

何时应该杜绝“搞变通”¹

——一个非正式授权理论解释

李石强²

摘要 党的十九届四中全会提出要杜绝制度执行中“搞变通”的现象。本文构建了一个一般化的委托人具有事后私人信息的两期模型对此进行解释。模型表明,中央政府不会允许地方政府在无法观察社会环境冲击的情况下进行灵活变通。但是,如果社会环境冲击幅度足够大,中央政府会允许有能力观察社会环境冲击的地方政府在第1期进行变通,然后在第2期沿用第1期的实际决策,以实现对冲击最好的应对。反之,如果社会环境较为稳定,中央政府对社会最优政策的认识清楚,那么从一开始就不会允许地方政府灵活变通,而是始终坚持执行最初的既定决策。

关键词 变通;授权;执行;中央政府;地方政府

0 引言

2019年党的十九届四中全会通过的《中共中央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》中,要求“健全权威高效的制度执行机制,加强对制度执行的监督,坚决杜绝做选择、搞变通、打折扣的现象”。

这种“做选择”“搞变通”或者“打折扣”的行为,是指地方政府没有完全严格执行已经明确制定并颁布的中央政策,而是出现了走样或者相应的“土政策”(“上有政策,下有对策”)。这无疑会导致中央令不行禁不止,也经常使得地方政府自身陷入各类行政纠纷或者舆论的中心^①。有意思的是,在改革开放初期,地方政府的灵活变通(flexible adaptation)行为却经常得到中央的默许,有些甚至在后来获得正式文件或法律的承认。一个典型例子就是农村“包产到户”制

¹ 本文初稿曾在首届中国微观经济理论论坛(中国人民大学)上报告,感谢龚炯、柯荣住、董保民和张文章的有益评论;感谢两位匿名审稿人和副主编翁翕的宝贵意见。文责自负。

² 李石强,中国社会科学院大学经济学院副教授,E-mail:lishiqiang@cass.org.cn。

^① 概览式的介绍,可参见Pei(1997)、O'Brien and Li(2004)和包万超(2013)。

的命运。1956年5月,中共永嘉县委就在雄溪乡燎原社进行了农业生产产量责任制的试验,在中国首创“包产到户”。但是,在1957年中央高层领导获知此事后,责令浙江省委、温州地委、永嘉县委予以纠正。然而,安徽凤阳小岗村1978年实行的“包产到户”却获得了中央的容忍,最终被正式承认。

一些关注中国公共管理与政策的社会学家和法学家早已发现了这一现象。例如,Lampton(1977)在研究中国医疗卫生政策制定和执行的时候就观察到了地方政府的变通行为。王汉生等(1997)将变通视为在单一制决策体系运作失灵和上下级政府间讨价还价失灵并存的情况下出现的一种制度运作方式,并尝试在制度变迁的框架下描述其出现的条件。翟学伟(1997)从中国的社会结构转型去分析“土政策”的产生及其功能。刘鹏(2012)将变通具体分为四种不同形式,考虑中央政策是否与地方政府利益冲突以及地方政府是否具备实施变通的条件,对这些形式分别进行解释。刘骥和熊彩(2015)讨论了地方政府在计划生育工作中的变通行为。周雪光(2017)将地方政府在政策执行中的灵活变通视作中国的一统体制与有效治理矛盾的一个应对机制。王清(2021)则将地方政府在执行政策时出现偏差归因于在模糊的属地管理下中央对地方的压力和激励失效。

遗憾的是,经济学界虽然长期关注中国分权化改革过程中的中央与地方关系,却没有正式讨论过地方政府对中央政策的变通行为。本文主旨就在于此。通过构建一个非正式授权(informal delegation)模型,本文讨论了地方政府在执行中偏离既定政策的灵活变通行为,以及中央对该现象所采取态度的决定因素,从而在理论上解释当前为何要杜绝搞变通。

本文之所以从授权的角度进行讨论,是因为在中国的单一制政治体制下,中央掌握对正式制度的制定和解释权,拥有最终权威。地方政府无论拥有什么权力,原则上都是由中央自上而下授予的,必要时随时可以被收回。进一步地,之所以要构建一个“非正式”授权模型,则是因为已有的授权模型实际上讨论的都是委托人对代理人的“正式”授权。在政策制定和执行中,这种正式授权表现为中央在政策中留下一些空白或模糊之处,允许地方政府在执行中因地制宜或因时制宜,使其在一定范围内具有了裁量权(discretion)^①。这种为地方政府刻意准备的“留白”实际上是中央事先制定的政策框架中的一部分,一旦其范围被明确划定,就被中央在事前正式授予了地方政府,直到下一轮政策调整为止。相比而言,灵活变通行为则是地方政府对已有政策正式授权范围的偏离或突破,中央是可以当即予以纠正甚至处罚的。但是,如果中央打算容忍地方政府

^① 关于中国的政策试验传统,参见韩博天(2019)。

的这种做法,就相当于给予其以事实上的授权。不同之处在于,这种授权是临时的、事后的和非正式的。

具体地,本文沿用已有的授权文献关于委托人通过赋予代理人一定决策权来提供激励的基本思想。但是,为了模型化非正式授权行为,本文作出了三个方面的扩展^①:第一,将分析从单期扩展至两期,中央在每一期的决策都是是否容忍地方政府偏离上一期的既定决策(灵活变通);只有这样,才能回答中央为何对灵活变通做法的态度最多是默许,而不是将其正式化^②,进而解释中央对其态度的变化。第二,要考虑到地方政府一旦偏离既定政策,就存在被指责为与中央政策相抵触的风险;换句话说,执行既定政策对地方政府才是最安全的,灵活变通必须伴以相应的激励^③。第三,假设对于社会环境所受到的随机性冲击,中央和地方政府在每个时期初都处于无知状态;在期中,地方政府若付出努力,可以以一定的概率获知;时期结束后,中央自动得知。双方也不能就社会环境冲击进行沟通(communication),从而排除掉地方政府临时向中央请示汇报以寻求政策调整的复杂情况;这样,如果发生了变通行为,只能在期末留待中央依据社会环境冲击的事后实现值来加以评判。

本文模型表明,中央是否默许地方政府灵活变通,取决于社会环境随机性变化的幅度大小。在单期模型中,默许地方政府偏离既定政策以响应社会环境冲击的好处会被为了实现这一点而付出的激励成本所抵消,并不如坚持既定政策更为经济。但是,在两期模型中,社会环境受到的冲击会累积起来,相对于既定政策可能越来越远,导致始终要求地方政府坚持既定政策的代价高昂。如果在第1期默许并激励地方政府灵活变通而在第2期再转为沿用既定政策,在总效果上可能优于从一开始就采用既定政策。这样,中央与地方政府之间达成的默契是如下一套方案:如果社会环境冲击不够大,中央会要求地方政府自始至终坚持既定决策;反过来,如果社会环境冲击足够大,高能力地方政府在第1期付出努力去了解社会环境所受的冲击。如果成功,地方政府便灵活变通,否则就按照初始的既定政策行事;低能力地方政府则不付出努力,直接执行既定政策。到第2期,地方政府无论能力高低都只是执行各自在第1期的既定政策。

本文在理论上将偏离规则的非正式授权与包含在规则中的正式授权区分开来,对中央过去不时默许地方政府灵活变通的条件进行分析,解释了当前形势下中央要求杜绝搞变通的原因。由于模型中委托人在第1期期末可能会获

① 已有文献在授权的主题下,沿着决策规则本身是否可承诺以及信息是否可验证的路线,对这种正式权力的最优形式已有许多讨论。参见 Gibbons et al. (2012) 和李石强(2015) 的综述。

② 例如,颁布正式文件以允许地方试验。

③ 本文不考虑外部利益方通过贿赂方式诱使行政部门偏离既定决策的情况。

得对代理人的信息优势,本文针对 Gibbons et al. (2012)对委托人信息优势下的决策权配置问题鲜有讨论的评论作出了理论尝试。最后,由于地方政府灵活变通在成功后会得到中央的默许,下一期的政策便随之改变。在此意义上,本文也为环境不确定下的制度变迁研究提供了一个思路。

本文余下部分安排如下:第1部分是模型的基本设定;第2部分推导单期和两期情况下灵活变通作为一种治理方式的最优形式;第3部分将默许灵活变通与执行既定决策这两种治理方式进行比较,得出中央在两者之间选择的条件;第4部分是中国的例证;最后是结语。

1 模型设定

1.1 社会环境

考虑两个时期。用 B_t 表示社会环境在第 t 期的取值 ($t=1,2$)。鉴于本文的研究目的,我们关注那些难以验证从而不可契约化的社会环境信息,例如外交、军事、国际形势、气候或全国政治经济局势等。政府在制定政策时需要依据这类社会环境,却无法基于该信息设计一个状态依存的转移支付方案。

假设在每个时期初,社会环境都会受到独立的服从两点分布的随机冲击 b_t , $\text{Prob}(b_t=\bar{b})=\text{Prob}(b_t=-\bar{b})=1/2$, 其中 $\bar{b}>0, t=1,2$ 。这样,社会环境在第1期冲击 b_1 的基础上,在第2期会继续受到独立的冲击 b_2 , 在第 t 期末的最终位置为 $B_t=B_0+\sum_i b_i$ 。为方便起见,假设社会环境的初始位置为 $B_0=0$ 。于是可以得到, $B_1=b_1, B_2=b_1+b_2 \in \{-2\bar{b}, 2\bar{b}, 0\}$, 每个取值的概率为 $\text{Prob}(B_2=-2\bar{b})=\text{Prob}(b_1=b_2=-\bar{b})=1/4$, $\text{Prob}(B_2=2\bar{b})=\text{Prob}(b_1=b_2=\bar{b})=1/4$, $\text{Prob}(B_2=0)=\text{Prob}(b_1=\bar{b})\text{Prob}(b_2=-\bar{b})+\text{Prob}(b_1=-\bar{b})\text{Prob}(b_2=\bar{b})=1/2$ 。假设这些都是公开信息。

1.2 委托人(中央政府)

令 y_t 是关于某事项在第 t 期所采用的决策。鉴于社会环境处于不断变化之中,对于委托人或中央政府而言,最理想的情况是每一期的 y_t 都能对受到随机冲击后的当期社会环境 B_t 予以适当地响应,而不是简单地延续上一期的既定决策 y_{t-1} 。不妨假设 $y_0=0$ 。结合前面 $B_0=0$ 的假设,这意味着初始决策是与社会环境相匹配的。这样,委托人的单期(Bernoulli)效用函数可以设定为如下拟线性形式:

$$u_t^P = -(y_t - B_t)^2 - w_{it}(\cdot) \quad (1)$$

其中, $w_{it}(\cdot)$ 是在第 t 期向代理人支付的报酬, i 表示代理人的不同类型。

根据这一效用函数, 给定支付给代理人的报酬, 委托人在第 t 期最希望采取的决策(满足点) 是 $y_t^{\text{Optimal}} = B_t$, 亦即完全反映社会环境当前状况的政策。实际决策 y_t 与当前社会环境的差距越大, 带给委托人的当期效用损失就越大^①。

假设委托人在每期期初都不知晓社会环境在当期将要受到何种冲击, 因此在每期期初也就不知道应该采用什么样的决策才是最合适的。不过, 在每一期的期末, 委托人会自动(零成本) 得知当期冲击的具体值。在这样的简化假定下, 委托人就能在第 1 期期末对代理人的工作进行评价(并决定是否予以非正式的认可), 从而才有可能在第 1 期采取与第 2 期不同的治理方式。在现实中, 对于外交、军事、国际国内整体局势这类全局性的事项, 中央和地方可能在一开始都不清楚相关冲击的情况。但是, 由于中央政府对于这些事项比地方政府持有更高的关切度, 结合其综合全局的地位, 在经过一段时期(比如一年) 之后便往往能够获知其当期的真实值。

1.3 代理人(地方政府)

在每期期初, 代理人也不知晓社会环境在当期将要受到的冲击值 b_t , 与委托人之间关于社会环境是信息对称的。不过, 对于该冲击值, 代理人或地方政府可以在当期中通过付出努力 e_{it} 以概率 $\rho_i e_{it}$ 观察到。其中, ρ_i 是代理人的能力水平, $i \in \{L, H\}$ 。为了简化, 假设 $\rho_H > \rho_L = 0$, 亦即对于低能力代理人而言, 无论付出多大努力都无法得知环境冲击。同时, 为了保证得到定义良好的概率值, 假设 ρ_H 的绝对数值足够小, 足以使得 $\rho_H e_{it}^*$ 始终处于区间 $[0, 1]$ 之中。假设委托人在第 1 期初并不确切知道代理人的能力水平, 只知道代理人是高能力的先验概率为 α_1 , $0 < \alpha_1 < 1$ ^②。

假设代理人即便在每期间观察到了冲击值, 也不能就此与委托人进行信息交流^③。于是, 不管因为什么原因偏离(在上一期做出的) 既定决策, 代理人都会因为没有得到委托人的许可而承担相应的责任风险。这意味着, 除非获得足

① 可以等价地将 b_t 解释为委托人偏好函数乃至决策满足点受到的随机冲击。这是因为 b_t 不必是客观的, 而可能是委托人对外部环境变化程度的主观认识, 也就类似于偏好冲击。第 4 节对这两种解释的现实对应都进行了梳理。

② 后面会表明, 委托人在期末可以依据第 1 期的结果对代理人是高能力的概率 α_2 进行后验推断, 并且 $0 \leq \alpha_2 \leq 1$ 。值得提醒注意的是, 当 $\alpha_1 = 0$ 时, 委托人不会允许任何灵活变通。当 $\alpha_1 = 1$ 时, 问题退化为一个道德风险问题。

③ 这就避免了已有文献所关注的代理人通过策略性的信息传递对决策制定施加影响的情况, 将焦点集中在灵活变通上。Austen-Smith(1994) 讨论了允许这种交流情况下的信息传递机制。这时, 代理人可以通过信息操纵而影响委托人的决策制定。有关授权的文献对此已有充分讨论, 参见本文引言中的引用。

够的激励(转移支付),代理人在每个时期都是不愿意偏离既定决策的,上一期既定的 y_{t-1} 就是代理人在第 t 期最为偏好的决策^①。另外,假设不论能力高低,代理人付出努力而承担的成本都为 $c(e_{it}) = \frac{1}{2}e_{it}^2$ 。于是,代理人在第 t 期的(Bernoulli)效用函数可以写为:

$$u_{it}^A = -(y_t - y_{t-1})^2 - \frac{1}{2}e_{it}^2 + w_{it}(\cdot) \quad (2)$$

最后,假设代理人在每一期都只能独立地对当期冲击 b_t 进行观察,委托人(此时相对于代理人具有了信息优势)在第 t 期末无法就 b_t 与代理人有任何形式的交流。于是,如果代理人未能观察到 b_t ,即便自己在第 $t+1$ 期观察到了 b_{t+1} ,也无法将其和 b_t 相结合以得出委托人偏好在第 $t+1$ 期的实际位置(即冲击的累计值 $B_t = \sum_t b_t$)。后面会看到,这种对上一期结果的路径依赖将两个时期联系了起来,正是使得两期模型有意义的关键之处。只有如此,中央才会想要通过事后对高能力代理人进行非正式授权的方式来激励其在期中去努力探知社会环境冲击^②。

1.4 两种治理方式

无论在第1期或者第2期,委托人既可以要求代理人严格执行既定决策,也可以默许其灵活变通(亦即非正式授权)。为了叙述上的方便以及与现实的对,我们不考虑代理人的混合行动策略,并且假定不同能力的代理人在劳动市场上能够获得相同的保留收入 $\underline{w}$ ($\underline{w} > 0$)。两种治理方式的基本含义如下。

1.4.1 执行既定决策

委托人可以要求代理人简单地执行既定决策 y_{t-1} 。在第1期初,既定决策指的是初始状态外生给定的 $y_0 = B_0 = 0$ 。在第2期,既定决策是指在第1期实际采取的决策 y_1 。因此,执行既定决策就是一个简单的委托人要求代理人执行某项已知的明确命令。可以预想,执行既定决策虽然简单易行,但会由于无法适应变化中的社会环境而给委托人造成效用损失,并且该损失可能随相同既定政策执行的时期增加而增大。具体地,结合委托人和代理人的效用函数以及随机

① Merton(1940)和Thompson(1961)将组织中的执行者对既定规则的偏好称为“官僚人格”(bureaucratic personality)。Bozeman and Rainey(1998)运用美国的数据检验了这一点,发现这种官僚人格在公共部门里尤为明显。

② 否则,如果代理人在与第1期末也自动得知环境信息,在第2期初面临的场景就会与第1期初完全相同,委托人在第1期也就没有任何必要允许高能力代理人灵活变通了(根据命题4,要求代理人执行既定决策即可),本文也就失去了意义。

冲击的概率分布,分别考虑两个时期的情况,我们关于这种治理方式有如下结论(所有命题的证明见附录):

命题 1: (1) 若委托人要求代理人执行既定决策,则不管代理人属于何种类型,一律支付 $w_{it, y_t=y_{t-1}}=w$, 代理人付出努力 $e_{it, y_t=y_{t-1}}=0$, 选择 $y_t=y_{t-1}$;

(2) 在第 1 期初,委托人对在当期执行既定决策($y_1=y_0=0$)的无条件期望效用为 $Eu_{1, y_1=0}^P = -w - \bar{b}^2$, 对第 2 期继续执行既定决策($y_2=y_1=0$)的无条件期望效用为 $Eu_{2, y_2=0}^P = -w - 2\bar{b}^2$;

(3) 在第 2 期初,委托人对在当期执行既定决策的条件期望效用取决于第 1 期的契约选择及决策 y_1 : 若 $y_1=y_0=0$, 那么 $y_2=0$, $E(u_{2, y_2=y_1}^P | y_1=0) = -w - 2\bar{b}^2$; 若 $y_1=b_1$, 则 $y_2=b_1$, $E(u_{2, y_2=y_1}^P | y_1=b_1) = -w - \bar{b}^2$ 。

由该命题可见,从中央政府的角度来讲,若希望地方政府执行(上一期的)既定决策,就不会对地方政府施加任何搜寻社会环境信息的激励,只需向其支付固定的保留收入(例如财政转移支付等)即可。

1.4.2 默许灵活变通

如果委托人默许代理人在第 t 期灵活变通而偏离 y_{t-1} , 就有可能充分适应当期的社会环境冲击。要做到这一点,需要对代理人进行恰当的激励。根据显示原理(Revelation Principle), 委托人只须考虑“讲真话”机制: 代理人在付出一定努力后, 如果观察到社会环境受到的冲击, 就偏离既定决策加以正确响应(亦即选择 $y_t=y_{t-1}+b_t$, 获得 $w_{it}(b_t)$); 否则, 就继续执行在上一期确定的既定决策(亦即选择 $y_t=y_{t-1}$, 获得 $w_{it}(y_{t-1})$)。我们先在给定代理人灵活变通的报酬 $w_{it}^*(b_t)$ 、坚持既定决策的报酬 $w_{it}^*(y_{t-1})$ 以及激励代理人付出努力 e_{it}^* 的情况下, 给出计算委托人在两个时期预期效用的思路, 而将该方式最优形式如何确定的问题留到下一节。

在第 1 期期初, 委托人通过在当期默许代理人灵活变通所能获得的预期效用为:

$$Eu_{1, FA}^P = \text{Prob}(b_1 = \bar{b}) E(u_1^P | b_1 = \bar{b}) + \text{Prob}(b_1 = -\bar{b}) E(u_1^P | b_1 = -\bar{b})$$

其中, 考虑到 $\rho_L=0$, 可得 $E(u_1^P | b_1 = \bar{b}) = \alpha_1 [-\rho_H e_{H1}^* w_{H1}(\bar{b}) + (1 - \rho_H e_{H1}^*) (-\bar{b}^2 - w_{H1}(y_0))] + (1 - \alpha_1) [-\rho_L e_{L1}^* w_{L1}(\bar{b}) + (1 - \rho_L e_{L1}^*) (-\bar{b}^2 - w_{L1}(y_0))] = \alpha_1 [-\rho_H e_{H1}^* w_{H1}(\bar{b}) + (1 - \rho_H e_{H1}^*) \cdot (-\bar{b}^2 - w_{H1}(y_0))] + (1 - \alpha_1) (-\bar{b}^2 - w_{L1}(y_0))$ 。同时, $E(u_1^P | b_1 = -\bar{b})$ 的形式与之相同, 仅仅将 $w_{i1}(\bar{b})$ 换成 $w_{i1}(-\bar{b})$ 。于是可将上式进一步写为:

$$Eu_{1, FA}^P = -\alpha_1 \left[\frac{1}{2} \rho_H e_{H1}^* (w_{H1}(\bar{b}) + w_{H1}(-\bar{b})) + (1 - \rho_H e_{H1}^*) (\bar{b}^2 + w_{H1}(y_0)) \right] - (1 - \alpha_1) (\bar{b}^2 + w_{L1}(y_0)) \quad (3)$$

在第2期,社会环境会在 b_1 的基础上继续受到冲击,其最终位置为 $B_2 = b_1 + b_2$ 。在第2期期初,委托人可能面临两种情况:

第一种情况是,代理人在第1期成功观察到了 b_1 。由于委托人知道低能力代理人是无法做到这一点的,便能够推断代理人一定是高能力。除了这一点之外,委托人在第2期初面临与第1期类似的场景,对当期效用的条件期望类似于式(3)(仅须改变时间下标并且令 $\alpha_2 = 1$):

$$E(u_{2,FA}^P | y_1 = b_1) = -\frac{1}{2} \rho_H e_{H2}^* (w_{H2}(\bar{b}) + w_{H2}(-\bar{b})) - (1 - \rho_H e_{H2}^*) (\bar{b}^2 + w_{H2}(y_1))$$

第二种情况是,代理人在第1期未能观察到 b_1 。这时,委托人对其能力的推断会依据贝叶斯推断而向下修正: $\alpha_2 = \frac{\alpha_1 (1 - \rho_H e_{H1})}{\alpha_1 (1 - \rho_H e_{H1}) + (1 - \alpha_1) \cdot 1} < \alpha_1$ 。在第2期,代理人只能在 $y_1 = b_0 = 0$ 的基础上去尝试观察 b_2 。如果成功,灵活变通意味着选择 $y_2 = b_2$; 如果失败,就选择 $y_2 = y_1 = 0$ 。于是,委托人在第2期初对当期效用的条件期望为:

$$E(u_{2,FA}^P | y_1 = 0) = \sum_{b_2} \text{Prob}(B_2) \{ \alpha_2 [-\rho_H e_{H2}^* [(b_2 - B_2)^2 + w_{H2}(b_2)] + (1 - \rho_H e_{H2}^*) (-B_2^2 - w_{H2}(y_1))] + (1 - \alpha_2) (-B_2^2 - w_{H2}(y_1)) \}$$

对以上两式求全期望,便可得到委托人在第1期期初所计算的对自己在第2期默许代理人灵活变通中所能获得的无条件期望效用:

$$Eu_{2,FA}^P = \text{Prob}(y_1 = b_1) E(u_{2,FA}^P | y_1 = b_1) + \text{Prob}(y_1 = 0) E(u_{2,FA}^P | y_1 = 0) \quad (4)$$

其中, $\text{Prob}(y_1 = b_1) = \alpha_1 \rho_H e_{H1}^*$, $\text{Prob}(y_1 = 0) = 1 - \alpha_1 \rho_H e_{H1}^*$ 。

最后,委托人的目标就是通过设计支付给代理人灵活变通的报酬 $w_{it}^*(b_t)$ 和坚持既定决策的报酬 $w_{it}^*(y_{t-1})$, 最大化在第1期期初计算的两期总期望效用(不考虑时间贴现问题,亦即假设贴现率为0):

$$\max_{w_{it}^*(b_t), w_{it}^*(y_{t-1})} EU^P = Eu_{1,k_1}^P + Eu_{2,k_2}^P$$

其中, k_t 表示在第 t 期采用的契约类型,亦即默许灵活变通(FA)或执行既定决策($y_t = y_{t-1}$)。

博弈的时间顺序如下:(1)第1期初,委托人选择要求代理人执行既定决策还是默许其灵活变通;(2)第1期期中,代理人付出努力以了解社会环境所受的冲击,并根据对冲击的了解决定是否灵活变通;(3)第1期末,委托人了解到社会环境冲击,进行转移支付;(4)进入第2期。

2 灵活变通

现在,我们讨论如果委托人决定在两个时期都默许代理人灵活变通,需要

支付的灵活变通的报酬 $w_u^*(b_i)$ 、坚持既定决策的报酬 $w_u^*(y_{i-1})$ 以及代理人所付出努力 e_u^* 的最优形式。我们先考虑只有一个时期的情形,再将其引申至第2个时期。至于灵活变通方式如何运用以及该方式与执行既定决策之间的比较选择,留待第3节讨论。

2.1 单期

先考虑只有一个时期(不妨设为第1期)的情况。“讲真话”机制要求中央政府在设计灵活变通契约的最优形式时,要同时满足地方政府的激励相容约束、自选择约束和参与约束。

第一,委托人希望代理人在成功观察到社会环境冲击后愿意主动进行响应,而非默守成规地坚持既定决策:

$$u_{il,FA}^A(y_1=b_1) = -\bar{b}^2 + w_{il}(b_1) - \frac{1}{2}e_{il}^{*2} \geq u_{il,y_1=y_0}^A = w_{il}(y_0) - \frac{1}{2}e_{il}^{*2}$$

整理得

$$w_{il}(b_1) - w_{il}(y_0) - \bar{b}^2 \geq 0$$

在响应社会环境冲击时,代理人还应“正确地”选择 $y_1=b_1$ 。这便要求当 $b_1=\bar{b}$ 时,代理人不会选择 $y_1=-\bar{b}$,亦即 $-\bar{b}^2 + w_{il}(\bar{b}) - \frac{1}{2}e_{il}^{*2} \geq -\bar{b}^2 + w_{il}(-\bar{b}) - \frac{1}{2}e_{il}^{*2}$; 当 $b_1=-\bar{b}$ 时,代理人不会选择 $y_1=\bar{b}$,亦即 $-\bar{b}^2 + w_{il}(-\bar{b}) - \frac{1}{2}e_{il}^{*2} \geq -\bar{b}^2 + w_{il}(\bar{b}) - \frac{1}{2}e_{il}^{*2}$ 。整理可得:

$$w_{il}(\bar{b}) = w_{il}(-\bar{b}) \equiv w_{il,FA}$$

这样,给定 $w_{il,FA}$ 和 $w_{il}(y_0)$,代理人的努力投入 e_{il}^* 由其预期效用最大化决定:

$$e_{il}^* = \operatorname{argmax}_{e_{il}} \left\{ E u_{il,FA}^A = \rho_i e_{il} (w_{il,FA} - \bar{b}^2) + (1 - \rho_i e_{il}) w_{il}(y_0) - \frac{1}{2} e_{il}^2 \right\} \equiv \rho_i \Delta_{il} \quad (5)$$

其中, $\Delta_{il} \equiv w_{il,FA} - w_{il}(y_0) - \bar{b}^2$ 。

上式表明,由于低能力代理人($\rho_L=0$)无论如何都无法观察到社会环境冲击,必然有 $e_{L1}^*=0$ 。这意味着,委托人没有任何必要对其进行激励。但是,为了激励高能力代理人付出努力并主动对社会环境冲击予以响应,需要让他们从这种灵活变通中得到的收入奖励($w_{H1,FA} - w_{H1}(y_0)$)至少能够弥补为此遭受的效用损失($-\bar{b}^2$)。这种收入奖励越大,高能力代理人愿意付出的努力也越多,成功探知社会环境冲击的概率就越高。

将式(5)代入式(3),可将委托人在单期灵活变通契约下的预期效用函数进一步写为:

$$Eu_{1,FA}^P = -\alpha_1 [\rho_H^2 \Delta_H^2 + w_H(y_0)] - (1 - \alpha_1) w_L(y_0) - \bar{b}^2 \quad (3')$$

值得注意的是,如果代理人未能了解社会环境冲击,最优契约还应该能够防止其为了骗取奖励而故意偏离既定决策。由于冲击 b_1 以相同概率取值为 $\bar{b}$ 或 $-\bar{b}$,而委托人在期末会自动知晓其实现值,便可以将相关信息加以比对,从而有 1/2 的概率察觉到代理人在未观察到冲击值的情况下也有偏离既定决策的欺骗行为。只要通过威胁施加足够严厉的惩罚,委托人就能够确保代理人回到均衡路径,主动选择既定决策 y_0 ^①。

第二,委托人希望不同低能力代理人主动选择执行既定决策,高能力代理人则选择尝试灵活变通,亦即自选择(self-selection)条件:

$$\begin{aligned} Eu_{i1,FA}^A(\rho_i, \Delta_{i1}) &= \rho_i e_{i1}^* (w_{i1,FA} - \bar{b}^2) + (1 - \rho_i e_{i1}^*) w_{i1}(y_0) - \frac{1}{2} e_{i1}^{*2} \\ &= \frac{1}{2} \rho_i^2 \Delta_{i1}^2 + w_{i1}(y_0) \\ &\geq \frac{1}{2} \rho_i^2 \Delta_{j1}^2 + w_{j1}(y_0) = Eu_{i1,FA}^A(\rho_i, \Delta_{j1}) \end{aligned} \quad (6)$$

其中, $i \neq j$ 。

需要注意的是,这一自选择是代理人在期间内做出的,其结果只能以一定的概率在期末观察到,因此无法在事前将不同能力的代理人区分开来。灵活变通作为一种非正式授予的权力,不存在像正式权力那样在事前让代理人选择契约这一能够彰显其类型的环节(例如,地方政府在期初都会宣称自己将坚决执行既定政策)。只有在事后观察到代理人的灵活变通行为并与社会环境的实现值相印证后,委托人才能后验推断出代理人的类型(参见上一节)。

第三,结合式(5),代理人还需要满足如下参与约束:

$$Eu_{i1,FA}^A = \frac{1}{2} \rho_i^2 \Delta_{i1}^2 + w_{i1}(y_0) \geq \underline{w} \quad (7)$$

这样,委托人的规划问题就是在(5)、(6)和(7)的约束下,通过设计 $w_{i1}(b_1)$ 和 $w_{i1}(y_0)$ 来最大化式(3')。求解该规划问题,可以得到委托人默许代理人单期灵活变通契约的最优形式:

命题 2: 在单期(只有第 1 期)情况下,若委托人默许代理人灵活变通,将为两类代理人分别设计不同的契约:

^① 惩罚作为对代理人偏离均衡路径的威胁,也保证了代理人不会在不付出任何努力的情况下就随意灵活变通以期“瞎猫撞到死耗子”。

$$(1) (w_{L1,FA}^*, w_{L1}^*(y_0)) = (\bar{b}^2 + \underline{w}, \underline{w}); (w_{H1,FA}^*, w_{H1}^*(y_0)) = \left(\bar{b}^2 + \frac{1}{\rho_H} \sqrt{2\underline{w}}, 0 \right);$$

(2) 低能力和高能力代理人分别付出努力 $e_{L1}^* = 0$ 和 $e_{H1}^* = \sqrt{2\underline{w}} > 0$, 委托人的预期效用为 $Eu_{1,FA}^P = -(1 + \alpha_1)\underline{w} - \bar{b}^2$ 。

根据该命题, 所谓的灵活变通契约只是中央政府为高能力地方政府设计的。中央政府仅仅希望高能力地方政府付出努力去了解社会环境冲击, 低能力地方政府则只用严格执行既定决策就好。为了激励高能力地方政府付出努力, 中央政府不惜将他们简单执行既定决策的报酬 $w_{H1}^*(y_0)$ 降为 0, 低于其保留收入, 同时将其灵活变通后的报酬提高到超过保留收入的水平。

2.2 两期

我们已经知道在单期(标记为第 1 期)灵活变通下, 代理人会以概率 $\text{Prob}(y_1^* = b_1) = \alpha_1 \rho_H e_{H1}^*$ 成功观察到社会环境受到的冲击 b_1 , 并由既定决策 y_0 变通至 $y_1 = y_0 + b_1 = b_1$ 予以响应。代理人也可能未能观察到 b_1 (概率为 $\text{Prob}(y_1 = 0) = 1 - \alpha_1 \rho_H e_{H1}^*$), 从而只能执行既定决策 $y_1 = y_0 = 0$ 。

现在考虑委托人在第 2 期继续默许代理人灵活变通。这时, 社会环境会受到冲击 b_2 。由于 b_2 和 b_1 相互独立, 社会环境状态会在 b_1 的基础上累积至 $B_2 \in \{-2\bar{b}, 2\bar{b}, 0\}$, 于是既可能回到初始的 0 点, 也可能偏离地更远。根据第二节的设定, 代理人只能对 b_1 和 b_2 (而非 B_2) 分别进行独立观察, 因此其决策能否及时跟上社会环境的变化就取决于两次观察的结果。如果代理人仅观察到 b_1 而未观察到 b_2 , 会有 $y_2 = y_1 = b_1$; 如果 b_1 和 b_2 都观察到, 会有 $y_2 = y_1 + b_2 = B_2$; 如果未观察到 b_1 但观察到了 b_2 , 则有 $y_2 = y_0 + b_2 = b_2$; 如果 b_1 和 b_2 都未观察到, 则有 $y_1 = y_2 = 0$ 。

因此, 不管对委托人还是代理人而言, 不论第 1 期是否观察到了 b_1 , 在第 2 期都将面临与第 1 期类似的场景, 仅仅需要改变相应参数的时间下标, 并且更新关于高能力代理人可能性的信念。具体地, 我们有:

命题 3: 在两期情况下, (1) 若代理人在第 1 期未能观察到 b_1 , 委托人对其为高能力的信念更新为 $\alpha_2 = \frac{\alpha_1(1 - \rho_H \sqrt{2\underline{w}})}{\alpha_1(1 - \rho_H \sqrt{2\underline{w}}) + 1 - \alpha_1}$ 。委托人要在第 2 期继续允许灵活变通的话, 仍将区别对待不同能力的代理人: $(w_{L2,FA}^*, w_{L2}^*(y_1)) = (\bar{b}^2 + \underline{w}, \underline{w})$, $(w_{H2,FA}^*, w_{H2}^*(y_1)) = \left(\bar{b}^2 + \frac{1}{\rho_H} \sqrt{2\underline{w}}, 0 \right)$; 低能力和高能力代理人分别付出努力 $e_{L2}^* = 0$ 和 $e_{H2}^* = \sqrt{2\underline{w}}$ 。在第 2 期初, 委托人对该期所能获得的条件期望效用为: $E(u_{2,FA}^P | y_1 = y_0) = -(1 + \alpha_2)\underline{w} - 2\bar{b}^2$;

(2) 若代理人在第1期成功观察到 b_1 , 委托人对其为高能力的后验推断更新为 $\alpha_2 = 1$ 。委托人若要在第2期继续允许灵活变通, 须提供契约

$$(w_{H2,FA}^*, w_{H2}^*(y_1)) = \left(\underline{w} + \bar{b}^2 + \frac{3}{8\rho_H^2}, \underline{w} - \frac{1}{8\rho_H^2} \right), \text{ 代理人付出努力 } e_{H2}^* = \frac{1}{2\rho_H}.$$

在第2期初, 委托人对第2期所能获得的条件期望效用为: $E(u_{2,FA}^P | y_1 = b_1) = -\underline{w} - \bar{b}^2 - \frac{1}{8\rho_H^2}$ 。

将命题3与命题2进行比较可以看出, 如果代理人在第1期未能观察到社会环境冲击, 委托人为不同能力水平的代理人提供的契约在结构和形式上是跨时稳定的: 第2期的契约与第1期完全相同, 代理人各自付出的努力也与第1期相同。只不过由于信念得以更新, 在第2期初对该期所能获得的条件预期效用发生了变化。如果代理人在第1期成功观察到社会环境冲击, 会使得委托人准确判断出代理人是高能力者, 从而在第2期专心考虑是否对其实施激励的问题, 契约形式便发生了较大变化。

3 治理方式选择

在前面两部分的基础上, 我们可以对委托人在要求代理人执行既定决策和默许其灵活变通这两种治理方式之间的选择条件进行讨论了。我们运用逆向归纳的思想, 先从单个时期下这两种治理方式带给委托人预期效用的比较入手, 讨论仅有一个时期或者两个时期中第1期结束后所余下的第2期的委托人选择。然后, 回到两期设定下的第1期初, 对两种治理方式进行完整的比较。首先, 结合命题1、2和3可得:

命题4: (1) 在单期(例如, 仅有第1期)情况下, 由于 $Eu_{1,FA}^P \leq Eu_{1,y_1=y_0}^P$, 委托人会要求代理人执行既定决策;

(2) 在两期情况下, 不论第1期结果如何, 委托人在第2期都将转而要求代理人执行第1期确定的既定决策。具体地, 在第2期初, 有 $E(u_{2,FA}^P | y_1 = b_1) \leq E(u_{2,y_2=y_1}^P | y_1 = b_1)$ 和 $E(u_{2,FA}^P | y_1 = y_0) \leq E(u_{2,y_2=y_1}^P | y_1 = y_0)$ 。委托人在第1期初对从代理人在第2期执行既定决策中所获得的无条件期望效用为: $Eu_{2,y_2=y_1}^P = -\underline{w} - (2 - \alpha_1 \rho_H \sqrt{2w}) \bar{b}^2$ 。

该命题告诉我们, 如果只考虑单个时期, 不论仅有一个时期还是在第1期结束后余下的第2期, 默许代理人(即便其为高能力)灵活变通在效果上并不比简单地要求其执行既定决策来得更好, 原因是通过代理人“恰当地”灵活变通带给委托人的好处会被所需花费的激励成本所抵消。在两期情况下第1期结束后的第2期初, 不管代理人在第1期是否观察到了社会环境冲击, 也不管第1期是否进行了灵活变通, 委托人在第2期都将不会允许代理人灵活变通, 而是要

求代理人执行在第1期确定的既定决策 $y_2 = y_1$ 。鉴于此,我们也就不必再去计算委托人在第1期初对从默许代理人在第2期灵活变通所能获得的无条件期望效用($Eu_{2,FA}^P$)了。

这个结论预示我们,允许代理人灵活变通只有在考虑两个时期预期总效用的动态场景中才可能具有优势,并且只可能是在第1期出现。原因在于,将命题4与命题1第(2)部分比较可以看出,如果代理人在第1期观察到了环境冲击并灵活变通,将会提高委托人要求第2期执行既定决策的预期效用,从而使委托人两期总预期效用可能超过从第1期开始就执行既定决策(即 $y_2 = y_1 = y_0$)的情况。这样,委托人就可能在第1期默许代理人灵活变通。

然后,结合命题1、2和4,假设贴现率为0,我们就能够对从第1期就开始坚持既定决策带给委托人的预期效用($Eu_{1,y_1=y_0}^P + Eu_{2,y_2=y_1}^P$)与第1期默许代理人灵活变通再在第2期执行既定决策这两种方式带给委托人的预期效用($Eu_{1,FA}^P + Eu_{2,y_2=y_1}^P$)进行比较,从而得到两期情况下委托人在第1期默许代理人灵活变通的条件。

命题5: (1) 当 $\bar{b}^2 \geq \frac{\sqrt{w/2}}{\alpha_1 \rho_H}$ 时,委托人会在第1期默许代理人灵活变通,然后要求代理人在第2期执行第1期确立的既定决策。为代理人提供的契约形式如下: $(w_{L1,FA}^*, w_{L1}^*(y_0)) = (w_{L2,FA}^*, w_{L2}^*(y_1)) = (w_{H2,FA}^*, w_{H2}^*(y_1)) = (\bar{b}^2 + \underline{w}, \underline{w})$, $(w_{H1,FA}^*, w_{H1}^*(y_0)) = \left(\bar{b}^2 + \frac{1}{\rho_H} \sqrt{2w}, 0 \right)$;

(2) 当 $\bar{b}^2 < \frac{\sqrt{w/2}}{\alpha_1 \rho_H}$ 时,委托人在第1期就要求代理人执行既定决策并延续到第2期,亦即执行 $y_2 = y_1 = y_0$ 。

命题5证实了命题4关于灵活变通作为一种治理方式在使用方法上的启示。即便存在两个时期,也仅仅是在第1期才可能会默许代理人偏离既定决策,第2期就会回到既定决策。只不过,第2期的既定决策是在第1期中确立的,而不是最初的 y_0 。促使委托人选择灵活变通契约的关键因素是社会环境(或者等价地,委托人偏好)所受到随机冲击幅度的相对大小。只要社会环境受到的冲击足够大,委托人就会在第1期末默许代理人的灵活应对。反之,如果社会环境受到的冲击比较小(极端地,没有冲击,即 $b_t = 0$),那么委托人就会要求代理人自始至终都坚持既定决策。

触发委托人采用灵活变通契约所需冲击幅度的临界值 $\left(\frac{\sqrt{w/2}}{\alpha_1 \rho_H} \right)$ 取决于代理人是高能力者的可能性、能力水平以及保留收入水平。高能力代理人能力水

平的初始信念期望值 $\alpha_1\rho_H$ 越高(代理人是高能力者的可能性越大,或者代理人的绝对能力水平越高),委托人默许代理人在第1期灵活变通的可能性就越高。代理人的保留收入水平 w 越低,就越不愿意安于现状,而是希望通过灵活变通获得奖励。这样,委托人激励其付出努力观察环境冲击并予以恰当应对所需的激励成本就越低,也就越愿意默许其第1期灵活变通。

4 中国地方政府的灵活变通

由于缺乏相应的数据,我们无法为本文的理论提供系统性的计量经济学证据。但是,从中国的历史经验来看,中央在处理与地方政府关系的实践中,关于灵活变通的态度及其变化都可以用本文模型来加以解释。当既有制度与现实变化出现严重错位的时候,坚持过去的既定政策可能造成较大的损失,中央对地方进行非正式授权的现象就会尤为突出。

根据命题5,促使中央在前期默许地方政府灵活变通的关键因素是社会环境(或者等价地,中央的政策偏好)所受随机冲击的幅度足够大。自1949年新中国成立以来,中国经历了急剧的社会和经济变迁。党和政府工作重心的每一次重大调整,都是在环境不确定性最严重或冲击幅度最大的时候。由于中央经常在事前并不清楚当期最优的政策应该是什么,便倾向于抱以走一步看一步的摸索态度。从“社会主义改造”到“以阶级斗争为纲”再到“以经济建设为中心”的每个转变期,中央都会有较强的动机允许部分具有较高水平的地方政府在执行全国政策时进行适当的灵活变通。一旦工作重心确定下来,在每个政策期内就会相对稳定,中央对灵活变通的容忍度就会下降,倾向于要求地方政府坚决执行既定政策。然而,随着时间的推移,社会环境遭受的随机冲击不断积累,始终坚持一开始的既定政策就可能越来越不合时宜,中央就可能再次允许灵活变通。不论何时,对灵活变通的默许都是事后的,而不是在事前进行正式制度化。

王汉生等(1997)为以上结论提供了一些生动案例。20世纪70年代,江苏无锡县委在允许农村创办“为农业服务”的“五小工业”的中央政策下,将社队企业发展空间扩展到以非农产品为原料生产的加工工业。在“割资本主义尾巴”的大环境下,中央对这种变通行为却保持了容忍(未正式承认),直到1978年《中共中央关于加快农业发展若干问题的决定(草案)》和1979年《国务院关于发展社队企业若干问题的规定》,才作出正式的政策调整。同样在20世纪70年代,无锡县委还通过将社队企业与附近大中城市的国有企业联营,突破中央对原材料、能源、设备等的计划分配体制限制。对此,中央也是到1980年《国务院关于推动经济联合的暂行规定》和1986年《国务院关于进一步推动

横向经济联合若干问题的规定》,才将横向经济联合转变为正式制度。另外,1994年分税制改革以前,乡镇企业通过核算结算利润和账面利润的“两本账”来降低企业所得税负,也是地方政府对中央规定的企业财务制度的普遍变通做法。不过,这一做法始终未被正式制度化,而是随着1994年分税制的实行而自然消亡。

更典型的例子是农村“包产到户”制的命运。1953年,中国完成了社会主义改造,开始全面建设中央计划经济。1956年5月,中共永嘉县委却在雄溪乡燎原社进行了农业生产产量责任制的试验,在中国首创“包产到户”,并迅速在温州地区推广。1957年11月19日,《浙南大众报》发表文章认为这是“倒退的做法”,并转给了中央高层领导。结果,浙江省委、温州地委、永嘉县委被迫下令停止试验包产到户,永嘉县委主要负责人受到处分,全县因参与包产到户而被判刑劳改的达20多人。然而到了1978年,安徽凤阳小岗村实行“包产到户”却取得了中央的容忍,并最终在1980年《关于进一步加强和完善农业生产责任制的几个问题》得到了正式承认^①。1983年8月,浙江永嘉县李桂茂、李云河、戴洁天等人获得平反。1961—1963年的“三自一包”,也遭遇了类似的命运。

1992年,中央宣布要建立“社会主义市场经济体制”,许多既不是经济特区也不是沿海开放城市的内陆城市,在事前未获中央同意的情况下,也建立了许多经济开发区(Qian and Weingast, 1997)。

中央对地方政府灵活变通的态度大致从1999年开始发生了逆转。国务院于该年颁布了《关于全面推进依法行政的决定》,《行政复议法》也开始实施,日益强调各级政府要依法照章办事。2000年,最高法院公布行政诉讼法的司法解释,对原告标准加以放松,力图缓解行政纠纷“起诉难”的问题(包万超,2013;何海波,2012)。同时,非诉执行案件的数量也开始逐渐下降^②。

2004年,中央政府颁布了《全面推进依法行政实施纲要》,强调严格依法办事,政府所有行政行为都要于法有据、程序正当。2010年8月27日,温家宝在国务院召开的全国依法行政工作会议中敦促各级政府和所有工作人员贯彻实施《全面推进依法行政实施纲要》。最高法院也在这段时期就行政诉讼“起诉难”问题密集表态,要求地方法院“不得随意限缩受案范围,不得额外增加受理条件”^③,坚决清除限制行政诉讼受理的各种“土政策”^④。

自2013年新一届政府成立以来,始终强调“讲规矩”、依法治国和依法行

① 对此所作的最新考察,见魏众(2021)。

② 非诉执行是指未经诉讼程序的非诉行政案件中由法院裁定予以执行。该方式从1993年开始采用,旨在维护行政机关行使行政职权。

③ 最高人民法院.《关于当前形势下做好行政审判工作的若干意见》(2009年7月5日)。

④ 新华网. 最高法院要求坚决抵制和清除行政诉讼受案“土政策”[N/OL], http://news.xinhuanet.com/legal/2010-05/22/c_12130201.htm。

政,要求地方政府坚持既定政策,对灵活变通的容忍度很低。可以看出,中央关于执政理念和发展方向的思路日渐清晰,顶层设计日益加强,中央心目中最优政策的不确定性程度降低。

2013年,党的十八届三中全会审议通过《关于全面深化改革若干重大问题的决定》,开始推行从中央到地方各级政府及其工作部门的权力清单制度,依法公开权力运行流程。2014年,中央在第十八届四中全会《关于全面推进依法治国若干重大问题的决定》中将“深入推进依法行政,加快建设法治政府”作为全面推进依法治国的重大任务之一,提出“把法治建设成效作为衡量各级领导班子和领导干部工作实绩重要内容、纳入政绩考核指标体系,把能不能遵守法律、依法办事作为考察干部的重要内容”。2015年,中央全面深化改革领导小组审议通过了《关于人民法院推行立案登记制改革的意见》。最高法院也开始改革案件受理制度,将立案审查制变为立案登记制,对依法应该受理的案件,做到有案必立、有诉必理。全国人大常委会关于修改《中华人民共和国行政诉讼法》的决定也正式实施。到了2019年,党的十九届四中全会通过的《中共中央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》中,进一步要求“健全权威高效的制度执行机制,加强对制度执行的监督,坚决杜绝做选择、搞变通、打折扣的现象”。

5 结论性评论

中国是一个单一制国家,地方政府可被视为中央政府的派出机构,执行中央制定的政策规章制度。然而,中国各级地方政府在实际运行中的角色远不限于对中央政策的执行,而是拥有相当大的自主性,表现出高度的灵活应变能力和主动精神,以至于被冠以“动员型”“能动型”“进取型”乃至“生产型”政府的称号(例如,姚洋,2011;周黎安,2014,2017)。

本文通过一个两期模型证明,在面对高度复杂多变的社会环境时,如果中央始终要求地方政府严格执行既定政策,可能造成每期政策与当期最优政策的偏离随时间推移而越来越远,成本巨大。为了避免这种情况,中央可能会暂时性地容忍地方政府的灵活变通,非正式地授予其在实践中“摸着石头过河”的权力,突破不合时宜的条条框框,待相关条件具备后再严格执行在实践中逐步确立起来的既定政策。鉴于中国大多数法律和地方性法规以及全部行政规章制度都是由各级政府来贯彻、执行和实施的,这一问题的现实意义显得尤为突出^①。反过

^① 在2014年10月23日通过的《中共中央关于全面推进依法治国若干重大问题的决定》中,中央仍然认为当前国内“有法不依、执法不严、违法不究现象比较严重”。

来,当前中央政府要求杜绝搞变通的行为,则反映出中央关于执政理念和发展方向的思路日渐清晰,日益强调顶层设计,中央对最优政策的认识日益成熟。

将偏离规则并获得容忍的非正式授权与规则中包含的正式授权区分开来进行分析,揭示委托人默许代理人偏离既定决策的条件,是本文的主要贡献。本文模型是一般性的,对企业组织中层级间权力配置以及制度变迁等问题也有借鉴意义。

需要指出的是,本文假设中央和地方政府之间在期初是信息对称的并且不能进行信息交流,仅仅是工具主义的,并非不能放松。该假设一方面使得中央无法事先将灵活变通权纳入既定政策中,从而与已有的最优授权文献区别开来;另一方面,这使得代理人不能观察冲击的累计值,于是让多期模型具有了意义。在实践中,中央可以通过会议、文件、领导讲话以及形势学习等方式向地方政府传递自己的偏好或有关社会环境变化的信息。只要这种信息传递不是完全准确的,仍然需要地方政府付出努力进行解读,我们做出的这种简化就是合理的。另外,两期模型的设定也只是为了简化。代理人无论能力高低,都在每一期独立地对随机冲击进行观察,若将分析扩展至第3期及以上,就会遇到时期越往后所需计算的情况越多的麻烦。但是,基本思路没有变化,随机冲击的幅度大小仍然是委托人治理方式选择的关键。

不无遗憾的是,本文只是列举了一些中国地方政府在执行中央政策时经常灵活变通的例证,无法提供系统的数据经验证据。另外,本文对激励与惩罚、地方政府能力水平以及高能力者比例在中央治理方式选择中的作用,也没有作出进一步的比较静态分析。我们也没有对其他非货币性的激励方式(例如,晋升)进行讨论。最后,本文的讨论限于中央-地方两个层级。如果引入多个层级,就需要考虑异质性的社会环境冲击和地方政府之间的关系,问题会变得更为复杂。这些都是未来可能的研究方向。

附录

命题 1 的证明: (1) 令 $w_{it, y_t=y_{t-1}} = \underline{w}$, 代理人的效用函数变为 $u_{it}^A = -(y_t - y_{t-1})^2 - \frac{1}{2} e_{it}^2 + \underline{w}$, 最优选择将是 $e_{it, y_t=y_{t-1}} = 0$ 和 $y_t = y_{t-1}$, 即不付出努力而单纯执行既定决策。

(2) $Eu_{1, y_1=0}^P = -\frac{1}{2}(b_0 + \bar{b})^2 - \frac{1}{2}(b_0 - \bar{b})^2 - \underline{w} = -\underline{w} - \bar{b}^2$ 。社会环境在第2期的位置为 $\text{Prob}(B_2 = -2\bar{b}) = \text{Prob}(b_1 = b_2 = -\bar{b}) = 1/4$, $\text{Prob}(B_2 = 2\bar{b}) = \text{Prob}(b_1 = b_2 = \bar{b}) = 1/4$, $\text{Prob}(B_2 = 0) = \text{Prob}(b_1 = \bar{b}) \text{Prob}(b_2 = -\bar{b}) + \text{Prob}(b_1 = -\bar{b}) \text{Prob}(b_2 = \bar{b}) = 1/2$ 。

在第1期初,委托人并不知道对随机冲击的观察结果,于是 $Eu_{2,y_2=0}^P = -\frac{1}{4}(2\bar{b})^2 - \frac{1}{4}(-2\bar{b})^2 - \underline{w} = -\underline{w} - 2\bar{b}^2$ 。

(3) 类似地,若 $y_1=0$,委托人在第2期坚持既定决策就意味着 $y_2=y_1=0$,这样做与在第1期初对第2期采用 $y_2=0$ 的期望效用相同;若 $y_1=b_1$,委托人在第2期初就面临与第1期初类似的情况,因此期望效用也与之相同。(证毕)

命题2的证明: 委托人的规划问题是在(5)、(6)和(7)的约束下最大化式(3')。就高能力和低能力执行者分别写出约束(6)和(7),结合 $\rho_L=0$,可将原问题简化为:

$$\max_{w_{l1}(y_0), w_{l1,FA}} Eu_{1,FA}^P = -\alpha_1 [\rho_H^2 \Delta_{H1}^2 + w_{H1}(y_0)] - (1 - \alpha_1) w_{L1}(y_0) - \bar{b}^2 \quad (A1)$$

$$\text{s. t. } e_{i1}^* = \rho_i \Delta_{i1} \geq 0 \quad (A2)$$

$$w_{L1}(y_0) \geq w_{H1}(y_0) \quad (A3L)$$

$$\frac{1}{2} \rho_H^2 \Delta_{H1}^2 + w_{H1}(y_0) = \frac{1}{2} \rho_H^2 \Delta_{L1}^2 + w_{L1}(y_0) \quad (A3H)$$

$$Eu_{L1}^A(\rho_L, \Delta_{L1}) = w_{L1}(y_0) = \underline{w} \quad (A4L)$$

$$Eu_{H1}^A(\rho_H, \Delta_{H1}) = \frac{1}{2} \rho_H^2 \Delta_{H1}^2 + w_{H1}(y_0) \geq \underline{w} \quad (A4H)$$

暂时忽略约束(A2)、(A3L)和(A4H)。首先注意,高能力代理人的自选择约束(A3H)和低能力代理人的参与约束(A4L)为紧。这是因为,若(A3H)为松(等号变为严格大于号),那么政府总是可以降低一点(A3H)的左侧部分(例如降低 $w_{H1}(b_1)$ 从而降低 Δ_H^2),这可以在不违背该约束的情况下提高目标函数(A1)的取值,同时还不会违背低能力代理人的自选择约束(A3L)。若(A4L)约束为松,政府也可以通过降低 Δ_L 中的 $w_{L1}(b_1)$ 来降低向低能力代理人支付的成本,同时也不会影响高能力代理人的相关约束。由(A3H)和(A4L),将 $\rho_H^2 \Delta_{H1}^2$ 和 $w_{L1}(y_0) = \underline{w}$ 代入(A1)得:

$$Eu_{1,FA}^P = -\alpha_1 \rho_H^2 \Delta_{L1}^2 + \alpha_1 w_{H1}(y_0) - \bar{b}^2 - (1 + \alpha_1) \underline{w} \quad (A5)$$

给定 $w_{H1}(y_0)$,由于 $0 < \alpha_1 < 1$,令 $\partial(Eu_{1,FA}^P) / \partial \Delta_{L1} = 0$ 即得最优契约下 $\Delta_{L1}^* = 0$ 。结合 $w_{L1}^*(y_0) = \underline{w}$,得到 $w_{L1,FA}^* = \bar{b}^2 + \underline{w}$ 。现在考虑约束(A3H)。对委托人而言,结合(A3L),成本最低的方式是令 $w_{H1}^*(y_0) = 0$,从而 $\Delta_{H1}^* = \frac{1}{\rho_H} \sqrt{2\underline{w}}$,于是 $w_{H1}^*(b_1) = \bar{b}^2 + \frac{1}{\rho_H} \sqrt{2\underline{w}}$,并且 $e_{H1}^* = \sqrt{2\underline{w}} > 0$ 。容易验证,约束(A3L)和(A4H)仍然是满足的。

最后,将以上契约代入式(A5),可得委托人在单期允许灵活变通的预期效

用为 $Eu_{1,FA}^P = -(1+\alpha_1)\underline{w}-\bar{b}^2$ 。(证毕)

命题 3 的证明: 首先, 类似于正文中推导式(5)的方式, 可以得到:

$$w_{i2}(B_2=0|y_1=0) = w_{i2}(B_2=2\bar{b}|y_1=0) = w_{i2}(B_2=-2\bar{b}|y_1=0) \equiv w_{i2,FA}$$

(1) 如果代理人在第 1 期未能观察到社会环境冲击(亦即 $y_1=0$),

$$\begin{aligned} \max_{w_{i2}(y_1), w_{i2,FA}} E(u_{2,FA}^P | y_1=0) &= \sum_{B_2} \text{Prob}(B_2) \{ \alpha_2 [-\rho_H e_{H2}^* ((b_2 - B_2)^2 + w_{H2,FA}) + \\ &\quad (1 - \rho_H e_{H2}^*) (-B_2^2 - w_{H2}(y_1))] + \\ &\quad (1 - \alpha_2) (-B_2^2 - w_{i2}(y_1)) \} \\ &= -\alpha_2 [\rho_H e_{H2}^* w_{H2,FA} + (1 - \rho_H e_{H2}^*) (2\bar{b}^2 + w_{H2}(y_1))] - \\ &\quad (1 - \alpha_2) (\bar{b}^2 + w_{i2}(y_1)) \\ &= -\alpha_2 [\rho_H^2 \Delta_{H2}^2 + w_{H2}(y_1)] - (1 - \alpha_2) w_{i2}(y_1) - 2\bar{b}^2 \end{aligned}$$

$$\text{其中, } \alpha_2 = \frac{\alpha_1(1-\rho_H e_{H1}^*)}{\alpha_1(1-\rho_H e_{H1}^*) + 1 - \alpha_1}.$$

该规划问题的目标函数与(A1)只相差一个常数项和更新后的信念, 约束条件与命题 2 完全相同, 因而最优解也相同。将最优解代回目标函数, 即得委托人在第 2 期初的条件预期效用。

(2) 如果代理人在第 1 期成功观察到社会环境冲击(亦即 $y_1=b_1$), 委托人将判定其必为高能力者, 因此 $\alpha_2=1$ 。这样, 委托人在第 2 期允许灵活变通的话, 需要求解如下问题:

$$\max_{w_{H2}(y_1), w_{H2,FA}} E(u_{2,FA}^P | y_1=b_1) = -\rho_H^2 \Delta_{H2}^2 - w_{H2}(y_1) - \bar{b}^2 \quad (A1')$$

$$\text{s. t. } e_{H2}^* = \rho_H \Delta_{H2} \geq 0 \quad (A2')$$

$$Eu_{H2}^A(\rho_H, \Delta_{H2}) = \frac{1}{2} \rho_H^2 \Delta_{H2}^2 + w_{H2}(y_1) \geq \underline{w} \quad (A4H')$$

暂时忽略约束条件, 直接就(A1')就 $w_{H2}(y_1)$ 求一阶偏导数并令其为零, 可得 $\Delta_{H2}^* = \frac{1}{2\rho_H^2}$, 进而 $e_{H2}^* = \rho_H \Delta_{H2}^* = \frac{1}{2\rho_H}$ 。代入(A4H')可得 $w_{H2}(y_1) \geq \underline{w} - \frac{1}{8\rho_H^2}$ 。当等号成立时, 代理人在灵活变通和既定决策之间无差异。假设此时他选择灵活变通, 于是将等号式代入 Δ_{H2} 的定义式即得 $w_{H2,FA}^* = \Delta_{H2}^* + w_{H2}^*(y_1) + \bar{b}^2 = \underline{w} + \bar{b}^2 + \frac{3}{8\rho_H^2}$ 。

最终, $E(u_{2,FA}^P | y_1=b_1) = -\underline{w} - \bar{b}^2 - \frac{1}{8\rho_H^2}$ 。(证毕)

命题 4 的证明: (1) 比较命题 2(2) 和命题 1(2), 由于 $\rho_H > \rho_L = 0$ 以及 $0 \leq \alpha_1, \alpha_2 \leq 1$, 可得 $Eu_{1,FA}^P = -(1+\alpha_1)\underline{w}-\bar{b}^2 \leq -\underline{w}-\bar{b}^2 = Eu_{1,y_1=0}^P$, 当且仅当 $\alpha_1=0$ 时等号成立;

(2) 比较命题 3(2) 和命题 1(2) 即得 $E(u_{2,FA}^P | y_1 = b_1) = -\underline{w} - \bar{b}^2 - \frac{1}{8\rho_H^2} < -\underline{w} - \bar{b}^2 = E(u_{2,y_2=y_1}^P | y_1 = b_1)$ 和 $E(u_{2,FA}^P | y_1 = 0) = -(1+\alpha_2)\underline{w} - 2\bar{b}^2 \leq -\underline{w} - 2\bar{b}^2 = E(u_{2,y_2=y_1}^P | y_1 = 0)$, 当且仅当 $\alpha_2 = 0$ 时等号成立;

(3) 注意到, $\text{Prob}(y_1 = b_1) = \alpha_1 \rho_H e_{H1}^* = \alpha_1 \rho_H \sqrt{2w}$ 。结合命题 1 可得:

$$Eu_{2,y_2=y_1}^P = \text{Prob}(y_1 = b_1) E(u_{2,y_2=y_1}^P | y_1 = b_1) + \text{Prob}(y_1 = 0) E(u_{2,y_2=y_1}^P | y_1 = 0) = \alpha_1 \rho_H \sqrt{2w} (-\underline{w} - \bar{b}^2) + (1 - \alpha_1 \rho_H \sqrt{2w}) (-\underline{w} - 2\bar{b}^2) = -\underline{w} - (2 - \alpha_1 \rho_H \sqrt{2w}) \bar{b}^2. \quad (\text{证毕})$$

命题 5 的证明: 由命题 1 可得 $Eu_{1,y_1=y_0=0}^P + Eu_{2,y_2=y_1=0}^P = -3\bar{b}^2 - 2\underline{w}$ 。根据命题 2 和 4, 在第 1 期初, 委托人两个时期的预期总效用为:

$$EU^P = Eu_{1,FA}^P + Eu_{2,y_2=y_1}^P = -(1+\alpha_1)\underline{w} - \bar{b}^2 - \underline{w} - (2 - \alpha_1 \rho_H \sqrt{2w}) \bar{b}^2 = -(2+\alpha_1)\underline{w} - (3 - \alpha_1 \rho_H \sqrt{2w}) \bar{b}^2. \quad \text{令 } EU^P \geq -3\bar{b}^2 - 2\underline{w}, \text{ 即得 } \bar{b}^2 \geq \frac{1}{\alpha_1 \rho_H} \sqrt{\frac{w}{2}}. \text{ 此时, } \frac{\bar{b}^2}{2} \geq \frac{1}{2\alpha_1 \rho_H} \sqrt{\frac{w}{2}}.$$

此时, 也可验证此时高能力代理人在第 1 期获得的报酬 $w_{H1,FA} > \underline{w}$ 。这是因为:

$$w_{H1,FA} = \bar{b}^2 + \frac{1}{\rho_H} \sqrt{2w} \geq \frac{1}{\alpha_1 \rho_H} \sqrt{\frac{w}{2}} + \frac{1}{\rho_H} \sqrt{2w} = \frac{1+2\alpha_1}{\alpha_1 \rho_H} \sqrt{\frac{w}{2}}$$

由于 $\rho_H e_{H1}^* < 1$ 并且 $e_{H1}^* = \sqrt{2w}$, 有 $\rho_H < \frac{1}{\sqrt{2w}}$, 我们有 $w_{H1,FA} = \frac{1+2\alpha_1}{\alpha_1 \rho_H} \sqrt{\frac{w}{2}} > \frac{1+2\alpha_1}{\alpha_1} \underline{w}$ 。

鉴于 $\frac{1+2\alpha_1}{\alpha_1} > 1$, 便一定有 $w_{H1,FA} > \frac{1+2\alpha_1}{\alpha_1} \underline{w} > \underline{w}$ 。(证毕)

参考文献

- 包万超. 2013. 行政诉讼法的实施状况与改革思考——基于《中国法律年鉴》(1991—2012)的分析[J]. 中国行政管理, (4): 48-55.
- Bao W C. 2013. The implementation and reformation of administrative litigation law in China—An Analysis Based on *Law Yearbook of China (1991—2012)* [J]. *Chinese Public Administration*, (4): 48-55. (in Chinese)
- 韩博天. 2019. 红天鹅: 中国独特的治理和制度创新[M]. 石磊, 译. 北京: 中信出版集团.
- 何海波. 2012. 困顿的行政诉讼[J]. 华东政法大学学报, (2): 86-96.

- He H B. 2012. Administrative lawsuit in the aporia situation[J]. *ECUPL Journal*, (2): 86-96. (in Chinese)
- 李石强. 2015. 最优授权理论: 一个文献综述[J]. 制度经济学研究, (1): 1-20.
- Li S Q. 2015. A literature survey on the optimal delegation within organizations[J]. *Research of Institutional Economics*, (1): 1-20. (in Chinese)
- 刘骥, 熊彩. 2015. 解释政策变通: 运动式治理中的条块关系[J]. 公共行政评论, 8(6): 88-112, 187.
- Liu J, Xiong C. 2015. Explaining China's policy deformation: Tiao/Kuai authority relations under campaign-style governance[J]. *Journal of Public Administration*, 8(6): 88-112, 187. (in Chinese)
- 刘鹏. 2012. 土政策与政策变通. 理解中国政治——关键词的方法[M]//景跃进, 张小劲, 余逊达. 北京: 中国社会科学出版社, 182-196.
- 王汉生, 刘世定, 孙立平. 1997. 作为制度运作和制度变迁方式的变通[J]. 中国社会科学季刊, (21): 85-100.
- Wang H S, Liu S D, Sun L P. 1997. Flexible adaptation as a way of institutional running and changing[J]. *Quarterly Journal of Chinese Social Sciences*, (21): 85-100. (in Chinese)
- 王清. 2021. 模糊的属地化管理: 政策执行偏差的一种解释——以流动人口服务供给为例[J]. 探索, (3): 129-141.
- Wang Q. 2021. Fuzzy localized management: A framework for public policy deviation—Case study on floating population in S city[J]. *Probe*, (3): 129-141. (in Chinese)
- 魏众. 2021. 从“责任田”实践到家庭承包制——基于安徽的考察[J]. 中国经济史研究, (4): 52-65.
- Wei Z. 2021. From responsibility farmland practice to rural household responsibility system—Anhui perspective[J]. *Researches in Chinese Economic History*, (4): 52-65. (in Chinese)
- 姚洋. 2011. 生产型政府[J]. 理论学刊, (8): 52-54.
- Yao Y. 2011. Productive government[J]. *Theory Journal*, (8): 52-54. (in Chinese)
- 翟学伟. 1997. “土政策”的功能分析——从普遍主义到特殊主义[J]. 社会学研究, (3): 88-95.
- Zhai X W. 1997. The functional analysis of “Tu Policy”—From universalism to particularism[J]. *Sociological Research*, (3): 88-95. (in Chinese)
- 周黎安. 2014. 行政发包制[J]. 社会, 34(6): 1-38.
- ZHOU L A. 2014. Administrative subcontract[J]. *Chinese Journal of Sociology*,

- 34(6): 1-38. (in Chinese)
- 周黎安. 2017. 转型中的地方政府: 官员激励与治理[M]. 2版. 上海: 上海人民出版社.
- 周雪光. 2017. 中国国家治理的制度逻辑: 一个组织学研究[M]. 上海: 生活·读书·新知三联书店.
- Austen-Smith D. 1994. Strategic transmission of costly information[J]. *Econometrica*, 62(4): 955-963.
- Bozeman B, Rainey H G. 1998. Organizational rules and the “bureaucratic personality” [J]. *American Journal of Political Science*, 42(1): 163-189.
- Gibbons R S, Matouschek N, Roberts J. 2013. *Decisions in organizations*[M]//Gibbons R S, Roberts J. *The Handbook of Organizational Economics*. New Jersey: Princeton University Press, 373-431.
- Lampton D M. 1977. *The politics of medicine in China: The policy process 1949—1977* [M]. Boulder: Westview Press.
- Merton R K. 1940. Bureaucratic structure and personality[J]. *Social Forces*, 18(4): 560-568.
- O'Brien K J, Li L J. 2004. Suing the local state: Administrative litigation in rural China [J]. *China Journal*, 51: 75-96.
- Pei M X. 1997. Citizens v. mandarins: Administrative litigation in China [J]. *The China Quarterly*, 152: 832-862.
- Qian Y Y, Weingast B R. 1997. Federalism as a commitment to preserving market incentives[J]. *Journal of Economic Perspectives*, 11(4): 83-92.
- Thompson V A. 1961. *Modern organization*[M]. New York: Knopf.

When to Forbid Flexible Adaptation —An Explanation from An Informal Delegation Model

Shiqiang LI

(School of Economics, University of Chinese Academy of Social Sciences)

Abstract In the Communique of the Fourth Plenary Session of the 19th Central Committee of CCP, it is announced that such phenomenon as flexible adaptation should be forbidden. In order to give an explanation to this decision, this paper sets up a general two-period model in which the principal has private information *ex post* in each term. It is showed by the model that the central government will ask the local government who cannot observe the shock to social environment to stick to the established decision all the time.

However, when the random shock to social environment is big enough, the central government would allow the local government who can observe it to bias from the determined decision in the first stage and take the decision in period 1 as established in the second stage, in order to adapt to the circumstances as well as possible. As the preference becomes more and more stable with the social and economic environment, the central government would acquire local governments to stick to the established policies from the beginning.

JEL Classification D23, H11, L22