

·公共管理与社会治理·

强化问责可以提升政府绩效吗? ——基于CFPS数据的断点回归分析



□李文彬 卢琳静

[华南理工大学 广州 510641]

[摘要] 【目的/意义】强化问责是否可以提升政府绩效?学界的答案仍存在争议,需要继续检验。中共十八大后党中央实施全面从严治党、强化问责,为这一问题的解答提供了自然实验的新机会。【设计/方法】基于2010~2016年的中国家庭跟踪调查数据,使用断点回归方法,以2013年中央强化问责前后中国的县级政府绩效变化为分析对象,可以检验强化问责对政府绩效的影响及其路径。【结论/发现】强化问责在短期内显著提高了政府绩效水平,政民互动过程的回应性、服务性、公平性和规范性发挥中介效应。但是,强化问责对政府绩效的影响力呈现倒U型的发展趋势,先升后降,可持续性较差。这一立足全国宏观数据和基于自然实验的分析结果,回应了学界关于问责“有效”和“无效”的争论,提供了争论成因的初步解释,并为理解强化问责对政府绩效的影响提供了时效视角。

[关键词] 强化问责;政府绩效;断点回归

[中图分类号] D630

[文献标识码] A

[DOI] 10.14071/j.1008-8105(2021)-5022

Can Strengthening Accountability Improve Government Performance? —A Regression Discontinuity Design Based on the Data from CFPS

LI Wen-bin LU Lin-jing

(South China University of Technology Guangzhou 510641 China)

Abstract [Purpose/Significance] Can strengthening accountability improve government performance? The answer is still controversial and needs to be tested. After the 18th National Congress of the Communist Party of China, the CPC Central Committee implemented comprehensive and strict governance of the party and strengthened accountability, which provided a new opportunity for natural experiment to answer this question. [Design/Methodology] Based on the tracking survey data of Chinese families from 2010 to 2016, this paper uses the breakpoint regression method to test the impact and path of strengthening accountability on government performance by taking the changes of county and district government performance in China before and after the central government strengthened accountability in 2013 as the analysis object. [Conclusions/Findings] It is found that the implementation of strengthening accountability significantly improves government performance in the short term, and the intermediary effects of responsiveness, service, fairness and standardization of government behavior are significant. However, the influence of strengthening accountability on government performance shows an inverted U development trend, and the sustainability is poor. Based on the national macro data, the analysis based on natural experiments and the interview data of grass-roots civil servants, this paper responds to the debate on the “effectiveness” and “ineffectiveness” of accountability in academic circle, provides a preliminary explanation of the causes of the debate, and a new perspective for understanding the impact of accountability on

[收稿日期] 2022-03-23

[基金项目] 国家社科基金项目(17BZZ032)。

[作者简介] 李文彬(1976-)男,博士,华南理工大学公共管理学院教授;卢琳静(1998-)女,华南理工大学公共管理学院硕士研究生。

government performance.

Key words strengthening accountability; government performance; regression discontinuity

引言

问责一直被视为规范政府管理和提升政府绩效的有效手段。中共十八大以来,党中央坚持全面从严治党、强化监督和问责,明确强调各级党政干部要做到“有权必有责、用权受监督、失职要问责”,将强化问责提升到了前所未有的高度。经过多年的发展,我国也逐渐形成行政问责、政治问责、党内问责三位一体的问责体系。然而,关于加强问责的实际效果,学界的观点泾渭分明。一种观点认为,强化问责可以增强政府官员的责任感,提升政府的回应力,进而提高政府绩效;另一种观点则指出,问责和避责为一体两面,高压态势下的问责也会导致官员避责和不作为,反而损害政府的创新活力和服务绩效。因此,强化问责是否及如何对政府绩效产生影响,仍有待进一步检验。目前,国内学界围绕问责效果的研究主要采用微观个案评估的方式,缺乏基于全国宏观数据和立足因果机制的检验。中共十八大以来,党中央实施全面从严治党、推行强化问责对政府形成的外生冲击刚好构成了自然实验,不仅为在全国范围的宏观层面揭示其影响政府绩效的因果逻辑提供了机会,也为检验强化问责对政府绩效的影响提供新的证据,更有助于拓展现有理论对两者关系的解释。

一、文献回顾与问题提出

关于问责与政府绩效的关系,新公共管理运动的基本主张是,加强问责将导致更高的绩效^[1-2]。然而,也有研究表明,问责与绩效的关系充斥着矛盾性和复杂性:问责的后果并不确定,难以把握其积极或消极的可能性^[3]。这种复杂关系,将直接引发问责对政府绩效产生不同甚至相反的影响,并使得学界的研究呈现问责“有效派”和“无效派”两种观点。

“有效派”认为,问责制因其塑造的规范性和传导的强制性压力能够有效提升政府绩效。问责机制是政府保持和改善政府绩效的最重要手段之一^[4],Anderso指出,问责制强化了民主治理中的责任性与透明性,进而能够提升整个社会的治理效能^[5]。Dubnic和Frederickson认为,精心设计的基于绩效测量的问责制原则上可以提高治理质量,从而改善公共部门实现其目标的前景^[6]。Han考察应用

于政策实施过程的问责机制时发现,美国联邦机构的三种问责机制同时对机构绩效产生了积极影响^[7]。Choi等学者也提出,更强的问责制为组织带来了更理想的结果,特别是法律问责与所有机构中的大多数绩效变量都呈正相关^[8]。Brewer和Walker则指出,问责通过保证行政过程的规范性、开放性和对利益相关者的回应性,来提升公共服务绩效^[9]。可见,这些研究揭示,问责制度有效提升了公共政策实施和公共服务供给绩效。然而,“无效派”则指出,问责削弱了组织的灵活度和创新性且与避责伴生,进而导致组织绩效的损失。Bozeman认为,强加给公共组织的一系列复杂规则、法规和程序通常会失灵并使组织运行僵化守旧,将对其绩效产生负面影响^[10]。Schillemans等指出,问责制消耗了行政人员的精力和分散其注意力并导致目标错位,引发了一系列适得其反的官僚行为进而可能损害组织绩效^[11]。倪星强调,问责的强度塑造着政府官员行为,当避责策略有效时,高强度的问责可能导致更为激烈的避责行为^[12]。陈汉宣和高洁的案例研究也发现,加强问责会阻碍政府绩效提升,目标责任制的推行会导致重视短期问责而轻视长期政府能力建设问题^[13]。因此,两派关于问责是否提升政府绩效的观点可谓泾渭分明,两者的关系有待进一步检验。

综合上述研究“盲人摸象”般的不同结果,可能的原因是研究者各自“摸”到了问责对政府绩效产生影响的侧面,如一方面是在实施问责的起始阶段由于问责客体“不敢顶风作案”或尚未发现避责策略而使问责对政府绩效产生正面影响,另一方面是严苛问责导致组织创新活力下降、问责与避责的共生、实施问责中后期阶段问责对象的“机会主义”避责策略等因素导致问责被消解甚至损害绩效,第三种可能就是有的研究使用个案来探寻问责的影响,而不同个案的属性又会对揭示两者的因果关系形成干扰。总之,关于问责影响政府绩效的因果关系,需要使用范围更广的宏观数据和因果机制更为明晰的方法予以验证。这也是本文的研究目的所在。中共十八大后的第一年,即以2013年为时间点,党中央围绕全面从严治党实施了一系列的措施和活动,如强化反腐倡廉、启动十八届中央委员会首轮巡视、发布党内法规、开展“反四风”等,具体如图1所示。这构成了强化党政官员的法律、政治、行政、伦理责任意识和负责行为的组合拳,也

成为高悬在官员头上的问责“达摩克利斯之剑”，强化了对官员行为和作风的失责、失范、失德的约

束。因此，这一自然实验为在宏观层面检验强化问责对政府绩效的影响及其路径提供了机会。

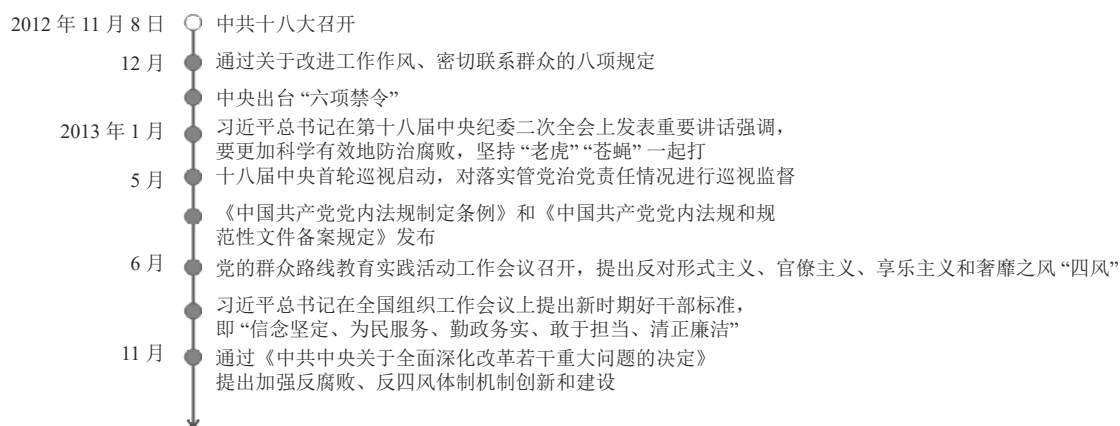


图 1 2013 年中央全面从严治政以强化问责的主要事件

二、强化问责影响政府绩效的路径假设

问责是现代国家治理中的重要制度设计，隐含着提升政府绩效的诉求。问责的含义至少包括回应和强制两个基本维度，回应代表“问责对象有义务告知和解释他们正在做什么”，而强制则代表问责对象需要为不当行为承担强制性的惩罚^[14]。因此，对问责广义理解是指符合这两重维度要求的责任追究形式，如行政问责、积极问责等概念^[15]。对此，

本文所指的强化问责并非指向某项针对具体行政事项或行为的问责制度的实施，而是指 2013 年党中央实行的强化党政领导干部廉洁自律、工作实绩、工作作风的一系列举措所形塑的综合问责环境的变化。这一变化改变了权力运行的诸多条件，在很大程度上重新塑造了政府履职的行为特征。从理论上说，这一制度环境的变化可以促进政府行为的公平性、回应性、服务性和规范性进而对政府绩效结果产生影响。具体假设路径详见图 2。

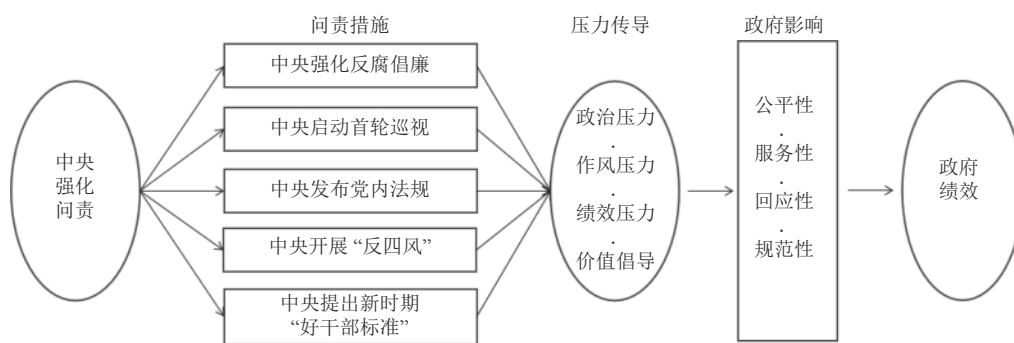


图 2 2013 年中央全面从严治政、强化问责对政府绩效影响路径图

首先是公平性机制。公平正义是社会的首要价值^[16]，也是问责制的主要捍卫对象和政府绩效的重要内涵。实行问责的目的涵盖绩效、公平和公正^[6]。确保官员干部不以权谋私，对待群众一视同仁、不偏私，以及在住房、教育、医疗等公共资源分配领域的公平性，往往是问责的主战场。如“八项规定”中明确要求切实改进工作作风，减少特权现象；在“六项禁令”中严禁官员违反规定收送礼品、礼金、有价证券等任何影响公正执行公务的物品。其次是服务性机制。秉持专业、良好的服务态

度，为公民提供优质公共产品和服务，是政府绩效的核心要求。行政问责制源于专业精神、专业知识和道德标准，希望通过应用专业知识和解决问题的方法来为公众利益服务^[17]。中央在“反四风”中明确要求，反对形式主义，着重解决工作不实的问题；反对官僚主义，着重解决在人民群众利益上不维护、不作为的问题。可见，问责是强化服务性以提升政府绩效的有效途径。再次是回应性机制。政府回应的及时性和有效性，是行政问责的主要维度和政府绩效的重要内容。而提升政府的回应性更是

绩效问责和政府绩效目标的应有之意。高强度的行政问责和高强度的民众问责下更容易驱动官员的有效回应行为^[18]。在2013年全党上下开展的党的群众路线教育实践活动中，中央要求各级领导干部着力解决群众反映强烈的突出问题，突出了在维护人民群众最关心最直接最现实的利益问题上失职失责问题的问责，提高政府回应公众诉求的及时性、有效性，最终提升政府绩效。最后，规范性机制。规范性强调政府权力的边界，问责以责任控制的方式减少官员的滥权、越权等不规范行为。规范性是政府履职增效的前提，强化问责可以通过规范权力运行来确保绩效产出、减少绩效摩擦和损失以及获得公民认同来提高政府绩效。《中共中央关于全面深化改革若干重大问题的决定》提出，推行地方各级政府及其工作部门权力清单制度，依法公开权力运行流程，明晰政府权力的界限，增强依法行政的规范性。因此，通过对2013年党中央强化问责这一时间断点政府权力运用规范性、公平性、回应性和服务性所产生的政民互动过程体验的前后比较，可以有效识别强化问责是否循这四条路径显著影响公

众对政府绩效的评价。

三、变量、数据与方法

（一）变量测量与数据来源

本文的数据源自2010年、2012年、2014年和2016年持续四期的中国家庭追踪调查（China Family Panel Studies, CFPS）对全国16岁及以上的成年人调查中获得追踪数据。该调查为全国性调查，采用内隐分层（Implicit Stratification）方法抽取的多阶段概率样本，覆盖25个省（直辖市、自治区）、162个县（市、区），调查对象包含样本家庭中的全部家庭成员。在成人数据库中，2010~2016年平均每期样本量约为38 000。CFPS的追踪数据为本研究提供了大量的数据点，增加了数据的自由度并降低了解释变量间的共线性程度，因此能够提高计量模型估计的有效性。在数据处理过程中，删除了没有连续四期受访的样本，并对缺失数据采用临近点的线性趋势方法和均值法进行填补。最后得到每期用于实证分析的9 117个样本，四期共36 468个观测值。各个变量设置及其具体含义见表1。

表 1 变量及其含义

变量	变量名称	含义
结果变量	政府绩效	公众对去年本县（市、区）政府工作的总体评价：1~5（1=比之前更糟了，5=有很大成绩）
处理变量	强化问责	2013年之后取值为1，2013年之前取值为0
中介变量	公平性	公众“没有受到政府干部的不公正对待”取值1，否则为0
	服务性	公众“与政府干部没有发生过冲突”取值1，否则为0
	回应性	公众“到政府办事时没有受到不合理的拖延、推诿”取值1，否则为0
	规范性	公众“到政府办事时没有遭到不合理的收费”取值1，否则为0
	性别	女=0，男=1
控制变量	年龄	根据出生日期计算当年年龄
	学历	高中以下=0，高中及以上=1
	收入	个人上一年税后工作总收入，取对数
	相对社会地位	自评社会地位：0~5（0分表示很低，5分表示很高）
	生活满意度	自评生活的满意程度：0~5（0分表示很不满意，5分表示非常满意）
	健康状况	自评健康状况：1~3（1分表示不太健康，3分表示健康）
	廉政评价	公众对政府腐败问题严重程度的感知：0~10（0分表示不严重，10分表示非常严重）

资料来源：CFPS调查问卷。

1. 结果变量。本文的结果变量为政府绩效。在政府绩效评价实践中，存在着重视成本效益分析和量化数据的客观测量和强调公众主观评价两种基本模式^[19]。公众对政府工作表现的评价体现了政府有效满足社会、市场需求和公众心理预期的程度，反映了以公民为中心的政府绩效评价导向^[20]。本文关于政府绩效的测量内容并非是针对某一政府的具体行为或政策（项目）结果的微观判断，而是公众对本地县（市、区）政府工作表现的总体评价，采用从“1”（很差）到“5”（有很大成绩）的五级李

克特量表。

2. 自变量。自变量即处理变量。中共十八大后党中央实施全面从严治党，开创了我国强化问责的新局面。本文以2013年为时间断点，代表实施强化问责的时点。相应地，2013年之前取值为0，2013年之后取值为1。

3. 中介变量。中介变量分别为政府工作人员与公众互动行为的公平性、服务性、回应性和规范性。这四个变量反映了政府与公民之间的日常接触和互动过程，而接触评价是影响公众主观评价的最

稳定因素^[21]。因此,可以选取政民互动的过程体验作为公众对政府工作绩效总体评价的中介变量。CFPS问卷收集了调查样本关于这些变量的切身体验数据。

4. 控制变量。以往的研究发现,公众对政府绩效的评价会受到人口和经济社会特征、心理状态的影响,因此本文控制变量包括个人的性别、年龄、学历、收入以及相对社会地位、生活满意度和健康状况^[22-25]。此外,反腐败也会影响公众对政府绩效

的评价^[26],因此本文将公众对政府的廉政评价列为控制变量。

(二) 描述性统计结果

表2报告了各变量的均值、标准差、最大和最小值,以及在强化问责前后的均值。可见,政民互动行为的公平性、回应性、服务性和规范性,以及公众对政府绩效评价的均值在实施强化问责后得到提高,强化问责可能起到了提升政府绩效的作用,不过其可靠性仍有待后续的检验。

表 2 变量的描述性统计

变量	观测值	均值	强化问责前的均值	强化问责后的均值	标准差	最小值	最大值
政府绩效	36468	3.60	3.55	3.65	0.85	1	5
性别	36468	0.52	0.52	0.52	0.50	0	1
年龄	36468	48.89	46.93	50.86	13.58	16	94
学历	36468	0.22	0.22	0.22	0.414	0	1
收入	36468	4.05	5.517	2.53	4.59	0	13.59
相对社会地位	36468	2.88	2.77	2.98	1.01	1	5
生活满意度	36468	3.62	3.45	3.79	1.04	1	5
健康状况	36468	2.46	2.40	2.52	0.75	1	3
廉政评价	36468	6.39	6.10	6.69	2.44	0	10
公平性	36468	0.89	0.89	0.90	0.31	0	1
服务性	36468	0.95	0.95	0.96	0.21	0	1
回应性	36468	0.85	0.85	0.86	0.35	0	1
规范性	36468	0.91	0.87	0.94	0.29	0	1

数据来源:作者统计。

(三) 模型估计与方法

本文运用断点回归方法检验政府绩效在党的十八大实施全面从严治党强化问责的2013年之后是否有向上跳跃。断点回归 (Regression Discontinuity) 分析方法近似于随机试验,同时也可以对其设计的有效性进行检验,从而为检验变量的因果关系提供可能性。其基本思想就是在外生变化环境中寻找合适的配置变量,使因变量发生“跳跃”,但不直接影响因变量与其他变量的关系^[27]。本文借鉴部分学者的研究,以政策制度实施的时间为断点,以时间作为配置变量来评估政策效应^[28-29]。如能观察到政府绩效在2013年前后发生显著跳跃,就可将这种变化归因于强化问责的施行。由于在断点后的问责政策实施是确定性的,因此本文根据Imbens和Lemieux^[30]采用的局部实验效应的估计方法,构造断点回归模型:

$$\text{Assess}_{i,t} = \alpha + \beta(t-c) + \delta D_{i,t} + \gamma(t-c)D_{i,t} + \sum_{k=1}^8 \mu_k X_{k,t} + \varepsilon_{i,t} \quad (1)$$

$\text{Assess}_{i,t}$ 是本文的结果变量,表示第 t 年第 i 个公众对政府绩效的评价,代表政府绩效; t 是年份, c 是断点2013年,当 $t \geq 2013$ 时,虚拟变量 $D_{i,t} = 1$ 否则 $D_{i,t} = 0$; k 是表示控制变量的序号,如 $X_{k,t}$ 为第 k 个控制变量在 t 年的值; α 、 β 、 δ 、 γ 、 μ_k 均为系数,其

中 δ 为政策系数; $\varepsilon_{i,t}$ 为随机扰动项。当 $t=c$ 时,此模型就可以转换成OLS回归模型,来初步检验强化问责是否会对政府绩效产生显著影响。

四、实证结果

(一) OLS基准模型

首先,本文使用OLS回归,将强化问责设置为二分处理变量(2013年以前年份为0,2013年以后年份为1)来评估强化问责对政府绩效的影响。结果见表3。如(1)(2)列所示,无论是否加入控制变量,强化问责均显著提升了政府绩效;而在(3)(4)列控制了时间固定效应之后,结果也表明强化问责对政府绩效产生显著的提升作用。不过,OLS模型可能会忽视变量内生性问题,并导致估计的结果出现偏差。另外,其他影响公众政府绩效评价的因素也可能随着问责的变化而发生变化,例如社会经济发展带来的个人收入、社会地位、生活满意度等因素。可见,如果仅依靠OLS模型则无法将强化问责的净效果准确评估出来,因果机制仍有待检验。因此,本文需要进一步使用断点回归模型解决变量内生性问题,并评估强化问责的政策效应。

表 3 强化问责对政府绩效影响的OLS回归结果

变量	政府绩效			
	(1)	(2)	(3)	(4)
处理变量 (强化问责)	0.092*** (0.009)	0.048*** (0.008)	0.256*** (0.097)	0.253*** (0.096)
性别		0.028** (0.121)		0.033*** (0.009)
年龄		0.005*** (0.000)		0.006*** (0.000)
学历		0.312*** (0.121)		0.048*** (0.009)
收入		0.004*** (0.001)		-0.003*** (0.001)
相对社会地位		0.048*** (0.005)		0.058*** (0.005)
生活满意度		0.072*** (0.004)		0.008*** (0.005)
健康状况		0.028*** (0.006)		0.040*** (0.006)
廉政评价		-0.004** (0.006)		-0.011*** (0.002)
控制变量	无	有	无	有
时间固定效应	否	否	是	是
观测值	36468	36468	36468	36468
R ²	0.0029	0.037	0.011	0.044

注：括号内为系数的标准误，***、**、* 分别表示在1%、5%、10%水平上显著。
数据来源：作者统计。

(二) 断点回归模型

为更可靠地评估强化问责对政府绩效的影响，此处采用断点回归的非参数方法来进行估计分析。首先，借助图形来初步判断结果变量与配置变量之间的关系。对配置变量进行去中心化处理之后，如图3和图4所示，无论是进行线性拟合还是二次拟合，政府绩效在2013年强化问责前后存在明显的跳跃，跳跃幅度在0.3左右，初步说明实施强化问责在一定程度上通过某种机制提高了政府绩效。虽然图形为强化问责与政府绩效提升之间的因果关系提供了部分证据，但还不能确定具体的影响程度以及显著性。同时，在进行估计之前，仍需要对断点回归模型进行有效性检验，以判断其是否满足模型的假设条件。

1. 有效性检验

(1) 配置变量不受人为操纵的检验

即检验断点附近是否存在操纵配置变量的问题，也就是检验配置变量在断点处是否连续。如果存在人为控制配置变量的情况，则会在断点左右导致明显的个体分布差异，系数估计值就会出现偏差。有学者认为，如果采用时间作为配置变量则无法人为操纵^[27-28]。本文亦以时间作为配置变量，调查样本无法对其进行操纵。同时，此处使用rdcont命令检验断点两侧的配置变量密度函数的连续性。断点回归在2013年处的样本非随机检验的p值为0.791，不能拒绝原假设，表明配置变量密度函数

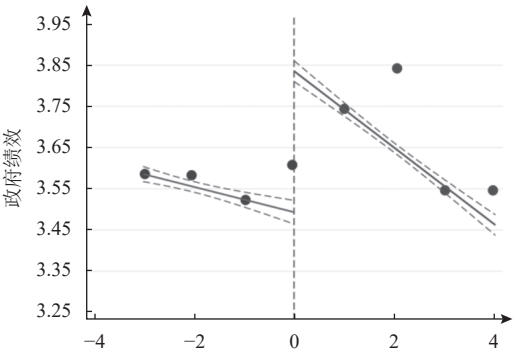


图 3 线性拟合结果

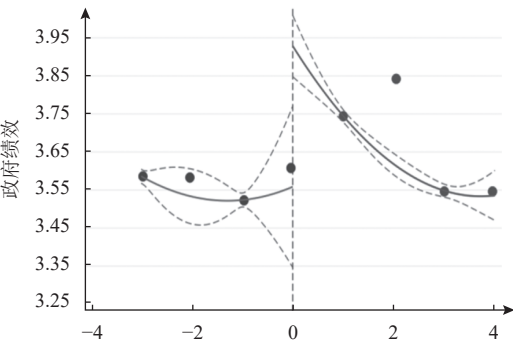


图 4 二次型拟合结果

注：横坐标数值表示各年份与2013年之间的差距。图3实线为强化问责对政府绩效的线性拟合线，图4实线为强化问责对政府绩效的二次方拟合线，虚线表示95%水平置信区间上下限，圆点为散点。
数据来源：作者统计。

在临界值处是连续函数。因此，模型的配置变量不存在人为干预，满足模型的随机性假设，通过了配置变量不受人为操纵的检验。

(2) 局部平滑性的检验

局部平滑性就是检验控制变量是否在断点处连续。本文采用李卫兵、刘生龙等学者^[31-32]使用的方法来检验控制变量的连续性。即使用局部线性回归的方法，将控制变量代替断点回归模型中的结果变量进行回归分析。由表4可知，控制变量在断点处的变化并不显著，断点效应系数较小且远低于均值，因此不会干扰本文的断点回归估计。

2. 断点回归估计

通过了有效性检验后，可以认为断点回归模型适用于本次的研究，使用强化问责作为自然实验并采用断点回归估计去考察其对政府绩效的影响是有效的。本文使用stata软件的rd断点估计命令来检验中共十八大后的2013年中央实施的强化问责对政府绩效的影响。此处使用Imbens和Kalyanaraman提出的MSE方法来计算出此次断点回归的最优带宽为3.341。如表5所示，强化问责的实施显著提高了政

表 4 断点回归模型的控制变量连续性检验

控制变量	性别	年龄	学历	收入
断点效应	-0.00 (0.01)	-0.13 (0.31)	0.00 (0.01)	-0.12 (0.02)
控制变量	相对社会地位	生活满意度	健康状况	廉政评价
断点效应	0.08 (0.14)	-0.04 (0.14)	0.04 (0.08)	0.51 (0.34)

注：括号内为系数的标准误，***、**、* 分别表示在1%、5%、10%水平上显著。

数据来源：作者统计。

府绩效，且加入控制变量后系数仍然显著，有无控制变量的政策效应相差不大。前文图1显示因变量在断点前后存在跳跃的初步发现，得到断点回归模型结果的证实。因此，OLS模型和断点回归模型的结果均表明，强化问责促使政府绩效水平得到提升，但呈现先升后降的倒U型趋势。

表 5 强化问责影响下政府绩效的断点回归估计

变量	政府绩效	政府绩效
	(1)	(2)
断点效应	0.343*** (0.019)	0.305*** (0.019)
控制变量	无	有
观测值	36468	36468

注：括号内为系数的标准误，***、**、* 分别表示在1%、5%、10%水平上显著。

数据来源：作者统计。

3. 稳健性检验

(1) 不同带宽下的断点回归

断点回归中所选带宽的大小会影响对断点回归结果的稳健性和有效性。一般认为，带宽越小，断点附近的各种因素就越相似，跳跃的识别就越准确，但也可能导致样本大量损失而估计失败。因此，本文基于样本量与最优带宽考虑，分别使用最优带宽的50%、200%作为带宽对模型进行回归，来检验估计结果的稳健性。结果如表6所示。使用0.5倍最优带宽进行断点回归，强化问责对政府绩效的积极影响在有控制变量情况下显著；使用2倍最优带宽进行回归，无论是否有控制变量，政策效应依然显著。由于有控制变量的情况更贴近现实情况，此时的系数均显著，与之前得到的结论一致，说明断点回归模型是稳健的。

表 6 稳健性检验：改变带宽的断点回归结果

变量带宽	对称带宽=1.524		对称带宽=6.095	
	(1)	(2)	(3)	(4)
断点效应	0.085 (0.09)	0.146* (0.084)	0.343*** (0.02)	0.268*** (0.02)
控制变量	无	有	无	有
观测值	21424	21424	32277	32277

注：括号内为系数的标准误，***、**、* 分别表示在1%、5%、10%水平上显著。

数据来源：作者统计。

(2) 断点的安慰剂检验

在2013年前后，我国也出台了一些规范官员行为、改进工作作风的规章制度，这些制度有可能对政府绩效造成影响，因此需要设置伪断点进行安慰剂检验。如果在其他制度实施节点里，结果变量也出现了明显的跳跃，就说明原来的断点回归模型得到的结论并不可靠。本文在原政策断点年份前后各取一个假定的断点，即2012年和2014年进行断点回归分析。结果显示，改变断点位置以后，无论是否有控制变量，回归结果均不显著，证明了本文的结论具有稳健性，如表7所示。

表 7 稳健性检验：伪断点回归

断点时间	2012年		2014年	
	(1)	(2)	(3)	(4)
断点效应	-0.06 (0.058)	-0.04 (0.058)	0.136 (0.088)	0.02 (0.091)
控制变量	无	有	无	有
观测值	9321	9321	9221	9221

注：计算2012年为断点时的带宽为1.6，2014年为断点时的带宽为1.26。

4. 异质性检验

(1) 不同地区省份的异质性分析

考虑到不同地区的政府对于强化问责的敏感度不同，可能有着不同的绩效水平。为了更可靠地检验本文的观点，限于数据库只开放省级编码，本文将样本按省份进行异质性分析。我国东部地区的经济发展水平比较高，与中西部地区相比对政策制度的执行情况并不一致，因而强化问责对不同地区的政府绩效的影响可能具有差异。按照国家发改委的解释，东部地区是指较早实行沿海开放政策的东部沿海省市，包括了北京市、天津市、河北省、广东省、广西壮族自治区、辽宁省、海南省、上海市、浙江省、江苏省、福建省和山东省。将东部与中西部地区进行分样本断点回归，结果如表8的（1）～（4）列所示。回归结果显示，在有控制变量的情况下，强化问责对东部和中西部地区政府绩效的效应均为显著，而且对东部地区的作用更为明显。这可能是因为东部地区有着更高水平的公共服务和更多的公共管理人才，对问责制度的执行效率更高，因而强化问责对政府绩效的影响更大。

(2) 不同学历样本的异质性分析

在以公众的主观评价作为政府绩效的测量中，公众的受教育程度显著影响绩效评价结果^[33]。不同的知识水平会带来差异化的个体认知，不同知识水平的个体在绩效信息的收集、识别、分析等方面表现出不同的能力，对政策环境的变化和实际服务质

表 8 样本的地区异质性检验

差异化样本	东部		中西部	
	(1)	(2)	(3)	(4)
断点效应	0.373*** (0.028)	0.254*** (0.028)	0.102 (0.097)	0.118*** (0.025)
控制变量	无	有	无	有
N	16640	16640	10145	10146

注：括号内为系数的标准误，***、**、*分别表示在1%、5%、10%水平上显著。
数据来源：作者统计。

量的感知也不一样,进而不同学历的公众在强化问责的影响下可能作出不同的绩效评价。根据收集到的数据,将所有样本划分为高中及以上学历和高中以下学历,然后分别进行断点回归估计。结果显示,对具有高中及以上学历的公众而言,强化问责对其绩效评价的提高作用要小于高中以下学历的居民,即强化问责对学历较低公众的绩效评价的正面影响更大,如表9所示。可能的原因是,学历越高的公民对政府的期望值也更高、获取绩效信息的能力也更强,因此更倾向于对政府绩效作出更审慎的评价。

表 9 样本的学历异质性检验

差异化样本	高中以下		高中及以上	
	(1)	(2)	(3)	(4)
断点效应	0.373*** (0.022)	0.309*** (0.022)	0.238*** (0.036)	0.165*** (0.036)
控制变量	无	有	无	有
N	28434	28434	8031	8031

注：括号内为系数的标准误，***、**、*分别表示在1%、5%、10%水平上显著。
数据来源：作者统计。

(3) 不同收入样本的异质性分析

公众对政府绩效的评价收到收入水平的影响^[34]。按照CFPS问卷中的收入划分区间,以年收入5万元作为分界点,将所有样本划分为税后年收入5万元及以下和5万元以上的样本分别进行断点回归估计。结果显示,强化问责对较低收入者的政府绩效评价的影响显著高于较高收入者。在无控制变量的情况下,较低收入者的政府绩效评价因强化问责而提升的幅度大约是较高收入者的1.7倍,而且在有控制变量的情况下强化问责对后者的效应在统计上并不显著,如表10所示。可能的原因是,较低收入的群体会因政府质量的提高而产生更高水平的幸福感受^[35],从而倾向于对政府绩效给予更高的评价。

5. 作用机制检验

构建中介效应模型,以检验政民互动过程的公平性、服务性、回应性和规范性是否在强化问责与公众对政府工作绩效的总体评价之间发挥中介作

表 10 样本的收入异质性检验

差异化样本	5万元及以下		5万元以上	
	(1)	(2)	(3)	(4)
断点效应	0.347*** (0.02)	0.283*** (0.02)	0.205* (0.112)	0.112 (0.116)
控制变量	无	有	无	有
N	35724	35724	744	744

注：括号内为系数的标准误，***、**、*分别表示在1%、5%、10%水平上显著。
数据来源：作者统计。

用。根据Baron和Kenny^[36]提出的中介效应模型和检验程序,本文研究的中介效应需要满足三个条件:首先,政府绩效在2013年的断点前后有显著跳跃,通过方程(1)评估,以测量强化问责对政府绩效的总效应。其次,公平性、服务性、回应性和规范性四个中介变量在断点前后也需要有显著跳跃,代表强化问责对四个中介变量的作用效果,具体操作为将方程(1)的因变量分别改为四个中介变量。最后,政府绩效与四个中介变量分别存在显著相关关系,具体中介效应计量模型如式(2)所示:

$$Assess_2 = \vartheta_0 + \vartheta_1 d + \varphi Mediator_n + \pi X_2 + \varepsilon_2 \quad (2)$$

Assess₂为此方程的结果变量,反映政府绩效;*d*为虚拟变量,强化问责前为0,强化问责之后为1; Mediator_{*n*}依次为中介变量政民互动的公平性、服务性、回应性和规范性, *n* = 1,2,3,4; *X*₂代表控制变量; φ 为控制*d*之后,中介变量对结果变量的作用效果; ϑ_1 是控制中介变量后,强化问责实施对政府绩效的直接效应; π 是待估计的参数, ε_2 为随机扰动项。

检验结果如下:(1)表5的回归结果验证了强化问责可以显著提升政府绩效,断点回归系数为正且在1%的水平上显著,满足中介效应检验的第一个条件。(2)对四个中介变量分别作断点回归的图形分析,从图5到图8可以看出,无论是公平性、服务性还是回应性和规范性,在95%置信区间水平下均发生了显著的跳跃。同时,模型结果显示,强化问责改善了公众对政民互动过程的公平性、服务性、回应性和规范性的体验感,回归系数显著为正,中介效应检验的第二个条件得以成立,详见表11。(3)运用OLS模型检验政府绩效与政民互动过程的公平性、服务性、回应性、规范性的关系。回归结果显示,四个变量对政府绩效的影响系数分别在1%的置信水平下显著为正,说明政民互动过程的公平性、服务性、回应性、规范性显著正向影响公众对政府工作绩效的总体评价,详见表12。同时通过以上三个步骤的检验,可以证明四个变量在强化问责对政府绩效的影响发挥了中介作用,即

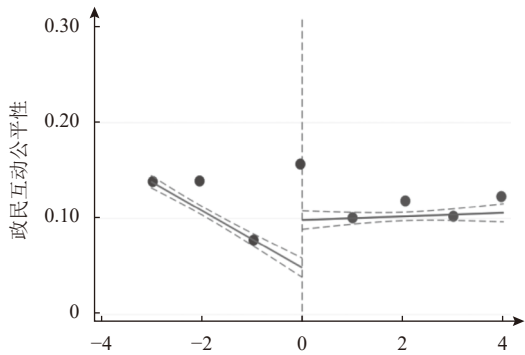


图 5 强化问责对政民互动公平性的影响

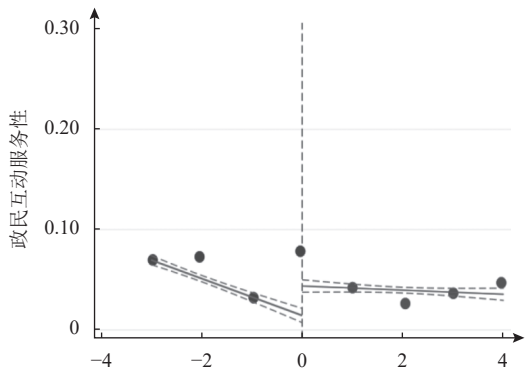


图 6 强化问责对政民互动服务性的影响

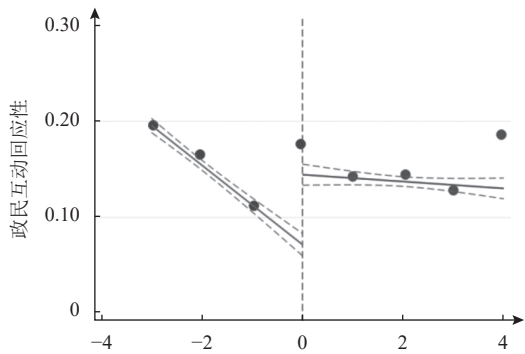


图 7 强化问责对政民互动回应性的影响

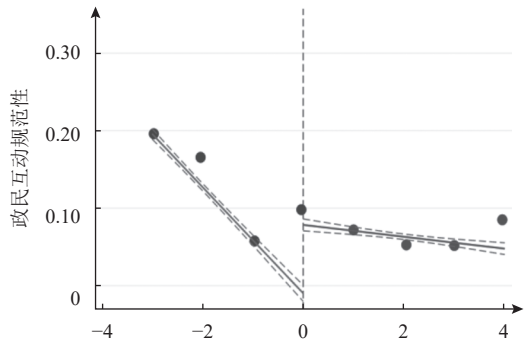


图 8 强化问责对政民互动规范性的影响

注：横坐标数值表示各年份与2013年之间的差距。实线为线性拟合线，虚线表示95%水平置信区间上下限。
数据来源：作者统计。

强化问责的实施可以通过改善公众对政民互动过程的公平性、服务性、回应性、规范性体验感来提升公众对政府绩效的总体评价。

表 11 强化问责对中介变量影响的断点回归估计

中介变量	公平性	服务性	回应性	规范性
	(1)	(2)	(3)	(4)
强化问责	0.072*** (0.007)	0.039*** (0.005)	0.096*** (0.008)	0.105*** (0.006)
控制变量	有	有	有	有
带宽	3.939	4.571	5.165	3.171
观测值	32 085	32 085	32 085	32 085

注：括号内为系数的标准误，***、**、*分别表示在1%、5%、10%水平上显著。
数据来源：作者统计。

表 12 强化问责对政府绩效影响机制的OLS检验

变量	政府绩效			
	(1)	(2)	(3)	(4)
公平性	0.350*** (0.014)			
服务性		0.400*** (0.021)		
回应性			0.332*** (0.012)	
规范性				0.335*** (0.015)
控制变量	有	有	有	有
常数项	2.427	2.317	2.460	2.426
样本数	36 468	36 468	36 468	36 468
R ²	0.053	0.046	0.056	0.050

注：括号内为系数的标准误，***、**、*分别表示在1%、5%、10%水平上显著。
数据来源：作者统计。

五、总结与讨论

本文使用CFPS连续四期的成人追踪数据，基于中共十八大后党中央实施全面从严治党、强化问责的自然实验，采用断点回归方法尝试在宏观上进一步检验加强问责对提升政府绩效的影响，形成以下研究结论：第一，实施强化问责在短期内显著提高了政府绩效水平。断点回归模型通过了有效性检验，并且在改变了带宽、选择伪断点之后的回归结果依然稳健。进一步的异质性检验发现，强化问责对东部地区政府的绩效影响更大，并对学历低或收入低的群体的政府绩效评价产生更大的影响。第二，强化问责通过政民互动过程的公平性、服务性、回应性和规范性这四个中介变量对政府绩效产生影响。中介效应模型检验结果显示，这四个变量的中介作用效应显著。第三，强化问责对政府绩效的影响力呈现倒U型的发展趋势。图1形象地反映了这一趋势。数据统计也发现，在实施强化问责后，公众对政府绩效评价的均值在2014年上升到

3.74,比2012年(均值为3.52)提高了6.25%;然而,2016年则下降至3.55,基本上恢复至2012年的水平。这说明,强化问责初期对提升政府绩效发挥了立竿见影的短期效应,但随着时间的推延,其作用逐步弱化并难以持续。究其可能的原因,一是问责与避责并非二元对立,而是相伴而生、互动一体的“一币两面”^[37],问责的不断强化驱使官员的行为走向“避责”,两者相互触发甚至可能相互抵消最终不断消解问责的效力;二是运动式治理的问责风暴,可能强化机会主义的行政作风并最终形成问责制的短暂生命周期^[38];三是问责制度本身可能存在全过程问责缺失、重政治方法轻管理技术的缺陷^[39]。另外,问责制度及其实施存在的“一刀切”“泛化”“运动化”问题,导致问责对象在问责过程中产生问责疲惫与问责规避心理和问责缓冲及问责对冲的应对策略,从而逐渐消减强化问责对政府绩效的积极作用。总言之,本文基于自然实验分析发现的倒U型趋势有效回应了学界关于问责“有效”抑或“无效”的争论:强化问责在实施初期显著提高了政府绩效水平,这与问责“有效派”的观点相一致;但随着时间的推延,问责对提升政府绩效的影响趋于下降甚至消失,这与问责“无效派”的观点相吻合。可见,两派关于问责效果讨论发生争议的原因可能是各自关注了问责发生的不同时间阶段。因此,本文立足全国宏观数据和基于自然实验的分析,为这一争议的成因提供了初步的解释,以及为理解强化问责对政府绩效的影响提供了时效视角。为打破强化问责提升政府绩效的“时差”(时效差别),强化问责的同时应该配套赋能、容错和构建动态问责的机制,关注问责的时间效应,针对问责对象的机会主义行为动态调整问责方式方法和控制好问责的时间节奏,持续发挥问责效力。

本文研究在时间长度和数据全面性存在不足。本文追踪的时间只是反映了强化问责前后六年间县级政府绩效的变化,还需要更长时间的数据才能更好地呈现问责的长远和持续影响;同时,囿于疫情期间难以获取CFPS调查的县级政府的名称并匹配客观绩效数据,本文仅采用公众评价的角度来测量政府绩效,具有一定的主观性。另外,影响政府绩效的因素有很多,其与政策的协同度可能会影响强化问责的效应,本文难以一一囊括。因此,未来的研究可立足更长时期、综合反映政府绩效的主客观数据,整合更多的影响政府绩效的变量因素,来进一步检验强化问责对政府绩效的长远和可持续影响。

参考文献

- [1] DUBNICK M J. Accountability and the promise of performance: in search of mechanisms[J]. *Public Performance and Management Review*, 2005, 28(3): 376-417.
- [2] CHRISTENSEN T, LÆGREID P. Performance and accountability: a theoretical discussion and an empirical assessment[J]. *Public Organization Review*, 2015, 15(2): 207-225.
- [3] 阎波, 吴建南. 绩效问责与乡镇政府回应行为——基于Y乡案例的分析[J]. *江苏行政学院学报*, 2012(2): 109-115.
- [4] SCHILLEMANS T. Calibrating public sector accountability: translating experimental findings to public sector accountability[J]. *Public Management Review*, 2016, 18(9): 1400-1420.
- [5] ANDERSON J. Illusions of accountability: credit and blame sense making in public administration[J]. *Administrative Theory & Praxis*, 2009, 31(3): 322-339.
- [6] DUBNICK M J, FREDERICKSON H G. *Accountable governance: problems and promises*[M]. New York: Routledge, 2014.
- [7] HAN Y, HONG S. The impact of accountability on organizational performance in the U. S. Federal Government: the moderating role of autonomy[J]. *Review of Public Personnel Administration*, 2019, 39(1): 3-23.
- [8] CHOI S, CHUN Y H. Accountability and organizational performance in the public sector: analysis of higher educational institutions in Korea[J]. *Public Administration*, 2021, 99(8): 353-370.
- [9] BREWER G, WALKER R. The impact of red tape on governmental performance: an empirical analysis[J]. *Journal of Public Administration Research and Theory*, 2009, 20(2): 233-257.
- [10] BOZEMAN B. *Bureaucracy and red tape*[M]. Upper Saddle River, NJ: Prentice-Hall, 2000.
- [11] SCHILLEMANS T, VAN TWIST M, VANHOMMERIG I. Innovations in accountability[J]. *Public Performance & Management Review*, 2013, 36(3): 407-435.
- [12] 倪星, 王锐. 从邀功到避责: 基层政府官员行为变化研究[J]. *政治学研究*, 2017(2): 42-51+126.
- [13] CHAN H, JIE G. Putting the cart before the horse: accountability or performance?[J]. *Australian Journal of Public Administration*, 2010, 68(S1): 51-61.
- [14] SCHEDLER A. Conceptualizing accountability[M]// DIAMOND L, PLATTNER M F, SCHEDLER A, eds. *The self-restraining state*. Boulder: Lynne Rienner, 1999: 13-28.
- [15] 王若磊. 问责的概念与模式[J]. *理论界*, 2012(7): 17-20.
- [16] 约翰·罗尔斯. 正义论[M]. 何怀宏, 何包钢, 廖申白, 译. 北京: 中国社会科学出版社, 1988: 1.

- [17] FRIEDLICH C J. Public policy and the nature of administrative responsibility[J]. *Public Policy*, 1940(1): 1-20.
- [18] 李华胤. 使回应运转起来: 线上互动中的多维问责与有效回应——以H村“互助问答”平台运作为分析对象[J]. *江苏社会科学*, 2021(1): 80-90.
- [19] 倪星, 李佳源. 政府绩效的公众主观评价模式: 有效, 抑或无效?[J]. *中国人民大学学报*, 2010, 24(4): 108-116.
- [20] MORGESON F V. Citizen satisfaction: improving government performance, efficiency, and citizen trust[M]. New York: Palgrave Macmillan, 2014.
- [21] WORRALL J L. Public perceptions of police efficacy and image: the “fuzziness” of support for the police[J]. *American Journal of Criminal Justice*, 1999, 24(1): 47-66.
- [22] 裴志军, 陶思佳. 谁会给政府“差评”: 社会资本和生活满意度对政府评价的影响——基于中国农村社会调查的数据研究[J]. *中国行政管理*, 2018(1): 98-103.
- [23] MORGESON F V, PETRESCU C. Do they all perform alike? an examination of perceived performance, citizen satisfaction and trust with US federal agencies[J]. *International Review of Administrative Sciences*, 2011, 77(3): 451-479.
- [24] 李文彬, 何达基. 政府客观绩效、透明度与公民满意度[J]. *公共行政评论*, 2016, 9(2): 93-111+206-207.
- [25] VAN RYZIN G G, MUZZIO D, IMMERWAHR S, et al. Drivers and consequences of citizen satisfaction: an application of the American customer satisfaction index model to New York city[J]. *Public Administration Review*, 2004, 64(3): 331-341.
- [26] 柳建坤. 从严反腐与中国政府绩效评价的优化——来自准自然实验的证据[J]. *公共行政评论*, 2019, 12(4): 44-61+190.
- [27] 段哲哲, 周义程. “凡进必考”可以提升基层公务员的专业自主性吗?——基于准自然实验的断点回归分析[J]. *甘肃行政学院学报*, 2020(1): 13-24+125.
- [28] 王钊, 王良虎. 税收优惠政策对高技术产业创新效率的影响——基于断点回归分析[J]. *科技进步与对策*, 2019, 36(11): 109-116.
- [29] 徐晔, 蔡奇翰. 高新技术企业认定对企业创新及财务绩效的影响——基于断点回归方法[J]. *复旦学报(社会科学版)*, 2019, 61(6): 139-150.
- [30] IMBENS G, LEMIEUX T. The regression discontinuity design — theory and application[J]. *Journal of Econometrics*, 2008, 142(2): 611-614.
- [31] 李卫兵, 邹萍. 空气污染与居民心理健康——基于断点回归的估计[J]. *北京理工大学学报(社会科学版)*, 2019, 21(6): 10-21.
- [32] 刘生龙, 周绍杰, 胡鞍钢. 义务教育法与中国城镇教育回报率: 基于断点回归设计[J]. *经济研究*, 2016, 51(2): 154-167.
- [33] FRANK J, SMITH B W, NOVAK K J. Exploring the basis of citizens' attitudes toward the police[J]. *Police Quarterly*, 2005, 8(2): 206-228.
- [34] 阮荣平, 周佩, 程郁. 多维度民生状态对政府满意度的影响及其动态效应——基于“中国民生调查”民生关切点入户调查的实证分析[J]. *世界经济文汇*, 2017(6): 88-105.
- [35] 陈刚, 李树. 政府如何能够让人幸福?——政府质量影响居民幸福感的实证研究[J]. *管理世界*, 2012(8): 55-67.
- [36] BARON R M, KENNY D A. The moderator-mediator variable distinction in social psychological research: conceptual, strategic and statistical considerations[J]. *Journal of Personality and Social Psychology*, 1986, 51(6): 1173-1182.
- [37] 谷志军, 陈科霖. 责任政治中的问责与避责互动逻辑研究[J]. *中国行政管理*, 2019(6): 82-86.
- [38] 徐晓林, 朱国伟. 解释与取向: 运动式治理的制度主义视野——以“治庸问责”风暴为背景的分析[J]. *学习与实践*, 2011(8): 86-94.
- [39] 尚虎平, 张怡梦, 钱夫中. 我国政府绩效问责的成就、不足与改进之路——面向20个改革案例的矩阵分析[J]. *中国行政管理*, 2016(2): 33-39.

编辑 朱娜