

# 劳动经济学实验室实验方法研究:理论与实践

代志新

中国人民大学财政金融学院

程鹏

中国人民大学财政金融学院

高宏宇

中国人民大学财政金融学院

**摘要：**本文分析了劳动经济学实验室实验研究方法的优势和科学性，介绍了在实验室研究劳动经济问题的实验设计原则和注意事项，并结合经典理论模型和具体研究案例，系统梳理了薪酬激励、社会偏好与互惠、团队竞争与合作互动、雇佣关系和最低工资制度以及劳动力市场歧视等具体研究领域的实验文献。劳动经济学实验室实验研究方法在国外经济学界已经非常流行，但在国内仍然处于起步阶段。本文结合我国文化、企业制度和劳动力市场结构等背景，为劳动经济学实验室实验研究方法在我国的应用提供了三个方面的研究思路和展望。

**关键词：**劳动经济学；实验室实验；实验经济学

## 一、引言

过去三十年里，实验劳动经济学(Experimental Labor Economics)迅速发展，涌现出一大批杰出学者和经典文献。实验劳动经济学领域的相关文献在经济学国际权威期刊上也逐渐占有一席之地，发表数量呈现逐年上升趋势，这也反映了使用随机实验的方法研究劳动经济学越来越得到经济学界的广泛认可。事实上，将实验方法引入劳动经济学研究当中，既开辟了新的研究方法和思路，同时也

---

作者：代志新：中国人民大学财政金融学院副教授，博士生导师，dai@ruc.edu.cn；程鹏：中国人民大学财政金融学院博士研究生，chengpeng97@ruc.edu.cn；高宏宇：中国人民大学财政金融学院博士研究生，hongyugao@ruc.edu.cn。  
\*本成果受到中国人民大学面上项目“行为财政学探索研究”（2022XNA001）支持。感谢匿名审稿专家提供的专业细致的审稿意见，文责自负。

为相关领域的研究增添了科学性。实验方法主要分为实地实验（field experiment）和实验室实验（lab experiment）。关于实地实验和实验室实验两种方法孰优孰劣存在非常激烈的争议：Levitt & List 质疑了实验室实验的外部有效性，即认为实验室实验得出的研究结论不一定能推广到具体环境中。<sup>①</sup>Falk & Heckman 则突出强调了实验室实验能够比实地实验更好地控制环境因素，并且实验室环境更有利于理论的仔细检验。<sup>②</sup>Charness & Kuhn 认为，实验室实验与实地实验方法各有优缺点，两种研究方法应该互相补充而非替代关系。<sup>③</sup>本文将重点介绍实验室实验方法在劳动经济学领域的理论进展及其应用。

劳动经济学领域的相关研究话题与我们日常学习、工作和生活息息相关，涉及劳动力市场中的工人、工会、企业以及政府等组织和个人，具体话题包括就业与失业、企业绩效薪资激励、团队合作与竞争、劳动力市场歧视等诸多方面。虽然经验数据研究已经对上述各类话题进行了详细深入的研究，但是由于无法排除内生性因素的干扰，研究结论饱受质疑和诟病。随着实验室实验方法的引入，这一问题便迎刃而解。实验室实验方法指通过利用大学生被试群体在受控实验室环境中模拟的各种具体场景下进行互动，再根据实验结果进行数据分析并将研究结论推广到现实环境中。与经验数据研究相比，实验室实验方法天然的优势在于，能够通过控制信息、生产技术、市场结构、经济发展趋势等环境因素从而避免内生性问题对结果的干扰。实验室方法在个体决策、群体决策、企业和组织管理以及劳动力市场等层面都有广泛且具体的应用。在个体层面，雇主的薪酬激励方案设计与职工的努力决策一直是劳动经济学领域的重点话题。关于个体的社会偏好和互惠行为的研究将传统委托代理理论推向了更深层次的“行为委托代理理论”。在群体层面，团队合作与竞争也一直备受关注，实验经济学家将“公共物品博弈”这一实验方法巧妙引入，更方便地研究“锦标赛理论”、团队生产等行为；在劳动力市场层面，雇佣关系、最低工资制度以及劳动力市场歧视也在实验室方法的应用下开启了新篇章。本文将在第三部分详细梳理实验室方法在劳动经济学诸多领域的具体文献。

本文剩余部分将按如下结构展开：第二节论证实验室实验方法的科学性和优势；第三节详细地总结当前劳动经济学各个热点领域实验室实验研究的发展现状；第四节介绍实验室实验的设计原则

<sup>①</sup> S. Levitt, and J. List. “What Do Laboratory Experiments Measuring Social Preferences Reveal about the Real World?”. *Journal of Economic Perspectives*, 2007,21 (2):153-174.

<sup>②</sup> A. Falk, and J. Heckman. “Lab Experiments Are a Major Source of Knowledge in the Social Sciences”. *Science*, 2009,326 (5952): 535-538.

<sup>③</sup> G. Charness, and P. Kuhn. “Lab Labor: What can Labor Economists Learn From the Lab?”. *Handbook of Labor Economics*, vol.4, Elsevier Inc, 2011.

和注意事项,并结合理论与实验设计,详细介绍分析实验劳动经济学领域的两个重点实验:礼物交换实验、团队竞争与合作实验;最后一节总结全文并结合我国现实背景讨论了劳动经济学实验室实验研究方法在国内的具体应用前景和研究展望。

## 二、为什么要在实验室研究劳动经济学

在理论层面,劳动经济学家们已经建立丰富而经典的理论模型来研究劳动市场、雇佣关系、工资合约等话题;在实证层面,有大量劳动市场的现实数据可用于计量分析和因果推断。实地实验(field experiments)方法利用真实情景下的受控实验,为检验理论提供现实证据,同时在很大程度上解决了自然观测数据因果推断的内生性问题。这样看来,似乎实验室实验并无“用武之地”。另外,相较于实地实验,实验室实验往往会因为外部有效性、被试自选择效应、实验者效应(又称“罗森塔尔效应”)等受到质疑。那么,为什么劳动经济学家们仍然关注实验室实验呢?

实验室实验作为一种科学的研究方法,其优势存在于诸多方面。实验室实验公认的第一个优势是能够严格地控制环境因素,避免其对因果推断产生干扰,这一点是经验数据研究不具备的。凭借这一优势,实验室实验一个非常关键的应用在于能够检验经典理论。例如,效率工资理论假设工资与工人努力之间存在正相关关系,Akerlof通过设计“礼物交换博弈”实验检验了这一假设。<sup>①</sup>Falk & Fehr在实验室控制参与锦标赛工人数量、工人的努力成本函数、奖品的价值以及工人的生产水平,成功地检验了“锦标赛理论”。<sup>②</sup>在检验理论的同时,实验室实验还能为经典模型提供补充并对原始模型进行拓展,例如“行为委托代理理论”。第二个优势在于实验的低成本和可复制性。在实验对象方面,实验室实验大多数被试是大学生,而实地实验的被试则是不同职业的从业人员。相较之下,大学生被试的机会成本更低,招募更容易,因此实验成本也更低。并且,只需要满足相同的实验设计和实验环境,实验室实验的结果则很容易被复制。<sup>③</sup>第三个优势在于能够研究动态博弈中的策略互动,<sup>④</sup>实验者通过在实验室环境下模拟公司内雇主和员工的委托代理行为,能够研究工资谈判、薪酬激励、团队合作等具体劳动经济学问题。第四个优势在于能够为政策制定提供建议,通过在实验室对现实

<sup>①</sup> George A. Akerlof. “Labor Contracts as Partial Gift Exchange”. *The Quarterly Journal of Economics*, 1982, 97(4): 543-569.

<sup>②</sup> A. Falk, and E. Fehr. “Why Labour Market Experiments?”. *Labour Economics*, 2003, 10(4): 399-406; E. P. Lazear, and S. Rosen. “Rank-order Tournaments as Optimum Labor Contracts”. *Journal of Political Economy*, 1981, 89(5): 841-864.

<sup>③</sup> G. Charness. “Laboratory Experiments: Challenges and Promise: A Review of ‘Theory and Experiment: What Are the Questions?’ by Yernon Smith”. *Journal of Economic Behavior & Organization*, 2010, 73(1): 21-23.

<sup>④</sup> J. D. Hey. “Search for Rules for Search”. *Journal of Economic Behavior & Organization*, 1985, 3(1): 65-81; M. Brown, C. J. Flinn, and A. Schotter. “Real-time Search in the Laboratory and the Market”. *American Economic Review*, 2011, 101(2): 948-74.

政策进行“预测试”，检验政策的可行性。<sup>①</sup>第五个优势在于能够研究在实地环境中难以观测的行为，因为这些行为大多是不道德的或者非法的，所以容易被当事人隐藏，例如妨碍同事工作，歧视和怨恨行为。在实验室环境下，使用货币激励措施能够有效激励被试展示这些行为，使得研究代理人的信念问题成为可能。<sup>②</sup>

实验室实验研究最容易受到的质疑主要体现在两个方面：一是实验结论的外部有效性不足；二是实验室环境可能无法很好地捕捉到现实中工人的“自选择”现象。首先，针对外部有效性不足的质疑包括多方面的内容，例如被试通常是在劳动力市场上缺乏经验的本科生，“劳动任务”大多只是涉及金钱分配以及“霍桑效应”对实验结果会产生影响。实际上，在科学的实验设计条件下，大学生被试的表现通常和真实职业的实验对象表现并无明显差异。<sup>③</sup>例如士兵、制造业工人、渔民等群体，实验人员发现大学生群体表现与这些职业工人的表现非常相似。此外，如果研究目标针对的是工人群体的一般性结论，大学生群体由于不同的专业背景可能比具体职业工人更具有吸引力。<sup>④</sup>其次，现实劳动力市场中工人可能存在自选择问题，例如利他主义动机较强的雇员倾向于自我选择参与团队合作，风险偏好的雇员更可能参与锦标赛竞争。在实验室，通过随机指派一些大学生被试扮演对应的雇员，显然不符合现实情况。然而，随着实验室实验的发展，劳动经济学家已经能在实验室解决这一问题，甚至能对自选择问题背后的机制提供解释。<sup>⑤</sup>并且，实验室解决自选择问题的优势在于能直接测量并控制混淆因素，如代理人的能力、信念、损失厌恶程度和风险偏好等。

### 三、劳动经济学在实验室的具体应用

#### (一) 薪酬激励与员工努力

薪酬激励是委托代理关系中最受关注的研究话题之一。最早的研究可以追溯到动物实验<sup>⑥</sup>，实验

<sup>①</sup> Deck, C., Farmer, A., & Zeng, D.-Z. "Amended final-offer arbitration over an uncertain value: A comparison with ca and foa". *Experimental Economics*, 2007, 10(4): 439-454.

<sup>②</sup> P. Battigalli, and M. Dufwenberg. "Dynamic Psychological Games". *Journal of Economic Theory*, 2009, 144(1): 1-35.

<sup>③</sup> E. Fehr, and J. List. "The Hidden Costs and Returns of Incentives-trust and Trustworthiness among CEOs". *Journal of the European Economic Association*, 2004, 2(5): 743-771; Snowberg, Erik, and Leeat Yariv. "Testing the Waters: Behavior across Participant Pools." *American Economic Review*, 2021, 111 (2): 687-719; Ernst Fehr et al. "When Social Norms Overpower Competition: Gift Exchange in Experimental Labor Markets". *Journal of Labor Economics*, 1998, 16 (2): 324-351; Abigail Barr and Pieter Serneels. "Reciprocity in the Workplace". *Experimental Economics*, 2009, 12 (1): 99-112; G. Charness et al. "Competition and the Ratchet Effect". *Journal of Labor Economics*, 2011, 29(3): 513-47.

<sup>④</sup> G. Charness, and P. Kuhn. "Lab Labor: What can Labor Economists Learn from the Lab?". *Handbook of Labor Economics*, vol.4, Elsevier Inc, 2011.

<sup>⑤</sup> S. Della Vigna, J. A. List, and U. Malmendier. "Testing for Altruism and Social Pressure in Charitable Giving". *The Quarterly Journal of Economics*, 2012, 127(1): 1-56; E. Lazear et al. "Sorting in Experiments with Application to Social Preferences". *American Economic Journal: Applied Economics*, 2012, 4(1): 136-63.

<sup>⑥</sup> R. C. Battalio et al. "Income-leisure Tradeoffs of Animal Workers". *The American Economic Review*, 1981, 71(4): 621-632.

人员通过给动物食物激励来引导动物进行对应的“劳动努力”（例如按压键盘或者杠杆），实验结果发现了动物向后弯曲的劳动供给曲线，证明了闲暇商品的存在。Sillamaa 在实验室进行了一项关于劳动所得税的实验，研究发现，相较于累进所得税情况，在线性所得税情况下工人的工作努力对工资增长的反应更积极。并且最高边际税率为零时，工人的工作努力程度会达到最大值。这一真实努力实验侧面反应了增加工资能增加工作努力<sup>①</sup>，然而越来越多的研究发现，薪酬激励与工作努力之间的关系并非单调的。Gneezy & Rustichini 的研究结论表明，绩效工资与努力之间存在“U”型关系，即较少的激励可能会挤出员工执行这些任务的内在动机。<sup>②</sup>Cadsby et al.进行了更深一步的实验测试，研究表明薪酬激励的效果与代理人的风险规避程度相关。<sup>③</sup>具体来说，当激励措施加强时，大约 25% 的被试反而减少了工作努力，并且这一概率随着风险规避程度增加而增加。因此，薪酬激励与工作努力之间的关系可能是高度非单调的。

关于绩效工资与工作努力之间的激励机制研究，一些劳动经济学家提出了另一种观点：绩效工资与生产率的提高很大程度上可以用工人的自选择行为来解释，即生产率高的工人会自我选择主动参与绩效工资任务，而非提高工人现有的努力程度。在理论层面，Lazear 提出的“绩效工资模型”很清晰地捕捉了这种自选择效应。模型假设工人的效用函数 $u(I, e)$ 满足如下条件：

$$u(I, e) = \begin{cases} w - \frac{c(e)}{a}, & \text{当选择计时工资时} \\ be - k - \frac{c(e)}{a}, & \text{当选择计件工资时} \end{cases}$$

式(1)中 $I$ 代表收入， $e$ 表示工人的努力程度， $c(e)$ 表示工人努力的成本函数， $a$ 表示工人的能力。

当工人选择计时工资时获得 $w$ 的工资收入，选择绩效工资时获得 $be$ 的工资收入， $k$ 表示参与绩效工资薪酬方案的工人需要支付的费用。式(1)隐含的条件是，当工人的能力超出一定水平，工人会倾向于选择绩效工资，即能力强的工人会自愿选择绩效工资方案。此后，许多实验室研究证实了绩效工资模型的预测。Dohmen & Falk 在控制了被试的风险偏好、过度自信、社会偏好等特征后，考察了固定工资、效率工资、锦标赛和收入分享四种类型合同下的员工的自选择行为，结果表明后三个可变薪酬方案的产出提高大部分是由于更有能力的工人参加了这些方案。<sup>④</sup> Cadsby et al.在实验室允许工

<sup>①</sup> M. A. Sillamaa. "How Work Effort Responds to Wage Taxation: An experimental Test of a Zero Top Marginal Tax Rate". *Journal of Public Economics*, 1999, 73(1): 125-134.

<sup>②</sup> U. Gneezy, and A. Rustichini. "Pay Enough or Don't Pay at All". *Quarterly Journal of Economics*, 2000, 115(3): 791-810.

<sup>③</sup> C. B. Cadsby, F. Song, and F. Tapon. "The Impact of Risk-Aversion and Stress on the Incentive Effect of Performance-Pay". *Experiments in Organizational Economics*, 2016, 19: 189-227.

<sup>④</sup> T. Dohmen, and A. Falk. "Performance Pay and Multidimensional Sorting: Productivity, Preferences, and Gender". *American Economic Review*, 2011, 101(2): 556-90.

人自我选择参与绩效工资任务，研究发现绩效工资能提高生产率，同时也发现了自选择行为增强了这种效果，并且生产率更高的员工更可能选择绩效工资任务。<sup>①</sup>在田野实验方面，Bandiera et al.也研究了自选择效应与薪酬激励问题，他们通过在一家水果采摘公司的田野实验发现，当经理人关心绩效工资任务时，更有可能选择生产率较高的工人进入他们的公司。<sup>②</sup>

除了绩效工资激励方式之外，作为现实公司中被广泛推行的激励方案——“锦标赛”，也在实验室被深入地研究。“锦标赛”的含义是公司经理人设置相对工资绩效，对于在工作中表现优异的员工给予奖金和晋升机会。关于锦标赛与绩效工资哪种激励方案更有效，争论由来已久。Lazear & Rosen认为，当经理人无法观察到工人的能力信息，锦标赛的效率可能会低于效率工资。在锦标赛制度下，通常工人需要考虑同事的努力水平来找到纳什均衡，因此可能会降低激励效果。<sup>③</sup>一个相反的观点认为，如果工人能从竞争中获得直接效用，锦标赛方案可能更加有效。<sup>④</sup>类似地，在锦标赛方案下也存在自选择问题，如果工人是风险偏好的<sup>⑤</sup>，或者高估自己的能力<sup>⑥</sup>，则更倾向于自我选择地参与锦标赛竞争。锦标赛制度的一个负面效果是，参与竞争的员工们可能会破坏和妨碍其他同事的工作。<sup>⑦</sup>

实验劳动经济学家关注的另一个员工激励问题是监督对于员工努力行为的影响。尽管从直觉上分析，加强监督能够督促员工提高工作努力程度，增加产出。然而，许多实验室研究发现经理人这种强制性“控制”监督行为可能会产生负面效果，即所谓的“隐性控制成本”。员工可能会对这种行为产生抵触，反而降低自己的努力程度，因为员工可能会把监督当成不被信任的信号。<sup>⑧</sup> Dickinson & Villeval 尝试探究这种挤出效应的背后机制<sup>⑨</sup>，他们开展了一个真实努力的实验室实验，其中经理选择监督强度，员工有一定概率面临被审查，一旦没有完成目标则会被处以罚款。研究发现，监督提高了员工在匿名环境下的努力程度，而在非匿名环境下，监督会产生挤出效应。这一结果说明了在

<sup>①</sup> C. B. Cadsby, F. Song, and F. Tapon. “Sorting and Incentive Effects of Pay for Performance: An Experimental Investigation”. *The Academy of Management Journal*, 2007, 50(2): 387–405.

<sup>②</sup> O. Bandiera, I. Barankay, and I. Rasul. “Incentives for Managers and Inequality among Workers: Evidence from a Firm Level Experiment”. The Field Experiments Website, Natural Field Experiments, 2006.

<sup>③</sup> E. P. Lazear, and S. Rosen. “Rank-order Tournaments as Optimum Labor Contracts”. *Journal of Political Economy*, 1981, 89(5): 841–864.

<sup>④</sup> M. O’Keefe, W. Viscusi, and R. Zeckhauser. “Economic Contests: Comparative Reward Schemes”. *Journal of Labor Economics*, 1984, 2(1): 27–56.

<sup>⑤</sup> M. Niederle, and L. Vesterlund. “Do Women Shy away from Competition? Do Men Compete too much?”. *The Quarterly Journal of Economics*, 2007, 122: 1067–1101.

<sup>⑥</sup> C. Camerer, and D. Lovo. “Overconfidence and Excess Entry: An Experimental Approach”. *The American Economic Review*, 1999, 89(1): 306–318.

<sup>⑦</sup> J. Carpenter et al. “Tournaments and Office Politics: Evidence from a Real Effort Experiment”. *American Economic Review*, 2010, 100(1): 504–17.

<sup>⑧</sup> B. Frey. “Does Monitoring Increase Work Effort? the Rivalry with Trust and Loyalty”. *Economic Inquiry*, 1993, 31(4): 663–70; A. Falk, and M. Kosfeld. “The Hidden Costs of Control”. *American Economic Review*, 2006, 96(5): 1611–1630.

<sup>⑨</sup> D. Dickinson, and M.C. Villeval. “Does Monitoring Decrease Work Effort? The Complementarity between Agency and Crowding-out Theories”. *Games and Economic Behavior*, 2008, 63(1): 56–76.

员工和经理互相了解的情况下，监督会使员工产生“不被信任感”，因此员工反而会“故意”不配合经理人的监督。

企业内部不同员工之间的生产率水平、努力程度存在异质性，因此工资也存在差异。一个与之相关的讨论是不同员工之间是否由于存在“同伴效应”（peer effect）而互相影响呢？Falk & Ichino 通过一项真实的努力实验发现<sup>①</sup>，当同事在场时，低生产率工人会提高工作努力，高生产率的工人却有可能减少努力。现实职场上公司往往规定不允许同事之间互相比较工资，那么，当员工们看到彼此的工资时，是否会影响自身的努力程度？Charness & Kuhn 研究发现，在员工知道彼此之间生产率存在差异但是不清楚差异大小和方向的前提下，如果工人能够观察到同事的工资，并不会改变自己的决策。<sup>②</sup>

## (二) 社会偏好与互惠行为

在理性人假设下，个人行为的目标是最大化自身利益。然而在实验室，经济学家们经常观察到与该假设相违背的现象。在独裁者博弈、公共品博弈等实验下，被试经常会选择减少自己的收入用于增加其他被试的收入，即所谓的“社会偏好”现象。社会偏好的本质是，人们会为了遵守承诺、避免内疚、维护社会形象、温情效应等目的主动牺牲自身利益来帮助他人。Güth et al. 最早在实验室通过最后通牒博弈实验证明社会偏好的存在<sup>③</sup>，此后发展出关于社会偏好的模型，大致可以分为互惠模型和分配模型。互惠模型<sup>④</sup>的核心假设是，博弈参与人的策略选择取决于三个因素：自己选择某种策略的回报，对其他人的恩惠以及对于其他人对自己的恩惠的信念。考虑如下效用函数形式：

$$U_i \left\{ a_i(t), \{b_{ij}(t), [c_{ijk}(t)]_{k \neq j}\}_{j \neq i} \right\} = \pi_i \{a_i(t), [b_{ij}(t)]_{j \neq i}\} + \sum_{j \in N \setminus \{i\}} \left\{ \left( \theta_{ij} \cdot \rho_{ij} \{a_i(t), [b_{ij}(t)]_{j \neq i}\} \right) \cdot \sigma_{iji} \{b_{ij}(t), [c_{ijk}(t)]_{k \neq j}\} \right\} \quad (2)$$

其中， $t$ 表示某一时期， $U_i(\cdot)$ 表示参与人 $i$ 的效用函数， $a_i(t)$ 表示参与人 $i$ 的策略， $b_{ij}(t)$ 代表 $i$ 对于其他博弈参与人的策略选择的信念， $c_{ijk}(t)$ 则表示 $i$ 对于其他参与人对自己选择的策略的信念 $b_{ij}(t)$ 的信念。参与人的效用取决于三部分： $\pi_i$ 表示参与人自己选择策略的回报， $\rho_{ij}$ 衡量参与人 $i$ 对其他参与人 $j$ 的恩惠， $\theta_{ij}$ 是互惠效应敏感系数，反映 $i$ 对于 $j$ 的互惠关系的敏感程度， $\sigma_{iji}$ 代表的含义

<sup>①</sup> A. Falk, and A. Ichino. “Clean Evidence on Peer Effects”. *Journal of labor economics*, 2006, 24(1): 39–57.

<sup>②</sup> G. Charness, and P. Kuhn. “Does Pay Inequality Affect Worker Effort? Experimental Evidence”. *Journal of Labor Economics*, 2007, 25(4): 693–723.

<sup>③</sup> W. Güth, R. Schmittberger, and B. Schwarze. “An Experimental Analysis of Ultimatum Bargaining”. *Journal of Economic Behavior and Organization*, 1982, 3(4): 367–388.

<sup>④</sup> M. Rabin. “Incorporating Fairness into Game Theory and Economics”. *American Economic Review*, 1993, 83(5): 1281–1302; M. Dufwenberg, and G. Kirchsteiger. “A Theory of Sequential Reciprocity”. *Games and Economic Behavior*, 2004, 47(2): 268–298.

是参与人*i*对于参与人*j*如何对自己施予恩惠的信念。分配模型假设在一个双人博弈中,参与人*i*的效用函数取决于自己的回报水平 $x_i$ 和对方*j*的回报水平 $x_j$ :

$$U_i(x_i, x_j) \equiv \begin{cases} \alpha x_j + (1 - \alpha)x_i, & \text{当 } x_i \geq x_j \text{ 时} \\ \beta x_j + (1 - \beta)x_i, & \text{当 } x_i < x_j \text{ 时} \end{cases} \quad (3)$$

其中,系数 $\alpha$ 和 $\beta$ 衡量参与人*i*的利他主义行为和对分配公平的偏好。具体而言,当对方收入比自己低时,参与人*i*更关心对方的收入;当对方收入更高时,则更关心自己的收入。两类模型的侧重点存在差异,互惠模型假设给予他人正向回馈取决于对方行为的公平程度,重点在于互惠;而分配模型假定个人厌恶差异并且关心获得低收益的群体,重点在于个人选择。

礼物交换实验(Gift-Exchange Experiments)是研究社会偏好最常用的实验之一,该实验旨在模拟雇佣关系,并测试工资和员工努力之间的关系。在两阶段博弈实验中,第一阶段雇主提出工资报价,第二阶段工人选择努力程度。如果工人完全是自利的,预计结果是无论工资是多少工人都会一直选择最低水平的努力,雇主预期到这一结果则会支付最低的工资。然而实验结果是,雇主支付的平均工资和工人选择的努力程度远高于预期,这体现了雇主和工人直接的社会偏好以及互惠行为。其中一项经典的研究是 Fehr et al.的公平感如何影响劳动力市场均衡。他们设计一个礼物交换实验研究雇主工资支付和员工努力选择之间的互惠行为,为公平工资—努力理论提供了实验证据。<sup>①</sup>

什么因素影响或者决定了社会偏好现象呢? Roth et al.通过实验发现<sup>②</sup>,制度和背景因素在社会偏好现象中发挥了重要作用。社会偏好是否发挥作用还取决于合约的完备性和可执行性。<sup>③</sup>在不完全合约的竞争市场中,互惠行为起主导作用。这是因为在不完全合约下,工人采取积极合作的工作态度能够得到雇主更慷慨的报酬支付。在完备的合约下,工人清楚工资标准,因此不会改变自身的努力水平。Fershtman 提出了“框架效应”对社会偏好的影响证据。<sup>④</sup>实验结果表明,在小型独裁者博弈和信任博弈中,在环境中引入框架效应会挤出社会偏好效应。此外,还有一些来自现实工作环境中的实验研究表明,工人的生产率与社会偏好之间存在正相关关系,即生产率越高的工人社会偏好越强。<sup>⑤</sup>

<sup>①</sup> E. Fehr et al. “Does Fairness Prevent Market Clearing? An Experimental Investigation.” *The Quarterly Journal of Economics*, 1993,108(2): 437–59.

<sup>②</sup> A. E. Roth et al. “Bargaining and Market Behavior in Jerusalem, Ljubljana, Pittsburgh, and Tokyo: An Experimental Study”. *The American Economic Review*, 1991, 81(5): 1068–1095.

<sup>③</sup> E. Fehr, and S. Gächter. “Fairness and Retaliation: The Economics of Reciprocity”. *Journal of Economic Perspectives*, 2020, 14(3): 159-181.

<sup>④</sup> Chaim Fershtman et al. “Equity Aversion: Social Norms and the Desire to Be Ahead.” *American Economic Journal: Microeconomics*, 2012,4(4): 131–44.

<sup>⑤</sup> Abigail Barr, and Pieter Serneels. “Reciprocity in the Workplace”. *Experimental Economics* 2009, 12 (1): 99–112; Jeffrey Carpenter, and Erika Seki. “Do Social Preferences Increase Productivity? Field Experimental Evidence from Fisherman in Toyama Bay”. *Economic Inquiry*. 2005, 49: 612-630.

### (三) 团队生产，合作与竞争

团队生产和内部工人之间的互动行为一直是劳动经济学家关注的热点。现实公司的业务开展离不开优秀的团队，许多项目需要团队成员互相合作、信息共享、优势互补，同时也需要设计激励机制使得团队内部员工之间形成良心竞争从而促进团队产出。<sup>①</sup>从经典的团队代理模型霍尔姆斯特罗姆模型开始<sup>②</sup>，已经有大量实验文献研究这一话题。霍尔姆斯特罗姆模型提出了“产出共享”规则，研究团队内部员工的努力贡献程度，毫无疑问，搭便车（free-riders）行为在团队生产中是无法避免的。Nalbantian & Schotter 在实验室进行了一项团队生产实验，他们创建了两个六人团队，比较收入平均分享方案、“强制”合同、按表现分配收入以及团队锦标赛四种薪酬方案下团队的产出和努力水平，实验结果与经典模型预测一致，收入平均共享方案下工人的努力水平偏低，而团队锦标赛方案导致了更高的效率产出。收入平均共享方案下普遍的搭便车行为难以避免，因此如何激励团队内部个人提高努力程度是经理人需要考虑的问题。<sup>③</sup> Croson et al.关于团队内部单个员工激励方案的做法是，根据成员表现将他们分为“优秀”“良好”“一般”和“较差”等类别，表现较差的成员在未来面临被剔除团队的风险。<sup>④</sup>此外，还有许多研究也讨论了如何设计相对个人激励问题。<sup>⑤</sup>

在实验室中，团队生产往往和公共物品博弈联系起来。这类文献将团队生产定义为自愿捐款机制中的参与者的捐款贡献，即用团队中成员自愿捐款数额代表努力程度和产出。<sup>⑥</sup>在标准的自愿捐款模型中，团队中的每个成员*i*自愿捐赠 $v_i$ 给团队，每个成员的禀赋为 $a_i$ ，则每位成员的最终收益为 $(a_i - v_i) + \beta \sum_i v_i$ ，其中 $\beta$ 表示公共收益的效率系数。与公共品博弈结果相似，在实验最初几轮团队每个成员的捐款额会高于个体理性水平，但随着实验轮次的增加，会收敛到个体理性水平。利用自愿捐款机制研究团队行为，需要在原有的框架下修改奖励计划，例如 Dickinson & Isaac 引入了“最高贡献奖”用于奖励捐款最多的员工，这一机制设立提高了员工贡献水平。<sup>⑦</sup>此外，还有一些文献提出其他

<sup>①</sup> A. Danilov, C. Harbring, and B. Irlenbusch. “Helping under A Combination of Team and Tournament Incentives”. *Journal of Economic Behavior and Organization*, 2019, 162(6): 120-135.

<sup>②</sup> B. Holmstrom. “Moral Hazard in Teams”. *The Bell Journal of Economics*, 1982, 13(2): 324-340.

<sup>③</sup> H. R. Nalbantian, and A. Schotter. “Productivity under Group Incentives: An Experimental Study”. *The American Economic Review*, 1997, 87(3): 314-341.

<sup>④</sup> R. Croson et al. “Excludability: A Laboratory Study on Forced Ranking in Team Production”. *Journal of Economic Behavior and Organization*, 2015, 114: 13-26.

<sup>⑤</sup> Shahla Ghobadi, John Campbell, and Stewart Clegg. “Pair Programming Teams and High-quality Knowledge Sharing: A Comparative Study of Competitive Reward Structures”. *Information Systems Frontiers*, 2017, 19(2): 397-409; M. Rahmani, G. Roels, and U. S. Karmarkar. “Team Leadership and Performance: Combining the Roles of Direction and Contribution”. *Management Science*, 2018, 64(11): 5234-5249.

<sup>⑥</sup> R. M. Isaac, and J. M. Walker. “Group Size Effects in Public Goods Provision: The Voluntary Contributions Mechanism”. *The Quarterly Journal of Economics*, 1988, 103(1), 179-199.

<sup>⑦</sup> D. L. Dickinson, and R. M. Isaac. “Absolute and Relative Rewards for Individuals in Team Production”. *Managerial and Decision Economics*, 1998, 19(4/5): 299-310.

类型的报酬形式, 如对最低贡献者给予罚款的机制。<sup>①</sup>

与绩效工资和锦标赛制度相同的是, 参加团队也存在自选择问题。例如, 在团队组建过程中, 团队成员更倾向邀请能力较强的同事加入团队。<sup>②</sup> Bandiera et al. 研究发现, 当团队面临的激励加强时 (报酬奖励增加), 按照能力进行匹配的团队数量会增加。<sup>③</sup> 与团队组建相关的问题是, 团队规模如何影响团队产出和决策。Carpenter 开展了一项自愿捐款实验, 结论表明团队规模影响团队成员的捐款数量, 当团队规模增加时, 团队内部成员之间的监督机制作用下降, 搭便车现象变得更加普遍, 捐款数量下降, 员工的平均努力水平下降。<sup>④</sup>

#### (四) 雇佣关系和最低工资

在宏观经济研究方面, 就业和失业问题一直备受关注, 经济学家们建立了丰富的理论模型用于研究就业和失业问题, 例如奥肯定律, 新凯恩斯菲利普斯曲线, 动态搜寻匹配理论等。<sup>⑤</sup> 搜寻匹配理论被广泛用于研究失业的决定因素和就业过程的工资谈判, <sup>⑥</sup> 劳动经济学家们尝试为该理论寻找微观证据。Schotter & Braunstein 最早在实验室测试搜寻理论, <sup>⑦</sup> 实验被试从已知的工资分布中随机抽取工资, 并且每次抽取都会花费固定的搜寻成本。根据理论预测, 工人的最优策略是恒定的保留工资。实验结果很大程度上符合理论的预测。但是实验的额外发现是, 随着搜寻次数的增加, 工人的保留工资会不断下降, 意味着风险厌恶程度会影响工人的求职行为。此后, 劳动经济学家们在实验室对搜寻匹配理论进行了丰富的拓展, 包括有限视野、可变搜索强度、未知的工资分布以及个人偏好异质性对搜寻行为的影响。<sup>⑧</sup> 搜寻匹配理论一个非常普遍的应用在于研究失业保险制度和工人的

<sup>①</sup> D. Dickinson. "The Carrot vs. the Stick in Work Team Motivation". *Experimental Economics*, 2001, 4(1): 107-124; B. Irlenbusch, and G. K. Ruchala. "Relative Rewards within Team-based Compensation". *Labour economics*, 2008, 15(2): 141-167.

<sup>②</sup> R. A. Weber. "Managing Growth to Achieve Efficient Coordination in Large Groups". *American Economic Review*, 2006, 96(1): 114-126; A. Hammermann, A. Mohnen, and P. Nieken. "Whom to Choose as a Team Mate? A Lab Experiment about In-Group Favouritism" IZA Discussion Papers No. 6286, *Institute of Labor Economics*, 2012.

<sup>③</sup> O. Bandiera, I. Rasul, and I. Barankay. "Team Incentives: Evidence from A Firm Level Experiment". *Journal of the European Economic Association*, 2013, 11(5): 1079-1114.

<sup>④</sup> J. P. Carpenter, "Punishing Free-riders: How Group Size Affects Mutual Monitoring and The Provision of Public Goods". *Games and Economic Behavior*, 2007, 60(1): 31-51.

<sup>⑤</sup> M. Okun. "Potential GNP: Its Measurement and Significance." American Statistical Association, Proceedings of the Business and Economics Statistics Section, 1962; Richard Clarida, Jordi Galí, and Mark Gertler. "Monetary Policy Rules and Macroeconomic Stability: Evidence and Some Theory". *The Quarterly Journal of Economics*. 2000, 115 (1): 147-180; Per Krusell, Toshihiko Mukoyama and Ayşegül Şahin. "Labour-Market Matching with Precautionary Savings and Aggregate Fluctuations". *The Review of Economic Studies*, 2010, 77 (4): 1477-1507.

<sup>⑥</sup> J. Nash. "Two-Person Cooperative Games". *Econometrica*, 1953, 21(1): 128-140; E. Kalai. "Proportional Solutions to Bargaining Situations: Interpersonal Utility Comparisons". *Econometrica*, 1997, 45 (7): 1623-1630.

<sup>⑦</sup> Yale M Braunstein, and Andrew Schotter. "Labor Market Search: An Experimental Study". *Economic Inquiry* 1982, 20(1): 133-144.

<sup>⑧</sup> James C. Cox, and Ronald L. Oaxaca. "Direct Tests of The Reservation Wage Property". *Economic Journal*. 1992, 102(415): 1423-1432; Glenn W. Harrison, and Peter Morgan. "Search Intensity in Experiments". *Economic Journal*, 1990, 100(401): 478-486; James C. Cox, Ronald L. Oaxaca. "Good News and Bad News: Search from Unknown Wage Offer Distributions". *Experimental Economics*, 2000, 2(3): 197-225; Daniel Schunk. "Behavioral Heterogeneity in Dynamic Search Situations: Theory and Experimental Evidence". *Journal of Economic Dynamics and Control* 2009, 33(9): 1719-1738.

求职努力行为, 虽然在这一领域已经有大量的经验数据研究<sup>①</sup>, 然而大部分的实证分析仍然存在难以解决的内生性问题, 如失业保险政策的内生性或者环境因素的干扰。相比之下, 实验室实验则具有严格控制环境和随机干预的独特优势。我们首先考虑一个基本的动态搜寻匹配模型框架: 假设劳动力市场上有  $i (i \in \{1, \dots, N\})$  种类型的工人, 他们在技能和劳动生产率上存在异质性。匹配函数为  $m(s_i n_i, v_i)$ , 其中  $s_i$  表示  $i$  类型工人的工作搜寻强度, 即找工作的频率和花费的时间,  $n_i$  表示  $i$  类型工人的数量,  $s_i n_i$  衡量  $i$  类型工人的总搜寻努力,  $v_i$  代表企业可以提供的总工作岗位数量。定义劳动力市场密集程度  $\theta \equiv v_i / (s_i n_i)$ ,  $\theta$  越大, 劳动力市场就业岗位越充足。  $i$  类型的工人选择求职努力程度最大化期望效用如下:

$$s_i f(\theta_i) u((1 - \tau_i) w_i) + (1 - s_i f(\theta_i)) u(b^l) - \phi(s_i) \quad (4)$$

其中,  $s_i f(\theta_i)$  表示工人找到工作的概率,  $\phi(s_i)$  代表工人求职花费的时间和金钱成本。对式(4)关于搜寻努力  $s_i$  求解一阶最优化条件, 则有:

$$f(\theta_i) [v((1 - \tau_i) w_i) - u(b^l)] - \phi'(s_i) = 0 \quad (5)$$

上式的经济含义是,  $i$  类型的工人选择搜寻努力来最大化自身效用。除了工资水平  $w_i$  之外, 工人的求职决策还受到劳动所得税税率  $\tau_i$  和失业补助水平  $b^l$  的影响。利用这一理论框架, 我们可以在实验室环境下考察工人找工作的努力程度对于工资率、劳动所得税税率以及失业保险金的反应。例如, Lechthaler & Ring 在实验室模拟了一个存在搜寻努力和劳动力参与决策的就业市场, 通过改变失业保险金的领取市场<sup>②</sup>, 研究失业工人寻找工作的努力程度。

劳动力市场中工人(工会)与雇主之间经常面临着工资合约的谈判情况。实验室中通常用简单的讨价还价博弈(如最后通牒博弈和独裁者博弈)来模拟现实的谈判行为<sup>③</sup>, 进而研究工人的罢工行为。例如, 集中谈判和分散谈判对罢工行为的影响。<sup>④</sup>劳动仲裁指的是工人和雇主之间因为谈判不成功导致的第三方(通常是政府部门)调解, 劳动经济学家在实验室进行了大量的实验来评估不同的

<sup>①</sup> H. S. Farber, and R. G. Valletta. "Do Extended Unemployment Benefits Lengthen Unemployment Spells? Evidence from Recent Cycles in the U.S. Labor Market". *The Journal of Human Resources*, 2015, 50 (4): 873-909; A.C. Johnston, A. Mas. "Potential Unemployment Insurance Duration and Labor Supply: The Individual and Market-level Response to a Benefit Cut". *Journal of Political Economics*, 126 (6): 2480-2522.

<sup>②</sup> Wolfgang Lechthaler, and Patrick Ring. "Labor Force Participation, Job Search Effort and Unemployment Insurance in the Laboratory". *Journal of Economic Behavior & Organization*. 2021, 189: 748-778.

<sup>③</sup> James Andreoni, Marco Castillo, and Ragan Petrie. "What Do Bargainers' Preferences Look Like? Experiments with a Convex Ultimatum Game". *American Economic Review*, 2003, 93(3): 672-685; Armin Falk, Ernst Fehr, and Urs Fischbacher. "Reasons for Conflict: Lessons from Bargaining Experiments". *Journal of Institutional and Theoretical Economics*, 2003, 159 (1): 171-187.

<sup>④</sup> Siegfried Berninghaus, Guth Werner, and Keser Claudia. "Unity Suggests Strength: An Experimental Study of Decentralized and Collective Bargaining". *Labour Economics*, 2003, 10 (4): 465-479; Fabienne Tournadre, and Marie-Claire Villeval. "Learning from Strikes". *Labour Economics*, 2004, 11 (2): 243-264.

仲裁方案，如传统仲裁方案和最终报价仲裁方案<sup>①</sup>，以及其他各种改进的仲裁方案（如联合仲裁）。因此，实验室实验不仅能研究员工和雇主的谈判行为，还能为解决劳动力市场争端方案提供试验场所。

“最低工资”制度究竟如何影响就业？背后的机制以及影响效果一直是经济学家们争论的焦点。最开始，经济学家普遍认为最低工资的设立会对就业产生负面影响，因为工人的工资水平上升增加了企业的用工成本，企业会减少对原先工资水平低于最低工资标准的工人的雇佣。Card & Kruger 运用自然实验方法，<sup>②</sup>对新泽西州和宾夕法尼亚州快餐店最低工资和就业数据的分析表明，最低工资制度设立会增加当地工人 13% 的就业率。关于最低工资制度的争议，还是体现在理论分析和经验数据推断之间的矛盾。当争议相持不下，实验室实验方法便大有“用武之地”。Brandts & Charness 最开始运用礼物交换博弈研究最低工资的影响，<sup>③</sup>实验中每个被试被赋予 10 个单位的代币，第一阶段，先行者可以交换至多 10 个单位的代币，而响应者则会收到 5 倍的代币。第二阶段响应者返还至多 10 个单位代币，原先的先行者也可以收到 5 倍的返还代币。当规定最低工资为 5 代币时，实验结果表明最低工资反而会降低工人们 30% 的平均努力程度。Owens & Kagel 在此基础上改进了实验设计<sup>④</sup>，研究发现最低工资的设立会减少原先工资水平在最低工资附近的工人的努力，但是不会影响高工资工人的努力程度。

## （五）劳动力市场歧视

劳动力市场歧视指具有相同生产率、教育背景和经验的工人，由于一些非经济特征不同（如性别、种族、地域等），存在不平等的就业机会和工资报酬。现实中劳动力市场歧视现象非常普遍，研究劳动力市场歧视行为，厘清歧视背后的机制，对减少收入不平等、促进社会公平具有重大意义。

性别歧视是最受重视的话题之一。早期文献使用基于观测数据的回归方法来测试劳动力市场中的歧视现象。最常见的方法是 Oaxaca 分解理论，即将男女之间的工资差异分解为两个部分：能够由可观察到的特征（如生产率、工作时长、教育背景、经验等）解释的部分和无法解释的部分，无法由

<sup>①</sup> Orley Ashenfelter et al. “An Experimental Comparison of Dispute Rates in Alternative Arbitration Systems”. *Econometrica*, 1992, 60(6): 1407–1433; Alexander S. Kritikos. “The Impact of Compulsory Arbitration on Bargaining Behavior: An Experimental Study”. *Economics of Governance*, 2006, 7(3): 293–315.

<sup>②</sup> D. Card, L.F. Katz, and A. B. Krueger. “Comment on David Neumark and William Wascher, ‘Employment effects of minimum and subminimum wages: Panel data on state minimum wage laws’”. *ILR Review*, 1994, 47(3): 487–497.

<sup>③</sup> Jordi Brandts, and Gary Charness. “Do Labour Market Conditions Affect Gift Exchange? Some Experimental Evidence”. *Economic Journal*, 2004, 114 (497), 684–708.

<sup>④</sup> M.F. Owens, J.H. Kagel. “Minimum Wage Restrictions and Employee Effort in Incomplete Labor Markets: An Experimental Investigation”. *Journal of Economic Behavior and Organization*, 2010, 73(3): 317–326.

可观察到的经济特征解释的工资差异部分则衡量了性别歧视的效应。<sup>①</sup>该理论假设男性工资决定方程为  $W_m = \beta_m X_m + e_m$ , 其中  $X_m$  表示可观察到的经济特征,如年龄、教育程度、工作经验、生产率等, $e_m$ 则表示无法观测的因素,  $m$ 和 $f$ 分别表示男性和女性。同样, 假设女性工资的决定方程为  $W_f = \beta_f X_f + e_f$ , 工资差异的 Oaxaca 分解可以用以下式子表明:

$$W_m - W_f = [\beta_m X_m + e_m] - [\beta_f X_f + e_f] = [X_m - X_f]\beta_m + [\beta_m - \beta_{nm}]X_{nm} \quad (6)$$

式(6)右侧第一项则衡量了由于经济特征导致的工资差异, 第二项衡量了不可观测因素导致的性别工资差异, 即工资性别歧视。

经验数据研究性别歧视的缺点体现在两个方面: 一是难以完全观察到工人生产率以及生产率的决定因素信息, 因此可能会产生遗漏变量偏误; 二是当家庭预期到未来劳动力市场的性别歧视时可能会选择改变人力资本投资计划, 例如减少对女性的教育投资, 这样工资差距可能会被生产率吸收, 歧视程度将被低估。由于严格控制环境因素和干扰变量, 实验研究方法不需要担心这些内生性偏误问题。已经有许多实验研究了性别歧视和就业不平等现象, 例如, Neumark 开展了一项实地实验研究美国餐饮业的雇佣性别歧视。<sup>②</sup>在实验中, 男性大学生和女性大学生被指派去 65 家餐馆申请服务业工作, 男性和女性大学生的简历基本一样, 并且性格和外貌差异非常小。研究结果证实了性别歧视会影响女性劳动力群体的就业。实验室实验方法通过使用更加严格的控制手段, 研究性别歧视以及歧视背后的决定因素, 为性别歧视和工资差异提供了新的思路。一系列实验研究发现, 性别歧视背后至少存在三个方面的决定因素: 一是女性更可能在工资谈判中退缩, 并且对于晋升缺乏偏好<sup>③</sup>; 二是女性风险厌恶程度更高, 导致很少参加锦标赛竞争以及难度更高的任务<sup>④</sup>; 三是劳动力市场存在对女性的刻板印象和偏见, 与其他因素无关。<sup>⑤</sup>SeEun et al. 建立了一个风险偏好与性别工资差异的理论模型, 并通过实验室实验分析对理论模型加以验证。<sup>⑥</sup>模型基本框架假设劳动力市场上存在两种工作类型: 稳定的工作和有风险的工作, 两种工作类型可以类比于国有企业与私营企业。稳定的工作假设工人  $i$  的绩效工资满足  $W_{si} = \gamma_s Y_i$ , 其中  $\gamma_s$  表示稳定的工作的计件工资率,  $Y_i$  衡量工人的生

<sup>①</sup> J. G. Altonji, R. M. Blank. "Race and Gender in the Labor Market". *Handbook of labor economics*, 1999, 3:3143-3259.

<sup>②</sup> David Neumark. "Sex Discrimination in Restaurant Hiring: An Audit Study." *Quarterly Journal of Economics* 1996,111(3): 915-41.

<sup>③</sup> Linda Babcock, and Sara Laschever. "Women Don't Ask: Negotiation and the Gender Divide". Princeton University Press: 2009.

<sup>④</sup> Niederle, and Lise Vesterlund. "Do Women Shy away from Competition? Do Men Compete Too Much?" *Quarterly Journal of Economics*, 2007, 122(3): 1067-1101.

<sup>⑤</sup> Ernst Fehr, and Armin Falk. "Wage Rigidity in a Competitive Incomplete Contract Market". *Journal of Political Economy*, 1999, 107 (1): 106-134.

<sup>⑥</sup> SeEun Jung, Chung Choe, Ronald L. Oaxaca. "Gender Wage Gaps and Risky vs. Secure Employment: An Experimental Analysis". *Labour Economics*, 2018, 52:112-121.

产率水平,并且满足  $Y_i = \mu_i + \varepsilon_i, \mu_i = E(Y_i)$  代表生产率均值,  $\varepsilon_i \sim i.i.d(0, \sigma_i^2)$ , 因此工人  $i$  选择稳定工作的均值和方差满足:

$$\begin{aligned} E(W_{si}) &= \gamma_s \mu_i \\ \text{Var}(W_{si}) &= (\gamma_s)^2 \sigma_i^2 \end{aligned} \quad (7)$$

工人选择风险工作则面临失业风险,失业概率为  $p$ ,失业状态下的收入为  $w_u$ ,可以理解为失业救济金。就业状态下的工资收入满足  $w_r = \gamma_r Y$ 。因此工人选择风险工作的工资均值和方差满足:

$$\begin{aligned} E(W_{ri}) &= p w_u + (1-p) \gamma_r E(Y_i) = p w_u + (1-p) \gamma_r \mu_i \\ \text{Var}(W_{ri}) &= p(1-p) [(\gamma_r)^2 (\sigma_i^2 + \mu_i^2) + (w_u)^2 - 2 \gamma_r \mu_i w_u] + (1-p)^2 (\gamma_r)^2 \sigma_i^2 \\ &= p(1-p) (\gamma_r \mu_i - w_u)^2 + (1-p) (\gamma_r)^2 \sigma_i^2 \end{aligned} \quad (8)$$

当两种类型工作期望工资相等  $E(W_{ri}) = E(W_{si})$  时,可以求解出补偿风险溢价如式(9)所示。补偿风险溢价的经济学含义是: 对于一个风险中性的工人来说, 工人在对两类工作选择偏好表现无差异时,两类工作的绩效工资满足的条件。

$$\begin{aligned} \gamma_s \mu_i &= p w_u + (1-p) \gamma_r \mu_i \\ \gamma_r - \gamma_s &= \left( \frac{p}{1-p} \right) \left( \gamma_s - \frac{w_u}{\mu_i} \right) \end{aligned} \quad (9)$$

补偿风险溢价  $\gamma_r - \gamma_s$  衡量了两类工作工资率之间的差异。由于实验中观察到女性比男性更有可能选择稳定的工作, 补偿风险溢价可以用来解释性别工资差异的 40%~77%。Kuhn & Villeval 在实验室开展了一项真实努力实验<sup>①</sup>, 研究发现女性更倾向于参与团队合作而不是竞争, 并且更加偏好于团队绩效与产出共享。Coffman et al.通过巧妙的实验设计<sup>②</sup>, 考察了信念对于性别歧视的影响。实验将 100 名工人根据出生月份和性别分成奇数月份出生的男性工人、偶数月出生的女性工人等四类, 其中, 出生月份被设计为与工人的生产率相关。雇主分成两类: 一类只能获得出生月份奇偶性信息, 另一类只能获得性别信息。对于两组雇主来说, 出生月份和性别相当于对于两组处理组贴上了“标签”, 雇主根据这些标签信息做出决策, 对于标签的选择反映了雇主对于性别和生产率的“信念”偏好。通过比较两组雇主的雇佣偏好, 实验结果发现雇主们更加倾向雇佣男性员工。

歧视现象也广泛存在于不同种族之间, 经常发生于多种族社会和移民国家。例如, 在美国对于黑人、亚裔、拉美裔等少数族裔的歧视现象层出不穷。与白人相比, 少数族裔群体往往在就业市场上面临不公平的待遇, 就业机会更少并且工资待遇更差。许多实验文献研究了种族歧视这一话题,

<sup>①</sup> Peter Kuhn, and Marie Claire Villeval. "Are Women More Attracted to Co-operation Than Men?". *Economic Journal, Royal Economic Society*, 2015,125 (582): 115-140.

<sup>②</sup> K B Coffman, C.L. Exley, and M. Niederle "The Role of Beliefs in Driving Gender Discrimination". *Management Science*, 2021.

Fershtman & Gneezy 在以色列对德系犹太人和东方犹太人进行了一项实验<sup>①</sup>，“信任博弈”和“最后通牒博弈”的结果均显示出东方犹太人与德系犹太人之间存在系统性的信任差异，该实验的结果可以用于解释以色列社会中存在的种族歧视现象。Chowdhury et al.在澳大利亚几所主要大学开展了一项关于种族歧视的实验研究。<sup>②</sup>在实验中将中国人和白人的姓名以及性别随机分配到真实的考试封面上，并另外招募大学生对这些试卷进行判分，结果发现，有中文名字的考卷平均得分要高于白人姓名的分数。实验结果说明人们对于中国学生具有某些刻板形象，即认为中国学生普遍会在考试中表现更优秀。除了性别种族之外，劳动力市场普遍还存在外貌歧视。Mobius & Rosenblat 运用实验室实验方法调查劳动力市场的“外貌溢价”<sup>③</sup>，实验结果表明，雇主在筛选简历时非常关注简历上的照片，求职者的外貌越突出，越容易受到面试邀请。

#### 四、如何在实验室研究劳动经济学

##### （一）实验基本注意事项

在实验室研究劳动经济学，需要考虑实验对象、实验环境、实验制度等诸多方面的问题。对于实验被试的选择要求，一般来说，需要挑选非经济专业本科学生作为标准的被试群体，因为经济专业大学生可能会学习过实验相关内容，了解实验目的，由此可能产生“霍桑效应”。实验设计一个最基本的问题是如何权衡“理论模型”和“现实环境”，即实验环境与现实环境的匹配程度。通常情况下，实验室实验需要考虑各种现实中可能出现的细节，但是也没必要完全模拟现实，这样会导致过高的实验成本。与实验环境相关的问题是如何向实验被试提供实验规则相关的信息帮助他们理解实验框架。实验中的解释术语包括中性和非中性的语言，两种语言都有自身的优点和缺点，应该依据具体的研究问题进行选择。关于实验框架效应的一个经典的例子是 Liberman et al.开展的“囚徒困境”博弈实验<sup>④</sup>，实验人员对于两个处理组分别给被试提供“华尔街游戏”和“社区游戏”的标签，实验结果显示在“华尔街游戏”组内合作率低于 30%，而在“社区游戏”组内合作率超过 70%。实验设计中第二个需要考虑的问题是被试内设计(within-subjects designs)与被试间设计(between-subjects designs)问题。被试内设计的含义是每个被试接受一系列的处理；被试间设计指的是不同被试只接受一种处理。两者各

<sup>①</sup> C. Fershtman, and U.Gneezy. “Discrimination in a Segmented Society: An Experimental Approach”. *Quarterly Journal of Economics*, 2001, 116 (1): 351–377.

<sup>②</sup> Chowdhury Shyamal, Klauzner Ilya, and Slonim Robert. “What's in a Name? Does Racial or Gender Discrimination in Marking Exist?” Institute of Labor Economics, Working Paper, 2020.

<sup>③</sup> Markus M. Mobius, Tanya S. Rosenblat. “Why Beauty Matters”. *American Economic Review*, 2006, 96 (1), 222–235.

<sup>④</sup> V. Liberman, S. Samuels, and L. Ross. “The Name of the Game: Predictive Power of Reputations Versus Situational Labels in Determining Prisoner's Dilemma Game Moves”. *Personality and social psychology bulletin*, 2004, 30: 1175–85.

有优势和不足：被试内设计可以减少个体异质性的影响，但是存在顺序效应的问题。而被试间设计虽然避免了顺序效应，但是成本较高。一个折中的办法是将两种设计方法组合起来，先在被试内设计中改变处理的顺序，然后使用首次处理的横截面数据作为被试间设计的实验结果，用于被试内设计的稳健性检验。实验设计中第三个需要注意的问题是如何支付报酬。通常有两种支付实验报酬的方式，一种是在每个环节为被试做出的决策支付报酬，另一种是在整个实验中随机挑选某一（些）环节的某个（些）决策给予报酬。前者会受到收入效应和参照点效应的影响，即被试在前两轮获得报酬后可能会对后面的实验环节产生懈怠，或者被试根据已经获得的报酬会对后面实验环节的收入目标存在预期。随机支付报酬的方法避免了收入效应，因此在实验报酬设计中越来越流行。在解决上述实验设计问题后，根据引致价值理论<sup>①</sup>，被试在实验中能够表现出与现实环境中相似的行动和策略。

在实验室研究劳动经济学需要考虑更具体的问题，例如委托代理关系、实验任务形式、劳动力市场模拟和角色扮演。在委托代理关系中，委托人（通常是实验者）指派一份“合同”，代理人（实验被试）选择行动和策略。如何在实验室模拟劳动市场呢？通常的办法是进行拍卖游戏。一个具体的例子是，Fehr et al.设计了一个两阶段游戏：第一阶段，公司（实验者）提出工资报价，如果一名工人（被试）接受了这份工资合约，则相当于与公司签订了劳动合同；第二阶段，工人选择努力程度进行生产。<sup>②</sup>关于工人如何选择努力程度，一些实验使用某种形式的真实努力(real effort)而另一些实验使用选择努力(stated effort)。真实努力指的是实验被试执行某些任务（如求解数学问题、迷宫游戏等）作为自己的努力水平，而选择努力指被试者在预先制定的范围内选择一个数字来代表他们的努力水平，较大的数字代表努力水平较高。两种设计各有优势，选择努力的优势在于能够允许实验者根据实验目的操纵被试努力的成本，从而排除被试个人特征（如风险厌恶程度、社会偏好等）对于实验结果的干扰，并且更容易衡量和测算努力水平。真实努力的优势在于，实验任务更符合大多数人所认知的劳动，因此往往被认为具有更强的外部有效性。最近的研究发现，两种努力设计并不会导致实验结果产生明显差异。<sup>③</sup>另一个需要考虑的实验设计问题是谁扮演委托人的角色。最简单且成本最低的方案是所有实验被试都扮演代理人，实验者充当委托人。考虑到实验被试更倾向于将实验者而不是实验室中的同学当作雇主，而且大学生缺乏当公司管理人员的经验，这种设计可能更加

<sup>①</sup> V. L. Smith. "Microeconomic Systems as An Experimental Science". *The American Economic Review*, 1982, 72(5): 923-955.

<sup>②</sup> E. Fehr, G. Kirchsteiger, and A. Riedl. "Does Fairness Prevent Market Clearing? An Experimental Investigation". *The Quarterly Journal of Economics*, 1993, 108(2): 437-459.

<sup>③</sup> E.G Dutcher, Timothy C. Salmon, and Krista J. Saral. "Is 'Real' Effort More Real?". Available at SSRN Working Paper, 2020.

贴近劳动市场的真实情景。另一种选择委托人的方案是选定一部分大学生被试扮演委托人的角色。这一方案在某些情况下也有优势，例如实验的目的是为了观察除实验者以外的人的社会偏好，或者是对“委托人”本身的行为感兴趣，这个时候指定学生被试作为委托人是必要的。

## (二) 礼物交换实验

Fehr et al.最早提出礼物交换实验用于测试公平工资努力假说，即解释员工付出超过经典经济理论预测的工作努力程度的行为。如今礼物交换实验被广泛应用于劳动实验经济学的各项具体研究领域，包括社会偏好、计件工资、委托代理关系等。我们首先介绍礼物交换实验的理论基础，然后讨论礼物交换实验的具体实施步骤。

考虑 Della et al.提出的一个基本模型，工人*i*通过选择最优努力程度 $e_i$ 使得终身贴现效用最大化：

$$\max_{e_i \geq 0} u(e_i) = w_0 + we_i - C_i(e_i) + \Theta(Gift, y, w)e_i \quad (10)$$

其中， $w_0$ 表示企业给工人的一次性转移支付，可以理解为基本工资。绩效工资收入取决于计件工资率 $w$ 和努力程度 $e_i$ ，工人的成本函数为 $C_i(e_i)$ ，可以看出，对于不同工人其成本函数形式存在差异，且成本函数满足 $C'(\cdot) > 0, C''(\cdot) > 0, \lim_{e \rightarrow \infty} C'(e) = \infty$ 。效用函数的第三部分捕捉了社会偏好的作用，其中系数 $\Theta$ 表示社会偏好系数，衡量工人对企业利润的关心程度，取决于雇主给予工人未预期到的“礼物”(Gift)。对式(10)求解一阶最优条件可得：

$$e_i^*(Gift, y, w) = C^{-1}(w + \Theta(Gift, y, w))$$

其中， $C^{-1}$ 表示成本函数的反函数，根据反函数性质可知 $C^{-1}$ 也是单调递增的， $y$ 表示单个工人的生产率，即产出。式(10)的二阶最优条件满足 $-C''(e^*) < 0$ 。根据式(11)，工人的最优努力 $e^*$ 随着社会偏好系数 $\Theta$ 和计件工资 $w$ 的增加而增加。考虑社会偏好的两种情况：纯利他主义和“温情效应”。在纯利他主义情况下，工人考虑到雇主的净回报 $y - w$ ，其社会偏好系数为 $\Theta = \alpha(y - w)e_i$ ，其中 $\alpha$ 表示的含义是，利他主义的工人将把雇主赚的每一单位收入都视作自身的 $\alpha$ 单位的收入。再根据互惠模型，有 $\Theta = \alpha + 1_{Gift}\alpha_{Gift}$ ，这说明社会偏好系数会受到“礼物”的影响。在“温情效应”情况下，工人只关注雇主的平均回报 $\bar{y}$ ，当工人收到礼物后，社会偏好系数为 $\Theta = (a + a_{Gift}1_{Gift})\bar{y}$ ， $a$ 的含义类似于纯利他主义情形。在结合利他主义和温情效应之后，工人的最优努力为：

$$\begin{aligned} e_{Contr}^* &= C^{-1}(\alpha p_E + a\bar{y}) \\ e_{Gift}^* &= C^{-1}((\alpha + \alpha_{Gift})y + (a + a_{Gift})\bar{y}) \end{aligned} \quad (12)$$

其中， $e_{Contr}^*$ 和 $e_{Gift}^*$ 分别表示控制组的工人最优努力和礼物交换实验组工人的最优努力。特别地，考虑具体的成本函数形式如 $c(e) = ke^{1+\gamma}/(1 + \gamma)$ 。则在纯“利他主义”情形下，工人的最优努力分别

为:

$$e_{\text{Contr}}^* = \left(\frac{\alpha y}{k}\right)^{\frac{1}{\gamma}}, \quad e_{\text{Gift}}^* = \left(\frac{(\alpha + \alpha_{\text{Gift}})y}{k}\right)^{\frac{1}{\gamma}} \quad (13)$$

通过测量生产率 $y$ 、成本函数参数 $k$ 和 $\gamma$ ，再结合工人最优努力值，可以估算出系数 $\alpha$ ，进而可以测算出社会偏好系数 $\theta$ 。

依据上述礼物交换实验的理论模型基础，我们简要介绍礼物交换实验的设计步骤：

(1) 招募被试：招募不同专业的大学生作为被试。为了避免顺序效应，对每位招募到的被试均只实施一次任务。招募到的被试被随机分在处理组(有礼物交换组)和对照组。

(2) 实验环境：为实验被试模拟一个与现实相似的工作情境框架，其中涉及公司雇主角色和员工角色的介绍与分工。

(3) 实验内容：考虑一个三人礼物交换实验，实验中一位参与者扮演公司雇主角色，另外两名参与者扮演员工角色进行匹配，两名员工的生产率不同，一个生产率较高，另一个较低。公司首先向员工提供工资 $w_i$ ，取决于员工的类型，然后员工选择一个努力水平。假设公司能够获得4个单位的收入，他可以选择给员工工资：0、1、2、3、4，相应地，员工可以选择的努力程度为：完全不努力、低、中、高。员工的报酬取决于雇主提供的工资和自己的努力程度投入，报酬在实验中用一定数量的代币衡量，实验结束后可以兑换为现实货币或者奖品。

(4) 数据分析：通过员工的努力程度和公司支付给员工的工资，来考察礼物交换的影响并识别其中的社会偏好行为。

### (三) 团队竞争与合作实验

现代组织的绩效越来越依赖于运作良好的团队。项目的顺利完成离不开每位团队成员的积极参与和团结协作。但是在现实环境中，由于工资歧视、晋升竞争等问题的存在，往往很难将团队内部的员工有效地组织协调起来，充分发挥团队优势从而为公司创造更高的利润和价值。因此，如何设计出适合团队竞争与协作的科学有效方案是至关重要的，既要做到激励团队内部每个员工积极表现，同时也需要协调团队内部积极提供互助。虽然研究企业内部团队竞争与合作的重要性不言而喻，但是我们难以通过简单的实证策略来考察相关绩效激励机制的作用和机理。一方面，企业内部的统计数据由于涉及企业隐私可能难以获得；另一方面，通过简单的计量方法识别绩效激励方案的因果关系难以排除内生性因素的干扰。相较之下，实验室实验方法可谓是尽展所长，既能在实验室环境下通过随机分组和严格的控制进而准确识别绩效激励对团队竞争与合作的因果关系；亦可灵活地设计

不同绩效激励组合，拓展相关理论；并且不需要担忧数据的可获得性。我们首先构建一个团队竞争合作的理论模型：

考虑一个由 4 个成员参与的团队，每个员工同时行动。<sup>①</sup>员工  $i \in \{1, 2, 3, 4\}$  可以选择给同事  $j \in \{1, 2, 3, 4\}$ ,  $j \neq i$  提供帮助或者阻碍，即积极配合或者故意不配合行为，用  $a_i \in [-a_0, a_0]$  表示， $a_i > 0$  表示协助，反之则为不配合。 $a_i$  会对其他团队员工业绩造成影响。此外，员工  $i$  选择努力程度  $e_i$  进行工作。因此，员工  $i$  的业绩(产出)  $y_i$  可表示为：

$$y_i = \sum_{j \neq i} a_j + e_i + \varepsilon_i \quad (14)$$

其中， $\varepsilon_i$  表示误差项，在区间  $[-\varepsilon_0, \varepsilon_0]$  上满足独立随机均匀分布。假设员工采取行动和工作努力的成本函数分别为： $C_a(a_i) = ma_i^2$ ， $C_e(e_i) = ne_i^2$ ， $m, n > 0$ 。员工的收入由三部分构成：基本工资  $w_0$ ，团队综合奖金  $t \sum_{j=1}^4 y_j$  (每位员工获得相同数量) 以及个人绩效奖金  $\sigma p_i(a_i, a_{-i}, e_i, e_{-i})$ ，其中， $p_i(a_i, a_{-i}, e_i, e_{-i})$  代表个人获得奖金的概率，取决于收到的其他同事的帮助（或阻碍）和自身对其他同事的  $a_j$  投入，以及其自身工作的努力程度和同事的努力程度。因此，员工  $i$  的净收入方程表示为：

$$\Pi_i = w_0 + t \sum_{j=1}^4 (3t_j + e_j + \varepsilon_j) + \sigma p_i(a_i, a_{-i}, e_i, e_{-i}) - C_a(a_i) - C_e(e_i) \quad (15)$$

对式 (15) 求一阶最优条件可得：

$$a_i = \frac{3t}{2m} + \frac{\partial p_i(a_i, a_{-i}, e_i, e_{-i})}{\partial a_i} \frac{\sigma}{2m}, \quad e_i = \frac{t}{2n} + \frac{\partial p_i(a_i, a_{-i}, e_i, e_{-i})}{\partial e_i} \frac{\sigma}{2n} \quad (16)$$

假设随机误差项  $\varepsilon_i$  满足随机均匀分布，则可证明得到个人奖金的边际概率仅取决于随机误差项区间大小。<sup>②</sup>因此有：

$$\frac{\partial p_i(a_i, a_{-i}, e_i, e_{-i})}{\partial a_i} = -\frac{1}{2\varepsilon_0}, \quad \frac{\partial p_i(a_i, a_{-i}, e_i, e_{-i})}{\partial e_i} = \frac{1}{2\varepsilon_0} \quad (17)$$

结合式 (16) 与式 (17) 可得，员工最优决策的两个选择分别为：

$$a_i^* = \frac{3t}{2m} - \frac{\sigma}{4m\varepsilon_0}, \quad e_i^* = \frac{t}{n} + \frac{\sigma}{4n\varepsilon_0} \quad (18)$$

根据式 (18)，团队综合奖金对于促进员工互相合作以及增加工作努力均起到促进作用，而个人奖金则会抑制团队互助行为，但会提高员工个人工作努力程度。因此，企业可以将团队奖励和个人

<sup>①</sup> Anastasia Danilov, Christine Harbring, and Bernd Irlenbusch. "Helping under A Combination of Team and Tournament Incentives". *Journal of Economic Behavior & Organization*. 2019,162:120-135.

<sup>②</sup> C. Harbring, and B. Irlenbusch. "Incentives in Tournaments with Endogenous Prize Selection", *Journal of Institutional and Theoretical Economics*, 2005,161, 636-663; A. Orrison, A. Schotter, and K. Weigelt "Multiperson Tournaments: An Experimental Examination", *Management Science*, 2004,50, 268-279.

奖励方案有机结合,设计出既激励员工个人提高工作努力又鼓励员工互相帮助、协调配合的绩效奖励方案,从而获得最大化的企业利润。依据团队竞争与合作的理论基础,我们可以设计相应的实验探究激励方案对员工行为和团队产出的影响。具体实验设计步骤如下:

(1) 招募被试:根据之前的理论基础,我们需要招募来自不同专业的大学生被试者,并随机分成处理组和控制组,每组4人。同时,为了避免顺序效应和学习效应,处理组每位被试都只进行一次实验处理。此外,为了使实验处理结果具有可比性,所有处理组被试均得到完全相同的实验指导。唯一的区别是团队综合奖金系数和个人奖金系数存在差异。

(2) 实验环境:我们需要为实验被试模拟一个与现实相似的工作情境框架,并为被试讲解术语“合作互助”与“不配合”的含义,将“努力”称为“工作”。

(3) 实验内容:实验被试根据程序设计的问卷测试问题模拟个人工作行为。例如,每个“员工”从集合{0, 1, ... 10}中选择一个数字代表其工作努力,同时还需要再选择一个数字用于对其同事进行帮助或者阻碍。被试回答所有有关个人决策和回报之间关系的测验问题之后,将计算实验报酬。实验被试的报酬分为三个部分:基本工资、团队综合奖金以及个人奖金。奖金用实验代币来衡量,实验结束后可以兑换现实货币或者奖品。

(4) 实验分析:根据实验处理组和控制组的数据进行分析,具体考察绩效激励对团队合作和竞争的影响。

## 五、总结与展望

本文首先详细论证利用实验室实验方法研究劳动经济学的优势和科学性,然后具体总结了实验室实验的设计原则和需要注意的问题,最后结合理论模型和具体实验研究案例,系统梳理了劳动经济学各个领域的实验室实验研究成果。得益于研究方法的科学性、精确控制环境因素,低成本且容易复制等一系列优点,实验室实验方法越来越受到劳动经济学家们的青睐。目前,国外劳动经济学家已经做了大量丰富的实验室实验,用于研究薪酬激励、社会偏好与互惠行为、团队竞争合作、雇佣关系和最低工资制度以及劳动力市场歧视等重要话题。而在国内,运用实验室实验方法研究劳动经济的实例还几乎是空白。实验劳动经济学的发展为我国相关领域的研究带来诸多机遇和挑战。一方面,由于政治体制、经济制度和文化背景的差异,可以预见,使用国内的大学生群体当作被试可能得到与国外研究不同的结论,实验室实验研究方法为丰富国内劳动经济领域相关问题的研究带来

契机；另一方面，如何通过科学合理的实验设计将制度、经济、文化等因素考虑在内，满足创新性的同时又紧密结合国内具体研究问题，仍然是一项具有挑战的任务。

针对目前劳动实验经济学在国内的发展现状，本文结合我国具体国情和现实背景，提出以下三点研究展望：（1）在雇佣关系和搜寻匹配理论研究主题上，可以结合我国劳动力市场的结构特征，重新检验和拓展动态搜寻匹配理论并探索最优失业保险制度。由于户籍制度的限制，我国的劳动力并非自由流动的，并且劳动力结构中存在大量的农民工群体，这些典型事实的存在使得重新构建我国劳动力市场的雇佣理论非常必要。另外，鉴于我国目前的失业保险体系现状，实验室可以充当失业保险制度改革的“试验田”，即提前在实验室对制度改革方案进行预实验测试，待评估后再推广到真实环境中。（2）自 1993 年起，我国才逐步建立现代化的企业制度。基于我国独特的“以公有制为主体、多种所有制经济共同发展”的基本经济制度，国有企业与民营企业之间存在着诸多差异，例如，员工激励、团队生产等。因此，研究我国不同企业类型内部的委托代理关系、薪酬激励、合作竞争以及社会偏好等话题，可能会产生许多有趣且重要的结论。（3）实验室实验为研究我国的劳动力市场歧视提供了一个方便的渠道。因此，可以在实验室探究民族、地域以及学历等背景因素对于劳动力市场歧视现象的作用，拓展相关领域的研究。总而言之，劳动经济学的实验室实验方法在国内“大有可为”，利用实验室实验研究方法，结合我国具体政治、经济、文化背景，研究劳动经济各个领域话题具有重大意义。

# Research of Laboratory Experimental Methods in Labor Economics: Theory and Experiment

DAI Zhixin, CHENG Peng, GAO Hongyu

(School of Finance, Renmin University of China)

**Abstract:** This paper introduces a scientific research method to study labor economics, that is, laboratory experiment method. We analyzed the advantages of this approach. We explain the principles of experimental design and matters needing attention in laboratory research of labor economics. Then, combined with classical theoretical models and specific research cases, we systematically review the experimental literature in specific research fields such as wage incentive, social preference and reciprocity, team competition and cooperative interaction, employment relationship and minimum wage, and labor market discrimination. The laboratorial experiment research method of labor economics has been very popular in foreign economic circles, but it is still in its infancy in China. Combining the background of Chinese culture, enterprise system and labor market structure, we provide three research ideas and prospects for the application of labor-economics laboratory experimental research methods in China.

**Key words:** Labor economics; Laboratory experiment; Experimental economics

## “乡财县管”改革、基层政府治理与经济发展

马光荣<sup>1</sup>

中国人民大学财政金融学院，中国财政金融政策研究中心

张玲<sup>2</sup>

北京大学经济学院

**摘要：**县乡两级政府处于我国政府层级的最末端，同时又是大量公共事务的最终承担者，科学合理的县乡财政管理体制，是提升基层治理能力的重要条件。2004 年以来，我国一些省份相继推行了乡财县管改革，由县级政府直接管理并监督乡镇财政收支，乡镇财政管理权被大幅度上收到了县。本文利用 2000-2019 年的县域面板数据和双重差分模型，评估了乡财县管改革对经济发展的影响。结果显示，乡财县管促进了县域经济发展。机制分析显示，乡财县管减少了行政管理支出比重，抑制了财政供养人员规模膨胀，降低了税费负担。这些结果表明，加强上级政府对基层政府的财政管理与监督，有助于改善治理水平、促进经济发展。本文也发现，在地形较为复杂的人口大县及经济较发达地区的县，由于改革提高了县级政府获取信息的难度以及不利于发挥乡镇政府的积极性，因此这些县实施乡财县管对经济发展的提升效果相对较弱，这说明乡财县管不宜一刀切地实行。

**关键词：**乡财县管 乡镇政府 政府治理 财政分权

**JEL 分类号：**H71 H3 H2      **文献标识码：**A      **文章编号：**

---

**作者简介：**马光荣，经济学博士，教授，中国人民大学财政金融学院、中国财政金融政策研究中心，Email: grma@ruc.edu.cn.

张玲（通讯作者），博士研究生，北京大学经济学院，Email: zzzl@pku.edu.cn.

\*感谢国家社会科学基金重大项目（批准号：21&ZD095）和国家自然科学基金面上项目（批准号：71973142）的资助。感谢匿名审稿专家的宝贵意见，文责自负。

## 一、引言

“基层强则国家强，基层安则天下安”。县乡两级政府处于我国政府层级的最底端，同时又是大量公共事务的最终承担者，县乡两级财政支出占全国的比重接近 50%，可谓“权力最小、责任最大”（周黎安，2017）。党的十八大以来，党中央提出，财政是国家治理的基础和重要支柱，科学的财政体制对于优化资源配置、维护市场统一、促进社会公平、实现国家长治久安具有重要作用。作为财政体制最基层的组成部分，县乡财政是国家治理“基础中的基础”。科学合理的县乡财政管理体制，是充分发挥基层政府积极性、提升公共服务水平、推动经济社会发展的重要制度保障。

改革开放后，我国县乡两级财政管理体制经历了两次大的改革。1983 年，随着人民公社体制的解体，乡镇政权重新确立，按照“一级政府、一级财政”的原则，乡镇一级财政也相应建立，乡镇政府拥有了自主的财政收支与预决算体系，形成了“乡财乡管”体制。但是，随着 1994 年分税制改革后财力大幅度上移，以及 2000 年后农村税费改革的实施，乡镇收支矛盾日益加剧，乡镇财政运行暴露出支出缺乏规范、财政供养人员膨胀、正常公共服务支出难以保障等问题。一些乡镇政府为了筹集资金还存在违规收费，甚至大规模举债的问题，加重了居民和企业的负担（刘克崧和贾康，2009）。为了化解这些问题，2003 年以来，从安徽省开始，一些省份开始探索实施“乡财县管”改革。2006 年财政部发布《关于进一步推进乡财县管工作的通知》，中央开始明确鼓励各地推进乡财县管改革。2022 年《国务院办公厅关于进一步推进省以下财政体制改革工作的指导意见》指出，进一步推动乡财县管工作提质增效，提升乡镇财政管理效率和水平。

乡财县管改革，是由县级财政部门直接管理并监督乡镇财政收支，主要内容是实行“预算共编、账户统设、集中收付、采购统办、票据统管、县乡联网”。改革后，通过“预算共编”，乡镇预算安排和预算调整都很大程度上由县级政府决定，乡镇政府财政收支的自主性大幅度下降。与此同时，县级政府大大加强了对乡镇预算执行过程的监督管理，取消了乡镇财政的账户，乡镇所有收支都通过县级国库集中收付，乡镇采购支出都通过县级统一办理。乡财县管后，尽管县乡财政体制仍保留“一级政府、一级财政”的形式，而实质上乡镇从独立收支的“一级财政”弱化为“半级财政”，乡镇财政管理权被大幅度上收到了县（杨发祥、刘楠，2012）。因而，从本质上看，乡财县管是一项上收财政权限的改革。随着乡镇财政的虚化，县级政府开始直接负担

大量公共服务的支出，从这个意义上看，乡财县管也是一项财政层级扁平化的改革（贾康，2007）。

本文试图回答的问题是，乡财县管改革是否改善了基层政府治理和公共服务提供能力，最终促进了经济发展？在实践中，尽管全国相当多的地区都已经先后实施了乡财县管改革，但是乡财县管改革是否取得了如期的效果，尚没有定论。乡财县管作为一项财政权限上收的改革，其对政府治理和经济发展的效应，从理论上看不确定的（Mookherjee, 2015）。首先，根据财政分权的传统理论，分权相比于收权有诸多优势：相比于县级政府，乡镇政府具有本地信息优势，存在更为激烈的辖区间竞争，独立自主的财政收入也激励乡镇政府发展经济。因此从这些角度看，乡财县管不利于发挥乡镇政府的积极性。但是，财政分权理论的另一一些文献则认为，由于对基层行政权力的监督机制通常较为薄弱，将权力过多的赋予基层政府，反而会恶化基层治理水平。乡镇作为中国最基层的政府，管理能力不足、监督机制失效等问题相比于县级政府更容易发生。因此从这些角度看，乡财县管有利于改善基层治理。事实上，很多地方推行乡财县管改革的出发点，正是解决乡财乡管体制下暴露出的乱花钱、乱进人、乱收费、乱举债等乡镇治理恶化问题。

本文利用 2000 年-2019 年的县域层面数据，使用双重差分模型，评估了乡财县管改革对经济发展的影响。结果显示，改革显著促进了经济发展。实施乡财县管改革后，县域人均 GDP 平均提升了 4.8%。乡财县管对一、二、三产业的发展都有提升作用，对第二产业的提升作用最强。不仅如此，我们也发现乡财县管对农村人均纯收入、人均储蓄、企业数量等经济发展指标均有正面效应。

进一步的机制分析结果显示，乡财县管通过规范财政支出与降低税负负担两个方面有利于经济发展：首先，改革降低了财政支出当中行政管理类支出的比重，抑制了财政供养人员的膨胀，从而保障了正常公共服务支出；其次，改革减少了政府收费，减轻了私人部门的税费负担。财政支出结构的优化和税费负担的降低，又进一步促进了当地投资的增长。这些结果表明，对于最基层的政府而言，通过财政权限上收，加强上级对下级的财政管理与监督，有助于改善政府治理水平，促进经济发展。

本文进一步进行了异质性分析。结果发现，在人口较多的大县和地形较为崎岖的县，乡财县管改革对经济发展的正面效果会减弱。我们认为，这些县实施乡财县管后，县级政府直接提

供公共服务时面临管理幅度大、获取信息成本高的问题。同时我们也发现,对于初始经济较为发达、基层财力较强的地区,由于不利于调动乡镇政府的积极性,乡财县管的正向效果也较弱。这表明,尽管总体而言,乡财县管改革带来的正面效应大于负面效应,但是对于部分县而言,改革存在一定代价。乡财县管改革不应一刀切地去推行,不宜在人口较多、内部情况较为复杂、经济较为发达的县推行乡财县管。事实上,近年来一些地区,如浙江省丽水市和辽宁省,正是鉴于乡财县管暴露出的问题,尝试部分地改回“乡财乡管”,扩大乡镇财政的自主权,发挥乡镇政府的积极性<sup>①</sup>。

本文的研究价值在于以下三个方面:首先,以往关于乡财县管的研究,大多数都集中于定性研究(夏杰长、陈雷;2005),李永友和王超(2020)利用安徽省数据研究了乡财县管改革对城乡收入差距的影响,本文则首次使用全国范围内的数据定量评估了乡财县管对经济发展和地方政府行为的影响,有利于正确认识这一改革的利弊得失,对于下一阶段县乡政府财政体制改革也具有参考意义。其次,以往关于中国财政分权经济效应的实证研究都集中于分析中央与省,或省与市、县之间的财政关系(Xu, 2011),但是却鲜有文章分析县乡两级财政关系。本文将政府间财政关系的研究拓展到了最基层的县乡两级财政关系。本文的研究的确表明,在最基层的县乡政府间财政关系上,适度的财政权限上收有利于经济发展,从而进一步揭示了政府间财政体制安排需要权衡分权与收权的相对成本与收益。最后,乡财县管与省直管县财政体制改革类似,均属于财政层级的扁平化改革。已有大量研究关注省直管县财政体制改革的效果(郑新业等,2011;贾俊雪等,2013;王小龙和陈金皇,2020),本文则从乡财县管改革的角度剖析了财政层级扁平化改革效果。我们认为,乡财县管与省直管县尽管同为财政层级扁平化,但是前者具有收权性,后者具有分权性。尽管大多数研究发现省直管县作为分权改革促进了经济发展,但是本文发现乡财县管作为收权改革,也促进了经济发展。本文认为,其原因在于,乡镇作为最基层政府,易存在政府失灵,因而适度的收权改革,有助于提高基层治理水平。

## 二、文献回顾

政府间财政关系的理论基础是财政分权理论,该理论发展大体可分为两个阶段,即第一代和第二代财政分权理论(Weingast, 2009)。第一代财政分权理论秉承了 Hayek(1945)和 Tiebout(1956)的思想,认为地方政府拥有本地信息优势,相比于上级政府更了解辖区内居民需求,

<sup>①</sup> 浙江省丽水市自2013年,辽宁省自2018年,启动将乡财县管改回“乡财乡管”。

由地方政府更多地承担公共服务职能可以改善社会福利、提高经济效率。Oates (1972) 将第一代分权理论总结为分权定理：当公共服务不存在跨地区外部性时，地方政府提供比中央政府提供更有效率。但是这一理论暗含了一个重要假设，即政府官员是完全仁慈的，不存在自私动机，他们愿意且能够依据本地居民偏好提供高效的公共服务。无疑，这一条件并不完全符合现实。

从 1980 年代开始兴起的第二代财政分权理论放松了地方政府官员完全仁慈这一假设，认为财政分权会改变地方政府官员面临的激励，从而影响其行为模式，进而影响社会福利和经济效率 (Mookherjee, 2015)。但是，第二代财政分权理论的文献，对分权的效应，存在截然相反的观点。第一种观点认为，分权对地方政府官员产生正向激励：分权可以促进地方政府间竞争，硬化地方政府的预算约束，分权带来的财政收入激励也会促使地方政府努力发展经济、保护市场发展 (Weingast, 2009)。第二种观点则认为，地方层面通常更缺乏有效的权力监督制约机制，舆论监督机制薄弱，且人数少、利益更为同质的人群更容易结成利益集团，从而地方政策更容易受利益集团 (Bardhan and Mookherjee, 2005)。

在关于中国分权体制的研究中，一些文献显示分权促进了中国经济增长 (林毅夫和刘志强, 2000; 沈坤荣和付文林, 2005; 张晏和龚六堂, 2005)，但也有文献发现财政分权抑制了经济增长 (Zhang & Zou, 1998)，或者不存在显著影响 (Davoodi & Zou, 1998; Xie et al., 1999)。既有研究结论大相径庭，主要是由于不同的研究使用了不同时间跨度的样本，以及不同的财政分权度量指标。

关于中国财政分权的经济效应，现有实证研究也存在多方面问题未得到充分解决：首先，大多数国内文献使用分省的人均财政收支与中央政府在全国的人均财政收支之比，度量分省的财政收支分权程度，但是同一年份所有地区该指标的分母都相同，因此该指标更多反映了各地财政收支的相对规模，而不是分权程度高低 (徐永胜和乔宝云, 2012; 陈硕和高琳, 2012)；其次，文献中度量的财政分权维度不同，绝大多数研究度量的是中央与省之间的分权程度，另外一些度量的是省与市、县政府的财政分权程度 (吕冰洋等, 2016)。但是鲜有文献研究县、乡级政府之间财政关系的经济效应。不同层级上的分权，对经济发展的影响可能不尽相同。正如 Bardhan and Mookherjee (2005) 等理论研究所强调的，越是基层地方政府，对政府权力的监督约束机制越是薄弱，将权力分给越基层的地方政府，其效果可能越差。因此，中央将财政收支分权到省，与县将财政收支分权到乡镇，效果可能存在差异。

已有关于“乡财县管”改革的研究大多从定性的角度，分析“乡财县管”改革的必要性和可行性、利弊梳理及改革方向 (吕连生, 2005; 张威和高瑞, 2010; 杨发祥和刘楠, 2012)。夏杰长、陈雷 (2005) 以安徽省 G 县的乡财县管改革为例进行案例研究，从社会学角度分析了乡财县管改革的成效和问题。李永友和王超 (2020) 利用安徽省数据研究了乡财县管改革对城乡收入差距的影响，但是，目前还没有研究利用全国大样本的数据，定量考察乡财县管改革的经济社会效应。

### 三、制度背景和理论分析

#### (一) 制度背景

改革开放后,随着人民公社体制的解体,乡镇政权重新确立。1983年,中共中央、国务院发出《关于实行政社分开建立乡政府的通知》,要求“随着乡政府的建立,应当建立乡一级财政和相应的预决算制度,明确收入来源和开支范围”。1985年,财政部根据中央建立乡级财政的精神,制定了《乡镇财政管理试行办法》,进一步明确了“一级政府、一级财政”、“统一领导、分级管理”的原则,“乡财乡管”体制开始形成。乡镇政府的财政收入由预算内资金和预算外资金组成,预算内资金来自于各类税收和上级的转移支付,预算外资金来自于村提留、乡统筹等收费以及乡镇企业利润等。乡镇政府承担了乡镇机构行政管理费、文教科学卫生事业费、农林水事业费等主要支出。

1994年分税制改革后,随着财政收入大幅度向中央集中,处于最底层的乡镇财政,财政收入更难以保障。2000年以来,随着农村税费改革和取消农业税政策的实施,以及乡镇企业的逐渐衰落,乡镇财政困难程度进一步加剧。由于乡镇政府自身治理水平有限,在资源向上集中、自身财力削弱、支出责任下移等多重压力下,一些地区乡镇财政运行产生了“乱花钱、乱进人、乱收费、乱举债”等若干乱象(陈锡文,2002;刘克固和贾康,2009)。首先,基层行政机构臃肿,人员超编严重,造成“生之者寡、食之者众”,乡镇财政沦为“吃饭财政”,难以保障教育、医疗、水利、基建等正常公共服务支出;其次,财政支出过程中科学化程度较低,预算制定与审查过程流于形式,支出难以满足当地的需要;第三,财政预算执行中存在管理漏洞,票据不合规、账户不规范、财务不透明等现象突出;第四,基层政府为了弥补财政收入不足,仍然存在乱收费、乱摊派的现象,加重了当地居民与企业的负担;第五,一些乡镇政府债务积累严重,影响了财政正常运行。

在上述背景下,一些地区为了规范乡镇财政收支行为,开始探索实施“乡财县管”财政体制改革。安徽省2003年率先选取9个县进行乡财县管改革试点,2004年在全省范围内全面推行改革,其他一些省份也纷纷跟进。财政部在2006年发布了《关于进一步推进乡财县管工作的通知》,鼓励各地区推行“乡财县管”。根据本文收集的各地方政府相关文件,2004年-2009年是乡财县管改革试点的大规模推广时期,试点县占比从2004年的7.3%逐步增加至2009年的79.2%。

乡财县管改革的主要内容是：县级财政部门在预算编制、账户设置、集中收付、政府采购和票据管理等方面，对乡镇财政进行管理和监督，实施“预算共编、账户统设、集中收付、采购统办、票据统管、县乡联网”（刘克崧和贾康，2009）。改革后，通过“预算共编”，乡镇预算安排和预算调整都很大程度上由县级政府决定，乡镇财政自主性大幅度下降。县级政府大大加强了对乡镇预算执行过程的监督管理，取消了乡镇财政的账户，乡镇所有收支都通过县级国库集中收付，乡镇采购支出都通过县级统一办理，乡镇行政事业性收费票据由县级统一管理。与此同时，“乡财县管”不仅上收了财权，也在很大程度上收了事权，原先很多由乡镇政府承担教育、医疗、基建等公共服务，更多的由县级政府或其派出机构直接承担，乡镇政府功能进一步虚化（夏杰长、陈雷,2005）。因此，乡财县管改革尽管从法律形式上还保留了乡镇财政收支，但是随着改革的推进，实质上乡镇财政收支占比也在逐渐下降，越来越多的财政收支向县本级集中。从图1可以看出，我国县本级财政支出占全国比重从2002年的30%上升到2010年的45%，而同期乡级财政支出占比则从10%下降到了6%。总的来讲，乡财县管后，县乡财政体制尽管仍保留“一级政府、一级财权”的形式，但实质上乡镇财权被很大程度上收了（杨发祥，2012），乡镇从独立收支的“一级财政”弱化为“半级财政”<sup>①</sup>。

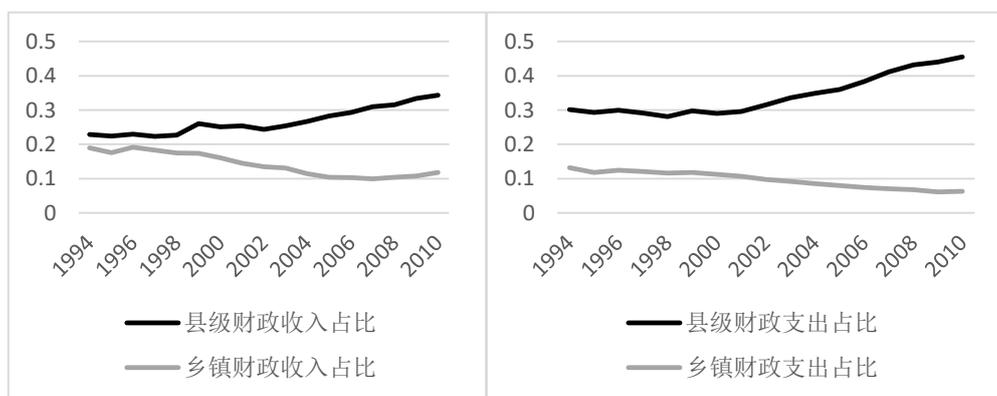


图1 县乡财政收入占全国财政收入的比重与县乡财政支出占全国财政支出的比重

数据来源：历年《地方财政统计资料》。

<sup>①</sup>以浙江省青田县为例，该县自2005年实施了乡财县管改革，撤销32个乡镇的176个账户，将乡镇财政收支集中到中心，建立“收入一个笼子，预算一个盘子，支出一个口子”的运营模式。实行乡镇建设项目审批制度，所有乡镇政府性投资建设的项目，须审核建设资金来源后才审批立项。对招待费支出排名前3名者给以警诫，促使乡镇压缩一般性支出。

## （二）理论分析

乡财县管将部分财政收支权力从乡镇上收到县，根据财政分权的理论，我们可以推测乡财县管对经济发展的效应存在正、反两类作用机制。首先，正面作用机制包括：

首先，乡镇政府比县级政府更容易发生政府失灵，乡财县管通过加强自上而下的监督机制，提升乡镇政府的治理水平。乡镇政府比县级政府更容易发生政府失灵主要是源于以下几个方面原因：第一，中国制度之下，对下级政府的监督主要来自于上级。由于乡镇处于我国五级政府的最末端，也处于自上而下监督链条的最末端。传递层级越多，“鞭长莫及”，监督的耗损和失控就更加严重。第二，到了乡镇这一层级，利益集团内部的人数更少、利益同质性越强，具有更深的宗族联系、地缘联系，更容易影响政府政策。第三，乡镇公务员处在最低的行政层级，政治仕途上升空间有限，薪酬较低，缺乏职业前景，也加重了基层腐败现象（肖滨和陈伟东，2019）。第四，进入新世纪后，取消农业税，加之乡镇企业衰落后，乡镇本级的自有财政收入大幅度萎缩，但是由于支出责任并未上移，乡镇政府承担了大量公共服务，乡镇财政支出更大程度上依赖于上级的转移支付，加重了软预算约束问题，导致乡镇财经纪律更为松弛，乡镇政府主动发展经济、培植税源的积极性也大大下降。

乡财县管之后，由于县级政府的治理水平高于乡镇，而且县级加强了对乡镇政府的监督，可以规范乡镇的财政收支行为，提升基层公共服务和治理水平。事实上，乡财县管所实施一些举措，就是将财权装进笼子里，以自上而下的方式，加强县级对乡镇财政进行管理和监督。改革后实施“预算共编、账户统设”，乡镇政府增加行政管理经费、扩充财政供养人员、擅自提高税费，都更为严格地受到县级政府的监督，从而有助于破解乱花钱、乱进人、乱收费等问题。改革后通过“集中收付、采购统办、票据统管、县乡联网”，县级政府大大加强了对乡镇预算执行过程的监督管理，有助于减少乡镇政府执行预算、政府采购等流程产生的腐败。

第二，根据组织经济学的理论，扁平化的组织结构，有助于减少信息传递的扭曲与执行上的延迟。乡财县管是财政层级扁平化改革，可以解决信息不对称和道德风险问题，减少行政管理成本，故有利于公共服务提供与经济发展。

第三，根据经典的财政分权理论，财政权限上收还有一个优势是可以更好地协调跨地区外部性问题。乡财县管之后，由县级政府直接管理，有助于处理乡镇政府之间在跨乡镇事务上的合作困难问题，更好地提供具有跨乡镇外部性、原本乡镇政府不愿意提供的公共服务，如跨乡

镇道路建设、跨乡镇环境治理等。与此同时，相比于乡镇政府，县级政府提供的一些公共服务可能具有规模经济效应，有助于降低公共服务提供成本。

在中国情境下，乡财县管前，县级政府原本就可以对外部性进行协调，乡镇的规模也还未小到产生明显规模不经济的现象，因此我们认为乡财县管的改革红利主要不是在解决跨乡镇外部性和利用规模经济上，而是在于建立了自上而下的监督机制，减少乡镇政府失灵。

尽管乡财县管带来了上述益处，但是它对经济发展也可能产生负向作用，具体如下：首先，乡财县管之后，县级政府直接提供公共服务，相比于乡镇政府，县级政府掌握本地信息的难度较大，公共服务提供的效率会降低。特别是，对于人口数量较多、情况较为复杂的大县而言，县级政府的直接管理难度更大，乡财县管改革的代价可能越大。其次，乡财县管之后，乡镇政府对财政收入的自主支配权大幅度下降，因而可能导致乡镇政府招商引资、培育税基、发展经济的积极性下降。尤其是，对于初始经济较为发达、基层财力较强的地区，调动乡镇政府的积极性去发展经济尤为重要，如果实施乡财县管，可能不利于发挥乡镇政府积极性。

实践当中，一些地区在乡财县管政策实施中也确实发现存在上述问题。近年来，部分地区也重新开始扩大乡镇财政的自主权，由乡财县管改回原本的乡财乡管。例如2013年起，浙江省丽水市，2018年起辽宁省，都恢复建立完整的乡镇一级财政，实施县与乡镇分灶吃饭体制。当然，乡财县管的正、反两类作用机制孰强孰弱，需要从实证上进行检验。

## 四、研究设计

### （一）模型设定

本文利用县域层面的面板数据，采用双重差分方法（DID）来研究乡财县管改革对经济发展的影响。具体实证模型设定如下：

$$Y_{it} = \beta PostReform_{it} + (X_i * \lambda_t)' \theta + \gamma_i + \lambda_t + \varepsilon_{it} \quad (1)$$

其中， $Y_{it}$ 表示*i*县在第*t*年的人均GDP对数值。除此之外，本文也使用夜间灯光亮度、人均第一、二、三产业增加值、人均工业总产值等作为经济发展的替代指标。

$PostReform_{it}$ 是一个虚拟变量，当*i*县在第*t*年已经实施乡财县管改革时，取值为1，否则取值为0。 $\gamma_i$ 是县域固定效应， $\lambda_t$ 表示年份固定效应。通过控制县域固定效应，可以控制所有不随时间变化的县域特征，将处理组和控制组的固有差异予以剔除。但是有些县域固有特征，

可能既影响了一个县域是否采取乡财县管改革，同时又对经济发展产生随时间变化的效应。为此，我们参照 Duflo（2001）的做法，在 DID 模型中控制了县域初始特征（ $X_i$ ）与年度固定效应（ $\lambda_t$ ）的交互项，从而允许这些县域特征在各年度对因变量产生异质性影响。 $X_i$ 包括 2000 年县域的人均 GDP、人口密度、平均海拔、平均坡度、是否是民族自治地方、人均粮食产量、所辖乡镇个数等指标。 $\varepsilon_{it}$ 为随机扰动项，这里将标准误聚类到（Cluster）县域层面。下文也将使用事件研究法检验处理组和控制组是否具有事前平行趋势。

## （二）数据来源

本文使用的数据主要包括县域层面的各类经济、财政与地理数据。由于市辖区的财政体制与县的财政体制有较大差异，且市辖区的经济数据大多不可得，这里将数据限定为县（含县级市、自治县），但不含市辖区。

具体而言，本文使用的县域数据来源如下：

第一部分是县域经济数据，来源于 2000-2019 年的《中国县（市）社会经济统计年鉴》，主要包括县域层面的 GDP、三次产业增加值、工业总产值、人口、面积、财政收入、财政支出、居民储蓄余额、固定资产投资等信息<sup>①</sup>。

第二部分是县域财政数据，数据来源于《全国地市县财政统计资料》，主要包括分类型的财政支出、政府性基金收支、财政供养人口等信息，数据覆盖区间为 2000-2007 年。

第三部分是县域的平均海拔、平均坡度等地理特征，该数据可以从中国科学院资源环境数据云平台获取。

第四部分是县域企业数量数据，数据来源于元素灵宝经济洞察平台，该数据库根据工商登记注册数据形成。我们从中摘取了每个县的历年小企业数量（注册资本在 50 万元以下），数据覆盖区间是 2005-2017 年。

第四部分是乡财县管改革的名单<sup>②</sup>是通过手动收集而来，主要来源于各省、市、县政府发布的文件。各地政府发布的文件主要通过北大法宝网、律商网、汇法网这三个网站，通过搜索“乡财县管”、“镇财县管”等关键词查找。除此之外，我们还收集了政府官方新闻报道作为补充来源。最终整理出各省份乡财县管实施名单与对应的改革文件名称、文件号、实施日期与资料

<sup>①</sup>农村人均纯收入数据覆盖 2000-2010 年，部分年份数据缺失，从分省统计年鉴当中补充。

<sup>②</sup>改革文件显示，绝大多数地区在推行乡财县管时，都是整县推进改革。极个别县是仅选取县内部分乡镇试点，对于这些县，我们也定义为改革县。我们也做了稳健性检验，如果样本中删去这些县，不影响本文的结果。

来源。

为了剔除极端值的干扰，本文将主要变量按最高和最低的1%样本进行缩尾处理。同时为剔除价格因素，将人均GDP、人均三次产业增加值、农村人均纯收入、人均储蓄、固定资产投资等变量以2000年为基期，利用GDP平减指数进行价格平减。

## 五、实证结果分析

### （一）基准回归结果

首先，本文报告乡财县管改革对县域人均GDP产生的影响。表1为基准回归结果。第（1）列在控制了年份固定效应和县域固定效应后进行回归，第（2）列进一步控制了县域初始特征与年份虚拟变量交互项，以剔除不同初始特征的县在随后各个年份所受到的差异性冲击。第（2）列系数0.0480，表明乡财县管改革的实施将使县域人均GDP提升4.8%。

在样本期间，除了乡财县管改革之外，其他改革措施也可能对结果产生干扰，为此我们在回归中控制了其他改革措施：首先，部分省份2004年起还实施了“省直管县”、“扩权强县”等改革。为了剔除这两项改革的影响，第（3）列额外加入了“是否省直管县”、“是否扩权县”两个控制变量。其次，2001年我国加入WTO，关税大幅度降低可能对县域经济也会产生影响，为了剔除进口关税降低对县域经济的影响，本文参照戴觅等（2019）的方法，利用各行业关税削减幅度的差异以及各县域“入世”前产业结构的差异（这一数据仅更新至2014年），构建了县域层面的关税削减效应指标，添加这一控制变量的结果报告在第（4）列。第三，2005年底我国全面取消农业税，对县域财政运行带来的较大压力，为此第（5）列我们进一步控制了县域层面农业税金占财政收入的比重。第四，上世纪90年代开始，很多地区还推行了乡镇撤并的改革。乡镇撤并精简了机构，降低了行政成本，因此也有可能促进县域整体的经济发展。乡镇撤并导致的直接结果是县域所辖乡镇个数会减少，因此，我们进一步控制了县所辖乡镇个数，以剔除乡镇撤并对县域整体经济的潜在影响。结果显示，添加上述控制变量后，乡财县管改革的回归系数都基本不发生变化。

表1 乡财县管对县域人均GDP的影响

	(1)	(2)	(3)	(4)	(5)	(6)
被解释变量	人均GDP	人均GDP	人均GDP	人均GDP	人均GDP	人均GDP
乡财县管改革	0.0894*** (0.0095)	0.0480*** (0.0101)	0.0432*** (0.0102)	0.0449*** (0.0101)	0.0448*** (0.0100)	0.0413*** (0.0106)
年份固定效应	Y	Y	Y	Y	Y	Y
县域固定效应	Y	Y	Y	Y	Y	Y
县域初始特征*		Y	Y	Y	Y	Y
年份虚拟变量						
财政省直管县改革			Y	Y	Y	Y
扩权强县改革			Y	Y	Y	Y
取消农业税冲击				Y	Y	Y
县域乡镇个数					Y	Y
WTO进口冲击						Y
观测值	36,235	33,741	33,573	31,839	31,023	23,455
R <sup>2</sup>	0.9925	0.9938	0.9938	0.9941	0.9942	0.9518

注：\*\*\*、\*\*、\*是指在1%、5%以及10%的置信水平下显著；括号内为聚类到（cluster）县域层面上的标准误。县域初始特征包括所辖乡镇个数、人口密度、人均GDP、平均海拔、平均坡度、是否民族自治地方、人均粮食产量，下同。

## （二）事件研究法检验平行趋势假设

从基准 DID 的结果来看，乡财县管改革对地方经济增长带来显著的正向影响。但 DID 方法成立的一个前提条件是，处置组（样本期内实施了改革的县）和控制组（样本期内最终未实施改革的县），在改革前，具有平行时间趋势。如果控制组与处置组两组县的经济增长趋势在改革前就存在差异，那么有理由怀疑处置组和控制组存在不可观测且随时间变化的差异，从而存在内生性问题。因此，还需要检验是否满足事前平行趋势。参考 Jacobson et.al（1993）的处理方法，本文利用事件研究法，检验乡财县管改革前后年份经济增长的动态变化。具体来说，模型设定如下：

$$Y_{it} = \beta_k \sum_{k>=-5}^k D_{t_{i0}+k} + \alpha X_{it} + \gamma_i + \lambda_t + \varepsilon_{it} \quad (2)$$

其中， $D_{t_{i0}+k}$ 代表实施乡财县管政策前后的事件窗口虚拟变量， $t_{i0}$ 是县域 i 开始实施乡财县管改革的年份， $t_{i0}+k$ 是指政策改革前后的各个年份，其中 k 取值分别是-5 及更早之前、-4、-3、-2、-1、0、1、2+，缺省组是 t=-1（即改革前一年）。其余控制变量和基准模型类似，标准

误仍然聚类到县域层面。结果显示，改革前两年之前的年份，处置组和对照组的县域人均 GDP 趋势没有明显差异，因此满足事前平行趋势。处置组改革之后，两组差异才开始显现。这很大程度上表明县域经济增长的确是由乡财县管改革带来的，而与其他不可观测变量或其他改革无关<sup>①</sup>。

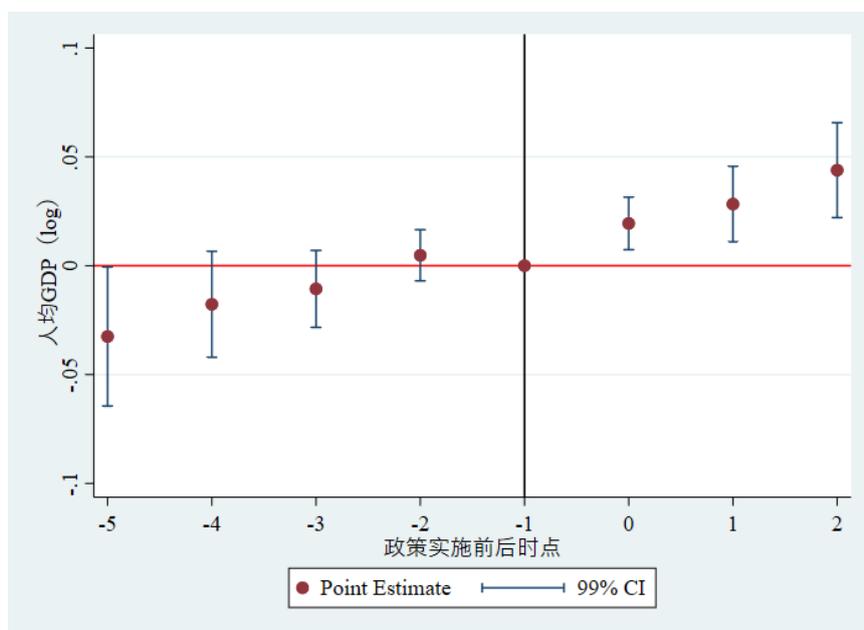


图 2 平行趋势图

### (三) 其它经济发展度量指标

接下来，本文用其他经济发展度量指标来衡量地区经济发展水平。本文考察了乡财县管改革对于三次产业是否均有正向影响，表 2 第 (1) - (3) 列的结果显示，乡财县管对一、二、三产业增加值均有显著正向促进作用。从系数大小上看，乡财县管对第二产业的正向作用大于对第一、第三产业的影响。我们推测，这是由于第二产业，尤其是当中的制造业，具有较高的跨地区流动性，对地方政府治理水平更为敏感。因此，乡财县管提升了一个县的治理水平，更有利于吸引制造业。除此之外，本文进一步使用人均居民储蓄指标，作为经济发展水平的度量指标。表 2 第 (4) 列的回归结果显示，乡财县管改革对人均储蓄提升具有正面效应。同时，第 (5) 列结果表明，乡财县管改革对县域内小企业数量增加也具有正面效应，由于小企业多数为民营企业，因此这一结果说明改革促进了民间创业。

<sup>①</sup> 限于篇幅，此处不再展示平行趋势假设检验的回归结果。

表2 乡财县管对其它经济发展指标的影响

被解释变量	(1) 人均第一产业 增加值	(2) 人均第二产业 增加值	(3) 人均第三产业 增加值	(4) 人均储蓄	(5) 小企业数量
乡财县管改革	0.0299*** (0.00892)	0.0712*** (0.0166)	0.0310** (0.0121)	0.0468*** (0.0087)	0.0300** (0.0121)
年份固定效应	Y	Y	Y	Y	Y
县域固定效应	Y	Y	Y	Y	Y
县域初始特征*	Y	Y	Y	Y	Y
年份虚拟变量					
观测值	33,859	33,858	33,595	33,805	20,173
R <sup>2</sup>	0.991	0.981	0.990	0.9923	0.9651

## 六、机制分析

上文的实证结果显示，乡财县管改革促进了经济发展。这一小节，我们分析乡财县管改革对经济发展的影响机制。“乡财县管”使乡镇政府财政管理权限减少，县本级的财政管理权限增加，因此改革不仅触及乡镇财政，也触及县级财政。乡财县管对于经济发展的效果，也是由于它同时对县级政府和乡镇政府的行为产生作用。县本级和乡镇本级各自的财政收支行为变化是乡财县管影响经济发展的重要机制。但由于数据所限，我们无法分别研究乡财县管对县本级财政支出结构（或财政收入）、乡镇本级财政支出结构（或财政收入）各自的影响，而只能分析改革对县乡两级整体的财政支出结构与税费负担的影响。由于对县域经济发展产生作用力的，是县域整体的财政支出结构和税费负担，因此我们分析县域整体财政收支行为，尽管有局限性，但是仍然对于机制分析是有价值的。

### （一）乡财县管改革对行政管理支出类占比和财政供养人口的影响

首先，本文检验了乡财县管改革是否改善了财政支出结构，降低了县域的行政管理支出占比。改革之前，县乡财政的一大乱象是乱花钱、乱进人，大量财政经费用于维持臃肿的财政机构，导致正常公共服务的支出无法保障，部分地区甚至成为典型的吃饭财政。在县域财政支出（含县本级与乡级）结构当中，行政管理类支出（含行政管理费用和公检法司支出）占财政支出的比重平均达到了21.9%。尽管行政管理支出是维持政府行政机构运转所必需的支出，但它属于“虚耗性支出”。一个县的财政支出结构当中，行政管理支出占比越高，越表明机构臃肿严重、资金使用效率低下，能够用于提供公共服务的资金相应被挤占。

本文使用的财政支出结构数据来自《地市县财政统计资料汇编》，其中行政管理类支出我们使用了两个口径，一个是仅包括行政管理费用，另一个是行政管理费和公检法司支出之和。

《地市县财政统计资料汇编》当中，各项财政支出是县域和乡镇级财政支出之和，以下简称县域财政支出。我们无法单独区分乡镇和县本级各自的财政支出。但是由于这里旨在检验改革对整个县域内财政支出结构的影响，因此无需区分乡镇级和县本级的各项财政支出。而且，由于改革后，乡级财政已经半虚级化，相当多数的公共服务也改为由县本级直接承担，乡级财政支出已不能反映出当地的公共财政支出倾向。

回归结果报告在了表3当中。结果显示乡财县管改革显著降低了行政管理类支出占比。由表3第(1)列的回归系数可知，乡财县管政策实施后，行政管理类支出占比降低了0.0066。在本文样本中，行政管理类支出（行政管理支出和公检法司支出之和）占一般公共预算支出比例的均值为0.219。相较于均值来说，乡财县管使行政管理类支出的占比下降了3.0%（即0.0066除以0.219）。这表明乡财县管改善了县域财政支出结构，县域内有更大比例的财政资金用于了公共服务支出，因而有利于经济发展。

表3 乡财县管对财政支出结构和财政供养人员负担的影响

被解释变量	(1) （行政管理费 用+公检法支 出）/财政支出	(2) 行政管理费 用/财政支出	(3) 财政供养人 数/县域财力	(4) 一般公共预 算收入/GDP	(5) 政府性基金 收入/GDP
乡财县管改革	-0.0066*** (0.0020)	-0.0060*** (0.0016)	-0.0239*** (0.0041)	-0.0021*** (0.0007)	-0.0007** (0.0004)
年份固定效应	Y	Y	Y	Y	Y
县域固定效应	Y	Y	Y	Y	Y
县域初始特征*年份 虚拟变量	Y	Y	Y	Y	Y
观测值	12,011	12,011	16,858	32,283	17,239
R <sup>2</sup>	0.7847	0.8157	0.9009	0.6366	0.5332

基层政府行政管理支出比重较高，通常与财政供养人口的膨胀高度相关，朱镕基同志在“搞好农村费改税试点工作”的谈话中也提到，“精简机构、减少农民供养的人员，是从根本上减轻农民负担的必要条件”<sup>①</sup>。本文进一步检验了改革是否降低了财政供养人员负担。这里不直

<sup>①</sup>参见《朱镕基讲话实录》（第三卷），2011年人民出版社。

接使用财政供养人数作为被解释变量,原因在于乡财县管改革提升了县域经济发展水平和可用财力,当地公共服务的增加,本身就需要增加财政供养人数来提供这些公共服务。财政供养人数,除了包括公务员外,也包括教科文卫等事业机构的人数。教师、医生数量的增加,一定程度反映县域公共服务质量的改善(如师生比、医务人员比等)。为此,这里使用财政供养人数与县域政府可用财力之比,可用财力包括县乡本级收入和上级给与的转移支付。这一指标衡量的是,一个县财力当中的财政供养人员负担,即单位财力需要支撑的财政供养人数。表7的回归结果显示,乡财县管改革后财政供养人员负担下降,这表明乡财县管改革一定程度减轻了基层行政机构的臃肿程度。

## (二) 乡财县管改革对税费负担的影响

上文从财政支出端发现,乡财县管改善了财政支出结构,提升了县域财政用于正常公共服务支出的比重。进一步地,本文考察了乡财县管改革对财政收入端的影响。在乡财县管改革之前,县乡财政运行的另一大乱象是基层政府的乱收费现象,加大了私人部门的负担。因此,改革是否降低了私人部门的税费负担,从而有利于经济发展?

为了衡量私人部门的税费负担,这里使用县域一般预算财政收入与GDP之比、县域预算外政府性基金收入与GDP之比这两个指标。县域一般预算财政收入包括税收和纳入预算的收费项目。预算外的政府性基金收入包括工交部门基金、商贸部门基金、文教部门基金、农业部门基金、土地有偿使用支出、其他部门基金、地方财政税费附加等项目。表3第(4)-(5)列的回归结果显示,乡财县管改革既降低了一般预算的税费负担,也降低了预算外收费的负担。由表3第(4)列的回归系数可知,乡财县管政策实施后,一般公共预算收入占GDP比例下降了0.0021。相较于样本均值(0.051)来说,一般公共预算收入占GDP比例下降了4.1%。由表3第(5)列的回归系数可知,乡财县管政策实施后,政府性基金收入占GDP比例下降了0.0007。相较于样本均值(0.007)来说,政府性基金收入占GDP比例减少了10%。因此乡财县管通过减轻了私人部门的税费负担,有利于当地经济发展。

## (三) 异质性分析

根据第三部分的理论分析,乡财县管对经济发展同时存在正向和反向两类作用机制。上面的实证研究表明,总体来看,乡财县管改革促进了经济发展,因此正向机制的作用力要大于反向机制。乡财县管对经济发展的平均净效应为正,并未产生某种具有普遍性的代价。但是,

改革对不同县的影响，可能存在异质性。表4第（1）列放入了乡财县管与县域人口规模的交互项，交互项系数显著为负，表明人口众多的大县，由于县级政府管理幅度过大，乡财县管改革对县域经济发展的效果要较弱。表4第（2）列放入了乡财县管与县平均坡度的交互项，交互项系数也显著为负。说明县域内地形较为复杂，县级政府直接提供公共服务时获取信息的成本较高，此时乡财县管改革对经济发展的效果也较弱。这表明乡财县管改革的一项成本是，在人口大县和内部情况较为复杂的县，乡财县管增加了县级政府获取信息的难度。

表4 异质性分析

	(1)	(2)	(3)	(4)
被解释变量	人均GDP	人均GDP	人均GDP	人均GDP
乡财县管改革	0.1277*** (0.0172)	0.0760*** (0.0153)	0.1837*** (0.0290)	0.8257*** (0.1412)
乡财县管*初始县域人口数量	-0.0016*** (0.0003)			
乡财县管*初始县域平均坡度		-0.0118*** (0.0040)		
乡财县管*初始财政自主度			-0.3110*** (0.0653)	
乡财县管*初始人均GDP				-0.0930*** (0.0170)
年份固定效应	Y	Y	Y	Y
县域固定效应	Y	Y	Y	Y
县域初始特征*年份虚拟变量	Y	Y	Y	Y
观测值	33,741	33,741	33,723	33,741
R2	0.9938	0.9938	0.9938	0.9938

另外，乡财县管之后，乡镇政府招商引资、培育税基、发展经济的积极性下降。那么对于初始经济较为发达、基层财力较强的地区，乡镇政府积极性较为重要，如果实施乡财县管，可能会带来一定的成本。为了对此进行检验，我们在表4第（3）-（4）列中分别加入了乡财县管与县域初始财政自主度、县域初始人均GDP的交互项，结果显示，交互项的系数均显著为负，支持了我们的假说。这反过来也说明，乡财县管在经济欠发达、财政自主度较低的地区，对经济发展具有更大的促进作用。因为在这些地区，乡镇政府对上级转移支付的依赖程度更高，自上而下的监督机制将“财权入笼”尤为重要。

## 七、结论及政策建议

本文利用2000-2019年的县级层面数据,使用双重差分模型,评估了乡财县管改革对经济发展的影响。结果显示,乡财县管促进了改革县的经济发展水平。实施乡财县管后,县域人均GDP平均提升4.8%。进一步地,我们检验了乡财县管促进经济发展的机制。我们发现,乡财县管通过规范财政支出与降低政府收费两个方面有利于经济发展:首先,乡财县管改革降低了财政支出当中行政管理支出的比重,抑制了财政供养人员规模膨胀,保障了正常公共服务支出;其次,乡财县管改革减轻了私人部门的税费负担。

乡财县管改革的经验告诉我们,中国政府间财政管理体制的改革,需要权衡权力下放与集中统一管理的相对成本与收益,激励与约束并重,既充分发挥基层政府的灵活性、积极性,又同时加强自上而下的管理与监督机制。特别是,对于欠发达地区而言,越到基层政府,事权、支出责任与自主财力不匹配越严重,基层财力很大程度上依靠上级转移支付,此时过分强调通过分权加强激励,反而可能导致激励失灵。通过分权给予基层过多的资源,而不加以监督和约束,反而导致更多的资源浪费。反而是,通过适度的财政管理权限上收,加强自上而下的财政管理与监督,有助于改善公共服务和政府治理水平。2022年《国务院办公厅关于进一步推进省以下财政体制改革工作的指导意见》要求将财政收入难以覆盖支出需要、财政管理能力薄弱的乡镇纳入乡财县管范围,具有很强的改革针对性。

在推进“乡财县管”工作提振增效上,上级政府应该从以下几个方面进一步加强自上而下的监督机制,提升乡镇财政管理效率和水平。第一,结合预算管理一体化系统建设和应用,强化县级财政对乡镇预算编制、预算执行、国有资产管理等监督。第二,继续完善国库集中支付、集中采购制度、收支两条线等制度,推动实现“财权入笼”。第三,要加大对基层政府转移支付资金使用的监督,确保转移支付资金达到基层后能够高效使用。第四,要更好地发挥审计作用,建立资金违规使用责任追究制度。第五,推进财政支出标准化,加快建立财政支出标准体系,为财政资源的分配与管理提供目标和依据,增强预算执行的刚性和约束力。党的十九届四中全会《关于坚持和完善中国特色社会主义制度 推进国家治理体系和治理能力现代化若干重大问题的决定》指出,要健全党和国家的监督体系,强化对权力运行的制约和监督。强化除了加强自上而下的管理与监督机制外,也应该健全人大监督、群众监督、舆论监督制度,加强预算公开,探索参与式预算,推进基层治理体系和治理能力的现代化。

本文的研究结论也表明,乡财县管改革尽管总体上有益于改善基层治理、推进经济发展,但不应一刀切地去推行:首先,在人口多、情况较为复杂的大县内,实施乡财县管改革后,县级政府直接提供公共服务时获取信息的难度较大。其次,在初始经济发展水平较高、具有良好工业基础的地区,由于乡镇政府自身在招商引资过程当中扮演了重要的角色,缺乏财政管理自主性不利于调动乡镇政府的积极性。2022 年《国务院办公厅关于进一步推进省以下财政体制改革工作的指导意见》也指出,改革应该坚持因地制宜原则,鼓励各地区因地制宜采取差异化措施,本文基于乡财县管改革的研究也为因地制宜原则提供了重要启示。

#### 参考文献:

- [1] 陈硕、高琳,2012,《央地关系:财政分权度量及作用机制再评估》,《管理世界》第 6 期,第 43~59 页。
- [2] 陈锡文,2002,《全面建设小康社会的关键在农村》,《经济研究》第 12 期,第 10~12 页。
- [3] 戴觅、张轶凡和黄炜,2019,《贸易自由化如何影响中国区域劳动力市场?》,《管理世界》第 6 期,第 56~69 页。
- [4] 贾俊雪、张永杰和郭婧,2013,《省直管县财政体制改革、县域经济增长与财政解困》,《中国软科学》第 6 期,第 22~29+52 页。
- [5] 贾康,2007,《财政的扁平化改革和政府间事权划分》,《中共中央党校学报》第 6 期,第 42~48 页。
- [6] 李永友、王超,2020,《集权式财政改革能够缩小城乡差距吗?——基于“乡财县管”准自然实验的证据》,《管理世界》,2020 年第 4 期,第 113~130 页。
- [7] 刘克崧、贾康,2009,《中国财税改革三十年亲历与回顾》,经济科学出版社。
- [8] 林毅夫、刘志强,2000,《中国的财政分权与经济增长》,《北京大学学报(哲学社会科学版)》第 4 期,第 5~17 页。
- [9] 吕冰洋、马光荣和毛捷,2016,《分税与税率:从政府到企业》,《经济研究》第 7 期,第 13~28 页。
- [10] 吕连生,2005,《安徽省“乡财县管”改革的研究》,《经济研究参考》第 60 期。

- [11] 沈坤荣、付文林, 2005, 《中国的财政分权制度与地区经济增长》, 《管理世界》第1期。
- [12] 王小龙、陈金皇, 2020, 《省直管县改革与区域空气污染——来自卫星反演数据的实证证据》, 《金融研究》第11期。
- [13] 夏杰长、陈雷, 2005, 《“乡财县管”改革的社会学分析——以安徽省G县为例》, 《经济研究参考》第77期, 第4~23页。
- [14] 肖滨、陈伟东, 2019, 《基层腐败问题的缘起:默契性容忍——基于A市镇街“一把手”的48个案例研究》, 《广东社会科学》第5期, 第14~22页。
- [15] 徐永胜、乔宝云, 2012, 《财政分权度的衡量:理论及中国1985—2007年的经验分析》, 《经济研究》第10期, 第4~13页。
- [16] 杨发祥、刘楠, 2012, 《“乡财县管”:理论视角与经验反思——一个跨学科的探索性研究》, 《华东理工大学学报(社会科学版)》第4期, 第90~98+116页。
- [17] 张威、高瑞, 2010, 《“乡财县管”的理论探讨与实践分析》, 《东北财经大学学报》第3期, 第49~52页。
- [18] 郑新业、王晗和赵益卓, 2011, 《“省直管县”能促进经济增长吗?——双重差分方法》, 《管理世界》第8期, 第34~44+65页。
- [19] 周黎安, 2017, 《转型中的地方政府》, 格致出版社。
- [20] 张晏、龚六堂, 2006, 《分税制改革、财政分权与中国经济增长》, 《经济学(季刊)》第1期, 第75~108页。
- [21] Bardhan, Pranab and Dilip Mookherjee, 2005, “Decentralizing Antipoverty Program Delivery in Developing Countries”, *Journal of Public Economics*, 89(4):675~704.
- [22] Davoodi, H., and Zou, H., 1998, “Fiscal Decentralization and Economic Growth: A Cross-Country Study”, *Journal of Urban Economics*, 43(2):244~257.
- [23] Duflo, Esther, 2001, “Schooling and Labor Market Consequences of School Construction in Indonesia: Evidence from An Unusual Policy Experiment”, *The American Economic Review*.
- [24] Hayek, F.A., 1945, “The Use of Knowledge in Society”, *The American economic review*, 35(4):519~530.

- [25] Jacobson, Louis S., Rober J. LaLonde, and Daniel G. Sullivan, 1993, “Earnings Losses of Displaced Workers.”, *American Economic Review*, 83(4):685~709.
- [26] Mookherjee, Dilip, 2015, “Political Decentralization”, *Annual Review of Economics*, 7(1): :150403165810008.
- [27] Oates, W. E., 1972, “Fiscal Federalism”, *Edward Elgar Publishing*.
- [28] Tiebout, Charles, 1956, “A Pure Theory of Local Government Expenditures”, *Journal of Political Economy*, 64(5):416~424.
- [29] Weingast, B. R., 2009, “Second Generation Fiscal Federalism: The Implications of Fiscal Incentives”, *Journal of Urban Economics*, 65(3):279~293
- [30] Xie, Danyang, Zou, Heng-fu and Davoodi, H., 1999, “Fiscal Decentralization and Economic Growth in the United States”, *Journal of Urban Economics*, 45(2):228~239.
- [31] Xu, C., 2011, “The Fundamental Institutions of China’s Reforms and Development”, *Journal of Economic Literature*, 49(4):1076~1151.
- [32] Zhang, T. and H. Zou, 1998, “Fiscal Decentralization, Public Spending, and Economic Growth in China”, *Journal of Public Economics*, 67(2):221~240.

## County-Managing-Town Fiscal Reform, Local Governance and Economic Development

MA Guangrong<sup>a</sup>, ZHANG Ling<sup>b</sup>

(a. The School of Finance, Renmin University of China, Peking, 100872

b. The School of Economics, Peking University, Peking, 100871)

**Summary:** The county and township governments are at the bottom of China's government hierarchy, they are also the ultimate undertakers of a large number of public affairs. The proportion of county and township fiscal expenditure in the country is close to 50%, which can be described as "the least power and the most responsibility" (Zhou, 2017). After the Reform and Opening up, the two hierarchies has experience two major reforms. In 1983, with the disintegration of the People's Commune system, the the township government and financial system had been re-establishment. However, with the Tax Sharing Reform in 1994 and the Rural Tax Reform after 2000, the contradiction between revenue and expenditure in villages and towns had become more and more serious. The township finance had exposed some problems, such as the lack of constraints on financial expenditure, the expansion of finical-supporting population, the difficulty in guaranteeing normal public service expenditure, and even illegal charges and large-scale debt raising. In order to solve these problems, since 2003, some provinces began to explore the "county-managing-town fiscal reform". In 2006, the central government began to explicitly encourage all provinces to promote the reform.

The county-managing-town fiscal reform means, direct management and supervision of township financial revenue and expenditure by county-level financial departments. Although the county and township financial system still remain, in fact, the township has been weakened, and the township financial management power has been greatly received by the county (Yang&Liu, 2012). Therefore, county-managing-town fiscal reform is a reform of fiscal centralization. At the same time, with the township finance becoming inappreciable, the county-level government began to directly bear a large

amount of public service expenditure. In this sense, the reform is also a flat reform of financial hierarchy (Jia, 2007).

This paper attempts to answer the question: has the county-managing-town fiscal reform improved the governance and public service provision capacity of the grass-roots government, ultimately promoted economic development? Based on the county level data from 2000 to 2014, this paper uses the DID model to evaluate the impact of the reform on economic development. The results show that the reform promoted the economic development, and the average per capita GDP of the county increased by 4.8%. Mechanism analysis shows that the county-managing-town fiscal reform reduces the proportion of administrative expenditure, inhibits the expansion of the scale of financial-support population, and reduces the tax burden. These results show that strengthening the financial management and supervision of the superior hierarchy to the grass-roots government will help to improve grass-roots governance and promote economic development. At the same time, this paper also finds that in the counties with large population and more developed economy, the promotion effect of reform is relatively weak. The reason is it's difficult for the county government to obtain information after the reform, or it's not conducive to inspire the enthusiasm of the township government, which indicates that reform should not be implemented across the board.

The contributions of this paper lies in the following three aspects: firstly, this paper quantitatively evaluates the impact of the county-managing-town fiscal reform on economic development and local government behavior for the first time. Secondly, this paper tentatively extends the research of inter-governmental financial relationship to the county and township level. Finally, this paper analyzes the effect of financial level Flattening Reform from the perspective of county-managing-town fiscal reform.

**Keywords:** County-Managing-Town Fiscal Reform; Township Government; Governance of Government; Fiscal Decentralization

**JEL Classification:** H71; H3; H2

**附：中文长摘要**

县乡两级政府处于我国政府层级的最底端，但同时又是大量公共事务的最终承担者，县乡两级财政支出占全国的比重接近 50%，可谓“权力最小、责任最大”（周黎安，2017）。改革开放后，我国县乡两级财政管理体制经历了两次大的改革。1983 年，随着人民公社体制的解体，乡镇政权重新确立，乡镇一级财政也相应建立，乡镇政府拥有了自主的财政收支与预决算体系，形成了“乡财乡管”体制。但是，随着 1994 年分税制改革后财力大幅度上移，以及 2000 年后农村税费改革的实施，乡镇收支矛盾日益加剧，乡镇财政运行暴露出财政支出缺乏约束、财政供养人员膨胀、正常公共服务支出难以保障等问题，甚至违规收费、大规模举债，加重了居民和企业的负担（刘克崧和贾康，2009）。为了化解这些问题，2003 年以来，从安徽省开始，一些省份开始探索实施“乡财县管”改革。2006 年财政部发布《关于进一步推进乡财县管工作的通知》，中央开始明确鼓励各地推进乡财县管改革。

乡财县管改革，是由县级财政部门直接管理并监督乡镇财政收支。乡财县管后，尽管县乡财政体制仍保留“一级政府、一级财政”的形式，而实质上乡镇从独立收支的“一级财政”弱化为“半级财政”，乡镇财政管理权被大幅度上收到了县（杨发祥、刘楠，2012）。因此，从本质上看，乡财县管是一项财政集权式的改革。与此同时，随着乡镇财政的虚化，县级政府开始直接负担大量公共服务的支出，从这个意义上看，乡财县管也是一项财政层级扁平化的改革（贾康，2007）。

本文试图回答的问题是，乡财县管改革是否改善了基层政府治理和公共服务提供能力，最终促进了经济发展？本文利用 2000 年-2014 年的县域层面数据，使用双重差分模型，评估了乡财县管改革对经济发展的影响。结果显示，乡财县管促进了经济发展，使县域人均 GDP 平均提升了 4.8%。机制分析显示，乡财县管减少了行政管理支出比重，抑制了财政供养人员规模膨胀，降低了税费负担。这些结果表明，加强上级政府对最基层政府的财政管理与监督，有助于改善政府治理、促进经济发展。同时本文也发现，在地形较为复杂的人口大县及经济较发达地区的县，由于改革后县级政府获取信息难度高，或不利于发挥乡镇政府的积极性，因此乡财县管对经济发展的提升效果相对较弱，这说明乡财县管也不宜一刀切地实行。

本文的研究价值在于以下三个方面：首先，本文首次定量评估了乡财县管对经济发展和地方政府行为的影响。其次，本文探索性的将政府间财政关系的研究拓展到了最基层的县乡两级财政关系。最后，本文则从乡财县管改革的角度剖析了财政层级扁平化改革效果。