

标尺竞争：政府管制与购买的激励效应

王 哲 顾 昕*

【摘要】在西方，发轫于1980年代的标尺竞争理论，现已逐渐成为公共管理（尤其是政府管制和政府购买）研究中的一个重要学术分支。作为一种重要的激励性政府管制工具，标尺竞争扩展了传统的个体管制思路，基于被管制企业与其同行之间的成本对比情况，来决定政府对企业的定价（或购买），从而人为地在那些存在垄断性、信息不对称性、强外部性等市场失灵的领域引入了竞争，极大地改善了委托代理关系中的激励机制，提升了政府公共管理的绩效。论文系统考察了标尺竞争理论的产生、发展及其应用条件，着重梳理了近年来该理论在四个方面的新进展，包括如何克服异质性、合谋的影响、服务质量下降和动态效率变化情况。

【关键词】标尺竞争 激励 管制 信息不对称 政府购买 公共服务

【中图分类号】D63

【文献标识码】A

【文章编号】1674-2486(2015)06-0009-16

一、导言：标尺竞争与公共服务中的激励机制

由于自然垄断、强外部性、信息不对称等原因，很多物品和服务（以下仅提及服务）的提供受到政府的管制，另有一些服务则需要政府购买，或两者兼而有之。提供这些服务的行业，例如公用事业（包括水、电、气的生产和配送）、交通运输、医疗服务（包括药品）、教育、社会服务等，往往事关国计民生，影响面广。这些行业常被笼统地称为公共服务。这些服务的供方既包括营利性组织（即企业），也包括民办的非营利性和公立组织。长期以来，政

* 王哲，北京大学政府管理学院，博士研究生；通讯作者：顾昕（guxin@pku.edu.cn），北京大学政府管理学院，教授；北京大学国家治理协同创新中心，高级研究员。感谢匿名评审人的意见。

基金项目：北京大学、复旦大学、吉林大学、中山大学、财政部财政科学研究所国家治理协同创新中心和国家社科基金重大项目“中国特色现代社会福利制度框架设计研究”（15ZDA050）。

府面临的难题是：如何为这类服务定价（即政府仅定价格，并不限定服务量），或如何购买（即政府不仅定价，而且也确定购买量），以及在定价或购买基础上如何监管服务质量。一方面，政府必须从社会福利最大化角度出发，争取要么限制甚至降低相关服务的价格，要么限定其年度总支付额；另一方面，政府还必须保证服务提供方有内在的动力提升服务的质量，或至少不降低质量，因此激励机制的设计至关重要。此外，公共服务由于多重难以避免因素本身就存在着激励扭曲（要么不足、要么过度）的现象（周碧华，2015），政府管制和购买如何避免进一步加剧激励扭曲，更是棘手。

传统的施政方法往往是基于成本加成法来确定价格或支付额（限于篇幅，以下在多数情况下仅论及价格，但其论述对购买也适用），但这种方法的问题重重。首先，政府需要对相关服务提供的成本信息进行实时监控，但监控成本有可能高昂；其次，监控者何以长期保持有效监控力度的积极性，也就是监控的激励机制，也成问题；再次，政府与被监控者之间存在着严重的信息不对称；最后，既然政府定价或购买基于成本，那么供方自然就会丧失降低成本的积极性，因而有可能导致服务费用不断上涨，而质量却不一定提升，甚至有可能下降。这类问题在医疗、教育和社会服务等社会福利领域尤为突出，因为这些服务的品质往往具有低度量性的特征。

这样的问题在世界各国普遍存在，中国也不例外。例如，中国政府对绝大多数医疗服务和药品实施政府定价以及（通过公立医保机构）实行部分的政府购买，价格倒是固定的，而且也不高，但中国医疗行业中过度医疗现象普遍存在，导致医疗费用每年都节节上升，百姓叫苦不迭，政府头痛不已。实际上，针对这类问题，新的政府管制和政府购买方法在发达国家不断涌现，标尺竞争（Yardstick Competition）就是其中的重要代表之一。

标尺竞争一开始是企业理论这一大研究领域的一个小分支。众所周知，企业理论的核心是委托代理关系，即委托人如何设计合约，激励代理人（如经理）的行动符合委托人的利益，而对代理人的绩效考评是激励合约的重要内容之一（Baiman & Demski, 1980）。尤其是当企业规模较大时，委托人面对多个代理人，最优激励机制的设计更是棘手。标尺竞争模型就是针对这种情形设计出来的一种激励机制。

这种涉及到多代理人的委托代理关系，当然不止存在于企业中，而且还存在于许多经济社会政治领域。因此，如2014年诺贝尔经济学奖获得者让·梯若尔（Jean Tirole）（Tirole, 1988: 42）在其1988年出版的名著《产业组织理论》（*The Theory of Industrial Organization*）中所说：

标尺竞争的潜在应用是很多的。面对相似成本或需求条件的部门经理可以由 CEO 们进行比较。类似地，对一个公司经理的奖励可以根据竞争公司经理的绩效做出。更一般地，管理报酬可以基于行业的平均利润。国防部和许多私人公司有时会使用两个外包商来供应零部件，即便在规模收益上有所损失也在所不惜。美国老年人医疗保险（Medicare）对同一个疾病诊断组中的所有病人向医院支付固定的费用。这笔费用的大小基于可比较医院同组病人治疗的平均费用。

事实上，自 1985 年以来，西方学术界就开始将标尺竞争理论的适用范围，从企业拓展到政府对各种公共服务的定价、购买与管制。现在，标尺竞争理论中的委托人不仅仅是股东或总经理，也包括政府机构（特别是负责政府管制和购买的政府部门），理论中的代理人已经不限于企业，也包括非营利组织（如慈善医院）（Newhouse, 1996）、公立组织（如公立医院）（Bloom et al., 2015），甚至包括地方政府本身（Besley & Case, 1995; Revelli, 2006; Bordonon et al., 2004; Hatfield & Kosec, 2013）。标尺竞争理论从组织经济学和管制经济学的狭窄视界进入了从公共经济学到公共管理学的宽广领域，成为福利国家治理变革的重要理论基础之一。

限于篇幅，本文仅对政府管制（及购买）中的标尺竞争理论，进行综述，以期从理论溯源、演变和发展的历史中，探寻未来可能的学术突破空间及其在中国的应用场景。对于方兴未艾的地方政府之间的标尺竞争问题，即所谓“政治标尺竞争”，我们已经另撰文加以详述。同时，限于篇幅，我们对中文学术期刊上相关文献，暂且不加评述，这方面的综述有待我们在将标尺竞争理论应用于中国具体问题的研究时再另行开展。

二、标尺竞争理论的产生与发展

关于标尺竞争的研究发轫于 1980 年代，其思想最初源于一些学者对企业委托代理关系中相对绩效考评（Relative Performance Evaluation）所进行的研究（Lazear & Rosen, 1981）。委托代理关系既存在于股东与经理之间，也存在于管理者与员工之间。以股东和经理人的关系为例，如果企业规模不大，委托人（股东）仅对单一代理人（管理层）的绩效进行考评，这被称为“绝对绩效考评”；如果企业规模较大，常常会出现多代理人的情形，例如集团公司中的总公司作为委托人需要管理若干分公司经理，此时委托人一般采用“相对绩效考

评”，即将承担类似职责的诸多代理人进行比较，按照比较的结果逐一对各代理人进行奖惩（Mookherjee, 1984）。

相对绩效考评的激励办法可以过滤掉代理人共同面临的外部不确定性和震荡因素，使得代理人的报偿基于他们自身可控制的因素，从而激励代理人提升自己的努力程度（Nalebuff & Stiglitz, 1983），从而不断提升企业的业绩。此外，对于委托人来说，实施相对绩效考评只需要收集相对信息，这既可以极大地节约信息搜集和处理成本，也可以让代理人绩效信息显示真实与否的重要性降低，从而克服信息不对称带来的麻烦；尤其是在那些个体绝对绩效的好坏难以度量的环境中（Holmstrom, 1982），相对绩效考评的优势将更为明显（Gibbons & Murphy, 1990）。

1985年，著名经济学家、哈佛大学经济学系教授安德列·施莱弗（Andrei Shleifer）将企业理论中相对绩效考评的思想应用于政府管制研究（Shleifer, 1985），从而一方面正式提出了标尺竞争理论，另一方面开启了这一理论从管制经济学向其他领域的拓展之路。

在管制经济学领域开展这一研究的背景是，在世界各地，以服务成本管制（Cost-of-service Regulation）为代表的传统方法存在着严重激励不当的问题。在这种旧管制方法下，政府主要基于供方上报的平均成本，并参考一段时期的历史运营成本，同时按照预先设定的投资回报率和估计的投资折旧率，设定被管制服务的单位价格或购买总量。由于在一定时期内投资回报率是固定的，因此这种管制方法又称“收益率管制”（拉丰、梯若尔，2014：11）。如果提供者真实上报其平均成本，或者政府部门通过比较低的成本可以甄别出平均成本，那么这种管制方案就是有效的。但是，由于信息不对称，管制者可能很难甚至根本无法了解供方的真实成本信息，也就无法真正实现“科学的”定价。更为严重的是，在这种管制规则下，供方根本没有动力降低成本，甚至缺乏动力显示真实的成本，反而会有提升成本或虚报成本的积极性，并据此不断地游说政府提高价格。由于供方拥有信息优势，因此最终的价格或补贴总能让供方弥补其成本，供方仅凭信息优势就可以获得超额租金，这将导致民众福利水平的降低（史普博，1999）。简言之，由于政府与服务提供方之间存在着信息不对称，以成本为基础的政府管制和购买，无论如何改善，都有可能产生诸多不良的后果。

基于这种情况，施莱弗提出，管制者没有必要殚精竭虑去设法调查被管制者的成本信息，而是可用同行业其他类似企业的信息来模拟该企业的成本信息，从而提升管制者的信息水平，降低企业的信息租金。他假设在一个博弈周期内

该行业存在若干完全相同的企业，同时假设企业经理风险中性，管制者对代理人进行支付是没有成本的，企业之间没有合谋，管制者没有被俘获，此时，如果对该企业产品的定价不是基于其自身成本，而是基于其他企业的平均成本，则企业将被迫与其“影子企业”竞争。一旦企业的实际成本高于其“影子企业”的平均成本，企业将亏损。施莱弗通过数理模型证明，这种定价方案将迫使那些成本较高的企业努力降低成本至平均水平；而那些运营成本本来就低于平均水平的企业，将会有超额利润。只要标尺竞争型管制维持着，这些企业将至少保持原有的成本水平，甚至还有可能降低，否则会导致企业利润降低。

通过这种标尺竞争管制，管制者不必知道企业具体的生产和成本条件，通过“人为地”引入竞争，可以降低企业的信息租金，并促使落后企业降低成本。同时，在面临共同的外部环境不确定性时，标尺竞争可以过滤掉企业共同面临的不确定性环境，将企业获得的奖惩建立在自身可控的因素上，从而提升企业努力的激励水平。很显然，这种人造竞争的思路对于自然垄断型行业的管制，是至关重要的。值得注意的是，施莱弗的经典性论文是针对公用事业行业写的，但其思路显然可以拓展到诸多公共服务领域，因此在下文，被管制对象既可以是企业，也可以是非营利组织，无论是公立的还是私立的。在下文中，当我们使用“企业”一词时，仅仅是为了行文方便，实际上有关讨论完全可以适用于非企业组织。

施莱弗的理论模型建立在一系列抽象的假设上，而这些假设在现实中是很难满足的。值得一提的是，理论模型中的假设具有抽象性并且在现实世界中几乎不可能得到满足，这一点往往被很多不了解理论特质的人视为研究不完善的表现，甚至贬斥为“黑板经济学”；但这恰恰是这一模型具有理论基础性的表现，因为唯有这样的基础模型才能为后续研究奠定逻辑分析的基础，而后续研究往往都建立在基础模型中抽象假设的放松或现实化之上。

施莱弗标尺竞争基础模型中两个最重要的假设大有放松的余地。（1）同质性企业（Identical Firms）假设在现实世界中找不到实例：即便是处于相同环境中的企业，也很难保证生产条件的相同，遑论现实中不同企业所处的环境很可能大相径庭；（2）无合谋假设也可放松，事实上一些激励方案本身就引致企业间合谋出现（Potters et al., 2004）。一旦合谋发生，标尺竞争方案将很难达成合意的结果（Tangerås, 2002）。

此外，标尺竞争的实施虽然将促进成本水平的降低，但也可能导致其他非合意结果的出现，这意味着在决定是否引入标尺竞争时，管制者必须在不同目标间进行权衡。例如，标尺竞争在促使成本控制的同时很可能导致质量的下降

(Tangerås, 2009)；或者说，这一管制导致供方提供了并不相同的东西。在某些质量攸关且质量难以度量的领域，例如医疗和教育，质量的降低意味着管制的失败 (Bloom et al., 2015)。另外，在实施标尺竞争后，由于供方通过信息优势获得的租金减少了，供方很可能会选择降低研发的投入 (Sobel, 1999)，导致创新不足，这也是管制者所不愿意看到的。

三、标尺竞争理论的主要争论

如上所述，对施莱弗模型中假设条件的放松及其可能结果的继续探讨，以及对标尺竞争可能带来的非合意结果的分析，就构成了标尺竞争理论后续争论的主题。从现有文献来看，这些探讨与争论集中于以下四个主题：

(一) 如何克服异质性

现实中受管制的企业并不是完全相同的。要实施标尺竞争，管制者就必须考虑到企业之间的异质性，将企业成本变为真实可比的，其目的是过滤掉企业所面临的共同外部不确定性。施莱弗的论文实际上已经考虑到这一点，并提出可以通过计量回归方法，控制住外生异质性变量。很多学者对施莱弗提出的思路进行了深入研究，并总结出在实践中通过计量方法控制异质性的一整套技术 (Coelli et al., 2005; Burns et al., 2005)。一些公司 (例如 Frontier Economics) 运用先进的工程学方法，已经开发出了一系列模型，来确定到底加入何种控制变量，以及在不同环境下究竟采用哪种计量方法 (Frontier Economics, 2010)。

我们用一个案例来说明如何使用计量手段克服异质性。20 世纪末，加拿大安大略省的管制机构决定对其管辖下的 81 处配电站实施标尺竞争管制 (Yatchew, 2001)。但难题是，分处不同区域的配电站彼此间差异巨大，有的位于城市，有的位于农村，还有的位于山区；此外，不同电站覆盖人数也不同，最小的电站仅覆盖 600 人，最大的则要为近 22 万人供电。不同电站所处环境的差异意味着施莱弗模型的同质性假设无法成立，简单采用其他企业的平均成本进行定价将丧失科学性。

对此，管制机构决定采取计量方法对企业真实成本进行估算：第一步，根据行业经验确定那些真正影响或决定配电成本的关键变量和外生变量，包括用户数量 (cust)、每公里电线成本 (pcap)、是不是公共事业委员会旗下配电所 (PUC)、人均每小时千瓦时 (kWh)、设备平均剩余寿命 (life)、负荷量 (If)、人均电线长度 (wire)，等等。

第二步，基于上述变量，管制者建立起如下 OLS 计量模型（ tc 代表总成本）：

$$tc = f(cust) + \beta_1 wage + \beta_2 pcap + \beta_3 PUC + \beta_4 kWh + \beta_5 life + \beta_6 lf + \beta_7 wire + \varepsilon$$

第三步，将所管辖的 81 处配电站的各项数据带入模型，从而得到各系数的估计值，保留那些系数显著的变量，删去不显著的变量，继而得到最终模型。

第四步，将每个配电站的各项数据带入方程，即可得到各自的预期成本 tc' 。管制者即以 tc' 为主要依据，再适当考虑其他因素，为各个企业定价。

相比简单地采用其他企业的平均成本定价，使用计量模型估算的定价方式无疑更加科学，大大提升了标尺竞争的应用性和准确性。

（二）标尺竞争中的合谋

企业理论的研究已经表明，相对绩效考核可能导致代理人之间的合谋（Gibbons & Murphy, 1990）。作为相对绩效考核的一种，标尺竞争也有此问题。施莱弗同样意识到了这一问题，他写道，“标尺竞争中一个重要的可能的限制就是合谋”（Shleifer, 1985: 327）。通过合谋，企业可以虚报成本以使管制者提高标尺价格，从而获取额外利润，也可以骗取来自管制者的补贴却不用投入足量资本以降低成本（Sobel, 1999）。有学者证明，当企业的生产率主要取决于共同知识（Common Knowledge）时，即由行业内生产率而不是企业独有生产率来决定，合谋的诸多企业会变为事实上的一体，这使得标尺竞争与个体管制（Individual Regulation）变得没有差别。此时，标尺竞争与个体管制相比的好处也就丧失殆尽了（Tangerås, 2002）。

既然合谋的结果是管制者不愿看到的，那么合谋的可能性又是多大呢？

依照现有文献的挖掘，有四点因素会影响企业间合谋的可能性：

1. 惩罚力度：施莱弗提出，管制者可以对可被证实的合谋行为施以惩罚，从而降低合谋的可能性（Shleifer, 1985）；当然，在这里，可证实性（Verifiability）特别是可以被法庭或第三方证实是问题的关键，否则无法使惩罚具有真正的可置信性；此外，惩罚的实施也并非无成本。

2. 企业数量：如果企业数量比较多，合谋的可能性就比较低（Burns et al., 2005；拉丰、梯若尔, 2014），由此，管制者应该对可能发生的企业兼并行为加以限制（Sawkins, 1995）；

3. 企业间合约：企业间的合谋契约本身也会影响合谋的可能性。其中第一个关键点是契约是否有约束力，如果契约没有约束力，合谋将自然失效（Tangerås, 2002）。

4. 企业间行为的可观测性问题：如果企业之间的行为容易被其他企业观测到，这将降低合谋的成本，其中，企业所处的地理位置是一个重要影响因素（Gibbons & Murphy, 1990）。

另外，由于标尺竞争环境下，被管制企业之间的信息存在比较高的相关性，因此只要有一家企业有激励不参与合谋，就很可能导致合谋被管制者识破（Sobel, 1999）。管制者也可以通过设计特定的激励机制使得其中的部分企业拒绝合谋，例如，智利政府曾就电力供应在每个区选出一个模范公司作为效率参考参与标尺竞争，这可能会限制该领域的合谋行为（Rudnick & Donoso, 2000）。

（三）质量下降问题

标尺竞争虽然促进了企业成本的下降，却也可能导致副作用的出现，其中最为引人关注的就是质量的下降。实施标尺竞争后，作为代理人的被管制者面对不断减少的私人信息租金，会通过降低质量的方式来间接降低成本。在某些质量至上的领域，例如医疗和教育，这一现象的出现意味着管制的失败。很显然，如果质量因素很容易被管制者所度量，那么标尺竞争管制将会取得良好的效果，至少不会有质量下降的副作用。质量下降的原因在于，被管制者所提供的产品或服务的质量很难度量，或者由于其质量要在很长一段时期后才能显现出来，致使管制者在管制期内根本无法测量，这在医疗和教育领域尤甚。

应对这一棘手问题的方法有二：（1）构建有效的质量指数（Index of Quality），这种指数可以基于消费者的评价，即所谓“基于消费者的质量指数”（Consumer-based Index of Quality），也可以基于市场份额，即所谓“基于市场的质量指数”（Market-based Index of Quality），但这两种方法也都面临着前文所述的可证实性问题（Laffont & Tirole, 1991；Dalen, 1997）；（2）弱化对可度量性指标（例如成本）的激励，但这种做法会导致标尺竞争激励效应的下降。总而言之，两难困境难以避免。

第一种思路很容易流行。在中国，我们经常会看到政府通过构建复杂的绩效指标自上而下实施绩效管理，也经常听到强化绩效管理的呼吁，而消费者满意度是经常被提及的指标之一。有些行业的质量指标相对容易测量，例如电力配送和自来水等行业，这样可以采用第一种方法，构建起质量指数，消费者的评价也足以管制者提供所需的信息。但在另外一些行业，如医疗和教育领域：一来服务质量的影响因素较多，而且很多因素与服务提供者无关（即所谓“外生变量”）；二来质量监测成本高昂；三来由于信息不对称、质量呈现的跨期性等因素，消费者究竟在何种意义或程度上作为服务质量的合适评价者，这本身

也成问题（莱恩等，2012：91-114）。

面对低度量性问题，管制者此时的应对是第二种思路，即对容易度量的绩效（如成本之类）采用弱激励。弱激励型应对之策的理论依据是多任务委托代理模型的研究成果。正如霍尔斯特罗姆（Bengt Holmstrom）与米尔格罗姆（Paul Milgrom）对多任务委托代理关系的经典性研究所揭示的，当代理人面临多个任务且各任务完成绩效之间存在可度量性差异时，代理人可能将主要努力都花到绩效容易度量的任务上，而减少甚至忽略在其他任务上的努力。例如在医疗领域，作为代理人的公立医院有可能热衷于创收，因为经济绩效指标可度量性高，而政府期望最大化的公益性指标可度量性低；在研究型大学，教师们倾向于投入更多的精力于科研而不是教学，因为发表论文的数量（影响因子加权后的论文数）和质量（如影响因子加权后引证率）相较于教学质量，更具有可度量性；在义务教育中，教师们会更多地倾注于应试教育而不是素质教育；等等。在这种情况下，委托人的理性选择是对高度可度量的任务（如经济性任务），采取弱激励（Holmstrom & Milgrom, 1991）的治理模式。当然，弱激励之策极有可能是无奈之举，而且终将成为无效之举，因为如果委托人对高度可度量任务采取“大锅饭”式弱激励的绩效评估，那么代理人就会在完成此任务上采取“干多干少都一样”的对策；与此同时，对于难以度量的其他任务，反正委托人难以评估绩效，于是代理人依然没有动力在这类任务上费心尽力。弱激励的最后结果极有可能是，代理人的无论哪一种任务的完成情况，都不能符合委托人的利益。

总体来说，关于引入竞争是否会促进质量提升，学界有不同的意见。一些学者认为，竞争强度的提高必然意味着引入了质量上的竞争，因此可以间接提升服务质量。例如在医疗领域，通过提高患者对供方的选择权、降低交易（转换）成本（例如强制性病例共享之类）、增加医院数量等方式提升供方间的竞争强度，会间接促进医疗服务质量的提升（Bloom et al., 2015; Ma & Burgess, 1993; Brekke et al., 2008）；但也有学者认为，竞争强度提高会导致“激励扭曲”问题，即供方在那些容易度量的指标上展开激烈竞争，而在那些不易度量但却十分重要的指标上缺乏竞争，因此加强竞争的效果基本上是换汤不换药（Werner et al., 2009; Feng Lu, 2012; Dalen, 1997）。这些情况既适用于一般性竞争，也适用于标尺竞争，因此这可以解释为何现实中标尺竞争的应用没有想象的那么简单，尤其是在医疗和教育领域（Tangerås, 2009; Dalen, 1997; Newhouse, 1996）。

（四）标尺竞争与动态效率

除质量下降外，标尺竞争导致的另一个副作用是可能出现的动态效率变化。

效率分静态效率（Static Efficiency）和动态效率（Dynamic Efficiency），后者涉及技术进步和相应的研发投入。如前所述，如果条件适宜，标尺竞争可以提升企业的静态效率，即促使其设法降低成本。而在动态效率方面，标尺竞争的作用如何呢？

一般来说，标尺竞争将促使企业设法降低成本，这在一定程度上会促使企业在有助于降低成本的技术研发上进行投资，但问题在于技术研发投资本身也是成本。随着管制者对成本降低信息的掌握，企业进行技术研发所能带来的好处将呈现递减之势，从而最终降低企业投资于技术研发的积极性（Sobel, 1999）。然而，情况或许并非如此绝对，因为技术与技术之间的差异有可能会让企业采取不同的对策。有学者在将技术进一步区分为行业专属性技术（Industry-specific Technology）和企业专属性技术（Firm-specific Technology）后发现，由于溢出效应的存在，标尺竞争将导致企业不愿意投资于前一种技术，而更愿意投资于后者，这是因为标尺竞争通过过滤掉同行企业共同面对的技术震荡所带来的不确定性，因而可以增大激励的强度，使得企业特有技术的价值变得更高，激励企业投资于此（Dalen, 1998）。认识到这一点，政府可以通过其他方式来激励企业投资于行业专属性技术创新，而不必因标尺竞争对此无促进作用而因噎废食（Burns et al., 2005）。

标尺竞争对公共服务创新可能会产生的抑制作用，或许不大会产生，因为在诸如医疗、教育、社会服务等领域，创新一般体现在服务模式和组织文化中，具有组织专属性特征，类似于上文所说的企业专属性。

三、标尺竞争的适用条件

标尺竞争通过降低信息不对称对管制者的重要性，可以在一定程度上提升管制效果，因而被认为是一种重要的管制工具。但是，标尺竞争的有效使用，必须满足一些基本条件。

第一，市场失灵条件（Market Failure Requirement）：只有存在市场失灵，引入标尺竞争这种“人为竞争”（Artificial Competition）的方式，才有意义。否则，市场竞争本身就能解决问题。因此，标尺竞争适用于自然垄断、法定垄断和严重信息不对称的行业（Canoy et al., 2000）。

第二，可比性条件（The Comparability Requirement）：必须存在多个受管制的代理人，而且代理人之间（通过一定的技术处理后）同质可比。这意味着代理人之间的个体特异值不能超过共同特征值（拉丰、梯若尔，2014）。比如说，

对于公共旅游胜地或自然保护区管理的政府监管，就很难采用标尺竞争，因为被监管对象真是各有千秋。

第三，私人信息条件（The Private Information Requirement）：在管制中存在一定数量的私人信息，包括真实成本、技术、努力程度等，代理人可以利用这些私人信息获取租金。与此同时，这些私人信息之间又是相关的，这意味着一旦有代理人利用这一信息从中获利（例如通过研发新技术降低成本），将直接反映到会计报表中的成本一栏中，管制者甚至不需要了解具体技术信息，就可以在某种程度上发现行业内的技术进步，从而在下一期加大标尺难度，促进代理人效率的提升。

第四，可证实条件（The Verifiability Requirement）：这一条件意味着管制中的相关信息不仅仅可以被管制者掌握，最好还应该可以被观察和可证实，因而可以作为对代理人实施处罚或奖励的（法律）证据。

从上述条件不难发现，标尺竞争特别适合于对区域垄断行业进行管制，例如电力配送、自来水供应、市内公共交通、公立医院、公立学校等行业。这类产品或服务的提供者在所属区域内是垄断的，因而缺乏竞争，此时引入标尺竞争有利于政府监管，也能促进其提升效率。在这些行业中，相比较而言，标尺竞争又更加适用于公用事业和公共交通，因为相关领域不同区域的企业之间尽管存在差异，但多多少少具有很高的类似性，这意味着可比性条件的满足。与此同时，在这些领域，即使企业有一些相关的私人信息，但由于存在信息披露制度，企业的成本信息在一定程度上也是可以被管制者所获取并可证实的。对于医疗和教育领域，跨地区同行的差异性要比公用事业和公共交通领域的差异性大，但也并没有大到足以使标尺竞争完全失效的地步。更何况，在医疗和教育行业的某些亚部类，例如医疗卫生领域的精神病院、职业病院和不少特定的专科医院以及教育领域中的中小学、职业教育等，跨地区服务提供者的相似性还是很高的，其运行成本信息的可获得性也是相对较高的，因此完全适用于标尺竞争。

当然，现实中的标尺竞争管制可以采用多种形式，既可以采用上文陈述的标尺竞争的形式，也可以使之与其他管制方法结合起来；既可以在定价环节采用标尺竞争的形式，按照其他企业的成本确定价格，也可以在预算环节采用标尺竞争，按照竞争结果划拨预算。现实中采用的标尺竞争管制计划远比理论上的复杂，“大多数激励计划都是多种模型的综合”（Jamash & Pollitt, 2001: 108）。此外，除去经济因素考虑外，政治因素很可能也是管制机制设计中的重要考量。

在公共管理的实践中，标尺竞争理论的应用如果未充分注意到上述的诸多适用条件，哪怕是其中之一，也会造成无效的结果。例如，英国在1980年代起开始在其医疗领域引入“内部市场制”，以图通过在公立医院之间引入人为的竞争而提升其效率，并在1990年代以单位成本为标尺，在国民健康服务（NHS）体系内部实施成本控制绩效评估。然而，这一标尺竞争的试验并没有取得预期的效果，原因有二：一是成本本来就具有强烈私人信息性质，而标尺竞争理论的初衷就是让管制者或绩效评估者规避对私人信息的监控，因此以单位成本为标尺进行标尺竞争，违背了标尺竞争理论自身的原理，也不满足于上述的“私人信息条件”；二是公立医院的相对单位成本受到多种因素的影响，可证实性较弱（Dawson et al., 2001）。

实际上，在中国也有类似情况，即卫生行政部门对下属公立医院进行绩效考核，考核指标包括诊疗次均费用和住院次均费用（某种形式的单位成本），以期公立医院能控制医疗费用的上涨，但由于行政管理者对公立医院如何达标的各种信息不可能充分了解，最后的结果是，在单位成本达标的情况下，公立医院发生的医疗总费用却节节上升。但可惜的是，由于对标尺竞争等基础理论研究的缺乏，有关中国医疗领域绩效管理的研究也缺乏学术深度。迄今为止，我们尚未看到有关中国公立医院绩效管理的中文文献达到相应国际文献的学术水准。

四、标尺竞争理论在公共管理中的应用及其对中国的启示

1994年，激励理论的领军人物之一、法国图卢兹学派领袖让-雅克·拉丰（Jean-Jacques Laffont）曾抱怨标尺竞争理论诞生了近十年却“依然在很大程度上被忽略了”（Laffont, 1994: 524）。此后又经历了20多年的发展，这一理论本身的完善及其应用已经取得了相当数量的成果。然而，标尺竞争理论发展及其应用依然还有很大的空间。我们认为，至少在三个方面，学者们还可以继续努力。

第一，围绕施莱弗模型的其他假设，包括单博弈周期、代理人风险中性、政府购买不涉及公平问题等条件，还需要更深入的研究。倘若存在多周期重复博弈，是否标尺竞争的结果会发生变化？将代理人设定为风险厌恶（Risk Averse）后，标尺竞争的结果是否失效？如果在管制者在政府管制或政府购买的社会福利最大化函数中纳入了公平因素，标尺竞争的最优解是否也会发生变化？现有文献对这些问题尚未开展深入探讨。这些理论发展对于其在公共管理中的

应用尤为重要，例如公共管理诸多领域中的代理人具有明显而强烈的风险厌恶特征。

第二，在发现、分析、确定可能导致标尺竞争失败的原因及其结果后，如何通过进一步的机制设计改进激励计划尚需要更深入的研究。目前，国际学术界已经大致明确了各种可能导致标尺竞争失败的原因及其可能后果，但对于实践中如何避免错误、改进结果还缺乏有针对性的贡献，或者说应用型研究滞后于理论性研究。这是阻碍标尺竞争理论更广泛地运用于实践的最主要原因。当然，这种现象的发生缘于西方学术界重理论轻应用的传统，而在中国正相反，很多应用型研究由于缺乏理论的支撑而常常止于就事论事。

第三，有效标尺竞争思想的提出，是在引入市场竞争要素的情况下依然保持行政机制自上而下进行管控的格局，质言之，是对行政机制与市场机制之间制度嵌入性的探索。但是，有关的学术探索忽略了社群机制的作用，而社群机制极有可能通过认同与合作的方式对服务质量的保障和提升提供正向的激励，例如医疗、教育和科学领域中专业共同体所发挥的作用。限于篇幅，本文对于有关社群机制的浩繁文献无法详述，但可以确定的是，现有文献对社群机制与行政机制、社群机制与市场机制之间的制度嵌入性尚未给出充分研究，涉及到标尺竞争更是如此。尤其是公共管理领域，标尺竞争的理论和实践只是在行政手段和市场机制的结合上做文章，对于社群机制在其中的作用则完全忽视了。

尽管如此，现实中已经有许多国家在若干管制领域引入了标尺竞争，取得了不错的效果。例如，美国联邦政府针对老年人的社会医疗保险 Medicare，自 1980 年代以来，在对各家定点医院的医保支付制度中采用了标尺竞争（Shleifer, 1985）。标尺竞争还被用于英国自来水行业的管制（Cowan, 1997；OFWAT, 1998）、瑞典配电业的管制（Kumbhakar & Hjalmarsson, 1998）、日本铁路行业的管制（Mizutani, 1997）、挪威公交行业管制（Dalen & Gómez - Lobo, 2003）等。但也应该看到，标尺竞争的实践还存在一定的限度：一是成功应用的范围还比较窄，主要集中在公用事业的公共管理中；二是不乏失败的教训，尤其是在除公用事业治理之外的诸多公共管理领域。

发达国家在部分行业和领域尝试引入标尺竞争的相关经验和教训，值得中国学者深入研究。事实上，在众多涉及到民生的行业，尤其是医疗卫生、公共交通、公用事业、社会服务和教育培训等，中国政府始终在实施价格管制或政府购买，但对这些领域公共服务提供数量和质量的管理手段上，却相当滞后，管制效果自然不佳，时常引发民众的不满，尤其是网民的“吐槽”。即便是在公用事业的公共管理中，政府定价和质量管制也存在很多问题，其中很多“槽点”

的存在，与相关公共政策的决策者和研究者在包括标尺竞争在内的基础理论上的素养有所欠缺，有一定的关系。因此，积极开展对包括标尺竞争在内的新政府管理理论和手段的大量实证与经验研究，对于中国国家治理体系的创新，尤其是公共服务领域的公共管理变革，是刻不容缓的。

参考文献

- 拉丰、梯若尔(2014). 政府采购与管制中的激励理论. 上海：上海人民出版社.
- 莱恩、菲施巴赫、霍格、史密斯(2015). 医疗卫生服务管理导论. 北京：中国人民大学出版社.
- 史普博(1999). 管制与市场. 上海：上海三联书店.
- 周碧华(2015). 公共部门激励扭曲问题研究. 公共行政评论, 2: 146 - 161.
- Baiman, S. & Demski, J. S. (1980). Economically Optimal Performance Evaluation and Control Systems. *Journal of Accounting Research*, 18(S): 184 - 220.
- Besley, T. & Case, A. (1995). Incumbent Behavior: Vote - Seeking, Tax - Setting, and Yardstick Competition. *American Economic Review*, 85(1): 25 - 45.
- Bloom, N., Propper, C., Seiler, S. & Van Reenen, J. (2015). The Impact of Competition on Management Quality: Evidence from Public Hospitals. *The Review of Economic Studies*, 82(2): 457 - 489.
- Bordignon, M., Cerniglia, F. & Revelli, F. (2004). Yardstick Competition in Intergovernmental Relationships: Theory and Empirical Predictions. *Economics letters*, 83(3): 325 - 333.
- Brekke, K. R., Siciliani, L. & Straume, O. R. (2008). Competition and Waiting Times in Hospital Markets. *Journal of Public Economics*, 92(7): 1607 - 1628.
- Burns, P., Jenkins, C. & Riechmann, C. (2005). The Role of Benchmarking for Yardstick Competition. *Utilities Policy*, 13(4): 302 - 309.
- Canoy, M., Hindriks, F. & Vollaard, B. (2000). Yardstick Competition: Theory, Design and Practice. *CPB Working Paper*, No. 133, The Hague: Netherlands Bureau for Economic Policy Analysis.
- Coelli, T. J., Rao, D. S. P., O'Donnell, C. J. & Battese, G. E. (2005). *An Introduction to Efficiency and Productivity Analysis*. New York: Springer Science & Business Media.
- Cowan, S. (1997). Competition in the Water Industry. *Oxford Review of Economic Policy*, 13(1): 83 - 92.
- Dalen, D. M. (1997). Regulation of Quality and the Ratchet Effect: Does Unverifiability Hurt the Regulator?. *Journal of Regulatory Economics*, 11(2): 139 - 155.
- Dalen, D. M. (1998). Yardstick Competition and Investment Incentives. *Journal of Economics & Management Strategy*, 7(1): 105 - 126.

- Dalen, D. M. & Gómez - Lobo, A. (2003). Yardsticks on the Road: Regulatory Contracts and Cost Efficiency in the Norwegian Bus Industry. *Transportation*, 30(4): 371 - 386.
- Dawson, D. , Goddard, M. & Street, A. (2001). Improving Performance in Public Hospitals: A Role for Comparative Costs? *Health Policy*, 57(3): 235 - 248.
- Feng Lu, S. (2012). Multitasking, Information Disclosure, and Product Quality: Evidence from Nursing Homes. *Journal of Economics & Management Strategy*, 21(3): 673 - 705.
- Frontier Economics. (2010). RPI - X @ 20: The Future Role of Benchmarking in Regulatory Reviews. A Final Report Prepared for OFGEM. Available at (November 12, 2015): <http://www.frontier-economics.com/documents-05/the-future-role-of-benchmarking-in-regulatory-reviews-frontier-report.pdf>.
- Gibbons, R. & Murphy, K. J. (1990). Relative Performance Evaluation for Chief Executive Officers. *Industrial & Labor Relations Review*, 43(3): 30S - 51S.
- Hatfield, J. W. & Kosec, K. (2013). Federal Competition and Economic Growth. *Journal of Public Economics*, 97(1): 144 - 159.
- Holmstrom, B. (1982). Moral Hazard in Teams. *The Bell Journal of Economics*, 13(2): 324 - 340.
- Holmstrom, B. & Milgrom, P. (1991). Multitask Principal - agent Analyses: Incentive Contracts, Asset Ownership, and Job Design. *Journal of Law, Economics & Organization*, 7(S): 24 - 52.
- Jamash, T. & Pollitt, M. (2000). Benchmarking and Regulation: International Electricity Experience. *Utilities Policy*, 9(3): 107 - 130.
- Kumbhakar, S. C. & Hjalmarsson, L. (1998). Relative Performance of Public and Private Ownership under Yardstick Competition: Electricity Retail Distribution. *European Economic Review*, 42(1): 97 - 122.
- Laffont, J. J. (1994). The New Economics of Regulation Ten Years After. *Econometrica*, 62(3): 507 - 37.
- Laffont, J. J. & Tirole, J. (1991). Provision of Quality and Power of Incentive Schemes in Regulated Industries. In Barnett, W. A. Eds. *Equilibrium Theory and Applications: Proceeding of the Sixth International in Economic Theory and Econometrics*. Cambridge, UK.: Cambridge University Press.
- Lazear, E. P. & Rosen, S. (1981). Rank - Order Tournaments as Optimum Labor Contracts. *Journal of Political Economy*, 89(5): 841 - 864.
- Ma, C. T. A. & Burgess Jr, J. F. (1993). Quality Competition, Welfare, and Regulation. *Journal of Economics*, 58(2): 153 - 173.
- Mizutani, F. (1997). Empirical Analysis of Yardstick Competition in the Japanese Railway Industry. *International Journal of Transport Economics*. 24(3): 367 - 392.
- Mookherjee, D. (1984). Optimal Incentive Schemes with Many Agents. *The Review of Economic*

- Studies, 51(3): 433 – 446.
- Nalebuff, B. J. & Stiglitz, J. E. (1983). Prizes and Incentives: Towards a General Theory of Compensation and Competition. *The Bell Journal of Economics*, 14(1): 21 – 43.
- Newhouse, J. P. (1996). Reimbursing Health Plans and Health Providers: Efficiency in Production Versus Selection. *Journal of Economic Literature*, 34(3): 1236 – 1263.
- OFWAT. (1998). *A Proposed Approach to Assessing Overall Service to Customers: A Technical Paper*. Birmingham, UK: Office of Water Services.
- Potters, J., Rockenbach, B., Sadrieh, A. & Van Damme, E. (2004). Collusion under Yardstick Competition: An Experimental Study. *International Journal of Industrial Organization*, 22(7): 1017 – 1038.
- Revelli, F. (2006). Performance Rating and Yardstick Competition in Social Service Provision. *Journal of Public Economics*, 90(3): 459 – 475.
- Rudnick, H. & Donoso, J. (2000). Integration of Price Cap and Yardstick Competition Schemes in Electrical Distribution Regulation. *Power Systems, IEEE Transactions on*, 15(4): 1428 – 1433.
- Sawkins, J. W. (1995). Yardstick Competition in the English and Welsh Water Industry Fiction or Reality?. *Utilities Policy*, 5(1): 27 – 36.
- Shleifer, A. (1985). A Theory of Yardstick Competition. *The RAND Journal of Economics*, 16(3): 319 – 327.
- Sobel, J. (1999). A Reexamination of Yardstick Competition. *Journal of Economics & Management Strategy*, 8(1): 33 – 60.
- Tangerås, T. P. (2002). Collusion – proof Yardstick Competition. *Journal of Public Economics*, 83(2): 231 – 254.
- Tangerås, T. P. (2009). Yardstick Competition and Quality. *Journal of Economics & Management Strategy*, 18(2): 589 – 613.
- Tirole, J. (1988). *The Theory of Industrial Organization*. Cambridge, MA.: The MIT press.
- Werner, R. M., Konetzka, R. T., Stuart, E. A., Norton, E. C., Polsky, D. & Park, J. (2009). Impact of Public Reporting on Quality of Postacute Care. *Health Services Research*, 44(4): 1169 – 1187.
- Yatchew, A. (2001). Incentive Regulation of Distributing Utilities Using Yardstick Competition. *The Electricity Journal*, 14(1): 56 – 60.