is on identifying the limitations as well as other issues related to the selection, analysis, and reporting of the control variables in these studies. The findings reveal a concerning pattern of methodological shortcomings. First, the selection of the control variables frequently lacks a solid theoretical foundation, with variables often included as a matter of routine rather than based on a clear rationale linked to the research question. Second, there is a notable homogenization in the choice of the control variables, with researchers tending to rely on a standard set of variables without adequately considering the unique context of their studies or exploring alternative variables that may yield deeper insights.

Moreover, this study highlights the frequent omission of multicollinearity tests, which

are crucial for ensuring that the control variables do not introduce noise into the model by being correlated with one another. The absence of such tests may lead to distorted results and incorrect inferences. In addition, many studies fail to provide comparative analytical results, which are essential for contextualizing the findings and understanding the relative impact of the various variables. The reporting of the control variables is also often incomplete, leaving readers in the dark about the rationale behind their inclusion and the methods used to analyze them.

To address these identified gaps, this study offers a set of comprehensive recommendations for the proper use of control variables. These recommendations are designed to help researchers navigate the complexities of quantitative research designs and to enhance the robustness of their studies. By adhering to these guidelines, researchers can significantly improve the reliability and validity of public administration research. This study underscores the importance of careful and thoughtful methodological practices in an era when quantitative research continues to play a pivotal role in advancing our understanding of public administration issues.

Keywords: public administration; control variables; homogenization; omission; bias; validity

投稿日期: 2024/7/30 送外审日期: 2024/8/15 首轮外审完成日期: 2024/9/14

录用日期: 2024/12/3 最终修回日期: 2025/5/3