

我国官员晋升锦标赛机制：理论与证据

乔坤元

(北京大学光华管理学院 北京 100871)

摘要：我对官员考核和晋升是否存在一个锦标赛机制进行了考察，通过比较上级官员的收益给出了三种可能：不存在晋升的锦标赛机制、存在以经济增长为主要考核内容的 GDP 晋升锦标赛机制和存在考核指标不局限于经济增长的晋升锦标赛机制。我们结合先验和后验的方法，发现只有经济增长符合考核指标的条件，并且显著正向影响官员的晋升概率，这说明我国的确存在一个以经济增长为主要考核内容的官员晋升的锦标赛机制。本文有助于重新审视我国的官员晋升体系、理解当前政绩考核机制的问题。

关键词：晋升锦标赛 考核指标 经济增长

一、引言和文献综述

在自然禀赋、物质和人力资本的积累以及技术创新能力与外国相比并无独特之处的情况下，我国经济自改革开放以来保持了三十多年两位数字的高速增长，对于维持经济增长原因的一个重要的解释是官员晋升的锦标赛机制。

官员晋升的锦标赛机制由周黎安(2004, 2007)、李宏彬和周黎安(Li 和 Zhou, 2005)等一系列研究成果给出，该理论认为在我国的集权型政治体制之下，上级官员主要依据经济增长来考核和提拔下级官员，因此下级官员有着很强烈的动力来发展经济以求能够获得政治上的升迁，地方政府官员越来越多的被认为是“政治人”而非由金和辉、钱颖一和温格斯特(Jin、Qian 和 Weingast, 2005)提出的“财政联邦主义”中的“经济人”(陶然等, 2010)。具体的，周黎安(2007)说明了我国锦标赛制度的前提(上级对人事权力的集中、有可衡量的客观竞赛指标、相对竞赛成绩可分离和比较、参赛官员可以控制和影响最终结果、参与人不易形成合谋)得到了满足，因而可以有效的实施。在一定条件下(比如参赛人的风险倾向是中性的)，锦标赛可以取得最优的激励效果(周黎安, 2007)。这一概念的提出引发了大量学者从地方官员的晋升激励这一视角就中国经济展开了分析，如财政竞争、公共品投入等等(见傅勇和张晏, 2007; 郑磊, 2008 等等)，并且逐渐形成一条重要的文献支流。

然而，陶然等(2010)对于上述官员锦标赛机制的解释提出了质疑，他们认为过多的文献直接将官员晋升锦标赛体制视作给定的事物，而没有探讨它真实的存在性。他们从人事任免和干部考核、GDP 增长率客观性和统计上的技术问题、政府层级选择问题以及官员锻炼和提拔的角度对这一机制提出了逻辑上的挑战，认为我国的上级对人事的自由裁量权的缩小、经济增长指标的掺水和 GDP 的核算问题、政府的层级、官员派遣的锻炼目的都会使得锦标赛的机制不能满足周黎安(2007)提出的实施前提，并且使用 1980-1995 和

1979-2002 这两个时期的样本数据对于这一机制进行了重估，得到了不存在官员晋升锦标赛体制的结论。

与此同时，一支逐渐发展的文献呼吁官员的考核指标应该不仅仅局限于经济的增长，还应该包含其它的民生指标，比如环境、卫生、科教、能源等等（徐风华和王俊杰，2006），一些学者更是强调了民众参与的重要性（如邱法宗和张霖星，2007；尤建新和王波，2005 等等），傅勇和张晏（2007）也说明了当前官员考核机制“重基本建设、轻人力资本投资和公共服务”的问题，这些研究说明了我国官员的考核指标多元化的可能和趋势。

那么，我国是否存在着一个官员竞争的锦标赛体制呢？官员的主要考核指标是一个、多个还是没有呢？我们将在本文中重新考察这一机制，使用一个道德风险模型为我国官员晋升锦标赛机制提供一个可能的理论基础。具体的，我们将上一级官员比作风险中性的委托人而将下一级官员比作风险厌恶的代理人，计算出上级官员对下级官员最优的激励机制，从而给出了三种可能性：不存在一个晋升的锦标赛体制、存在一个以经济增长为主要考核内容的 GDP 晋升锦标赛机制和存在一个考核指标不局限于经济增长的晋升锦标赛机制。我们结合现有数据的实际情况，先验的（prior）测算各个指标的波动性从而探讨这些指标是否适合做晋升的考核指标，并且后验的（posterior）使用晋升可能性对各个潜在指标进行回归来考察哪些指标影响官员的晋升/去职的概率。我们发现经济增长是当前我国的官员考核和晋升体制最恰当的指标，其波动性小于其他几个指标，并且在控制了其它可能影响官员晋升的变量之后，潜在的考核指标中只有经济增长会显著的正向影响官员晋升的概率而负向影响官员去职的概率，这些结果说明我国的确存在一个官员晋升的锦标赛机制，并且目前锦标赛的考核内容主要是经济增长。

本文从以下四点丰富和充实了以往的文献。首先，我们重新审视了官员晋升锦标赛体制，同样也回应了陶然等（2010）对于这一机制的质疑以及人们对于官员考核指标多元化的猜想（如邱法宗和张霖星，2007；徐风华和王俊杰，2006；尤建新和王波，2005 等等）。其次，我们构建一个道德风险模型为官员晋升锦标赛提供了理论基础。目前的研究过多的集中在实证领域，而理论研究并不常见，我们的文章为之后的理论发展提供了一个新的切入点。再次，我们整合了《新中国 60 年统计资料汇编》、《中国卫生统计年鉴》和人民网的相关数据，从时间跨度上来看比以往的文献都要长，如李宏彬和周黎安（2005）、陶然等（2010），并且同时使用了《中国统计年鉴》、中经专网数据库、国泰安数据库等辅助数据源进行了仔细的交叉检查（cross check），使得我们得到的结论更加可信，尤其是避免 GDP 核算有误的问题（陶然等，2010）。最后，我们的发现进一步说明了在我国存在着一个以经济增长为官员核心绩效内容的晋升锦标赛体制，为晋升锦标赛体制提供了新的证据。同时，我们还探讨了目前锦标赛机制存在的问题以及解决方案，以供参考。

本文接下来按如下的方式安排。我们在第二部分利用一个道德风险模型给出锦标赛的理论基础，之后在第三部分说明我们的实证检验的数据、方法和结果，最后在第四部分对文章进行总结并且给出政策建议。

我们在原稿中同时使用了市级的数据进行了检验，得到了和省级数据一致的结论，这样也同样回应了陶然等（2010）对于政府层级选择方面的质疑。碍于篇幅，市级样本的结果没有汇报出。

二、官员晋升锦标赛的理论基础：考核指标的最优数量

我们在这里通过一个道德风险模型 (moral hazard model) 构建官员晋升锦标赛的理论基础。具体的, 我们计算委托人 (也即上级官员) 的预期收益从而得出官员考核指标的最优数量: 如果考核指标的最优数量为 0, 那么就应和了陶然等 (2010) 的结论, 即我国并不存在一个锦标赛式的晋升机制, 官员的提拔靠“德能勤绩廉”五个方面综合考虑, 并且没有固定的标准; 如果最优的考核指标数量为 1, 在这里我们假设这个单独的指标为经济增长, 那么说明周黎安 (2007) 提出的官员 GDP 锦标赛的体制是适用的; 如果得出的最优考核指标的数量大于 1, 也即在经济增长之外还有其它的指标需要考虑, 那么我们认为在这样的锦标赛之下, 除了经济增长还有其它的指标, 这符合邱法宗和张霖星 (2007)、徐风华和王俊杰 (2006)、尤建新和王波 (2005) 等等对于多个考核指标的呼吁。

我们假设, 上一级是委托人, 而下一级官员是代理人。我们首先假设上级官员, 也即委托人是风险中性的, 而下级官员 (代理人) 的合同是线性的, 即 $w(x) = D + Ex$, 官员绩效 (产出) 服从一个正态分布, $x = a + \epsilon$, 其中 $\epsilon \sim N(0, V)$ 。我们进一步假设下一级官员的效用函数是一个指数型的, 绝对风险厌恶系数为常数 (Constant absolute risk aversion, CARA), 即 $u(w, a) = e^{-K(D + Ex - C(a))}$, 代理人的保留效用为 $\bar{U} = e^{-KR}$ 。对于委托人, 我们可以求出他的最大化收益的问题:

$$\begin{aligned} \max_{a, D} \quad & \mathbb{E}[D + Ex] - E \\ \text{s.t.} \quad & \mathbb{E}[e^{-K(D + Ex - C(a))}] \geq \bar{U} \quad (IR) \\ & a = \arg \max_a \mathbb{E}[e^{-K(D + Ex - C(a))}] \quad (IC) \end{aligned}$$

因为 $x = a + \epsilon$, 而 ϵ 服从一个正态分布, 因此我们有:

$$\mathbb{E}[e^{-K(D + Ex - C(a))}] = e^{-K(D + C(a))} \mathbb{E}[e^{-KEx}] = e^{-K(D + C(a))} e^{-\frac{1}{2} K^2 E^2 V}$$

那么原来的最大化问题变为:

$$\begin{aligned} \max_{a, D} \quad & (D + Ea) - E \\ \text{s.t.} \quad & e^{-K(D + C(a))} e^{-\frac{1}{2} K^2 E^2 V} \geq e^{-KR} \end{aligned}$$

我们按照一般的假设, 如霍尔斯特罗姆和米尔格罗姆 (Holmström 和 Milgrom, 1991), 令 $C(a) = \frac{1}{2}ca^2$, 将这一值代入, 可以得到最优的解为

$$a = \frac{1}{c(1 - cV)}, D = \frac{1}{1 - cV} (R - \frac{1}{2}cE^2V)$$

我们之前只讨论了一个任务, 也就是一般意义上, 官员考核仅仅看一个指标的情况 (周黎安, 2007), 如果认为官员考核不仅仅是一个指标, 而应该有多个的话, 那么官员就会面临一个多个任务的环境。

基于霍尔斯特罗姆和米尔格罗姆 (1991), 我们假设有 k 个任务, 也即有 k 个指标。那么代理人 (下级官员) 会为每一种任务付出 a_i 的努力水平。产出向量 $\mathbf{x} = \mathbf{a} + \boldsymbol{\epsilon}$, 其中 $\mathbf{x}^T = (x_1, x_2, \dots, x_k)$, $\mathbf{a}^T = (a_1, a_2, \dots, a_k)$ 并且 $\boldsymbol{\epsilon} \sim N(0, \boldsymbol{\Theta})$ 。其中:

$$V = \begin{bmatrix} V_{1,1} & \dots & V_{1,k} \\ \vdots & \ddots & \vdots \\ V_{k,1} & \dots & V_{k,k} \end{bmatrix}$$

努力成本函数依然使用 $C(\mathbf{a})$ 来表示, 我们依然假设工资是线性的, 即 $w(x) = \mathbf{D}^T \mathbf{x} + E$, 而委托人依然是风险中性的, 代理人的效用是指数型的绝对风险厌恶系数为常数 (CARA), 即 $u(w, \mathbf{a}) = e^{-\kappa(\mathbf{D}^T \mathbf{x} + E - C(\mathbf{a}))}$ 。我们进一步假设, 代理人使用一个 k 维向量 $\mathbf{B}$ 来给予各种产出权重, 并且为了使代理人的总产出与单一任务的总量一样, 我们进一步要求 $\sum_{i=1}^k B_i = 1$ 。这样, 我们的最大化问题变为:

$$\begin{aligned} \max_{\mathbf{a}, \mathbf{D}} & \mathbf{B}^T \mathbf{a} - \mathbf{D}^T \mathbf{a} - E \\ \text{s.t.} & \mathbb{E} [e^{-\kappa(\mathbf{D}^T \mathbf{x} + E - C(\mathbf{a}))}] = \bar{U} \quad (IR) \\ & \mathbf{a} = \arg \max_{\mathbf{a}} \mathbb{E} [e^{-\kappa(\mathbf{D}^T \mathbf{x} + E - C(\mathbf{a}))}] \quad (IC) \end{aligned}$$

由于 ϵ 服从正态分布, 我们有 $\mathbb{E} [e^{-\kappa(\mathbf{D}^T \mathbf{x} + E - C(\mathbf{a}))}] = e^{-\kappa(\mathbf{D}^T \mathbf{x} + E - C(\mathbf{a}))} \frac{1}{2} \sqrt{\frac{2\pi}{\kappa^2 \mathbf{D}^T \mathbf{D}}}$ 。我们的问题简化为:

$$\begin{aligned} \max_{\mathbf{a}, \mathbf{D}} & \mathbf{B}^T \mathbf{a} - \mathbf{D}^T \mathbf{a} - E \\ \text{s.t.} & \mathbf{D}^T \mathbf{a} - \mathbf{D}^T \mathbf{a} - C(\mathbf{a}) = \frac{1}{2} \sqrt{\frac{2\pi}{\kappa^2 \mathbf{D}^T \mathbf{D}}} \quad R(IR) \\ & \mathbf{a} = \arg \max_{\mathbf{a}} \mathbf{D}^T \mathbf{x} - C(\mathbf{a}) \quad (IC) \end{aligned}$$

由于多个任务和两个任务的原理基本一致, 如霍尔斯特罗姆和米尔格罗姆(1991), 我们仅考虑两个任务的情况。假设还有 GDP 增长率以外的一个指标, 那么假设 GDP 的增长率指标为 x_1 , 另外一个指标为 x_2 , 由于各种指标的波动率与 GDP 既可能正相关, 又可能负相关, 我们将其它任务的波动率与 GDP 的波动关系设为 0, 即 $\frac{1}{2} \sqrt{V_{11} V_{22}} = 0$ 。这一点假设也符合大多数研究我国官员晋升机制的文献, 如李宏彬和周黎安(2005), 周黎安(2007)和陶然等(2010), 因为在这几篇经典文献当中, 作者都没有加入其它的关于官员绩效的变量, 如果这些变量和经济增长率相关的话, 那么就会有遗漏变量偏差 (omitted variable bias), 那么得到的估计值就不准确了。

假设努力的成本为 $C(\mathbf{a}) = C(a_1, a_2) = \frac{1}{2}(c_1 a_1^2 + c_2 a_2^2) - k a_1 a_2, k \in [\sqrt{c_1 c_2}, \sqrt{c_1 c_2}]$ 这说明增加对一项任务的努力会对于另一项工作努力的边际成本的影响, 由于这一影响即可能是正的, 也可能是负的, 因为比如与 GDP 增长相关的任务, 如果增加对于促进 GDP 增长的努力的话, 那么这样的协同作用可能会降低为该任务的努力成本。反之, 如果两者任务的性质相互排斥, 那么增加 GDP 的努力可能会使得执行该任务的努力成本上升。通过一阶条件, 我们可以得到最优的努力水平:

$$a_1^* = \frac{c_2 D_1 - k D_2}{c_1 c_2 - k^2}, a_2^* = \frac{c_1 D_2 - k D_1}{c_1 c_2 - k^2}$$

我们将 (a_1^*, a_2^*) 代入, 可以得到, 得到:

$$\max_{D_1, D_2} B_1 \frac{c_2 - 1}{c_1 c_2} \frac{D_2}{k^2} - B_2 \frac{c_1 D_2}{c_1 c_2} \frac{k - 1}{k^2} - \frac{K}{2} (B_1^2 V_2 + D_1^2 \frac{1}{V_1} + \frac{c_2 - 1}{c_1 c_2} \frac{k - 2}{k^2} S^2) \quad (1)$$

$$\frac{1}{2} c_2 \frac{c_1 - 2}{c_1 c_2} \frac{D_1 S^2}{k^2} - k \frac{D_2}{c_1 c_2} \frac{k - 2}{k^2} - \frac{c_1 - 2}{c_1 c_2} \frac{D k - 1}{k^2} \quad (2)$$

一阶条件得到最优的分成比例：

$$D_1 = \frac{B_1 c_2 - B_2 k}{K_1 (c_1 c_2 - k^2)} D_2 = \frac{B_2 c_1 - B_1 k}{K_2 (V_1 c_2 - k^2)} \quad (3)$$

和之前的讨论一致，两个任务的产出可能相关也可能不相关，我们进一步假设两个任务的产出努力不相关，也即 $k=0$ 。这同样取自于李宏彬和周黎安（Li 和 Zhou，2005），周黎安（2007）和陶然等（2010）的假设。这时，我们有：

$$D_1^* = \frac{B_1}{c_1 V_1} D_2^* = \frac{B_2}{c_2 K_2^2} \quad (4)$$

我们给出了最简单的例子，我们同样比较了当 $V_2 \geq 0$ 同时 k 也不为 0 的情况，这样的比较只会从计算上增加复杂性，并不影响我们的讨论。

假设单一任务为 x_1 ，将单一任务和多任务的产出努力水平和合同（分成比例）带入，很容易得到委托人在安排单一任务和多个任务时的收益：

$$\mathbb{E}(\text{单一任务}) = \frac{1}{2c_1(1 - c_1 V_1)} K$$

$$\mathbb{E}(\text{多个任务}) = \frac{B_1}{2c_1(1 - c_1 V_1)} + \frac{B_2}{K 2c_2(1 - c_2 \frac{2}{2})} V + K$$

$$\frac{1}{2c_1(1 - c_1 V_1)} - B_2 \frac{c_1(1 - c_1 V_1) - c_2(1 - c_2 \frac{2}{2})}{2c_1 c_2(1 - c_1^2)(1 - c_2 \frac{2}{2})} \quad (5)$$

比较这两个收益，当 $c_1(1 - c_1 V_1) < c_2(1 - c_2 \frac{2}{2}) < 0$ 时，说明任务 2 的产出成本和产出的波动性的一个加权要高于任务 1，这时任务 2 的权重 B_2 取 0，多任务的合同退化为单一任务合同。否则，任务 2 的权重大于 0，那么超过 1 个任务的合同会给委托人带来更多的收益，因此多任务的合同会更好。

进一步，我们假设一种极端的情况。如果 $V_2 = 0$ （对称地，可以讨论 V_1 ），那么我们可以看到：

$$D_1 = \frac{B_1 c_2 - B_2 k}{K_1 (c_1 c_2 - k^2)}$$

说明如果 $B_1 c_2 - B_2 k < 0$ （此时我们放松了 $k=0$ 的假设），那么 $D_1 < 0$ ，这在如霍尔斯特罗姆和米尔格罗姆(1991) 中被称为“低能量的激励”。此时，委托人会提供固定工资合同，也就是不给代理人提供激励，这种情况印证了陶然等（2010）的说法，即没有固定明确的干部考核指标，只是笼统的以“德能勤绩廉”来考察官员，这个考核体系是个“黑箱”。

因此，我们得到如下的理论预测结果：

- (1) 如果官员的考核指标数量为 0，那么不存在一个晋升锦标赛机制。
- (2) 如果官员的考核指标数量为 1，那么存在一个单一维度考核指标的晋升锦标赛机制，根据以往的文献，我们假设为经济增长。
- (3) 如果官员的考核指标数量大于 1，那么存在一个多维度的考核指标体系。

三、实证检验

上一节给出了官员晋升锦标赛考核指标数量的理论基础中的一个难点,即我们无法知道下级官员为经济增长付出的努力成本到底是多少,因此我们必须另辟蹊径,进行实证检验。首先,我们先验的使用 F 检验对各个可能的考核指标的波动性进行检验,如果某个指标的波动性过大,那么这个指标就不适合作考核指标。接着,我们后验的按照李宏彬和周黎安(2005)、陶然等(2010)的方法,通过排序的 probit 模型考察各个可能的指标对于晋升/去职概率的影响。

(一) 数据和变量

我们的数据来自于 1978 年至 2010 年各个省级行政单位的党委书记的简历和《新中国 60 年统计资料汇编》。这些简历可以通过人民网以及维基百科检索得到。我们重点关注的是党委书记,因为党委书记正如其所在的辖区内的 CEO,见瓦尔德(Walder, 1995),如果有省长(直辖市市长、自治区人民政府主席)同时也是该辖区内的党委书记的话,我们就记他为党委书记,与李宏彬和周黎安(2005)的一致,以下简称“官员”。

我们的数据搜集了官员具体的个人信息,包括:(1)职位变动 y ,我们按照李宏彬和周黎安(2005)和陶然等(2010)的方法,当官员退休、结束任职或者降职时则 $y=0$,官员级别不变取 $y=1$,官员晋升则 $y=2$ 。(2)年龄。1982 年中央政府取消了干部终身制并且实行了强制退休的政策,这一政策贯穿于改革开放的进程中,见李宏彬和周黎安(2005)、陶然等(2010)。另外,为了更好的获取年龄对于晋升/去职可能的非线性影响,我们对官员加上一个指示 65 岁的虚拟变量(如果大于 65 岁,则取 1,否则为 0),因此我们预期年龄对于晋升概率的影响为负。(3)教育。教育取 1,如果这个官员有大学或者以上的背景,否则取 0。教育预示着官员的人力资本,见李宏彬和周黎安(2005),因此这一变量可能对晋升概率产生正向的影响。(4)与中央的联系,过去工作的中央背景会影响到官员晋升的可能性,见李宏彬和周黎安(2005)。(5)在位时间的长度。在位时间可能释放了官员在位能力不足的信号,见李宏彬和周黎安(2005),资历长的官员晋升的可能性反而会更小,甚至“终老”于这一职位,我们预期这个变量负向影响晋升的可能性。(6)政治周期。李宏彬和周黎安(2005)认为每五年一届的全国人民代表大会召开之际的官员更迭的频率更高。

对于考核指标,我们有:(1)经济增长。李宏彬和周黎安(2005)和陶然等(2010)均使用的是 GDP 的增长率,我们同样使用 GDP 的增长率,但是同时考虑到周期性,我们使用年增长率和任期内平均增长率来衡量。(2)教育发展。因为教育体现着人们的知识水平(傅勇和张晏,2007),并且这一指标也越来越受到国内学者的重视,我们用教育支出占总财政支出的比重来衡量(见郑磊,2008)。(3)医疗水平。联合国强调了预期寿命,而寿命延长的保障则是医疗,我们用医疗卫生支出占总财政支出的比重来衡量,如伊万诺娃、阿塞留斯和斯里尼瓦善(见 Ivanova、Arcelus 和 Srinivasan, 1999)。

表 1 汇报了基本统计量。只有频繁的工作调动才可以使得我们更好的观察这些官员的晋升轨迹和他们相应的考核指标,而我们的数据显示的确是如此,与李宏彬和周黎安(2005)一样,我们发现官员交流的频率比较高,每年平均有 21.2%的省级官员的工作发生了变动。这一频率和美国 CEO 的更换频率(20%)相当,详见丹尼斯和丹尼斯(Denis 和

Denis, 1995)。另一个很明显现象是,我国各省在 1978-2010 年间的经济增长率较高,平均每年增长均超过了 10%。省级官员中获得大学及其以上学历的比重占了 79%,这显示了中央政府对于受过高等教育的人才的重视。28.8%的省级官员在党中央/国务院担任过职务,平均任期为 3.46 年,最长的省委书记任期为 16 年。

表 1 基本统计量

变量	观测数	均值	标准差	最小值	最大值
是否升迁 (是=1, 否=0)	994	0.078	0.254	0	1
是否卸任 (是=1, 否=0)	994	0.105	0.344	0	1
年龄	994	58.017	5.101	42	75
年龄大于 65 (是=1, 否=0)	994	0.170	0.355	0	1
是否大学毕业 (是=1, 否=0)	994	0.793	0.370	0	1
是否有中央工作背景 (是=1, 否=0)	994	0.288	0.453	0	1
任期	994	3.460	2.183	1	16
GDP 增长率	994	0.113	0.066	-0.159	0.504
GDP 平均增长率	994	0.104	0.045	-0.159	0.296
教育支出占财政支出的比重	994	0.243	0.107	0.118	0.427
医疗支出占财政支出的比重	994	0.166	0.109	0.081	0.228

注:数据来自于《新中国 60 年统计资料汇编》《中国统计年鉴》《中国卫生统计年鉴》和人民网。海南从 1988 年开始进入省级行政行列,之前算广东的一个行政区,按地级计算。重庆从 1997 年开始计算。

(二) 实证检验方法

我们先验的使用 F 检验考察潜在指标的波动性的相对大小,根据我们模型的推断,波动性比较大的指标不适合考核指标,而波动性并不是很大的变量则很有可能是官员晋升的考核指标。通过计算官员任期内各个指标的每年增长率的方差和平均增长率的方差从而得到各个指标的波动性。具体的,我们首先使用 F 检验考察他们的波动性是否相等,进而检验各个指标相对波动性的大小,最后通过求 F 检验的分位数求出波动性较大的指标的方差在统计意义上相当于波动性小的指标的方差的倍数。

与此同时,我们后验的考察各个潜在的考核指标对于晋升可能性的影响,如见李宏彬和周黎安 (Li 和 Zhou, 2005),陶然等 (2010),实证模型为:

$$y_{i,t} = \beta_1 GDP_{i,t} + \beta_2 Education_{i,t} + \beta_3 Health_{i,t} + \beta_4 X_{i,t} + u_{i,t} + \mu_i + \epsilon_t$$

其中, $y_{i,t}$ 为一个排序的离散随机变量,对应着提拔 (=2)、平级调动 (=1) 和去职 (=0), $GDP_{i,t}$ 、 $Education_{i,t}$ 和 $Health_{i,t}$ 分别对应着我们提到的经济增长、教育发展和医疗水平。 $X_{i,t}$ 是一系列的控制变量,包括年龄、教育背景、中央工作的背景以及任期。残差项中, μ_i 是省份的个体固定效应 (provincial/city fixed effect), 不随时间变化而变化; ϵ_t 是一个共同的冲击,刻画时间的固定效应 (time fixed effect), 不随个体的变化而变化,我们使用排序的 probit 模型进行估计。

(三) 实证结果

1、各个指标的波动性比较

我们首先考察各个指标波动性的大小,表 2 给出了假设检验的结果。首先,我们假设各个指标的波动性在统计上是相等的,通过 F 检验,我们发现经济增长的波动性与教育发

展和医疗水平的波动性并不相等，我们进一步假设经济增长的波动率小于教育发展和医疗水平的波动率，表 2 的列（1）和列（3）给出了检验结果，可以看到 p 值均远大于 0.10，因此在 10%水平下不能拒绝“经济增长的波动性小于教育发展/医疗水平”的原假设。

我们进一步探究在 10%水平下，教育发展和医疗水平在统计上相当于经济增长率的多少倍。具体的，我们通过求出 F 分布的 10%分位数，得到在 10%水平下与教育发展/医疗水平在统计上无差异的最小的波动率数值，再使用这一数值除以我们计算出的经济增长的波动率，得到这一倍数。表 2 的列（2）和列（4）汇报了相应的值，可以看到教育发展和医疗水平的波动率从统计上而言是 GDP 年增长率的 5.5 倍，是任期内年平均增长率的 10 倍多，这说明教育发展和医疗水平的波动率相对于经济增长率而言非常大，令 V_1 是经济增长指标的波动率， V_2 表示教育发展或者医疗水平的波动率，那么 $c_1(1 - c_1 V_1^2) - c_2(1 - c_2 V_2^2)$ 的可能性更大，因此教育发展和医疗水平不宜作为官员竞争的考核指标，那么上级官员为下一级官员设计的最优晋升“合同”应该仅包含经济增长这个指标从而最大化上级官员的收益。

然而，我们需要注意的是，官员对于各个指标的努力成本不可测度，因此我们的结果只能说明，在给定努力成本的情况下，波动性小的经济增长指标作为考核重点的可能性更大。

表 2 各个指标波动性的比较

列	(1)	(2)	(3)	(4)
波动性指标	教育发展		医疗水平	
GDP 年增长率	小于	N=5.423	小于	N=5.518
	P ≈ 1.0000	P=0.100	P ≈ 1.0000	P=0.100
GDP 平均增长率	小于	N=10.212	小于	N=10.409
	P ≈ 1.0000	P=0.100	P ≈ 1.0000	P ≈ 1.0000

注：数据来自于《新中国 60 年统计资料汇编》、《中国统计年鉴》、《中国卫生统计年鉴》和人民网。

2、各个指标对于晋升/离职可能性的影响

上一节我们从先验的角度探究了哪些指标适宜作考核指标而哪些不适合，我们接下来从后验的角度检验经济增长、教育发展和医疗水平是否会影响到晋升/离职的概率，同时也是对于李宏彬和周黎安（2005）、陶然等（2010）的争论进行重新的考察。

我们同时使用了全样本和 1994 年之后的样本进行分析，因为 1994 年的分税制改革之前“财政联邦主义”可能会对这一锦标赛机制产生重要的影响，见金和辉、钱颖一和温格斯特（2005），而李宏彬和周黎安（2005）的数据跨度为 1979 年到 1995 年，并没有包含 1995 年之后的数据，陶然等（2010）的数据从 1979 年到 2002 年，也忽略了最近些年的情况，因此我们有必要对 1994 年之后的样本进行考察可以更加充分的评估这一机制的存在性。与李宏彬和周黎安（2005）、陶然等（2010）的一致，我们考虑到各个省的异质性，使用稳健性标准误进行估计。表 3 汇报了使用排序 probit 模型的回归结果。

表 3 的前两列给出了使用全样本时的估计结果，可以看到只有 GDP 的增长率会正向影响官员晋升的概率而会负向影响官员去职的概率，并且在统计上显著，而教育和医疗对于晋升/去职的影响在统计上并不显著。具体的，GDP 的年增长率对于晋升/离职的可能性

我们同样对不包含控制变量的模型进行了估计，得到的结果和全模型的一致。

的影响在 5%水平下显著而任期内的 GDP 平均增长率在 1%水平下显著，并且这一影响有足够大的边际值：在经济增长率均值附近，GDP 的年增长率对于晋升概率的边际影响为 0.195，对于离职概率的边际影响为-0.266，那么 GDP 的年增长率每增加一个标准差(0.066)，晋升的概率会增加 0.0129，相当于其均值的 16.5%，同时离职概率会下降 0.0175，相当于离职概率均值的 16.8；任期内的年均 GDP 增长率每增加一个标准差对于官员升迁和离职概率影响分别为 0.0259 和-0.0347，相当于升迁和离职概率均值的 32.7%和-33.3%。这些结果说明经济增长对于官员晋升的影响不仅仅在统计上显著，在经济上也是不可忽视的。

表 3 的后两列给出了对 1994 年之后的样本进行检验结果，经济增长、教育发展和医疗水平的显著性和估计值的大小都与全样本几乎完全一致：仅有经济增长这一个指标会显著的正向影响官员晋升的概率，负向影响官员离职的概率，而其他两个指标均在统计意义上不显著。并且，1994 年之后的样本由于受到金和辉、钱颖一和温格斯特（2005）提出的财政联邦主义的影响更小，因而得到的估计更可信，可以看到经济增长的估计系数更大，边际效应也在经济意义上更加显著，这说明我们的估计是稳健的，结果并不是由财政联邦主义引起的。

对于控制变量，我们发现年龄对于晋升的影响都是负的，并且在 10%水平下显著，而是否大于 65 岁的虚拟变量则在 1%水平下显著：官员的年龄每增加 1 年，其晋升的概率下降 0.41%，而离职的概率增加 0.47%；年龄达到 65 岁意味着官员晋升的概率下降 0.057，将近总晋升概率的 73%，离职的概率会增加 0.093，相当于总离职率的 85.7%。此外，教育水平也是影响官员晋升的一个特质因素，这一变量在 10%水平下显著：给定其他条件不变，拥有大学本科文凭会使得官员的晋升概率增加 1.4%，去职概率则降低 1.47%，这些结果既说明了我国对于年轻干部的重视，也反映了教育对于官员人力资本水平的反映，与见李宏彬和周黎安（2005）的结果一致。同时，官员过去在党中央/国务院工作的经验将会增加晋升的概率并且降低去职的概率，这一变量不仅仅在统计意义上显著，其边际影响也是不容忽视的，中央工作背景会使得官员的晋升概率增加 3.56%，离职概率降低 3.65%。我们的结果支持了上一级党委/政府的工作经历会帮助建立与主管官员任命权的上级领导的非正式关系，详细的论述见李宏彬和周黎安（2005）。另外，我们发现官员任期的长度会负向影响其晋升的可能性而正向影响官员去职的可能性，并且这一影响在 10%水平下是显著的，这说明在位时间过长释放了官员在位能力不足的信号，同样验证了见李宏彬和周黎安（2005）的结论。

表 3 对于升迁、去职可能性的考察

变量	全样本	全样本	1994 年之后	1994 年之后
GDP 增长率	1.616** (2.027)		1.868*** (2.897)	
GDP 年均增长率		4.700*** (2.595)		4.977*** (2.993)

我们在原稿中也使用了 1985 年之后的样本进行考察，因为我国的干部年轻化和老干部离休可能会对我们的结果产生影响，详见李宏彬和周黎安（2005），我们依旧得到了一致的估计，结果是稳健的。

M	†'''	†'''	1994 M - ^a	1994 M - ^a
•â]•ÿ ¥1×	2.414 (1.023)	2.553 (1.018)	3.136 (1.355)	2.799 (0.893)
D]•ÿ ¥1×	2.675 (1.075)	2.206 (1.070)	2.659 (1.086)	2.643 (1.345)
M -	-0.0212* (-1.922)	-0.0200* (-1.918)	-0.0330* (-1.903)	-0.0311* (-1.903)
^Žvč 65 ß 8	-1.090*** (-2.857)	-1.081*** (-2.855)	-1.244*** (-2.788)	-1.245*** (-2.789)
^ŽvĐ8<	0.178* (1.848)	0.154* (1.745)	0.177** (1.996)	0.186* (1.999)
^Žμī /8¾ ß	0.441*** (2.862)	0.454*** (2.860)	0.566*** (2.965)	0.578*** (2.998)
©ùÉ	-0.0425* (-1.770)	-0.0403* (-1.781)	-0.0658* (-1.818)	-0.0697* (-1.824)
ī Ä .1	-2.860*** (-2.914)	-2.380*** (-2.907)	-3.052*** (-2.659)	-3.001*** (-2.877)
ī Ä .2	0.166 (1.202)	0.489 (0.895)	1.777 (1.277)	0.585 (1.021)
4 © "	994	994	523	523
À R ²	0.107	0.099	0.222	0.223

ÿ " Ÿ1çu•İS 60 Md9' ølv uİSd9Mkv uİSİ3d9Mkv„|ö©b

áì¥ 9 ç©T y_{i,t} D₁GDPgr_t E₂Education_t Hea_t X_t u_{i,t} u_{i,t} P_t E_{i,t} O H

t'ø ç"|=ø ç"|= *** V U p<0.01 ** V U p<0.05 * V U p<0.1b

¹a²,„ÿ|çl

'Óæ" Bñi£„B ~ó áS2ôĖ6ÆS× Å¥BñØ, \$ iO²
†ás8)ªě¥ěð"), fBÆS× Å¥i ŸÙ5 £L ás'-¥2ô
Ė6ÆS×8Å¥i iOªüfB Å^[Ü69É¹ö1¥I,=,b

ál¥ûîμùç×•,j' -9>¥2ôl, Å 9'ál 'Ó£L¥[Ü6
9É¹,•¥Ė6ÆS× V7÷z¥Ø³f"¥Ÿ l,8"b[Ü69É¹ \$¥Ė
6 ã Bç¥âNHùØî 2ô?Z¹ZÜ6¥ Ÿ -7Û-7Ÿ¥¹uséa
¹Z ¹ölpau×¥sý„æ†4[LC Bt¹u.À C 0Ÿés¥C` Û
Ó½ 2004 2007 b Q f"¥ ãöPÜ(©ù,žØMö¥¹Z2ô¥ ù>¹
÷Fø× ðìöÜ÷ ¥Úīb V© a? ù= î ¥Ÿ 'ī 1Â1ÿ

'y!„g' 7-j |ī''g'„ > ¥Ù5 °s„f 2007 b Q

Ė6ÆS× Å9] ás³E,>!¥+Ä 8èf"L" ã!ê¥ÆS× ç

<¥“û# Áf¥ ¹i, !Eprī yN¹ZŸ©]Hò îô„*u=g

ÆÜ6Š'ô¥~ä f" l¥~ä¹ásŸ©-?¥ M@F Bç¥Eīb

ë ÆS× Å{Ÿ¥Ù5 Ÿ©<¾Â... €³%;\$Bñ°4¥XE^~¶ás2

