

政府连环抉择机制：行政委托第三方评估实践差异现象的一种解释

张 圣*

【摘要】近年来，受行政委托的第三方评估已成为提升政府治理绩效的重要工具。对此，现有研究主要聚焦于探讨第三方评估实效及其内外部影响因素，却忽视了对第三方评估实践差异现象的对比分析，导致所得结论略显局限。基于对X市A区慈善超市和Y市文明城市测评过程的对比分析，可提炼出一种作用于行政委托第三方评估实效的政府连环抉择机制。研究发现，政府连环抉择机制是一种由行政委托逻辑、政社关系结构、政社合作模式和第三方评估实践类型等关联性选择构成的复杂运作过程。该机制始于“行政督导”与“同谋共治”两类行政委托逻辑。这两类逻辑又分别决定着“委托人—代理人”和“委托人—管家”两种政社关系结构。二分后的政社关系结构又相应决定了被动式和主动式两种政社合作模式。最终，被动式政社合作模式产生了形式化第三方评估，与之相对的主动式政社合作模式则形成了实质性第三方评估。立足政府连环抉择机制，政府应树立同谋共治型委托逻辑，科学对待第三方评估过程；第三方评估机构则需提升专业能力、增强责任意识，并以“管家”身份参与其中。在政社共同努力下，优化行政委托第三方评估的实践过程，从而提升其实效。

【关键词】行政委托 第三方评估 治理绩效 政社关系 政府连环抉择机制

【中图分类号】D63

【文献标识码】A

【文章编号】1674-2486 (2022) 03-0119-21

一、引言

党的十八大以来，新一届政府大举推进“放管服”改革，并将第三方评估视为改革的一项重要举措。希望借助第三方评估机构的专业力量系统考察政府

* 张圣，上海交通大学中国城市治理研究院、国际与公共事务学院博士研究生。感谢匿名评审专家、编辑部对本文提出的修改意见。感谢华东理工大学社会与公共管理学院李心怡，上海交通大学国际与公共事务学院陈鼎祥、季曦对本文修改作出的贡献。

基金项目：国家社会科学基金重大项目“慈善组织的治理和监督机制研究”（项目编号：20 & ZD182）。

治理绩效,及时发现政府治理问题,有效规避因政府失灵导致的治理困境(陈潭,2017),从而提升政府决策科学性、民主性和有效性。

反观现实,行政委托第三方评估却有着不同的实践表现,而且对政府治理绩效而言也并非都起到了正向的改进作用。基于形式主义与实质主义范式特征辨析(马良灿,2013),可将行政委托第三方评估的实践表现分为两类:一是仅注重表面化的形式运用,本质上仍受制于行政约束,偏向于依据行政统一化标准而非基于现实合理性指标开展的形式化第三方评估。二是将行政依附和专业自主充分融合,将行政经验通过专业能力融入现实的复杂性和不确定性考量之中,设计出顺应国家要求和社会诉求的综合测评指标,并以此开展实效性更强的实质性第三方评估。同时,综合已有研究,形式化第三方评估实属一种纯依附式评估过程,易出现在行政委托任务指示下,第三方评估机构采取曲解指标定义、选择性关注和干扰指标目标值设定等操作策略,变相生产政府所期绩效的伪评估行动(李佳源,2021),直接制约着专业优势对行政工作的有益补充。并且,这类形式化第三方评估也助长了有失公允的政社利益合谋行为(尚虎平、孙静,2020),造成“劣币驱逐良币”“利益链滋生”等负面后果(郑佳斯、卜熙,2020),致使第三方评估饱受诟病,演化为受到行政控制、不断围绕行政逻辑内卷的治理工具(高丽、徐选国,2019)。相较而言,实质性第三方评估则构成一种依附式自主过程(王诗宗、宋程成,2013),能够通过评估因地制宜地为政府提供新的视角、观念、方法和模式(Carman & Fredericks, 2010),督促政府有针对性地提高行政效率和治理效能(王程伟、马亮,2021),更好地开展与国家要求和社会诉求紧密贴合的适宜市政建设(Goyal & Howlett, 2019)。鉴于行政委托第三方评估存在的实践差异现象,该领域研究者也随之二分出担忧第三方评估失灵的审慎选择派和重视第三方评估优势的积极推进派,进而难以达成研究共识。故此,需重新审视行政委托第三方评估实践差异现象,给予其公允评判,从而尽可能推动研究共识的达成并推进相关实践和理论的发展。

聚焦研究过程,本研究主要采取双案例比较方式,对形式化和实质性两类第三方评估实践过程进行了对比分析,提炼出影响行政委托第三方评估实践选择的复杂机制,进而阐释其实践差异现象。由此体现的理论与实践意义主要有以下四个方面:一是通过双案例比较方式弥补了审慎选择派和积极推进派“一边倒”结论的不足,对不同类型的第三方评估进行了较为合理公正的评判。二是立足政社契约关系,构建了融合行政和专业两类要素以及“理念调和—行动调适—结果调整”三个维度的第三方评估新分析框架,有助于对行政委托第三方评估实践差异现象进行对比分析。三是提炼出一种由行政委托逻辑、政社关系结构、政社合作模式和第三方评估实践类型等关联性选择构成的政府连环抉择机制,从理论上深化了对行政委托第三方评估实践差异现象的理解。四是立足政府连环抉择机制,可对政社双方提出富有建设性的优化策略,以期通过二

者努力共同推进第三方评估朝向实质性方向发展，更好地发挥其实效。

二、研究基础、分析框架与研究设计

（一）研究基础

回顾行政委托第三方评估的已有研究成果，多数侧重于分析影响第三方评估的内外部因素。其中，内部影响因素主要涉及评估能力和指标设计两项内容；外部影响因素则多考究制度环境和关系网络两方面要点。

在内部影响因素方面，评估能力往往会受到更多学者关注，统合来看，大致包含了评估组织合法性、评估技术熟练性、评估信息全面性、评估方法专业性和评估结果反馈有效性等多方面特征（李晓方等，2019；王春业、邓盈，2017；徐双敏、崔丹丹，2016），将直接影响其实效的发挥。当然，为提升第三方评估能力，有学者倡导培育多元化评估主体，并建议选用“竞争上岗”方式从中择取最有能力的第三方评估机构去承接政府治理绩效评估事宜（郭渐强、严明，2017）。此外，进一步聚焦于评估实践过程，学者们又会关注到具体的指标设计，如针对慈善组织公信力建立的普适性评估标准指南（石国亮，2012）、以精准扶贫统一性为目标的社会评估体系（莫光辉、张菁，2018）和反映高校职能部门绩效的第三方评估指标设置（刘帮成、张宗贺，2017）等。同时，提出第三方评估指标应立足于政社诉求的客观判断来制定（李文彬等，2019），从而使得第三方评估能够与行政评估有所区别，展现其民主性、公正性和全面性的一面（林鸿潮，2014）。

在外部影响因素方面，已有研究指出制度环境是保障、关系网络是依托，二者缺一不可。在制度环境上，第三方评估应软硬兼施，将涉及主体资质管理制度、信息公开制度、声誉激励制度、流程管理制度和法律保障制度等硬性制度与评估宣传制度、绩效文化培育制度和绩效观念普及制度等软性制度充分结合（李春、王千，2014），注重制度内容与时俱进、不断调整完善，保障第三方评估效能发挥。同时，若将制度转化为实践，还需设置较为合理的评估程序，确保评估环节操作有序（孟志华、李晓冬，2018），并配以必要的监督手段予以纠偏（李善同、龚璞，2019）。规范的评估制度和有效的评估程序会使第三方评估能够保持一定的独立性和客观性（程燕林，2017），更能做出全面、深入的评估考核，从而充分发挥第三方评估的监督、问责和激励等多方面作用（卓越、张红春，2016）。再观之关系网络塑造，有学者认为第三方评估机构与政府和被评估方之间应构建较为公正合理的关系结构（阎波，2018），同时可引入公众评估环节（马佳铮，2016），甚至可呼吁构建多方共同参与的评估模式（张玉磊，2014），以期实现评估过程民主化，提高评估结果科学性。

总之,已有研究较好地阐释了行政委托第三方评估的实效及其内外部影响因素,为本文积淀了良好的研究基础。不过,由于这些研究忽视了对第三方评估形式化和实质性两种实践类型的对比分析,导致所得结论略显局限,也难以合理引导第三方评估规避形式化歧途,不断朝向实质性道路迈进。因此,本文拟建立行政委托第三方评估分析框架,基于双案例比较,结合代理理论、管家理论和理性主义等多维理论以及中国社会治理实践现实,提炼出一种解释第三方评估实践差异现象的政府连环抉择机制。同时,以该机制为依据,提出行政委托第三方评估的改进建议,以期更好地发挥其实效。

(二) 分析框架

行政委托第三方评估从本质上仍是政社契约关系的体现,因而需回溯契约学说,找寻构建分析框架的理论基础。基于契约学说可知,契约达成的前提,一方面是交易各方都预期有利;另一方面还应源于交易各方的自由理性行为,过程中双方都有对交易行为的选择权(苏力,1996)。然而,这种理想契约关系的达成必然要建立在契约关系主体身份对等的前提下,但这在政社间就很难成立。宏观上看,透过社会契约论所揭示的人类社会现象可知,在国家中达成诸多契约是一种维持国家秩序和维护民众权利的有效保障,也说明政社契约双方均能在某些方面达成共识,从而彼此获益。不过,由于政社关系的非对等性,使得获益最大者为契约制定者本身,即政府,社会则需要服从(卢梭,2005)。将宏观视角融于微观观察,会发现行政委托第三方评估中的政社契约关系亦是如此,它并非建立在政社身份对等的基础之上,属于非对等性因素造成的行政对社会的制约抑或嵌入(徐家良、张圣,2020)。这就迫使第三方评估机构从理念、行动和结果等多个层面上优先选择政府期待的评估行为,进而达成政府目标预期,产生形式化第三方评估。不过,随着社会意识的觉醒和社会参与能力的提升,政社契约关系往往会因第三方评估机构所持专业资源之于政府治理绩效提升的重要意义,而在二者关系对等性上发生一定偏移。第三方评估机构逐渐开始占据话语权和主动权,成为尽可能与政府对等交流的社会主体,不再囿于行政控制,能够在理念、行动和结果等多个层面上自主强化行政诉求和专业科学要求之间的内在联系。由此开展的第三方评估将会在实践中逐渐由形式化向实质性转变,从而更好地发挥其在提升政府治理绩效和社会治理水平上的关键作用。立足该变化,为提炼出便于比较的分析框架,可择取两类实践之共性,会发现无论是形式化第三方评估,还是实质性第三方评估,均会在理念、行动和结果等层面融入政府渴望的评估目标,只不过二者在兼顾第三方评估机构专业性和科学性上的程度有所不同。若将该过程具体细化到行政委托第三方评估实践之中,则体现为在政社缔结契约时,会将行政意志与专业认知从理念上相融合。之后根据行政偏好调适专业行为,开展符合政社预期的评估行动。

最后在评估结束时调整出顺应行政期待和符合专业评定的评估结果，做好彰显国家和社会发展需求的综合化评估。由此构建了行政委托第三方评估“理念调和—行动调适—结果调整”过程分析框架（如图1所示），用以对比形式化和实质性两类第三方评估的实践差异。

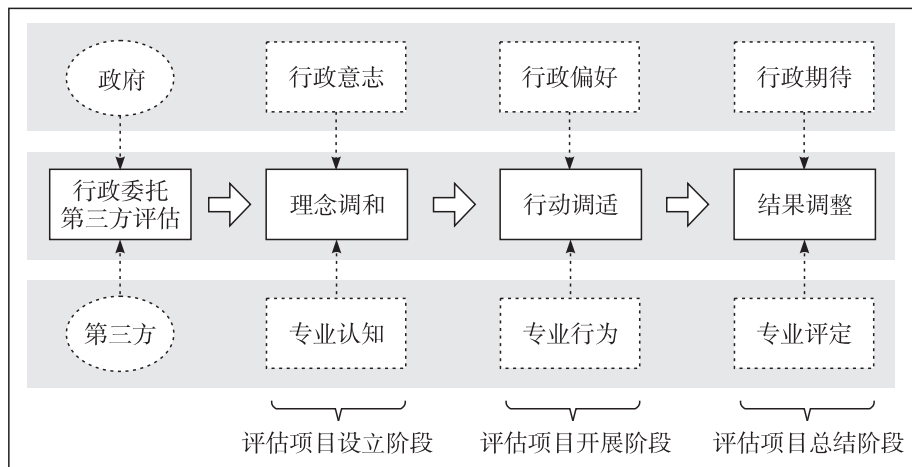


图1 第三方评估“理念调和—行动调适—结果调整”过程分析框架

资料来源：作者自制。

（三）研究设计

1. 方法选择

采用双案例比较法，运用“理念调和—行动调适—结果调整”过程分析框架，针对X市A区慈善超市和Y市文明城市的测评过程，分评估项目设立阶段、评估项目开展阶段和评估项目总结阶段三个阶段进行叙事性再现和动态对比分析，以此探究形式化和实质性两类第三方评估的产生原因和实践机理。

2. 案例选择

本文基于以下理由选取X市A区慈善超市和Y市文明城市的测评过程作为研究案例。

一是二者均为第三方评估机构受行政委托开展的政府治理绩效第三方评估，符合研究主题。

二是两个案例中的第三方评估机构均为专业化智库，能够熟练掌握政策评估、数据收集和数据分析等必要技术（陈振明、黄元灿，2019），排除了因第三方评估机构专业因素造成第三方评估实践差异的可能，使研究更聚焦于政社互动过程的分析。

三是两个案例同为行政委托的合法性行动却形成了不同的实践类型，符合

实践差异现象的探索要求。X市A区慈善超市测评接受行政控制较多，专业自主空间较少，属于形式化第三方评估；与之相对的Y市文明城市测评是在行政对第三方评估机构充分信任的前提下全权委托第三方评估机构开展的评估行动（何文盛、姜雅婷，2020），具有较强的专业自主性，构成了实质性第三方评估。

3. 资料收集

采用参与式观察方式收集资料。一方面，对于X市A区慈善超市测评，作者跟随第三方评估机构G（以下简称“机构G”）完整经历了评估项目设立、评估项目开展与评估项目总结三个阶段，与A区民政局负责慈善超市运营的处室领导进行了多轮交流，获知了A区民政局对于本区慈善超市测评的诸多想法，在密集交流中获得了丰富真实的第一手资料，可较好阐释形式化第三方评估的实践特征。另一方面，对于Y市文明城市测评，作者主要经历了评估项目开展阶段，另外通过与第三方评估机构H（以下简称“机构H”）在Y市的研究基地成员交流，了解到评估项目设立、评估项目总结两个阶段的详细情况，充分获知了Y市文明办对于本市文明城市测评工作的要求和评价，同样能够较好反映实质性第三方评估的实践状况。

三、案例对比：形式化第三方评估与实质性第三方评估

根据作者的参与观察，X市A区慈善超市测评实属受制于政府约束的形式化第三方评估，而Y市文明城市测评则属于获得政府信任和充分赋权的实质性第三方评估。具体可基于“理念调和—行动调适—结果调整”过程分析框架，对比二者在评估项目设立阶段、评估项目开展阶段和评估项目总结阶段的实践表现，从而概括出行政委托第三方评估实践差异现象的基本特征。

（一）形式化第三方评估：X市A区慈善超市测评

1. 评估项目设立阶段：行政意志控制专业认知

在2020年10月，机构G因在慈善超市研究方面有所专长，故被邀请至X市A区民政局对接了有关本区慈善超市的评估事宜。机构G与A区民政局签订了第三方评估协议，达成了政社契约关系。在签订协议前，机构G为彰显其慈善超市评估的专业性，制定了一套有关于A区慈善超市评估的预设方案。不过，当接触到A区民政局后，对方对预设方案不太认同，同时提出A区慈善超市不用做全面考察评估，仅需围绕市里对慈善超市的建设期待，满足A区民政局对各街道慈善超市运营提出的规范化要求即可。在A区民政局看来，此次评估重心在于督促慈善超市改进慈善服务功能，而非考察其是否能够可持续运营，这也是X市市级领导更为关心的。由此表明了A区民政局对待评估的基本立场：

通过“以评促建”方式让X市市级领导看到A区民政局在慈善超市建设上作出的努力，同时认可A区慈善超市的慈善服务绩效。

另外，在协商评估过程中，A区民政局系统盘点了各个慈善超市情况，向机构G提供了一份他们仿效市里要求制定的官方版A区慈善超市评估指标（该指标在半年前A区民政局评估本区慈善超市时就已使用），并提出了符合该指标的A区慈善超市评估方案，严格规定机构G要选用他们的指标和方案去评估。对于机构G提到的关于慈善超市制度化、多样化和创新型发展的评估指标，与A区民政局几经沟通后才勉强答应加入其中的某几条。这在较大程度上打击了机构G开展第三方评估的热情，感觉无论立场、标准还是最终目标都是为了赢得X市打造的慈善超市规范化建设“绩效锦标赛”（周黎安，2007），满足科层制下官员对上级褒奖的渴求。

因此，评估项目设立阶段的机构G属于较为被动的一方，较难调和行政意志和专业认知之间的关系，完全被行政意志左右，评估设计理念和标准已然成为行政化产物。此外，根据A区民政局的安排，在评估项目开展的7天中，A区民政局有关领导将全程跟随和监督机构G的评估工作，预示着下一阶段会有更强的行政干预。

2. 评估项目开展阶段：行政偏好干扰专业行为

在开展评估时，由于A区民政局相关领导的加入，使得评估过程步入到一个较为尴尬的境地。

一方面，机构G在评估时很难有独立空间。鉴于全程有A区民政局相关领导在场，机构G在表达上十分慎重。当被评估方表达对政府，尤其是区级政府的意见或者不满声音时，很难进一步追问，即便追问后也会被在场的区领导以“条件有限”或者“情况所迫”等诸多词汇打断，无法从被评估方那里获得更多信息。

另一方面，被评估方所言内容也相对有所保留。由于被评估方属于街道干部或慈善超市运营方，面对区领导在场也会在表达上较为谨慎，担心评估结果对自身工作评定产生影响。在回答问题时，针对做得好的方面会大肆宣扬，不好的地方要么含糊其词、要么避实就虚。若想刨根问底，在场街道干部乃至区领导还会打圆场，这也会影响机构G的判断。

以上两点以及其他类似干扰情况，导致整个评估过程较难按照机构G对评估的理解去操作。另外，在场的区领导有各自关心的事宜，会更多在意一些与个人行政绩效相关的内容，比如完成了多少次活动、活动开展的效果和有没有宣传展示等，对于慈善超市具体营业状态、经营状况等造血能力上的困难，或者慈善超市在多大程度上对旧模式予以突破均没有给予较多关注。由此可见，整个评估项目开展阶段实际上是受到了较强的行政控制，就连评估交流内容也

需要顺应行政偏好。相对地，机构 G 也在评估环节中默认了行政偏好对其造成的干扰，潜移默化中接受了政府在评估行动上的引导，调适出一种弱化专业行为、自觉接受政府控制并富有行政化色彩的评估方式。

3. 评估项目总结阶段：行政期待操纵专业评定

2020 年 11 月下旬，评估工作进入了收尾阶段，当机构 G 预根据前期评估结果确定 A 区每家慈善超市运营绩效分值时，A 区民政局通知机构 G 在确定前需向他们作汇报。因而机构 G 在独立评判慈善超市运营绩效前又多出一个政社商议环节。在该环节中机构 G 逐个汇报了 A 区每家慈善超市运营绩效状况。此时，A 区民政局相关领导基于先前了解以及对某些慈善超市发展的不满，做出了评估分数调整，变更了评估结果并增加了一些关键事项分值，希望推进慈善超市朝着上级政府期待的目标发展。A 区民政局做出的这些调整行为均说明行政期待正在操纵专业评定，使第三方评估机构的评定客观公正性大打折扣。

之后，在 A 区民政局要求下，机构 G 撰写了每家慈善超市的评估报告。评估报告在 2020 年 12 月 A 区召开的慈善超市点评会议上发布。在当日点评会议上，机构 G 汇报了评估操作细节，对于评估方案的设计、评估方案的实施过程以及评估结果的解读更多是由 A 区民政局负责，当然也只有 A 区民政局能够精准把握评估背后的行政寓意。之后，A 区民政局未采取与此次第三方评估相配套的优化举措，仅寄希望于评估后各街道和慈善超市运营方能够有针对性地自我完善，优化慈善超市的慈善服务单方面功能。

可以说，X 市 A 区慈善超市测评表面由机构 G 负责，实质还是受制于 A 区民政局，属于形式化第三方评估。另外，在该过程中也发现了一则很有意思的现象，即此次第三方评估与半年前 A 区民政局自行开展的评估相比，除了个别慈善超市有分值和排名变动外，两次评估整体变化幅度不大，并且这些慈善超市也多是按照 A 区民政局规划的方向作出了一定改进，进一步印证了形式化第三方评估实质上是一种变相行政评估，也可以说是政府强化监管的一种加压手段。

（二）实质性第三方评估：Y 市文明城市测评

1. 评估项目设立阶段：行政意志融入专业认知

机构 H 作为全国文明城市测评体系创制组织，有着较为丰富的测评和指导经验。因而在 2017 年 Y 市文明办就已邀请机构 H 以开展文明城市测评的方式，帮助其找寻本市文明城市建设上的不足并给予针对性指导。为确保测评后的指导效果，同年，机构 H 于 Y 市设立了研究基地，便于跟进了解 Y 市文明城市建设进度，并随时为其答疑解惑。最终，在机构 H 的专业支持下，Y 市以优异表现于 2020 年第六届“全国文明城市”评选中荣获了这块来之不易的“金字招

牌”。长期政社合作也使Y市文明办和机构H建立了较为良好的政社互信关系。到2021年9月，为了迎接“全国文明城市”复检“大考”以及检测Y市各区县文明城市建设绩效，Y市文明办再次与机构H签订了第三方评估协议，开启了基于政社契约关系的新一轮实地测评工作。

在签订协议期间，双方就新时代全国文明城市建设要求展开了诸多探讨，就如何按照Y市实际情况改良测评标准进行了多轮商议。在商议过程中，机构H展现出较强专业性，能够将中央文明办对于全国文明城市建设的关键阐述同Y市文明城市建设实地观察状况相融合，编制成一套适用性更强的新时代文明城市测评规划。如此周全的考虑也使机构H再度赢得了Y市文明办的信任。Y市文明办只是在商议时基于本市五年发展规划提出了一些新的评估构想。之后，这些构想内容也均在分类、加工和提炼后融入测评规划中，进一步优化了Y市文明城市测评的评估指标设计和评估策略选择。

总体而言，在评估项目设立阶段主要是以机构H的专业认知为主，融入的部分行政意志仍然属于专业加工后的产物。这样产出的第三方评估规划更能够体现机构H的专业性，也易于实现专业对行政的有益补充。

2. 评估项目开展阶段：行政偏好依托专业行为

在评估项目开展阶段，Y市文明办信任机构H的专业性，赋予其独立评估的相关权限。赋权后的机构H针对公共文化设施、政务大厅、建筑工地、中小学校、集贸市场和居住小区等被测点位开展了自主评估。评估成员接受了专业训练，在评估中采取先拍照记录、后集中打分方式，避免过程中因打分举动暴露身份，规避区县突击迎检情况，让评估所反映的情况更接近日常。同时，要求评估成员非必要情况尽量减少对被测点位工作人员的打扰，降低寻租行为发生的可能。

鉴于机构H的评估专业性，Y市文明办在其实施文明城市测评期间基本上未作干扰，仅是通过城市监控设施以及派人乔装跟从等方式实时监督机构H的测评行动。一旦有评估成员出现违规现象或者有一些新的有关文明城市测评的行政指示，Y市文明办会第一时间通知机构H的C主任，由其进行警告或传达，指导评估成员及时改进或调整测评行动。但多数情况下行政干预现象出现的频次较低，评估成员表现出的专业行为已经满足了Y市文明办的行政偏好。甚至，当某些行政偏好与专业行为冲突时，Y市文明办会因机构H的专业权威性而不加以干涉。除非遇到十分必要的行政要求，那也会在双方协商一致后方才补充到专业行为之中。由此可见，在Y市文明城市测评项目开展阶段表现出的行动调适，是在赋予专业自主基础上又增添了一定的行政偏好。不过，实际操作过程中主要还是以专业行为为主，行政偏好依托专业行为践行。

3. 评估项目总结阶段：行政期待依赖专业评定

机构 H 在评估项目总结阶段会预先给出一个基于专业测量数据的评估报告，随后会持评估报告与 Y 市文明办进行多轮沟通。沟通后的 Y 市文明城市测评结果仍主要以机构 H 专业评定为准。若 Y 市文明办对评估报告里面部分分值情况有异议，会在与机构 H 评估成员反复核实下作出微调，使之既能遵从专业评定，也可以更加符合行政期待。

总体上，Y 市文明办的行政期待是依赖专业评定去实现的，也是将机构 H 给出的专业评定视为一种督促下级政府工作的利器，依靠奖惩机制推进 Y 市区县文明城市建设工作。据了解，在机构 H 与 Y 市文明办确定评估报告最终版本后，Y 市文明办会召开大会予以发布，通报各区县文明城市建设绩效状况。之后，各区县拿到各自评估报告后又会对其辖区内街道（镇）以发布红黑榜的形式进行奖惩。所谓红榜，是指在此次评估中表现优异的街道（镇），会给予相应物质或荣誉奖励；相对地，黑榜则是用来惩戒那些不注重文明城市建设的街道（镇），会让其在各区县大会上发言反省、以求改进。这就使得评估结果对于各区县乃至其所辖街道（镇）具有较为强烈的触动效果，使其在今后治理环节中更加重视文明城市建设要求，轻易不敢出现懒政怠政情况。

鉴于评估结果上行政期待与专业评定达成了一致并在激发 Y 市文明城市建设上效果突出，后期 Y 市文明办继续邀请机构 H，结合此次评估结果，针对不同区县所反映的情况进行细致指导，为各区县提供个性化改良策略，以期将测评中暴露出的诸多问题逐一解决。这也使第三方评估不再仅是“走走过场”，而是赋予了其实质性的绩效改良价值。最后，机构 H 帮助 Y 市顺利通过了“全国文明城市”复检“大考”，守住了“全国文明城市”称号，也推动了本市文明城市建设工作向更大范围、更深层次和更高水平迈进。

整体来看，Y 市文明城市测评属于行政色彩较弱、专业成分较强的实质性第三方评估。在该评估中，政府所持有的行政意志、行政偏好和行政期待均合理地融入专业化认知、行为和评定之中，产生出良性政社互动关系，最终发挥出第三方评估预期的实效。

（三）两类实践辨析：塑造了被动抑或主动的政社合作模式

X 市 A 区慈善超市和 Y 市文明城市的测评过程均是建立在政社契约关系基础上的第三方评估。不过，二者却在评估项目设立阶段、评估项目开展阶段和评估项目总结阶段呈现出不一样的实践表现，形成了不同的政社合作模式，如表 1 所示。

表1 两类第三方评估的实践比较

比较项	X市A区慈善超市测评	Y市文明城市测评
第三方评估阶段	①评估项目设立阶段 行政意志控制专业认知	①评估项目设立阶段 行政意志融入专业认知
	②评估项目开展阶段 行政偏好干扰专业行为	②评估项目开展阶段 行政偏好依托专业行为
	③评估项目总结阶段 行政期待操纵专业评定	③评估项目总结阶段 行政期待依赖专业评定
政社合作模式	被动式政社合作模式	主动式政社合作模式
第三方评估实践类型	形式化第三方评估	实质性第三方评估

资料来源：作者自制。

在X市A区慈善超市测评中，机构G和A区民政局之间属于被动式政社合作模式。无论是机构G的评估理念、评估行动还是评估结果无不受制于A区民政局，全程较难有专业发挥空间乃至最后的测评结果都是经过行政干预后的产物。这种被动式政社合作模式造就了一种形式化第三方评估，使其较难发挥出预期的实效。与之相对，在Y市文明城市测评中，机构H同Y市文明办形成了主动式政社合作模式。Y市文明办赋予了机构H较为充分的自主空间，并以协商方式与机构H建立了良好互动关系。机构H也会主动考虑到Y市文明办的行政诉求，立足专业角度，在评估项目设立阶段、评估项目开展阶段和评估项目总结阶段分别进行理念调和、行动调适和结果调整，以此回馈Y市文明办的信任并展现出专业对行政的促进作用。相比于被动式政社合作模式，这种主动式政社合作模式更具实际价值，以至在政社契约基础上形成了一种实质性第三方评估，也使得第三方评估发挥出了预期的实效。

不过，分析至此仍是停留于案例本身，揭示的也仅是第三方评估实践差异的表象。那么在表象之内是否还潜藏一层诱因，诱发了不同政社合作模式选择，直至出现第三方评估实践差异现象？这可能需要回到行政委托的理论反思之中由表及里地去寻求答案。

四、回到行政委托：关系、逻辑与机制

一般而言，区别于专业评估机构的自发评估行为，在这种行政委托第三方评估中，政府往往是占据主导地位的一方，具有对政社合作模式以及第三方评估实践类型的选择权。第三方评估机构若想获取政府资源，必然需要遵照政府要求操作，只不过在行政委托下第三方评估机构会具有相对独立性并被赋予了

一定程度的评估自主权。鉴于此,决定第三方评估走向形式化和实质性的主要因素还是要追溯到对行政委托的思考上,考虑蕴含其中的政社关系、委托逻辑及其复杂机制。另外,这一思考过程实则是一种层层递进的“剥笋式”分析过程,符合实证主义中的“深度模式”新理论总结策略。像斯宾塞、涂尔干等持有有机观念的研究者以及涂尔干以后的许多结构主义者,还有巴斯卡、吉登斯等英国新社会实在论者,抑或是以帕金斯为代表的分析实在论者均采用这种研究策略。他们力图挖掘表面现象背后的深层原因或深层力量,逐层发现事件内在的关联性,作出埃里克松所谓的“系统的或辩证的分析”,以此将布劳代尔所言的“像灰尘一样”的事件梳理清晰,以至建构结构、系统或因果联系的理论模型(谢立中,2010)。为此,本文构造了立足实践和理论两个层面,包括第三方评估实践差异现象、合作模式、政社关系和行政委托逻辑等各圈层意涵的圈层递进分析图(如图2所示),以期由表及里进行现象阐释,再从内到外进行机制梳理。其中,案例比较部分针对实践层已作了相关剖析,之后就需要从实践到本质,循于理论依据揭示政社关系结构及其更深一层的行政委托逻辑,以期由行政委托逻辑出发,从内向外提炼造成行政委托第三方评估实践差异现象的内在机制。

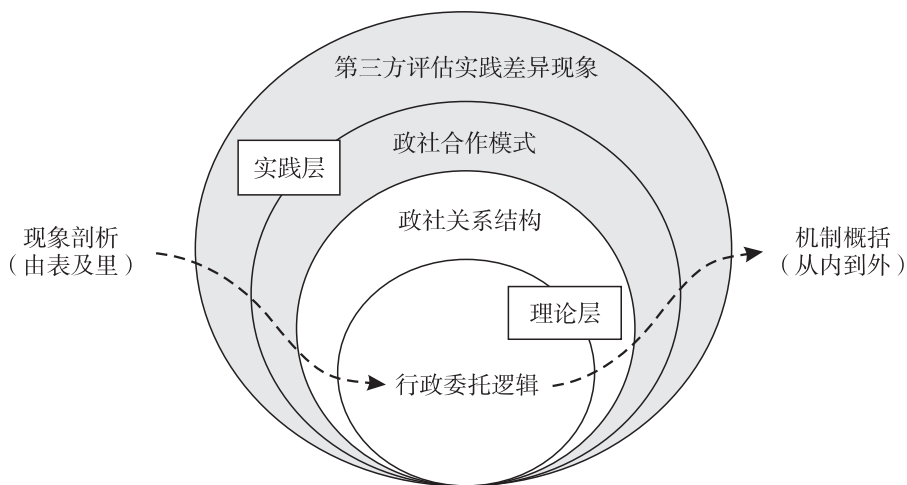


图2 第三方评估实践差异现象的圈层递进分析

资源来源:作者自制。

(一) 政社关系结构:“委托人—代理人”还是“委托人—管家”

辨析政社关系结构是在深挖第三方评估实践表象基础上,从结构角度出发对第三方评估实践差异性作出的较深层解析。鉴于实践中政府普遍采用与第三方评估机构以签订契约的方式委托其参与,塑造了基于政社契约关系的第三方评估过程,这即意味着解析政社关系结构是无法脱离契约学说理论范畴的。细

数契约学说诸多理论著述，择取了具有明显关联性且较适用于解析政社契约关系不同走向的代理理论和管家理论，借以揭示造成第三方评估实践差异现象的政社关系结构。其中，代理理论主要基于经济人假设界定了委托人和代理人之间的关系（Cribb，2006），提出由于双方追求的是各自利益最大化（Eisenhardt，1989），外加存在彼此信息不对称的情况，易触发信任危机、造成冲突博弈（Kluvers & Tippett，2011），最后在无法协调中做出保有各自收益的妥协行为，表现为委托人严加管控和代理人消极应对（Claessens et al.，2002），以此保证委托代理后的工作能够按照委托人意愿完成。私立至上成为整个代理环节中的重要标准。然而，管家理论则是基于社会人假设基础上的思考，强调雇主与管家之间的合作关系应建立在互信基础上，注重自我实现而非私利至上（Van Slyke，2007；Wasserman，2006），打造出一种无缝隙合作模式，而蕴藏其中的非正式信任关系也成为弥补正式合同不完全性的关键（敬父嘉，2011）。在管家模式下雇主与管家的相互信任一定程度上也减少了因二者利益追求不一致、彼此信息不对称等因素引发的冲突博弈，形成了一种较为理想的政社良性互动局面（敬父嘉、崔杨杨，2015）。

基于代理理论和管家理论，可辨别在 X 市 A 区慈善超市和 Y 市文明城市的测评过程中政社契约关系的差异性。如表 2 所示，在 X 市 A 区慈善超市测评中 A 区民政局与机构 G 属于“委托人—代理人”关系。在机构 G 接受 A 区民政局委托时，由于和 A 区民政局对测评目标的理解并非完全一致，因而未获取 A 区民政局的完全信任，几经博弈也无法从 A 区民政局获得自主权，只能饱受行政控制、被动参与。再观之 Y 市文明城市测评，Y 市文明办与机构 H 则属于“委托人—管家”关系，在契约合作中机构 H 无论是评估专业性还是评估目标设定均得到 Y 市文明办肯定，并在互信基础上获得 Y 市文明办赋权，可以直接介入评估决策、自主开展评估工作，表现出主动性参与特征。

表 2 第三方评估中政社契约关系比较

对比项	X 市 A 区慈善超市测评	Y 市文明城市测评
关系类型	“委托人—代理人”	“委托人—管家”
政社目标设定	不完全一致	完全一致
政社信任状态	不充分信任	充分信任
行政委托策略	管控	授权
第三方评估机构参与状况	被动	主动

资料来源：作者自制。

综上，机构 G 与机构 H 均具备专业能力，也同样接受了行政委托，却与政府分别构成了“委托人—代理人”和“委托人—管家”两种政社关系结构。这

反映出X市A区慈善超市和Y市文明城市的测评过程具有不一样的实践机理。对此，鉴于政社契约关系中政府的主导地位，理应回到政社关系结构背后的行政委托逻辑上去思考。

（二）行政委托逻辑选择：“行政督导”与“同谋共治”

在X市A区慈善超市和Y市文明城市的测评过程中，A区民政局与Y市文明办的行政态度实则透露着鲜明的科层制导向。所谓科层制是一种非人格化的多层级政府结构（韦伯，1997），具体体现为下级政府落实上级政府要求、地方政府达成中央政府要求的功利主义实践过程，多由组织激励手段和压力性任务环境共同作用而成（颜昌武，2018）。由此委托第三方评估机构参与，必然会基于科层制导向对其提出一定行政要求，彰显政府选择高于社会选择的特质（王名等，2001），让上级或中央政府看到下级或地方政府在推进政策落实上的努力。若科层制导向是行政委托第三方评估的必然要求，那为何在X市A区慈善超市和Y市文明城市的测评过程却展现出不一样的实践机理？欲想解答这一问题，仍需回归到行政委托逻辑本身的探讨上。当然，思考行政委托逻辑难以脱离对中国社会治理实践探索的整体反思。

纵观中国社会治理实践变迁，大致展现出三方面演进特征，即治理思路从集中管理转向多元治理、治理结构从国家统合转向政社融合、治理策略从总体支配转向局部放权（辛自强，2018），同时衍生出“国家创制社会”（吴晓林、谢伊云，2020）、“行政借道社会”（黄晓春、周黎安，2017）和“行政向社会渐进式赋权”（张圣、徐家良，2021）等各式治理探索。只不过传统意义上的社会参与是有限的，政府仅向社会赋予了部分治理执行权，而治理决策权依然把持在政府手中，同时配有对社会强有力的行政监管，由此达成的社会共治格局实则是一种限制在执行层面且表现为社会有限参与的多元协作过程。基于此，不难理解同为行政委托的第三方评估为何形成了形式化和实质性两类实践。这是因为第一类形式化实践仍具有传统社会治理特征，即政府与第三方合作中始终将其视为行政代理执行者，决策依然由行政把控并表现出鲜明的科层制导向和工具理性价值倾向，坚持“自上而下”高效地落实行政任务，借助第三方评估将行政命令机械发包给下级（周黎安，2014）。同时，这种实践无论对第三方还是对社会治理对象均施加较强行政监管，注重按照行政意志予以纠偏，形成了行政督导型委托逻辑。与之相比，后一类实质性实践已显现出对传统社会治理的突破，出现了决策权分享与行政空间让渡。这就是本文要提出的新社会共治格局，能够视社会参与为契机，不仅要引导社会共治，还应动员社会同谋，毕竟“谋”在“治”前，从而在决策、执行等全流程权力分享的基础上实现社会的充分参与。同时，在该过程中将科层制导向融于工具、价值和沟通等多重理性之中，因地制宜地规划治理行动，使之既能确保“自上而下”地落实行政

任务，也能“自下而上”地满足社会诉求，不断改良政府政策执行和优化社会发展路径，如此则塑造了一种同谋共治型委托逻辑。具体如图3所示。

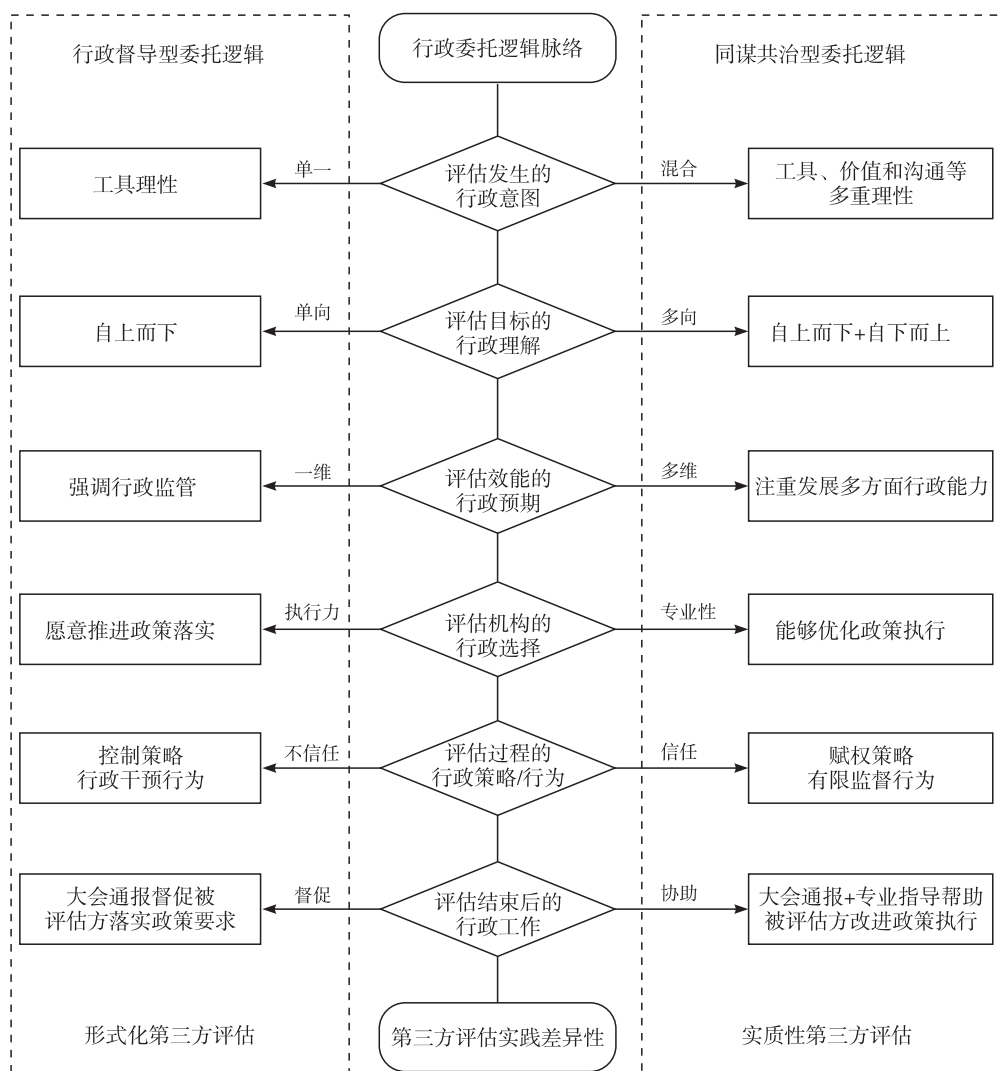


图3 两类行政委托逻辑示意

资源来源：作者自制。

聚焦实践，X市A区民政局对待第三方评估的行政态度，反映出工具理性至上的行政督导型委托逻辑，主张评估效率至上，并强调评估后产生的行政监管作用。由此动机出发，A区民政局在选择有合作意向的第三方评估机构时，相比于其专业性，会更看重其在运用评估方式推进政策落实上的执行力。并且，由于对第三方评估机构信任程度较低，A区民政局不仅不愿赋予第三方评估机构参与决策的权力，还由于担心第三方评估机构自主行动会偏移行政要求，故

而采取一些行政控制策略干预第三方评估过程,限制其自主行动空间。在第三方评估结束后也仅有大会通报环节,后期更多是依凭评估奖惩方式让被评估方自发落实政策要求。在这种行政委托逻辑基础上形成的政社关系属于“委托人—代理人”关系,政社合作过程多为被动模式,最终呈现的也是形式化第三方评估。相对地,Y市文明办则是在纳入多重理性的同谋共治型委托逻辑下,与第三方评估机构开展合作,主张信任和依赖第三方评估机构的专业性,借助第三方评估提升决策、执行等多方面能力,以至优化政策落实效果和社会治理效能。此外,在第三方评估结束后不仅有大会通报,更会继续邀请第三方评估机构给以指导,这使第三方评估既顺应了行政要求,也推进了社会治理优化,凸显出第三方评估的专业优势。故此,基于该逻辑形成的政社关系表现为“委托人—管家”关系,反映的是主动式政社合作模式以及由此产生的具有测评实效的实质性第三方评估。

(三) 政府连环抉择机制：第三方评估实践差异现象的系统阐释

从行政委托第三方评估谈起,本研究重新梳理行政委托逻辑、政社关系结构、政社合作模式与第三方评估实践类型之间的关联性选择关系,提炼出一种政府连环抉择机制,以此解答同为行政委托的第三方评估何以行于“殊途”,同时概括出两条截然不同的实践脉络(如图4所示)。

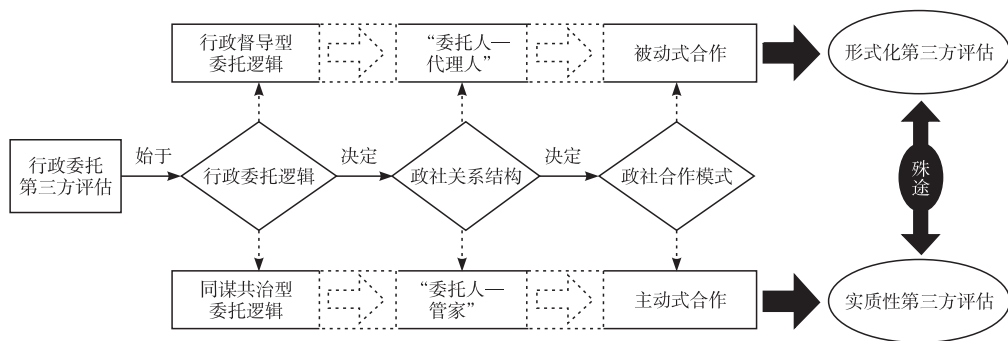


图4 政府连环抉择机制示意

资源来源：作者自制。

具体而言,政府连环抉择机制是立足本土化的行政委托第三方评估的实践机理概括,涉及行政理念、权力配置和流程设计等一系列关联内容,反映了行政委托第三方评估实践过程的复杂性及其分离至殊途的合理性。在中国行政机构难以脱离科层制导向的状况下,多数政府采取第三方评估自然会有行政督导型委托逻辑体现,过度强调对上级政策贯彻执行的效率,不会给予第三方评估机构太多余地。但还有少部分地方政府较为了解和认可第三方评估机构,选择

了同谋共治型委托逻辑，在决策、执行等全流程中会充分向第三方评估机构赋权并给予其高度自主的空间。如此逻辑差异直接决定了契约签订后政社关系结构的不同类型归属。其中，行政督导型委托逻辑产生了偏向于执行效率的“委托人—代理人”政社关系结构，同谋共治型委托逻辑则生成了更具责任观和信任属性的“委托人—管家”政社关系结构。进而，二分的政社关系结构又直接作用于政社合作模式选择，分别演化为以评估项目设立阶段行政意志控制专业认知、评估项目开展阶段行政偏好干扰专业行为和评估项目总结阶段行政期待操纵专业评定为特征的被动式政社合作模式；以及将评估项目设立阶段行政意志融入专业认知、评估项目开展阶段行政偏好依托专业行为和评估项目总结阶段行政期待依赖专业评定归为特点的主动式政社合作模式。最终，被动式政社合作模式产生了形式化第三方评估，而主动式政社合作模式则形成了实质性第三方评估。

总之，行政委托第三方评估到底属于形式化还是实质性实践类型，其实是由行政委托逻辑、政社关系结构、政社合作模式和第三方评估实践类型等一连串关联性选择系统作用的结果，由此呈现的政府连环抉择机制也更具全面性，能够给予第三方评估较为公允的评判，同时引导其合理发展。

五、结论与讨论

本文运用双案例比较方式，结合多维理论，由表及里分析了行政委托第三方评估实践过程及其差异现象，由此主要得出三方面研究结论：其一，同为行政委托的第三方评估会在政府连环抉择机制作用下形成形式化和实质性两种实践类型，且相比之下前者实效性较弱、后者实效性较强。其二，政府连环抉择机制作为一种由行政委托逻辑、政社关系结构、政社合作模式和第三方评估实践类型等关联性选择构成的复杂运作过程，可以系统解释行政委托第三方评估实践差异现象的演进脉络。其三，基于形式化第三方评估和实质性第三方评估的实践辨析，提出形式化第三方评估由内而外展现出行政督导型委托逻辑、“委托人—代理人”政社关系结构和被动式政社合作模式三方面特征，与之相对的实质性第三方评估更多体现为同谋共治型委托逻辑、“委托人—管家”政社关系结构和主动式政社合作模式三方面特点。

本文立足研究结论，从政府和第三方评估机构两方视角出发提出推动第三方评估由“形”转“实”的优化建议。

一方面，政府应树立同谋共治型委托逻辑，科学对待第三方评估过程。立足正式和非正式制度建设，依靠制定标准化第三方评估运行程序、传递第三方评估规范性价值观念以及通过培训增强行政对第三方评估认知意识等（斯科特，

2010),使政府认同第三方评估机构,并愿意与之结成伙伴、同谋共治,实现由单纯追求政策落实效率的工具理性,向兼顾政策落实效率和推行效果,充分融合工具、价值和沟通等理性意涵的多重理性方向转变。此时,政府所选的第三方评估机构也是看重其专业性,而非“听话照做”的执行力,同时在评估决策和执行等各环节予以信任和赋权,向第三方评估机构释放一定自主空间并尊重其给出的专业评定结果,之后还需依靠第三方评估机构进行专业指导,以期实现专业促建。

另一方面,对于第三方评估机构而言,专业素养与参与策略也同样重要。第三方评估机构应充分掌握所评领域专业知识和所评地域基本情况,这些是获取行政信任的关键。之后才需要第三方评估机构明确的自身角色定位,不应做一个仅为获取行政资源而行动的“代理人”,而应成为一个肩负行政期待和社会责任的“管家”。作为“管家”的第三方评估机构可以积极参与政社协商并主动在理念调和、行动调适和结果调整中将行政要求融入专业实践,使第三方评估既能顺应行政期待,又能彰显专业自主,从而具有较强的实效性。

本文也留下了三方面有待进一步完善与深化之处:其一,对于如何深刻理解介于形式化和实质性之间的第三方评估过渡类型的问题,本文较难予以回答,需要后期引入其他类型第三方评估案例再作阐释。其二,由于对比的第三方评估项目本身有一定差异,使所得研究结论有可能会受到评估项目的重要性、目标设计和实施策略等方面的影响,故而在未来研究中对同一评估项目予以跟进观察,探究其实践差异现象,进而深化对政府连环抉择机制的理解。其三,政府连环抉择机制除了解释第三方评估外,是否可以用于解答其他行政委托现象;换言之,这个机制的应用边界是什么,这就需要更为全面和深入的论证,明确政府连环抉择机制的适用范围。

总之,围绕造成第三方评估实践差异化的政府连环抉择机制的研究仅是起步,后续要继续深挖和延展,当然,对第三方评估研究亦然。期待为提升第三方评估实践质量和推进中国社会治理实践创新发展作出一定的学术贡献。

参考文献

- 陈潭 (2017). 第三方治理:理论范式与实践逻辑. 政治学研究, 1: 90-98+128.
- Chen, T. (2017). Third-party Governance: Theoretical Paradigm and Practical Logic. *CASS Journal of Political Science*, 1: 90-98+128. (in Chinese)
- 陈振明、黄元灿 (2019). 智库专业化建设与公共决策科学化——当代公共政策发展的新趋势及其启示. 公共行政评论, 3: 104-117+192.
- Chen, Z. M., & Huang, Y. C. (2019). Professional Construction of Think-Tanks and Scientificization of Public Decision-Making: New Trends in Contemporary Public Policy and Their Enlightenment. *Journal of Public Administration*, 3: 104-117+192. (in Chinese)
- 程燕林 (2017). 如何保证第三方评估的独立性. 中国科技论坛, 7: 14-18.
- Cheng, Y. L. (2017). How to Ensure Independence of Third-party Evaluation. *Forum on Science and Technology in China*, 7: 14-18. (in Chinese)
- 高丽、徐选国 (2019). 第三方评估组织发展的结构性困境及其生成机制——基于对S市的经验观察. 中国第三

- 部门研究, 2: 98-118+183-184.
- Gao, L., & Xu, X. G. (2019). The Structural Dilemma and Generating Mechanism of the Third-party Evaluation Organization Development: Based on the Empirical Observation Analysis of S City. *China Third Sector Research*, 2: 98-118+183-184. (in Chinese)
- 郭渐强、严明 (2017). 地方政府重大行政决策第三方评估机制研究. 湘潭大学学报(哲学社会科学版), 5: 35-37+70.
- Guo, J. Q., & Yan, M. (2017). The Research of "Third Party Assessment" Mechanism in the Major Administrative Decision-Making of Local Governments. *Journal of Xiangtan University(Philosophy and Social Sciences Edition)*, 5: 35-37+70. (in Chinese)
- 何文盛、姜雅婷 (2020). 政府绩效评估“三权”视角下第三方评估权力组织模式与评估结果偏差——基于多案例的阐释. 江苏行政学院学报, 4: 98-105.
- He, W. S., & Jiang, Y. T. (2020). Organization Model and Deviation in Evaluation Results by Third-party Evaluation from the Perspective of "Three Powers" in Government Performance Evaluation: An Interpretation Based on Multiple Cases. *The Journal of Jiangsu Administration Institute*, 4: 98-105. (in Chinese)
- 黄晓春、周黎安 (2017). 政府治理机制转型与社会组织发展. 中国社会科学, 11: 118-138+206-207.
- Huang, X. C., & Zhou, L. A. (2017). Transition in Governmental Governance Mechanisms and Development of Social Organizations. *Social Sciences in China*, 11: 118-138+206-207. (in Chinese)
- 敬义嘉 (2011). 社会服务中的公共非营利合作关系研究——一个基于地方改革实践的分析. 公共行政评论, 5: 5-25+117.
- Jing, Y. J. (2011). A Case Study of Public-Nonprofit Collaboration in Social Service Delivery. *Journal of Public Administration*, 5: 5-25+117. (in Chinese)
- 敬义嘉、崔杨杨 (2015). 代理还是管家：非营利组织与基层政府的合作关系. 中国第三部门研究, 1: 14-31.
- Jing, Y. J., & Cui, Y. Y. (2015). Agent or Steward: Understand the Collaborative Relationship Between Local Government and Nonprofit Organizations. *China Third Sector Research*, 1: 14-31. (in Chinese)
- 李春、王千 (2014). 政府购买养老服务过程中的第三方评估制度探讨. 中国行政管理, 12: 38-42.
- Li, C., & Wang, Q. (2014). On Third-party Evaluation System in the Process of Outsourcing Government-financed Social Service for the Elderly. *Chinese Public Administration*, 12: 38-42. (in Chinese)
- 李佳源 (2021). 公共部门绩效评估中的博弈行为：分类框架、识别策略与研究进路. 公共行政评论, 3: 176-195+200.
- Li, J. Y. (2021). Gaming Behavior in Public Sector Performance Evaluation: Typology, Identification Strategies and Future Directions. *Journal of Public Administration*, 3: 176-195+200. (in Chinese)
- 李善同、龚璞 (2019). 规划评估：实践与前瞻. 中国行政管理, 8: 14-16.
- Li, S. T., & Gong, P. (2019). Planning Evaluation: Practice and Prospect. *Chinese Public Administration*, 8: 14-16. (in Chinese)
- 李文彬、沈涵、李雅婷 (2019). 客观绩效、信任度与公众满意度——基于调查—实验法的探索. 公共行政评论, 4: 27-43+189-190.
- Li, W. B., Shen, H., & Li, Y. T. (2019). Objective Performance, Trust and Citizen Satisfaction: Based on Exploration of Survey Experiment. *Journal of Public Administration*, 4: 27-43+189-190. (in Chinese)
- 李晓方、孟庆国、王友奎 (2019). 绩效信息公开与政府响应——基于政府门户网站建设第三方评估数据的断点回归分析. 公共行政评论, 5: 115-132+214.
- Li, X. F., Meng, Q. G., & Wang, Y. K. (2019). Performance Information Disclosure and Government Response: A RDD Analysis Based on the Third Party Evaluation Data of Government Portal. *Journal of Public Administration*, 5: 115-132+214. (in Chinese)
- 林鸿潮 (2014). 第三方评估政府法治绩效的优势、难点与实现途径——以对社会矛盾化解和行政纠纷解决的评估为例. 中国政法大学学报, 4: 25-32+157.
- Lin, H. C. (2014). Advantages, Difficulties and Approach of the Third-party Evaluation of Government Under the Rule of Law. *Journal of CUPL*, 4: 25-32+157. (in Chinese)
- 刘帮成、张宗贺 (2017). 基于第三方评估的高校职能部门绩效体系优化研究：一个三年跨度的纵向设计. 公共管理与政策评论, 4: 41-49.
- Liu, B. C., & Zhang, Z. H. (2017). The Third-party Evaluation and Improving University Functional Department Performance System: A Three-Year Longitudinal Survey. *Public Administration and Policy Review*, 4: 41-49. (in Chinese)
- 马佳铮 (2016). 政府绩效第三方评估模式的实践探索与优化路径——以中国(上海)自贸区为例. 上海行政学院学报, 4: 17-25.
- Ma, J. Z. (2016). The Practice Exploration and Optimization Path of the Third Party Evaluation Model of Government Performance: Using Shanghai Free Trade Zone as an Example. *The Journal of Shanghai Administration*

- Institute, 4: 17-25. (in Chinese)
- 马克斯·韦伯 (1997). 经济与社会(上卷). 林荣远, 译. 北京: 商务印书馆.
- Weber, M. (1997). *Economy and Society (The First Volume)*. (Lin, R. Y. Trans.). Beijing: The Commercial Press. (in Chinese)
- 马良灿 (2013). 从形式主义到实质主义: 经济社会关系视域中的范式论战与反思. 北京: 社会科学文献出版社.
- Ma, L. C. (2013). *From Formalism to Substantivism: The Controversy and Reflection of Paradigms in Social-Economic Relationship*. Beijing: Social Sciences Academic Press. (in Chinese)
- 孟志华、李晓冬 (2018). 精准扶贫绩效的第三方评估: 理论溯源、作用机理与优化路径. 当代经济管理, 3: 46-52.
- Meng, Z. H., & Li, X. D. (2018). Third-party Assessment for Precision Poverty Alleviation Performance: Theoretical Origin, Mechanism of Action and Optimization Path. *Contemporary Economic Management*, 3: 46-52. (in Chinese)
- 莫光辉、张菁 (2018). 精准扶贫第三方评估长效机制建构策略——2020年后中国减贫与发展前瞻探索系列研究之一. 苏州大学学报(哲学社会科学版), 6: 113-120.
- Mo, G. H., & Zhang, J. (2018). The Construction Strategy of Long-term Mechanism for Third-party Evaluation of Targeted Poverty Alleviation. *Journal of Soochow University (Philosophy & Social Sciences Edition)*, 6: 113-120. (in Chinese)
- 让·雅克·卢梭 (2005). 社会契约论. 何兆武, 译. 北京: 商务印书馆.
- Rousseau, J. J. (2005). *Discourse on Political Economy and the Social Contract*. (He, Z. W. Trans.). Beijing: The Commercial Press. (in Chinese)
- 尚虎平、孙静 (2020). 失灵与矫正: 我国政府绩效“第三方”评估的效能评估. 学术研究, 7: 50-58+177.
- Shang, H. P., & Sun, J. (2020). Failure and Correction: Effectiveness Evaluation of “Third-party” Evaluation of Government Performance in China. *Academic Research*, 7: 50-58+177. (in Chinese)
- 石国亮 (2012). 慈善组织公信力重塑过程中第三方评估机制研究. 中国行政管理, 9: 64-70.
- Shi, G. L. (2012). Research on the Third-party Assessment Mechanism in the Course of Remodeling the Credibility of Philanthropy. *Chinese Public Administration*, 9: 64-70. (in Chinese)
- 苏力 (1996). 从契约理论到社会契约理论——一种国家学说的知识考古学. 中国社会科学, 3: 79-103.
- Sun, L. (1996). From Contract Theory to Social Contract Theory: A Knowledge Archaeology of State Theory. *Social Sciences in China*, 3: 79-103. (in Chinese)
- W. 理查德·斯科特 (2010). 制度与组织——思想观念与物质利益. 姚伟、王黎芳等, 译. 北京: 中国人民大学出版社.
- Scott, W. R. (2010). *Institutions and Organizations Ideas and Interests*. (Yao, W. & Wang, L. F. Trans.). Beijing: China Renmin University Press. (in Chinese)
- 王程伟、马亮 (2021). 绩效反馈如何影响政府绩效?——问责压力的调节作用. 公共行政评论, 4: 83-104+198.
- Wang, C. W., & Ma, L. (2021). How Does Performance Feedback Promote Performance Improvement? Exploring the Moderating Role of Accountability Pressure. *Journal of Public Administration*, 4: 83-104+198. (in Chinese)
- 王春业、邓盈 (2017). 重要立法事项第三方评估机制研究. 中南大学学报(社会科学版), 6: 27-33.
- Wang, C. Y., & Deng, Y. (2017). Research on the Third Party Evaluation Mechanism of Important Legislative Matters. *Journal of Central South University (Social Sciences Edition)*, 6: 27-33. (in Chinese)
- 王名、刘国翰、何建宇 (2001). 中国社团改革——从政府选择到社会选择. 北京: 社会科学文献出版社.
- Wang, M., Liu, G. H., & He, J. Y. (2001). *China Society Reform: From Governmental Selection to Social Selection*. Beijing: Social Sciences Academic Press. (in Chinese)
- 王诗宗、宋程成 (2013). 独立抑或自主: 中国社会组织特征问题重思. 中国社会科学, 5: 50-66+205.
- Wang, S. Z., & Song, C. C. (2013). Independence or Autonomy: A Reflection on the Characteristics of Chinese Social Organizations. *Social Sciences in China*, 5: 50-66+205. (in Chinese)
- 吴晓林、谢伊云 (2020). 国家主导下的社会创制: 城市基层治理转型的“凭借机制”——以成都市武侯区社区治理改革为例. 中国行政管理, 5: 91-98.
- Wu, X. L., & Xie, Y. Y. (2020). The State Initiates and Regulates Society: “Bridging Mechanism” for the Transformation of Urban Grassroots Governance in China: A Case Study of Community Governance Reform in Wuhou District, Chengdu. *Chinese Public Administration*, 5: 91-98. (in Chinese)
- 谢立中 (2010). 结构-制度分析, 还是过程-事件分析? 北京: 社会科学文献出版社.
- Xie, L. Z. (2010). *Structural-Institutional Analysis or Process-Event Analysis?* Beijing: Social Sciences Academic Press. (in Chinese)
- 辛自强 (2018). 观察中国社会治理的理论视角及其超越. 中州学刊, 5: 68-73.

- Xin, Z. Q. (2018). The Theoretical Perspectives of Observing China's Social Governance and Its Transcendence. *Academic Journal of Zhongzhou*, 5: 68 - 73. (in Chinese)
- 徐家良、张圣 (2020). 中国疫情防控多主体关系的动态诠释. 上海交通大学学报(哲学社会科学版), 5: 47 - 58 + 144.
- Xu, J. L., & Zhang, S. (2020). A Dynamic Interpretation of the Multi-participant Relationship in the COVID-19 Pandemic Prevention and Control in China. *Journal of Shanghai Jiao Tong University(Philosophy and Social Sciences Edition)*, 5: 47 - 58 + 144. (in Chinese)
- 徐双敏、崔丹丹 (2016). 完善社会组织第三方评估工作机制研究——基于5市调查数据的分析. 中南财经政法大学学报, 6: 52 - 57.
- Xu, S. M., & Cui, D. D. (2016). Research on Improving the Third-party Evaluation Mechanism of Social Organizations: Analysis Based on the Survey Data of Five Cities. *Journal of Zhong Nan University of Economics and Law*, 6: 52 - 57. (in Chinese)
- 阎波 (2018). 公共组织第三方评估何以可为? 一项双案例比较研究. 经济社会体制比较, 3: 125 - 135.
- Yan, B. (2018). How Does Third-party Evaluation Work? A Comparative Case Study from China. *Comparative Economic & Social Systems*, 3: 125 - 135. (in Chinese)
- 颜昌武 (2018). 行政国家: 一个基本概念的生成及其蕴涵. 公共行政评论, 3: 120 - 136 + 212.
- Yan, C. W. (2018). The Rise of the Concept of the Administrative State and Its Implication. *Journal of Public Administration*, 3: 120 - 136 + 212. (in Chinese)
- 张圣、徐家良 (2021). 政府慈善赋权何以走向有序? ——探寻渐进之道. 学习与实践, 3: 77 - 88.
- Zhang, S., & Xu, J. L. (2021). How Does the Government's Charity Empowerment Run in Order? Exploring the Way of Gradual Development. *Study and Practice*. 3: 77 - 88. (in Chinese)
- 张玉磊 (2014). 多元主体评估模式: 重大决策社会稳定风险评估机制的发展方向. 上海大学学报(社会科学版), 6: 124 - 132.
- Zhang, Y. L. (2014). Multi-subject Assessment Model: Development of Major Policy Social Stability Risk Assessment Mechanism. *Journal of Shanghai University(Social Sciences Edition)*, 6: 124 - 132. (in Chinese)
- 郑佳斯、卜熙 (2020). 失效的第三方: 组织自利性下的社会组织评估. 华南师范大学学报(社会科学版), 5: 98 - 112 + 191.
- Zheng, J. S., & Bu, X. (2020). Failure of the Third Party: Social Organization Evaluation Motivated by Self-interest. *Journal of South China Normal University(Social Sciences Edition)*, 5: 98 - 112 + 191. (in Chinese)
- 周黎安 (2007). 中国地方官员的晋升锦标赛模式研究. 经济研究, 7: 36 - 50.
- Zhou, L. A. (2007). Governing China's Local Officials: An Analysis of Promotion Tournament Model. *Economic Research Journal*, 7: 36 - 50. (in Chinese)
- 周黎安 (2014). 行政发包制. 社会, 6: 1 - 38.
- Zhou, L. A. (2014). Administrative Subcontract. *Chinese Journal of Sociology*, 6: 1 - 38. (in Chinese)
- 卓越、张红春 (2016). 绩效激励对评估对象绩效信息使用的影响. 公共行政评论, 2: 112 - 133 + 207.
- Zhuo, Y., & Zhang, H. C. (2016). An Empirical Analysis of the Performance Incentive's Effect on Public Manager's Performance Information Use Behavior. *Journal of Public Administration*, 2: 112 - 133 + 207. (in Chinese)
- Carman, J. G., & Fredericks, K. A. (2010). Evaluation Capacity and Nonprofit Organizations: Is the Glass Half-Empty or Half-Full? *American Journal of Evaluation*, 31(1): 84 - 104.
- Claessens, S., Djankov, S., Fan, J. P. H., & Lang, L. H. P. (2002). Disentangling the Incentive and Entrenchment Effects of Large Share Holdings. *Journal of Finance*, 57(6): 2741 - 2771.
- Cribb, J. (2006). Agents or Stewards? Contracting with Voluntary Organizations. *Policy Quarterly*, 2(2): 11 - 17.
- Eisenhardt, K. M. (1989). Agency Theory: An Assessment and Review. *Academy of Management Review*, 14(1): 57 - 74.
- Goyal, N., & Howlett, M. (2019). Combining Internal and External Evaluations Within a Multilevel Evaluation Framework: Computational Text Analysis of Lessons from the Asian Development Bank. *Evaluation*, 25(3): 366 - 380.
- Kluvers, R., & Tippet, J. (2011). An Exploration of Stewardship Theory in a Not-for-Profit Organization. *Accounting Forum*, 35(4): 275 - 284.
- Van Slyke, D. M. (2007). Agents or Stewards: Using Theory to Understand the Government-Nonprofit Social Services Contracting Relationship. *Journal of Public Administration Research and Theory*, 17(2): 157 - 187.
- Wasserman, N. (2006). Stewards, Agents, and the Founder Discount: Executive Compensation in New Ventures. *Academy of Management Journal*, 49(5): 960 - 976.