

评估地方政府层面的制度改革： 一个有待研究的领域

萨宾·库尔曼 赫尔穆特·沃尔曼

陈雄兵 刘 方 译

内容提要:在地方政府层面,对制度改革政策的评估是目前政策评估领域中很少涉及的内容。文章将相关的方法运用到制度改革政策的评估中,从而对该领域的研究做出一些探索。文章将制度方面的政策(如地域性和功能性改革)概括为一类特殊的公共政策,并从概念上和方法上将其与普通的实体政策相比较。文章探讨了评估过程中存在的诸多问题,包括交易成本的确定和找寻制度改革的度量指标等。文章最后提出了关于功能性和地域性改革评估的分析框架,以便区别制度评估和绩效评估。此外,该框架还可以应用于比较性的评估研究中。

关键词:评估 制度方面的政策 地域性和功能性改革 公共行政 政府间体制

中图分类号:D52 **文献标识码:**A **文章编号:**1003-3947(2012)03-0113-09

一、问题的提出

最近 25 年以来,欧洲国家在地方政府间制度环境方面经历了巨大而深刻的变革。具体来看,纵向的改革表现为分权,即中央政府将政治和行政职能下放到地方政府,横向的改革则表现为当地政府的地域性和组织性改革。虽然上述两个方面的改革是近年来政治学领域的热点问题,但学术界关注的重点主要集中在政治上的决策制定以及改革的实施过程,而确认改革影响和效果的评估研究则严重缺乏。与之形成鲜明对比的是,中央政府层面的公共部门改革评估问题则受到了学者和国际组织的广泛关注。

在此背景下,本文迈出关键的一步,旨在从概念上促进对地方政府层面的制度改革政策的评估,并对目前仍然被政策评估研究所忽视的该领域做出一些探索。我们希望构建相关的分析框架,并以此来整合包括单个国家和多个国家在内的比较研究文献。

作者简介:萨宾·库尔曼(Sabine Kuhlmann),德国行政科学大学教授。赫尔穆特·沃尔曼(Hellmut Wollmann),德国洪堡大学教授。**译者:**陈雄兵,经济学博士,中南财经政法大学金融学院教师。刘方,中南财经政法大学金融学院博士生。

原文为“The Evaluation of Institutional Reforms at Sub-national Government Levels: A Still Neglected Research Agenda”,载于《地方政府研究》(Local Government Studies)2011 年,第 37 卷,第 5 期,第 479—494 页。

二、对制度改革政策的评估:定义和概念

为了弄清制度改革政策评估这一概念的特征,我们需要强调制度改革政策和一般政策之间的区别。一般政策是指实际的政策,比如社会政策、就业政策和住房政策等,其基本目的是引起社会与经济的变革。相比之下,制度改革政策的目的则是影响政治和行政结构。政策科学研究领域强调政策(政策制定的内容)、政治(政策制定的过程)和政体(政策制定所在的制度环境)三者之间的区别,以此为基础,本文用政体政策来指代制度方面的政策。

在奠定制度改革政策之基础的模型中,存在三个相互依存、互为因果的阶段。在第一个阶段,制度改革政策借助合适的政治和其他措施来改变相关制度环境。在第二个阶段,制度环境的改变带来了制度绩效的改变。在第三个阶段,制度改革措施将在更广泛的政治、社会、经济和其他层面产生进一步的影响,包括更深远和更系统的影响,以及对政治和民主环境所产生的影响。

相应地,对制度改革政策的评估也包含三个序贯相连的阶段。

第一个阶段考察的是由制度改革措施所推动的制度改革。其问题是:预期的制度变革是否会发生,以及为什么会发生(或没发生)。在此阶段,预期的制度变革被设定为因变量,而影响目标成功和失败的因素被看作自变量(解释变量)。这个评估阶段与经典的关于制度和组织的研究以及其问题——什么塑造了组织——是一致的,它倾向于识别出那些促进或者阻碍制度变化的过程和冲突等因素。就这一点而言,它与由普瑞斯曼和威尔达夫斯基(Pressman and Wildavsky, 1973)所提倡的“执行研究”存在概念上的重叠。

第二个阶段是绩效评估,即在所考察的制度环境中,业已发生的制度改革是否以及为什么会引起制度运行绩效的变化。此时,运行绩效被当作因变量,而制度改革则被当作自变量,这与第一步中制度改革被设定为因变量完全相反。

第三步是结果评估,旨在确认制度改革的预期目标是否达到并分析其原因。此时,结果被当作因变量,而制度变化和绩效变化则被当作自变量。

三个评估阶段的具体内容见图1。

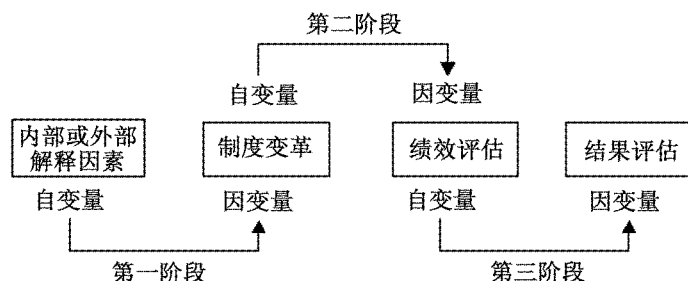


图1:评估的三个阶段

三、评估制度改革的交易成本

正如在制度改革政策的效果中所讨论的那样,目前对于改革措施的交易成本以及成本和收益之间的平衡和权重关系的研究仍然很少,而这正是本文所强调的。本文试图在此领域内展开创新的研究,从而有助于构建综合的评估体系。

交易成本最近频繁出现在政治学和行政学的争论中,其目的在于更好地理解组织活动所产生的物质成本和人力成本以及行政活动的发起、协调和控制成本。基于此,可以通过交易成本分析来为商业计划(谁能最有效地承担职责)提供决策参考,特别是当这个仍然由公共部门承担的职责被外包出去或者由私人部门来承担时更是如此。就本文所讨论的政府间改革而言,交易成本以不同形式出现(见图2)。

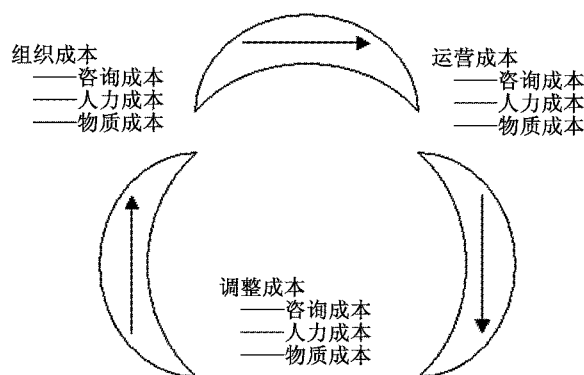


图2:分析制度改革交易成本的三阶段模型

首先,交易成本包括那些由于改革措施的发起、准备和构思而积累的成本,这是实际过程中首先发生的人力和物力开支。为使改革措施到位而配置人力资源、组织资源和财政资源等形成的机会成本也应该被详细记载,因为这些无法再配置到其例行的地方。由此导致的服务减少而产生的额外成本也应被看作机会成本。

其次,常规成本则包括新制度安排的营运成本以及协调、控制和反馈成本等。另外,由于监督、审计、汇报和反馈(在州政府与地方政府等不同级别的机构之间以及同一级别的不同机构之间)等而产生的狭义上的交易成本也应该被确认。

最后,必须识别未预期的结果。可能会产生被称作“休克疗法”的结果,意味着改革疗法最终是有害的而且会带来额外的沉淀成本。机构膨胀和功能重叠所导致的后果与错误构思和糟糕执行的政策所导致的雇员士气低落同样严重。

上述的交易成本分析应该出现在制度改革成本和收益综合平衡分析的第一阶段。只有搞清楚相互之间的平衡关系,我们才能获得关于改革(从旧制度系统到新制度系统的转变)是否成功以及是否值得等相关问题的结论。

四、政策评估的概念和方法论问题

对制度改革政策和措施的评估会碰到比普通政策评估更棘手的概念和方法论问题。

第一,确定改革措施的目标和目的,并以此作为评价改革成败的标准,即所谓的概念问题。

第二,使用既有的方法来确定因果关系,即所谓的因果问题。

就概念问题而言,其困难在于:

1. 很难确定作为测量标尺的目标与目的,尤其是因为制度改革措施往往伴随多重标准,它们之间有时甚至是冲突和矛盾的;
2. 很难确认改革的副作用和未预期的结果,虽然它们很重要;
3. 很难用可操作和度量的指标去判断目标与目的达成与否,对于长期效果(比如结果和系统性效果)而言尤其如此;
4. 很难获得足够稳健的实证数据去度量指标。

就因果问题而言,制度改革政策和措施的评估碰到的难题大致可以概括为如下经验法则:变量多但样本少。因此,依赖大样本、少变量和定量数据的研究方法只能有限地应用到制度改革的评估中,后者往往求诸于案例研究。

众所周知,单一案例研究在方法上存在缺陷(主要是其有限的概括性),而多案例的比较研究则能提供一个方法可行且富有成果的研究框架。在对案例进行有目的的选择后,这个研究框架将接近于“准实验”设计。这可能会导致在特定假说的支持下接受或者拒绝因果关系,这些假说来源于理论推导且被后续的实证分析所证实。因此,广泛的调查研究(比如对所有或者部分的研究对象进行书面的民意调查)有助于更好地选择研究对象,或者有助于确认、修改和拒绝案例研究所获得的结论。

制度改革评估使用来自其他渠道的相关信息和第二手数据来缩小信息差距,这将其与学术研究以及政府、国会和监督改革措施的其他委员会的报告联系起来。正如之前指出的那样,这些研究集中于制度改革在何时、如何以及为何开始并实施等方面,对于评估问题则有所忽视,它们还包含关于改革的收益与成本以及优势与不足的信息,后者来源于理论假设和实证观察。学术研究上的制度改革评估是如此,政府、国会和其他官方机构的报告也是如此,这些报告往往使用来自受委托的研究提供的经验证据。但无论如何,考虑到其内在的有限性、选择性和有偏性,使用这些二手的分析资料时必须小心谨慎。

最后来单独讨论一下关于交易成本评估分析的方法论问题。首先,行政部门很少系统地收集和记录关于改革成本的信息。评估者必须深入公共行政部门的内部(不是说人际关系的内部)去收集尽可能多的相关信息。一方面,这是因为在自身利益未能得到保证的情况下,改革的支持者缺乏激励去公布改革成本;另一方面,除非总体的目标转化成为了具体的计划和措施,否则无法计算改革的成本。这意味着在评估交易成本时,在改革初期偶尔被公布的官方的报告、文件和传单等记录的成本很少与实际发生的成本一致。

另外,成本—收益分析应该以不同的时间环境为基础,这个问题也阻碍了数据的收集。必须基于历史视角去看待评估成本,而且严格来说应该包括从计划的构思到实施的整个阶段,但实际上对于结果的评估主要关注的只是 $t+1$ 这个时间点,因此,这种评估更多的是事后的评估。在同一个经验研究中将上述两个视角联系起来,这对评估者和被评估对象都提出了更高的要求。

五、制度改革评估

(一) 制度改革的纵向和横向路径

在开始本部分的讨论之前,我们应该记住,本文主要是在广泛的公共部门改革范围内关注地方政府的两类制度改革。

纵向的制度改革是指中央政府(通过联邦化、准联邦化以及区域化的方式)将公共职能和职责移交给地方政府,或者是从上级政府层面移交到地方层面。此时,需要注意一个关键的区别,一方面是政治上的分权,即低级别政府借此获得政治自主权;另一方面是行政上的分权,即行政功能从一个行政级别转移到了另一个行政级别。本文主要关注政治上的分权。

横向的地域和组织改革在欧洲国家以两种形式出现。其一,在一些属于北欧模式的国家,地域和机构的合并是由于(历史上成长起来的小规模的)较低级别的市镇和较高级别的县合并所引起的。改革战略是通过扩大市镇(和县)的地域范围,从而扩展和加强其空间和人口统计基础,发挥其作为民主选举的多功能的当地政府的职能。通过合并而开始的第一波地域性改革浪潮出现在1960~1970年代,目的在于提升当地政府的行政和运行能力,从而在全民福利国家的建设和扩张中发挥更加重要的作用。英国和瑞典是这方面的典型(英国较低级别的地方行政当局平均拥有14万人,瑞典则有3.4万人)。最近的大规模合并浪潮以丹麦为代表,由于合并,2007年其地方当局的平均人口规模达到5.5万人。

由于历史原因,南欧国家仍然在很大程度上保留了小规模且高度分散的地方政府结构,但与此同时,也成立了一类新的市政当局间机构去支持单一功能或多功能的市政当局。法国作为这种地域和组织发展模式的典型代表,拥有3.6万个市政当局(平均人口1600人)和超过1.3万个市政当局间机构,由大约1.2万个单目的和多目的的联合体和2600个社区(享有税收自主权)组成。意大利也类似,包括8100个市政当局(平均人口规模为8000人)和大量的市政当局间机构。由于市政当局间机构在民选的当地政府之外运作,其目的在于履行特定(有时是单一)的职能,因此它们从制度上体现了“功能性原则”。

在德国,联邦州有权决定地方政府地域和机构的发展模式,不同州采取的战略有所不同。大部分州偏好于北欧模式而追求地域的合并(最著名的北莱茵—威斯特法伦州的地方当局的人口规模为4.5万人),但与此同时也存在建立地方当局间机构的战略,目的在于从运营和功能上支持人口规模较小的地方当局。

(二) 制度改革评估的分析框架

上面讨论了挑选和确定一定的标准和指标来作为制度改革评估——其三个阶段分别是制度评估、绩效评估和结果评估——的测量标尺,与这种需求(和概念上的挑战)一致,概念上的灵感和指导可以来自于新公共管理评估(NPM—evaluation)的既有研究,也可以来自监管和管制研究,还可以来自更一般的当地政府的比较研究^①。与此同时也应该记

^① 在下文中,我们主要考虑前两个评估环节,即制度评估和绩效评估。结果评估涉及大量的干预因素,在将结果指标(比如下降的失业率)归因于具体的改革措施时会遇到严重的因果关系问题(Kuhlmann et al., 2008),因此,评估的第三个阶段不是本文的重点。

住,制度改革政策——特别是其对于绩效和效果的影响——往往被部分冲突的多重目标所引导。因此,评估的挑战之一在于确认改革目标之间的妥协和权衡关系。另外,评估工作需要制定关于众多复杂的改革措施的平衡表。最后,在平衡表中,改革措施的成本和收益必须对比明显而且相互制衡。此时,考虑到制度改革的交易成本至关重要。

1. 制度评估的效果

为了确定制度改革(纵向的改革是职责的转移,横向的改革是地域和组织的改革)效果的评估标准,我们利用行政学领域中著名的政府组织分类方法来区分多目的的地域性组织和单目的的功能性组织。在理想化的多重目的模型中,地方政府作为政治上负责的全能组织承担了地方层面的所有职责,而且拥有广泛的权利去制定决策和提供服务。根据胡奇和马克斯(Hooghe and Marks, 2001, 2003)的研究,多目的的地方政府模式被称为“第一类型(Type I)”模式,其基本特征如表1所示。

表1:第一类和第二类当地层面参与者安排的特征

特征	第一类	第二类
运行空间的定义	地域	功能
合法性	直接选举(委员会)	非(直接)民主合法性
功能	总体目的/多功能	特定(可能)单一功能
行动定位	追求公共利益	追求特定利益
协调	委员会多数票决制、等级制	协商、折中以及同等地位参与者的交流
理性	政治理性	功能理性
理论框架	政治理论(民主和其他)	功能主义者,经济和其他理论(如公共选择理论)

(1) 当地政府以地域为基础而在固定的当地区域运作,后者构成较低级别的联邦州和单一州的政府间结构^①;

(2) 基于(委员会的)民主选举,当地政府是民主上合法和政治上负责任的政府;

(3) 当地政府有一个总体上的目标(多功能)框架;

(4) 面对众多职责、单目的、特定的利益集团以及利益相关者时,当地政府有权定义、支持和维护公共利益以及当地社区的最大利益;

(5) 多重职责和利益之间的协调最终由等级体制中层级最低的当地委员会以多数票来裁决;

(6) 通过扩展当地委员会的决策范围,地域性改革的目的是将进一步的功能和参与者内部化;

(7) 当地政府基于政治理性而存在;

(8) 多目的政府的理论基础主要是政治理论。

相比之下,在单目的模型中,政治控制和操作执行在功能和制度上是分离的。因此,当地政府的职能横向分散,并且被转移到单目的的公共、私人或者非盈利部门手中,后者

^① 政府间组织的两个基本原则是地域性和功能性,相关的讨论参考(Wollmann, 2004; Wollmann and Bouckaert, 2006)。

在完成这些任务时追求的是单目的而非多功能理性。这种制度安排的特点如下：

- (1) 它的运作空间由特定的功能所决定,因此可能存在重叠;
- (2) 它并非基于直接的民主合法性,而是依赖于参与者之间可能达成的(自愿)协议;
- (3) 参与者追求各自的特殊(自私的)利益;
- (4) 参与者地位对等,他们倾向于通过说服、协商及妥协等方式来平衡不同的利益;
- (5) 总的来说,它是基于功能理性而存在;
- (6) 它的理论基础是理性选择理论,当然还包括其他理论。

再次借鉴胡齐和马克斯(Hooghe and Marks, 2001、2003)的研究,这些当地层面的参与者可被称为第二类型(Type II)。在分析—描述的视角中,由于他们在当地政府之外运作,他们能进一步被当作治理网络和结构^①。

在评估改革对制度安排的影响时,表1中的标准能用作测量标尺。

2. 绩效评估的效果

对于绩效评估而言,存在三个不同的领域,见表2。

表2:行政改革影响分析的标准与指标

绩效标准	可能的指标
产出合理性	
资源、成本和产出	开支、资源(人力、时间、金钱)
	实现的储蓄
	产出
	投入与产出之间的关系
专业质量、目标达成	专业质量、坚持政策标准
	法律诉讼
	接近民众/客户定位/服务质量
	效力/效率、目的客户覆盖率
协调与掌控	
横向与纵向的协调	跨部门协调
	市政当局间的协调
	跨层级协调、摩擦损失
	来自上层的干预
	来自下层的服从/破坏/反对
投入合理性	
民主控制	当地委员会的参与
	民众参与、用户民主
	外在透明性

^① 关于治理的分析—描述—分析观点和规范概念观点之间存在区别,后者关注的是治理的导向和指引功能,相关讨论参考(Marcou, 2006:9; Wollmann, 2006:118)。

- (1)产出合理性的改变:效力、效率,包括改革的交易成本、生产率;
- (2)协调与控制的改变:纵向和横向的改变;
- (3)投入合理性的改变:民主控制、政治责任感、公民参与、透明性。

总体来看,伴随着这三种效果,绩效差距的增加和减少以及跨机构和跨地域的差异可被归纳到影响研究之中,并以此来判断改革(比如职责的下放)在多大程度上影响了不同组织单元之间的差距。

度量绩效影响的分析框架能以下面简化的形式体现出来(Kuhlmann, 2010; Reiter et al., 2010)。

基于多目的模型(第一类)和单目的模型(第二类)的分类,本文提出如下假说。民主控制和横向协调(基于跨政策和地域)的程度可能随着政治上负责任的当地自治政府(多目的/第一类模型)获得决策制定权的程度的增加而增加。但是,由于较低的专业化水平,政策的有效性和单一部门提供服务的质量可能因地域性原则的使用而下降,当地社区之间的不平等程度也将有所增加。反过来也一样,可以假设,处于地方当局之外的单目的模型会减少民主责任、公共行动的透明性以及接近民众的程度。虽然在给定的政策区域内,更大程度的纵向协调有助于增加单一政策的有效性并减少服务供给中的区级异质性,但它也可能削弱横向的跨政策协调,这将损害当地的社区发展和以地域为基础的地方治理。在根据多目的模型而构建的地方政府系统中,财政责任和服务供给在制度上的整合将促使政府提高效率,节约成本,也能鼓励当地全体选民监督公共开支。相反,根据单目的模型(第二类模型)而创建的组织分离了融资和服务功能,而且将公共开支置于直接的民主监督之外,因此会导致政策利益的最大化,后者又会使得单目的组织的规模远远超过正常履行其功能所必须的规模。这一趋势不仅增加了开支和部门的分散性,而且削弱了横向的协调性,最终有损于总体的制度效率。接下来作者将从三个国家比较分析的角度探求当地的机构和制度改革对相关绩效参数的影响。

六、未来的制度改革评估和地方政府研究的展望

本文旨在构建必要的分析框架去评估政府间背景下的制度政策,从而迈出该研究领域的第一步。与实体政策评估领域相比,制度政策评估领域的理论和方法发展缓慢且严重滞后。因此,制度政策评估领域应该大力借鉴实体政策评估领域在过去几十年中不断完善的专业化基础和概念基础。

为进一步阐明制度政策评估的研究方向,应该注意以下几点。首先,本文所讨论的横向和纵向的改革都需要更多的比较性评估。另外,应该扩大被比较的国家样本并在国际合作的背景下进行国别研究。而且,为了最大限度地发挥国际比较研究的分析潜能,国别研究应该由共同的理论框架所指引。

就指导制度评估的理论与概念而言,新制度主义方法似乎颇具成效。他们认为,制度能够影响、塑造和约束而非决定参与者的行为和产出绩效,这一观点引发了如下关键的评估问题:哪一个变量能够解释制度政策和绩效之间的实际关系。特别地,国别特征(政治、制度、历史、文化等方面)、政策因素和参与者因素等可能会影响到制度改革的效力。

这些因素通过不同方式影响制度政策和绩效之间的因果关系,因此必须对其解释效果进行实证检验。新制度主义者的理论关注的是制度改革和绩效的不同因素(历史、文化、参与者特征和政策相关等方面),因此它们能够从理论上指导上述检验过程。

下一步的研究方向是确认欧洲地方行政的不同形态和结构,并找出其典型代表,然后比较其绩效。多层次系统内行政结构的评估对于比较研究而言似乎颇具价值,它有助于回答下面的问题:对特定的目标和公共任务而言,哪一个行政解决方案更合适?在评估设计的制定过程中,需要特别注意平衡制度改革的不同绩效标准,因为它们往往存在冲突和权衡取舍关系。

The Evaluation of Institutional Reforms at Sub-national Government Levels: A Still Neglected Research Agenda

Sabine Kuhlmann & Hellmut Wollmann

(German University of Administrative Sciences, Speyer, Germany; Humboldt University, Berlin, Germany)

Abstract: This article is meant to take a noteworthy step towards conceptually promoting the evaluation of institutional reform policies in the sub-national space. It aims to apply pertinent evaluation approaches to the field of institutional reform policies in the intergovernmental setting and thus to contribute to a research field that arguably has so far been a 'missing link' in policy evaluation. The authors conceptualize institutional policies (territorial and functional reforms) as a particular type of public policy and contrast them analytically, conceptually and methodologically with 'normal' substantial policies. They reveal particular problems of measurement and of finding relevant indicators to evaluate the results of institutional reforms, one of which is the assessment of the transaction costs of reforms. Finally, an analytical framework for the evaluation of functional and territorial reform policies is suggested that makes a distinction between 'institution evaluation' and 'performance evaluation', and that can be applied in comparative evaluation studies.

Key words: Evaluation; Institutional Policy; Territorial/Functional Reform; Public Administration; Intergovernmental System

(责任编辑:刘承礼)