

“你合作，我也合作”

——森的确保规则在中国公社内部合作行为的应用

刘民权*

摘要 在社会交往中，个人的行为可能会被各种各样的承诺所限制。与标准的囚徒困境博弈中“完全自私自利”的参与者的行为相反，这些承诺有时会使人们暂时放弃对个人利益无休止的追逐。承诺形形色色，多种多样。在一个需要两个参与者进行合作的场合，当一个参与者确保自己会采取合作时，另一个参与者也会答应采取同样行动（Sen, 1969, 1974）。当合作需要在超过两个参与者之间（ $n > 2$ ）进行时，一个参与者会答应合作只要另外一个，另外两个……或所有其他（ $n - 1$ ）个参与者答应合作。最终，一参与者可能会答应单方的无条件合作。本文旨在分析中国人民公社（1958—1978）内部集体成员承诺的不同形式。由于成员承诺结构的多样性，可能会产生不同的合作结果并且实际上也观察到了这些结果。这种对公社内部合作问题性质的解释与仅仅承认个人利益是集体成员选择努力水平的惟一决定因素的另外两个模型（Lin, 1990, 1993；Dong 和 Dow, 1993）形成了鲜明的对照。

关键词 承诺，确保原则，人民公社

一、引言

在社会交往中，个人的行为可能会被各种各样的承诺（commitments）所限制，与“完全自私自利”者的行为相反，这些承诺有时会使人们暂时放弃对个人利益无休止的追逐。一个“完全自私自利”者的最好例证是标准的囚徒困境博弈中互相告发对方的参与者，因为告发对方是这些参与者的“占优策略”，也就是说，不管对手采取何种行动，每一个参与者都无休止地追逐他的个人利益。承诺可以有多种形式。在一个需要两个参与者互相合作的场合，假如一个参与者确保自己合作的话，另一个参与者也会答应合作（尽管该参与者在对方合作时采取不合作行动会获得更多好处）。这就是 Sen（1969, 1974）提出的确保规则（assurance principle）。当合作需要在超过两个参与者（ $n > 2$ ）之间进行时，一个参与者可能会答应合作，只要另外一个，另外两个

* 霍普金斯—南京中心。通讯地址：南京市南京大学中美中心 210093；电话：(025)3592832；Email：mliu@nhc.nju.edu.cn。本文主要是基于1991年我在牛津大学的博士论文“中国公社内部的工作激励和劳动配置”中的一章内容而成。Sudhir Anand, Keith Griffin 和 A. K. Sen 对论文早期工作的审慎评论以及他们的热情鼓励，作者表示衷心的感谢。这篇论文2000年10月在北京大学中国经济研究中心的一次专题讨论会上作过介绍。与会者对论文的提问和讨论，作者也表示感谢。北京大学中国经济研究中心的一位学生和南京大学学生俞建拖为本文的翻译和校对提供了帮助，谨在此一并致谢。

……或所有其他($n-1$)个参与者也合作。这是对 Sen 的标准确保规则的推广。最后,一参与者可能会答应单方的无条件合作。

本文旨在分析中国人民公社(1958—1978)内部集体成员承诺的不同形式。笔者认为,一般说来一些集体成员确实遵循了某些确保规则而不是完全自私自利。但通常并非所有的成员都是如此。一个或一些“完全自私自利”的成员的存在常常使得所有成员都采取不合作行动。由于成员承诺结构的多样性,可能会产生不同的合作结果,并且事实上也常常观察到了这些结果。

文章的第二部分引入和定义了承诺的概念,并把承诺和完全自私自利的行为作了对比,这个对比在标准的囚徒困境博弈背景下得到了最好的体现。文章的这一部分还讨论了各种可能的承诺行为以及策略行为。文章的第三部分讨论了囚徒困境博弈的一个例子——公社内部努力付出博弈,这个博弈的均衡结局是每个成员都偷懒(不合作)。文章的第四部分引入了建立在承诺或确保基础上的行为以阐明公社内部成员不合作的性质和原因。笔者假定了集体成员有多种多样的承诺,其中一部分成员遵循某些确保规则而另一部分成员则无视这些规则。文章的第五部分把以承诺为基础的解释公社内部合作失败的观点和最近提出的建立在完全利己主义理性假定基础上的另外两个模型(Lin, 1990; Dong 和 Dow, 1993)作了简要的对比。文章的第六部分总结全文。

本文的一个更大的目的是把承诺或道德因素重新引入当前关于中国经济改革和发展的讨论中。中国以市场为导向的经济改革不可避免地要给合法追逐个人利益以应有的地位。由于追逐个人利益是市场驱动和市场调节经济的一个根本支柱,所以这样做既是正确的也是恰当的,但它不是惟一的支柱。更加发达和更加成功的市场经济是依靠个人利益因素和非个人利益因素相互交织所构成的一个集合体来使其系统发挥作用的。然而,为了理解这个各种因素相互交织在一起的集合体是如何运作的,我们首先需要对个人是如何进行决策和选择自己行为的问题有一个深刻的理解。在此我们把公社内部成员的合作问题作为一个实例进行分析。¹

二、个人利益和承诺

(一)不合作和囚徒困境(PD)博弈

让我们再次回顾标准的两个参与者囚徒困境博弈。囚徒困境博弈中每一参与者在合作(C)与不合作(NC)之间进行选择,两个参与者都合作对于双方来说结果都优于两个参与者都不合作。然而,对于其中任意一个参与者来说,其最大的收益产生于对方合作而自己不合作的时候,并且其最小的收

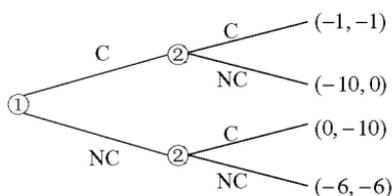
¹ 参见由一著名中国经济学家(Fan, 1998)写的一篇很有影响的题为“非道德的经济学”的文章。在文中,作者认为经济学家在自己的职业生涯中不应关注道德规范。

益产生于对方不合作而自己合作的时候。下面的双矩阵给出的收益反映了这种排序关系。

	参与者 1		
参与者 2		合作	不合作
合作		-1	0
	-1		-10
不合作		-10	-6
	0		-6

假定两个参与者均拥有关于该博弈的完全信息（每一个参与者都了解对方，并且还了解双方的策略空间以及双方对应于各种可能博弈结局的收益），并且假定博弈为同时博弈。如果两个参与者仅仅最大化自己的个人利益，则博弈的惟一均衡结局是双方都不合作。对于两个参与者来说，这样做确实有着令人信服的逻辑：不管对手如何行动，不合作对两个参与者来说都是一个更好的策略。也就是说，对于另一个参与者的任意一个可能的策略选择来说，不合作都优于另一可供选择的策略——合作，而且这一点对两个参与者都成立。换句话说，不合作是每一个参与者的“占优策略”。博弈中两个参与者都采取占优策略而得到的均衡称为“占优策略均衡”。

事实上，即使我们放松上面的部分假定，仍可得到两个参与者都不合作为惟一的均衡结局这一结果。首先，让每个参与者仅仅知道对应于每一个可能博弈结局的自己的收益，但却不知道相应的对手的收益，结果是均衡没有发生改变。其次，让参与者序贯行动并且假定参与者 1 首先行动；参与者 1 知道参与者 2 的收益，但反过来参与者 2 并不一定知道参与者 1 的收益；参与者 1 知道参与者 2 是理性的并仅仅追逐个人利益，但反过来参与者 2 并不一定知道参与者 1 也是如此。为简单起见，假定参与者 2 可以观察到参与者 1 所采取的行动（也就是说博弈为完美信息博弈）。则博弈的惟一结局仍是双方都不合作（子博弈完美均衡）。



上述对假定的放松不但进一步强调了双方都不合作存在着令人信服的逻辑，而且还扩大了结局是双方都不合作的博弈例子的种类。然而我们可以清楚地看到，对两个参与者来说，双方都不合作的结果比双方都合作要差。尽管每一个参与者都试图最大化自己的利益，但他们行为的总结果却是给双方都带来一个更坏的结局。

在上述的囚徒困境博弈中,可以说每个参与者都在无休止地追逐个人利益,除了纯粹的个人利益之外没有任何其他因素决定着他们的行为。例如,其中没有包含任何类型的道德考虑。然而,这些例子显然是适合于道德考虑的,并且参与者在作决策时不可能对道德考虑置若罔闻。当对手不合作时,自己也不合作也许可以被参与者看作是“合理的”或符合道义的。然而,当对手合作时,一参与者不可能再次把采取不合作看作是“合理的”和符合道义的行动,特别是当该参与者清楚地知道自己的行动会给对手带来痛苦(带来收益的急剧下降)。当然,对于一个仅仅从个人利益角度进行考虑和推理的人来说,这两种情形是毫无区别的。我们称一个仅仅考虑和追逐个人利益的人为“完全自私自利者”。

(二)个人利益与承诺

现在让我们引入承诺的概念。根据 Sen (1977) 的定义,承诺指的是在特定的社会状态或博弈情形中那些影响或决定一个人行为的考虑、原则或社会准则。也就是说,承诺指的是那些不是基于个人利益而是反映比如说一个人的道德观、意识形态信仰或者甚至爱国和宗教情操的考虑、原则和价值观。尤其是,在某些情形下,承诺也许会压倒个人利益考虑。但是,承诺并不否认在其他情形下个人利益在决定个人行为时所起的重要作用。只有当个人利益考虑和承诺发生冲突时,承诺才会压倒个人利益考虑。实际上,承诺不过是对个人利益考虑施加了一定的限制条件,在这些限制条件下个人利益考虑仍然起支配作用。

通常承诺包含对他人利益以及自身行动对他人利益所造成的影响的考虑。在此有必要对承诺和另一个重要的概念——“同情”作一区分。同情仍属于个人利益的狭隘范畴,但承诺超出了这个范畴。

在经济学中,人们的效用通常被专门地定义为一个人的“个人”效用。也就是说,人们从他自己对商品和服务的消费中所获得的满足感。Sen 把上述观点称为自我中心的个人利益观。然而,他人的福利或者幸福也会从心理上直接影响个人自己的福利或幸福。通常,人们为他人痛苦而痛苦,为他人高兴而高兴,这就是“同情”的例子;或者人们会因他人痛苦而高兴,因他人高兴而痛苦,这就是嫉妒或负同情的例子。在这两个例子中,一个人的福利直接影响到另一个人的福利。个人利益的一个更为广泛的定义可以包含这些情况。²

² 令嫉妒为负同情。更加一般的包含同情的个人利益可以定义为: $W_i = U_i + \sum_{j=1, j \neq i}^n a_{ij} U_j$, 其中: W_i 是个人 i 的福利, U_i 是 i 的个人效用, U_j 是 j 的个人效用 ($j = 1, 2, \dots, n, j \neq i$) a_{ij} 是个人 i 对个人 j 的效用所赋予的权重, a_{ij} 反映的是 i 个人对 j 个人的关心或同情的程度。Sen (1966) 分析了同情的例子,并且对在“按劳分配”和“按需分配”体制下农村集体中社员同情与工作激励之间的联系作了分析。

无论是出于承诺还是同情，人们也许都会采取旨在有利他人的行动。然而，在同情的情形下，人们这样做只是因为他人的效用直接影响到他自己的效用。而在承诺的情形下，人们是出于信念而这样做；或者说，人们是按照他们所遵循的非自利的价值观和原则而这样做的。³

尽管我们把承诺定义为任何人们所保证会遵循并且在某些社会场合下确实实会影响和决定他们行为的非自利原则或价值观，但最普遍的承诺类型还是道德承诺。在日常生活中，我们经常会遇到人们不是为了自己的个人利益而仅仅是出于道义来选择个人行为的例子。譬如说，一店主本可以欺骗一个偶然光顾的顾客，但他并没有这样做。一个粗心大意的顾客付款后把钱包遗失在商店，人们完全可以捡起钱包并将之占为己有，却选择叫回顾客并送还钱包。一个路人完全可以置身边落水者不顾却选择冒着生命危险去拯救他。这样的例子举不胜举。我们所提到的都只是一些平常的例子（也正是因为它们平常而被选用），但是可以说其中每一例都包含着道德规范的影响。确实，有时我们在做决策时是这样的基于道德规范，以至于没有给所做的选择以太多的考虑，因为它们是我们“应该做的”。在其他时候，我们则可能会就是否遵循某些道德规范而面临痛苦的抉择。道德规范对我们的行为所施加的影响之普遍在这里大概是不可能被夸大的。⁴

（三）确保规则和确保博弈

把承诺看作是人类行为的另一个决定因素与理解合作问题有着直接的联系。不用说，引入道德因素有可能会增加合作的可能性。然而，如果随意地引入承诺，这项工作就毫无意义了。我们只应该引入那些人们确实有可能会遵循或者那些断定已经得到遵循的承诺。这个问题最终是个经验问题，而且

³ 很显然，我们这里所定义的承诺和主流博弈论中所定义的承诺有很大的不同。在主流博弈论中，承诺指的是由参与者（比如说一个垄断者）采取的把自己逼入某种困境的威胁行为，以至于当相关情形发生时该参与者发现兑现这些威胁行为确实会为自己带来好处，而在其他情形下该参与者是不会实行这些威胁行为的（比如说，只要当一潜在进入者进入到自己的垄断行业，垄断者就与进入者打价格战）。显然，这样的承诺不过是一种极端的深谋远虑的自私自利行为。

⁴ 我们这里所定义的承诺的概念以及承诺在决定人们行为时所起的作用至今尚未被经济学家广为接受。把所有人类行为看作完全由个人利益所驱动的传统观念至今仍根深蒂固地存在于经济学这门学科中。比如，我们刚才引用的作为承诺例子中的每一个例子都可以被重新解释为简单的个人利益最大化的例子。正如 Sen（1977）所说，问题很大程度上取决于怎样对个人利益作精妙的定义和解释。一个人自愿选择去做的事情必定是他所偏好的事情，而他所偏好的事情必定是那些经过判断对他有利的东西，而这对于那些相信个人利益最大化具有普遍性的人来说是再自然不过的了。如果对个人利益作这样理解的话，那么任何人的行为都不可避免地在追逐自己的个人利益。不管一个人选择某种行为事实上是出于强烈的道德责任感或意识形态信仰，人们总是可以说这个人在最大化他自己的个人利益。

正如 Sen 所指出的，问题最终还是归结为一个“选择”：要么选择把所有的人类行为毫无例外地看作是仅仅为了个人利益，并给个人利益概念以一个足够广泛的定义，使得人们总是最大化自己个人利益的命题永远成立（但这样做也就使论证陷入同义反复）；要么选择把人们看作不光是为了自己的利益而采取某种行为，而且还按照某种道德的或其他关于“对”和“错”的基本准则而采取行为。后一种情况是我们所定义的“承诺”概念。

在不同的情形下无疑可能会有不同的价值观念起作用。

此处我们引入 Sen 的确保原则——承诺的一个相当弱的形式。在前述的囚徒困境博弈中，我们看到不合作是两个参与者的占优策略。Sen (1969, 1974) 把确保原则引入到这个博弈，其中每个参与者选择合作当且仅当另一个参与者确保也选择合作，但每个参与者选择不合作当且仅当另一个参与者也选择不合作。注意到这里两个参与者都不再完全按照个人利益来选择自己的行为。每一个参与者都把对手合作而自己合作看作是不符合道义的，而将对手不合作时自己也不合作视为是无可非议的。

假定博弈为完全信息博弈。博弈的均衡结局取决于博弈是同时博弈还是序贯博弈。在同时博弈情况下，博弈有两个可能的均衡结局：双方都合作或者双方都不合作。然而，当参与者序贯行动时，博弈的结局是双方都合作。此处我们假定，首先行动的参与者有一个关于对方会遵循确保规则的坚定信念。⁵

当参与者序贯行动时，博弈会得到双方都合作这一惟一的均衡结局。这时遵循确保规则以及参与者对追逐个人利益有条件的暂时放弃使两个参与者的收益都得到了提高。但是，即使参与者不序贯行动，也至少存在着两个参与者都合作并且双方都得到一个更好结局的可能。这种情形与囚徒困境博弈形成鲜明的对比，因为在囚徒困境博弈中两个参与者只追逐个人利益，所以采取不合作对双方都有着令人信服的理由。现在我们可以得出结论：至少在某些情形下，暂时中止对个人利益的斤斤计较而代之以遵循某些道德规则办事会使每个参与者的境况都得到改善。

(四) 承诺的多样性和策略行为

从纯粹个人利益的角度看，引入承诺因素使得每一个参与者的境况都得到了改善，这会增加策略行为产生的可能性。事实上，为了得到相互合作的博弈结局，我们仅仅需要第二个行动的参与者（假定为序贯博弈）遵守确保规则并且第一个行动的参与者知道这点，而第一个行动的参与者没有必要老老实实地遵循确保规则。第一个行动的参与者关于第二个行动的参与者会遵循确保规则的信息会告诉他只要他合作，则博弈的结局是双方相互合作。即使从第一个行动的参与者个人利益角度来看，这样的博弈结局也是一个更好的结局。

所谓策略行为，是指一参与者采取某种行动以使另一个参与者随后采取某种行动，导致对自己有利的结局。因此，策略行为本质上是自私自利的并

⁵ 当参与者同时行动时，由于参与者之间不可能存在沟通（行动的同时性排除了参与者之间沟通的可能性），参与者是否相互了解对方的承诺无关紧要。参与者 A 也许知道参与者 B 会遵循上述的确保规则，但是 A 不可能知道 B 究竟将如何行动，因为 B 不可能知道 A 将如何行动，等等。这种情形在参与者序贯行动时会得到改变。

且它与基于道德承诺的行为是截然不同的。策略行为产生的基础是参与者策略之间的相互依赖性，也就是说，一个参与者的策略选择会受到另一个参与者策略选择的影响。策略行为在参与者被假定为只追逐个人利益的一般博弈中较为常见。参与者策略的相互依赖性归因于这样的事实，即一个参与者的策略选择会影响到另一个参与者采用某一特定策略时的收益，因而后者会采用某一合适的策略给以回应。此处个人利益在上述参与者策略之间的相互依赖性中起到了中介作用。

在囚徒困境博弈中，虽然每一个参与者都追逐个人利益，但由于每一个参与者都有一个占优策略，从而使得上述策略行为失去了存在的基础，因而参与者策略之间的上述相互依赖性实际上并不存在。然而，当囚徒困境博弈转化为确保博弈时，参与者策略之间的相互依赖性就又重新产生了，这是因为由确保规则建立起了参与者策略之间的相互依赖性。这种相互依赖性是通过承诺而不是通过个人利益起中介作用的。因此，确保规则通过引入参与者策略之间新的相互依赖性而产生了策略行为，否则这种策略行为是不可能产生的。

注意我们在上面已经假定了参与者具有不同种的承诺（承诺的多样性）。在现实世界上，承诺的多样性看起来似乎是规律而不是例外。然而，当参与者具有上述形式的承诺多样性时，现实中要识别他们的真实承诺的确是非常困难的。例如，当参与者1仍然是专一的利己主义者，而参与者2遵循确保规则时，假如参与者1知道这点并且首先行动的话，博弈的结局是双方都合作。在这个例子中，参与者1的所作所为看起来好像他也遵循了确保规则；从另一方面上说，假如参与者2首先行动并且他知道参与者1是个专一的个人利益最大化者，在博弈第一轮中参与者2会选择合作，之后参与者1也选择合作。这个两个参与者都不合作的结局是博弈的惟一均衡结局。在这个例子中，参与者2的行为看起来好像他也是一个专一的个人利益最大化者。现实中，很难识别出参与者2实际上是遵循了确保规则的。⁶

（五）承诺和超排序

上述的确保博弈只有两个参与者，然而，许多社会状态涉及不止两个参与者。在一个有 n 个参与者的博弈中，确保原则会变得非常复杂并且会有许多可能的确保规则。比如， n 个参与者中的任意一个都有可能承诺合作只要其他一个，其他两个……或所有其他 $(n-1)$ 个参与者确保合作。这样，对于参与者 i 来说有 $(n-1)$ 种可能的确保规则，并且这对所有参与者也一样。所以，参与者的承诺原则有 $(n-1)^n$ 种可能结构。对所有这些可能的结

⁶ 经济学中惯用的方法是假定所有人都好像只最大化他的个人利益。可把这个例子看作是为这种惯用的方法提供了辩护。然而，在我们的例子中，显然它只仅仅是一种可能的博弈情况。

构都进行考虑是不现实的。在文章的第四部分我们会在公社内部努力付出问题的背景下考虑一些较为简单的结构。

到目前为止,我们还没有考虑一个参与者不管其他参与者如何行动自己都会选择合作的无条件承诺存在的可能性。Sen(1974)把这种无条件承诺称作“为他人考虑”的行为规则(other-regarding rule)。这样,完全建立在个人利益基础上的个人选择规则和上述“为他人考虑”的规则形成完全对立的两个极端。当我们加进这两种确保规则(这里应对“确保”和“承诺”两词作最为一般的理解)并且允许参与者之间存在差异时,则承诺会有 $(n+1)^n$ 种可能的结构。⁷

Sen(1974, 1977)认为对不同的个人行为规则可以按照某种带有普遍性的道德标准来进行排序。这里我们可以把这种带有普遍性的标准理解为“社会价值标准”(social values)。例如,更加无保留地对他人福利以更多的关注可以在道义上被认同为要优于不这么做。按照这个标准,“为他人考虑”的行为规则会被社会认同为是最优的行为规则,并且优于所有其他确保规则;并且在所有确保规则中,那些有着更弱限制条件的确保规则会比那些有着更强限制条件的确保规则得到社会更多的偏好。一心一意地追逐自己个人利益的规则会被社会认同为是处于最低层次的规则并且最为社会所不偏好。这样我们就能得到关于个人承诺的一个“社会排序”(social ranking),并可把这种“社会排序”称为一超排序(meta-ranking)或排序的排序(ranking of rankings)(之所以这样称是因为,在经济学中,人们的个人选择模型是建立在从最合意到最不合意的可能选择进行排序的基础上的)。超排序的概念允许我们把人们的个人承诺与范围更为广泛的社会价值观和社会规则联系起来,并使我们可以这样一个更为宽泛的背景下对个人承诺进行研究。⁸

三、公社和努力付出问题

人民公社是存在于20世纪50年代末到大约20世纪70年代末80年代初中国农村的一个众所周知的农业合作机构(自从大约20世纪70年代末80年代初起,中国农村实行了实际上属私有耕作的制度)。在这一部分,我们对公社内部的努力付出问题进行讨论并说明可把公社内部的努力付出问题作为 n 个参与者囚徒困境博弈的一个例子来加以建模。

⁷ 这里我们对“确保”和“承诺”两个概念作了最为一般的理解。注意,此处我们还没有考虑到参与者在他的确保行为中可能会存在歧视。例如,一个参与者只有在其中一个特定的对手而不是任何一个对手合作时才会承诺合作,等等。如考虑到这些可能性,承诺的可能结构集会进一步扩大。

⁸ Sen把超排序解释为代表一种特定的“道德规范”,这种“道德规范”可以是个人的而不一定是社会的。我们仅仅是从对个人承诺进行社会排序的角度对这个概念作解释。考虑到人们的个人道德规范通常是对社会价值标准的接受和内生化,从这个意义上说个人的道德规范和承诺与社会价值标准之间可能存在着密切的对应关系。但尽管如此,两者不一定是完全相同的,而且事实上两者也常常是不同的。

我们从公社的分配制度入手。公社的每一个农民为他的集体付出劳动并按照他的劳动付出从集体那里获取相应的报酬。这一分配制度可被写作：

$$y_i = (\omega_i / W) Y,$$

其中 ω_i 是对集体中第 i 个成员一年劳动付出的衡量； $W = \sum_n \omega_i$ ， n 是集体成员的数量； Y 是集体总的可支配收入。为简单起见，我们假定 Y 是一年中集体产出的总收入，尽管在实际中需要对总收入作一定的扣除才能得到总的可支配收入。

当成员的真实努力不能被完全监督时，就会产生激励问题。⁹ 尽管对代理人或集体成员努力贡献的监督被认为是任何委托—代理问题中的一个普遍问题，并且来自众多研究和文献报道的证据都表明公社内部对成员努力的监督严重不足，但公社内部确实存在对成员的努力付出进行准确监督的可能性，并且通常集体也确实对成员的努力付出进行了一定的监督。尽管如此，为了将注意力集中到不合作问题上，我们假定集体内部不存在监督。¹⁰ 这时公社成员的收入分配变为一个完全平均的体制，其中：

$$y_i = Y/n. \quad (1)$$

显然在这种分配体系下激励问题会变得相当严重。为了更加好地说明这个问题，首先让我们对生产技术和成员偏好进行描述。这里我们完全依照 Dong 和 Dow (1993) 的相关假定：

令集体成员 i 的效用函数为

$$U_i(y_i, e_i) = y_i - v_i(e_i), \quad i = 1, \dots, n. \quad (2)$$

其中 e_i 是第 i 个成员付出的努力， v_i 是第 i 个成员因付出努力而给自己带来的负效用， $v_i' > 0$ ， $v_i'' > 0$ ，且 $v_i(0) = 0$ 。 y_i 和 e_i 都非负。这种拟线性偏好假设在关于激励的文献中比较常见。

令集体总收入生产函数为

$$\begin{aligned} Y &= 0, & \text{若 } 0 \leq E \leq F/\theta, \\ Y &= \theta E - F, & \text{其他情况.} \end{aligned} \quad (3)$$

其中， $E = \sum_n e_i$ ， F 是集体在未生产任何正产出之前必须支付的固定设施成本（或者说，在生产出一个正的收入之前，集体必须预先支付 F/θ 单位的努力）。注意，对生产函数的这样描述意味着一个固定的正边际产量和递增

⁹ 事实上，在一个集体中，成员先是按他对不同任务所付出的劳动贡献（计件工资）或按他工作的天数（计时工资）从集体赚取“工分”。各集体成员通过这样的方式在一年里积累他的工分。上面式子中的 ω_i 指的就是集体成员 i 在一年里积累的总工分。

¹⁰ 放弃零监督假定并不会对我们下面的分析产生实质性的影响。

的平均净产量。

生产函数的线性形式确保了“偷懒”是每个成员的占优策略(即不管所有其他成员付出多大努力,偷懒都是自己的最佳选择)。为了说明这点,我们首先定义成员的“完全”努力水平和“偷懒”努力水平。考虑这样一个情形,在该情形下,成员的努力水平可以不费任何成本被观察到。这时集体就可以为每个成员直接选择努力水平。假定集体选择每个成员的努力水平以最大化集体福利:

$$\text{Max} \sum [\theta e_j - v_j(e_j)] - F.$$

一个有效的集体会选择满足:

$$\theta = v'_j(e_j^*), \quad j = 1, \dots, n. \quad (4)$$

此时有努力向量 $e^* = (e_1^*, e_2^*, \dots, e_n^*)$ 。我们称 e_i^* 为成员 i 的完全努力水平(即集体所要求的努力水平)。

然而,观察成员的努力水平不可能毫无成本,并且根据我们的假定,对努力的监督实际上是不存在的。这样,每一个成员就都可以选择自己的努力水平以最大化其个人收益:

$$\text{Max} (1/n) [\theta e_i + \theta \sum e_j - F] - v_i(e_i).$$

上述问题的解 e_i^S 满足

$$\theta/n = v'_i(e_i^S), \quad i = 1, \dots, n. \quad (5)$$

我们称 e_i^S 为成员 i 的偷懒努力水平。很显然,对所有 i 有 $e_i^S < e_i^*$ 。

我们只需对一成员分析 e_i^S 和 e_i^* 两个努力水平即可。根据我们对生产技术的假定, e_i^S 实际上是一成员不用考虑所有其他成员的努力选择时的最优努力选择(这从(5)式看很明显,因为只有成员 i 自己的努力水平出现在该式)。也就是说, e_i^S 是每个成员的占优策略。并且由于偷懒是所有 n 个成员的占优策略,努力向量 $e^S = (e_1^S, e_2^S, \dots, e_n^S)$ 或所有成员都偷懒就是这个努力付出博弈的占优策略均衡。¹¹

比较成员 i 分别采取 e_i^S 和 e_i^* 时的收益是有意义的。在比较这两个收益时,我们需要对其他成员选择的努力水平做出假定。当所有其他成员付出完全努力并且 i 也付出完全努力时,成员 i 的收益为

¹¹ 我们假定该努力付出博弈是一个一期同时行动博弈,并且每个参与者都掌握完全信息,即知道谁是其他参与者,自己的和所有对手的策略空间,以及对应于每一个可能策略组合的每个参与者的收益。然而,正如我们在第二部分关于两个参与者囚徒困境博弈的讨论所指出的,事实上,并不是所有这些假定对产生所有成员都偷懒的博弈均衡结局都是必要的。

$$U_i^C = (1/n) \left[\theta e_i^* + \theta \sum_{j \neq i} e_j^* - F \right] - v_i(e_i^*).$$

然而，如果成员 i 单方偷懒，他的收益为

$$U_i^S = (1/n) \left[\theta e_i^S + \theta \sum_{j \neq i} e_j^* - F \right] - v_i(e_i^S).$$

两者之差为

$$U_i^S - U_i^C = [v_i(e_i^*) - v_i(e_i^S)] - (1/n)(e_i^* - e_i^S)\theta. \quad (6)$$

考虑到 $v_i(\cdot)$ 的凸性假定，当 $n > 1$ 时，上述表达式的值应是正的。

上述的收益比较假定了所有其他成员都付出完全努力。但事实上，在我们关于偏好和技术的假定下，对于所有其他 $(n-1)$ 个成员努力选择的所有可能组合，成员 i 付出偷懒努力水平和付出完全努力水平的收益之差始终是一样的。也就是说，对于所有其他成员努力选择的所有可能组合来说，一个成员偷懒时的收益始终比付出完全努力时的收益大一个固定的数量。¹²

或许有人会反对说，上述结论是根据对偏好和技术所作的特殊假定才得出来的。拟线性偏好在文献中广为使用，但生产函数的线性形式更易遭致反对。然而，结论对生产函数的线性形式假定并不敏感。可以证明，在其他更一般的关于生产函数的假定下，偷懒在许多情形下是每一个成员的最优策略，尽管一成员来自偷懒的收益并不总是比来自完全努力的收益大一个固定的数量。我们这里对技术所作的特殊假定只不过是为了方便阐明在公社内部偷懒通常是成员的占优策略这样一个广为人知的事实。

因为偷懒是每个成员的占优策略，所以每个成员基于完全自私自利而都选择偷懒就有着令人信服的理由。然而，对所有参与者来说，所有参与者都偷懒的结果要差于所有参与者都付出完全努力的结果。当所有参与者都偷懒时，成员 i 的收益为

$$U_i^N = (1/n) \left[\theta e_i^S + \theta \sum_{j \neq i} e_j^S - F \right] - v_i(e_i^S),$$

显然，对所有参与者来说，这个结果要差于所有参与者都付出完全努力的结果（对所有 i 有 $U_i^N < U_i^C$ ）。

为了供后文的参照，我们定义成员 i 进行个体农作时的效用为：

$$U_i^P = \theta e_i^* - F - v_i(e_i^*).$$

通常，我们预期有 $U_i^S > U_i^C > U_i^P > U_i^N$ （参见下文）。显然，在相同的偏好和技术假定下，一成员在从事个体耕作时同样会付出完全努力 e_i^* 。

¹² 当 $n=1$ 时，上面的式子等于 0。在这种情形下，成员 i 实际上在进行个体农作。可以证明在成员从事个体农作时，他会像一个有效的集体所要求的那样选择努力水平 e_i^* 。

一次性努力付出博弈与重复努力付出博弈

我们上面描述的成员努力付出问题假定了博弈只进行一期。在一个一次性的一期博弈中,如果我们假定每个成员只从个人利益角度来决定努力付出水平,则偷懒或合作失败的原因是显而易见的:偷懒是每个成员的占优策略。如果事情果真如此的话,就没有合作结果产生的可能性。为了让合作成为博弈的一个可能均衡结局,人们引入了重复博弈。当一个一次性博弈重复进行时,则一个参与者本期的行为就可能会影响另一个或其他多个参与者下一期或随后所有期的行为。反过来,这也意味着一个参与者本期的行为不但会影响他本期的收益,而且还会(通过影响其他参与者在下一期或随后所有期的行为)影响该参与者在下一期或随后所有期的收益。而所有这些对一个参与者将来收益的影响,他在选择本期行为时是必须考虑到的。也就是说,重复博弈为参与者的本期行为和他在随后所有期的收益之间创造了一种跨期相关性,正是这种跨期相关性使得参与者本期的行为发生了改变,而这种跨期相关性在一期博弈里是不存在的。

在文章的第五部分,我们对 Lin (1990) 与 Dong 和 Dow (1993) 的两个模型进行了评论。在这两个模型中,上面提到的一次性努力付出博弈被重复进行了无数次。两个模型都在一个参与者的本期行为和其他参与者的随后期行为之间添加了某种相关性。在满足了某些条件后,博弈的一个均衡结局是每个参与者在每期都付出完全努力。这样看来,公社内部的合作失败可以归因于所假定的某种跨期相关性对成员偷懒失去了威慑作用。然而,在具体评论这两个模型前,我们首先对承诺是否可成为决定成员努力水平和合作可能性的一个决定因素进行探讨。这样做并不需要多期框架。与重复博弈通过使参与者行为具有跨期相关性而使得合作结果成为可能不同,承诺则通过引入确保规则使得不同参与者的行为之间具有相关性而使合作结果成为可能(而这种相关性在无确保规则作用时是不存在的)。后者可以在同期框架内展开,也可在多期框架内展开。

四、承诺和公社内部努力付出

当所有其他成员偷懒时,偷懒也许可以被成员 i 认为是“合理的”或“符合道义”的行为。但当所有其他成员付出完全努力时,偷懒就不可能再次被成员 i 简单地看作是“合理的”或“符合道义的”。然而,对一个仅仅追逐个人利益的人来说(我们称之为“完全自私自利者”),这两种情形是毫无区别的。因为按照假定,在这两种情形下(当所有其他成员都偷懒和当所有其他成员都付出完全努力),偷懒都占优于付出完全努力,而且在这两种情形下自己偷懒的收益比自己付出完全努力的收益恰好多出同样多。在这部分我们引入基于确保基础上的承诺。这里有必要指出,除了基于确保基础上的承诺

之外，公社成员也许还有过其他的承诺，但它们不是我们所要考虑的。根据一般观察，公社内部存在的最普遍的承诺类型还是建立在确保基础上的承诺。¹³

虽然建立在确保基础上的承诺在公社内部广为存在，但通常公社成员可能接受不同的确保规则。特别是，一个集体内部几乎总是存在着某些完全自私自利的成员。一般地，尽管公社部分成员会遵守确保规则，但最终的结局却是所有成员都偷懒。我们将研究为什么会有这样的结局，特别是在考虑到前述的策略行为后为什么仍有这样的结局（由于策略行为，即使有一些完全自私自利的成员存在，也有可能使合作成为博弈的均衡结局）。文章的第五部分给出了一些支持我们结论的证据，尽管我们不能提供支持这些结论的更强有力的证据。

（一）承诺和它的一些例子

我们在文章第二部分指出：在一个有 n 个参与者博弈的背景下，一个参与者也许会答应遵守某种行为选择当且仅当其他一个，其他两个……或其他 $(n-1)$ 个参与者也确保这样做。就公社内部的努力付出问题来说，这意味着一个成员可能会答应付出完全努力当且仅当其他一个，其他两个，……或其他 $(n-1)$ 个成员也付出完全努力。不同的成员可以有不同的承诺，因而参与者的确保规则有 $(n-1)^n$ 种可能的结构。此外，如前所述，成员还可以是完全自私自利或者完全为他人考虑的，这样参与者的确保规则有 $(n+1)^n$ 种可能的结构。

我们上面所给出的关于参与者确保规则的可能结构是从最一般的角度来说的。下面我们给出一些公社成员的可能承诺结构。应该说这些承诺都是属于相对保守的。我们考虑三个例子：

例1 集体中一个成员 A 是完全自私自利的，但所有其他成员遵守确保规则：付出完全努力当且仅当所有集体成员都付出完全努力。

例2 集体中两个成员 A 和 B 是完全自私自利的，但所有其他成员遵守确保规则：付出完全努力当且仅当所有集体成员都付出完全努力。

例3 集体中两个成员 A 和 B 是完全自私自利的，但所有其他成员遵守确保原则：付出完全努力当且仅当至少有 $(n-1)$ 个成员付出完全努力。

我们把上述三个例子中的每一例都作为确保博弈来进行分析，并考虑它们的同时行动和序贯行动两个变体。通过下面的分析，我们可以清楚地看到参与者同时行动还是序贯行动对博弈的结局至关重要。我们所作的其他标准

¹³ 根据可考的文献（例如 Unger, 1985）只要有涉及承诺的地方就会屡次提及确保原则。原因很简单，它本质上是互惠性（reciprocity）的一种形式：只有当部分或所有其他参与者答应采取某一行时，人们才会答应采取某一行。这种互惠性是制约人与人之间交往的最为根深蒂固的原则之一。

假定为：(1) 每一个成员对博弈都有完全信息（这些信息包括：参与者的入数、每个参与者的策略空间以及所有参与者对应于每个可能的博弈结局的收益）；(2) 当参与者序贯行动时，他们拥有完美信息（他们知道每一个参与者已经做出的每一个行动）；(3) 成员的承诺是“共同知识”（也就是说，他们每个人都互相知道他人的承诺；他们每个人都互相知道他人的承诺；他们每个人都互相知道他人的承诺；他们每个人都互相知道他人的承诺，如此直至无限）。

1. 例 1 的情形

乍看起来，少数几个完全自私自利的成员的存在会导致所有成员都偷懒的原因似乎是很显然的。根据确保规则的性质，假如一个成员偷懒，其他成员也会效仿。让我们考察例 1——一个最简单的例子。如果参与者对博弈有完全信息并且承诺是共同知识，则 $(n-1)$ 个遵守承诺的成员都知道 A 是完全自私自利的。首先假定参与者是同时行动的，这时博弈的均衡结局是所有成员都偷懒。这是因为 A 是完全自私自利的，所以 A 毫无疑问会偷懒，并且所有其他 $(n-1)$ 个成员能够推断出成员 A 的行为因而也会偷懒。

A 也许会出于自身利益考虑而希望其他成员不要因他而偷懒，并且只要其他成员会仿效他的行动，A 也许会考虑付出完全努力。然而，由于成员是同时行动的，A 不能够将他的选择信息传递给其他成员以影响他们的选择。即使 A 确实和其他成员进行了事前交流，考虑到所有其他成员对成员 A 承诺的了解，他们也不可能会相信 A 的言谈，当然他们就会选择偷懒。因此 A 也就不会费心去付出完全努力了。

在上述情形中，A 当然也希望采用我们前述的策略行为。然而，当参与者同时行动时是不可能采用策略行为的。一参与者能够采用策略行为的前提是他的选择应当被其他参与者明确无误地观察到。

下面我们考虑上述博弈的序贯行动变体，但保持所有其他假定不变。随着博弈行动顺序的不同，有可能产生策略行为并可能得到不同的均衡博弈结局。首先，假定 A 先行动。此时尽管 A 是完全自私自利的，他也会付出完全努力。A 会进行以下推理：如果自己付出完全努力并且这一选择能够被其他成员所观察到，由于成员的承诺是共同知识，所有其他参与者也会做出同样选择，从而导致全体成员付出完全努力。即使从 A 的个人利益角度看，这样的结局也好于所有成员都偷懒的结局。预料到这点，A 当然会付出完全努力，因此博弈的结局是所有参与者都付出完全努力。

有必要对上述策略行为产生的原因进行更深入的思考。事实上，与前述的囚徒困境博弈不同，此处 A 能够意识到自己的选择会通过其他参与者的承诺影响到他们的选择（在囚徒困境博弈中这种影响是不存在的）。而这些参与者的努力水平选择反过来又会影响自身的收益。具体地说，当其他任意一个

参与者将他的选择从偷懒变成付出完全努力时，参与者 A 的收益会增加 $(e_i^* - e_i^S) \theta / n$ 。A 能通过其他参与者的承诺行为而得到的收益增加总量取决于他能够施加上述影响的参与者的数量，这也意味着 A 可得到的收益增加量又取决于他在行动顺序中所处的位置。当成员 A 位于行动顺序的起点时，他的完全努力选择可以影响 $(n - 1)$ 个成员的选择，因此他可以从中得到 $(n - 1)(e_i^* - e_i^S) \theta / n$ 的收益增加量。这个收益增量会比他以 (6) 式衡量的因付出完全努力所遭受的损失大。注意在计算后者时完全没有考虑成员 A 的付出完全努力选择可能对其他成员选择所产生的影响。该分析还意味着，如果成员 A 在博弈中最后行动，则策略行为失去存在的基础，因为此时他的选择不能对其他成员的选择产生任何影响（此处我们忽略博弈重复进行的可能性）。

根据承诺是共同知识的假定，我们能够确定出在整个行动顺序中成员 A 将开始采取策略行为的截止点（设想我们在开始时把 A 放在行动顺序的终点，然后使他的行动次序越来越靠前）。假定在 A 之后行动的成员有 m 个，我们可求出 m 的临界值 m^* ， m^* 需满足如下条件：即 A 通过对在他之后行动的 m^* 个成员施加影响（促使他们选择付出完全努力）所得到的收益增量必须恰好等于他用 (6) 式衡量的收益损失。也就是说，临界值 m^* 满足等式 (7)：¹⁴

$$\sum_{k=n-m^*-1}^n (e_k^* - e_k^S) \theta / n = [v_i(e_i^*) - v_i(e_i^S)] - (1/n) \sum (e_i^* - e_i^S) \theta, \quad i = A, B, \dots \quad (7)$$

这里我们必须作一个说明。严格地说，我们现在还只能说 A 的选择具有影响 m^* 个成员选择的可能性。然而，这些 m^* 个成员只有当所有成员都付出完全努力时才会效仿 A 而付出完全努力，因此他们的选择同样也依赖于那些在 A 之前行动的成员的选择（当然也依赖于 A 以及那些在 A 之后行动的其他成员的选择）。因此，当在 A 之前行动的 $(n - m^* - 1)$ 个成员知道 A 是完全自私自利时，确定他们是否会确实选择完全努力水平就显得非常重要。然而，这 $(n - m^* - 1)$ 个成员会进行以下推理：假如他们付出完全努力，则处于截止点的自私自利的 A 也会付出完全努力，这将导致在 A 之后行动的 m^* 个成员也会付出完全努力，从而导致全体付出完全努力的结果。注意，根据共同知识假定，我们预期在 A 之前行动的每一个成员都会进行上述推理，并且他们中的每个成员也预期其他成员会进行同样的推理。基于以上分析，那些在 A 之前行动的成员会付出完全努力，A 付出完全努力，那些在 A 之后行动的成员也会付出完全努力，最终我们得到全体付出完全努力的结果。注意，

¹⁴ 我们忽略了整数问题。假如 (7) 式对任何正整数 m 都不成立，则 m^* 是使等式左边大于等式右边的所有正整数 m 中的最小值。

在本例中, A 不但存在着策略行为, 而且他的策略行为还必须被其他成员(特别是被那些在 A 之前行动的成员)所正确预期到。

在前面的分析中, 谁首先行动, 谁第二个行动, …… , 谁最后一个行动是完全任意的。这在分析行动顺序对博弈结局的影响时是可以接受的。现实中, 通常公社内部有制度或文化等因素决定着成员的行动顺序。在某些情况下, 成员可以自行选择自己行动的次序, 比如一些成员也许会希望在博弈的某一特定阶段行动, 而遵守承诺的成员可能会期待 A 首先行动。理论上, 只要 A 在行动顺序中的位置处于第 $(n - m^*)$ 个成员或之前, 都有望得到全体付出完全努力的均衡结局。但要得到这个结局, 不仅要求成员所进行的推理基于很强的共同知识假定之上, 而且还要求所涉及的成员具有完美无疵的逻辑思维。这显然比假定成员 A 首先行动且 A 的选择能够在其他成员行动前就确实无误地被观察到要更为苛刻。

2. 例 2 的情形

这里集体中有两个成员 A 和 B 一心一意地追逐个人利益, 但所有其他 $(n - 2)$ 个成员都承诺付出完全努力当且仅当所有成员都付出完全努力。我们再次假定博弈具有完全信息并且承诺是共同知识。如果成员同时行动, 则我们可以马上得出所有成员都会偷懒这一结论。¹⁵

在考虑这个例子的简单序贯变体之前, 我们首先假定 A 和 B 在博弈的起点同时行动, 但随后的所有成员序贯行动(后面我们会对这种行动顺序做出辩解)。这时 AB 之间产生了一个可称为“AB 博弈”的纯粹协调博弈, 它是整个博弈的一部分。¹⁶

		成员 B	
		完全努力	偷懒
成员 A	完全努力	5, 5	3, 0
	偷懒	3, 0	2, 2

出于说明的方便我们给出上面的收益, 其具体数值仅仅只是反映不同收益之间的排序关系。假如 A 和 B 都付出完全努力, 则所有其他成员跟着效仿, 这时的结局对 A 和 B 两人来说都比两人都偷懒从而使得所有成员都偷懒的结局好。当 B (A) 付出完全努力时, 则 A (B) 自己付出完全努力要比选择偷懒(从而导致所有其他 $(n - 2)$ 个成员都偷懒)好。当 A (B) 自己付出完全

¹⁵ 与例 1 不同的是这里多了一个自私自利的成员 B。显然, 例 1 中关于参与者同时行动而得到共同偷懒均衡的推理对例 2 更加适用。

¹⁶ 由于缺乏一个合适的术语, 我们把这个例子中 AB 两个参与者所进行的博弈简单称为“AB 博弈”。它与博弈论中传统意义上的“子博弈”不同。在本例中, 我们可以看到这两个完全自私自利的成员之间进行的 AB 博弈对寻找整个博弈的解起着关键作用。

努力而B(A)偷懒并因此导致除了A(B)外的所有成员都偷懒时，A(B)的收益最差。对A来说，如果其他 $(n-2)$ 个成员都偷懒，则这个结果最好是因A自己而不是因B偷懒所造成。并且对A来说，除了成员B外所有成员都偷懒这个结局要强于全体成员都偷懒时的结局。

这个协调博弈有两个可能的纳什均衡（偷懒，偷懒）和（完全努力，完全努力）。通过其他成员承诺的作用，这两个纳什均衡分别意味着整个博弈的两个均衡结局——全体偷懒和全体付出完全努力。我们并不清楚哪一个均衡结局会占上风。¹⁷

我们现在考虑例2的简单序贯行动变体。先假定A和B在博弈的起点序贯行动，则A与B之间存在着一个两个参与者序贯行动的AB博弈。很显然，此时任何一个第二个行动的参与者都会效仿第一个行动的参与者。因此第一个行动者就会愿意付出完全努力，因为这样做会导致第二个行动者付出完全努力，进而导致所有其他成员都付出完全努力。这样我们得到了全体完全努力的结果，这个结果对A和B两人来说都强于所有人都偷懒的结果。我们会发现这里并不存在上面出现的均衡多重性问题。

当成员序贯行动时，他们之间确切的行动顺序同样也变得至关重要。首先，如果A和B在整个博弈中最后行动，则他们两人都会偷懒。与例1中只有A一人是完全自私自利的情形一样，我们也可以确定A和B行动次序的截止点，在这些截止点他们各开始采用策略行为并因此导致全体付出完全努力的结局。令 m_A 和 m_B 分别为在A和B之后行动的成员数，临界值 m_A^* 和 m_B^* 的确定和例1中 m^* 的确定完全一样，则只要有至少 m_A^* 和 m_B^* 个成员分别在A和B之后行动，A和B两人各付出完全努力会影响到足够数量的成员选择付出完全努力，而这些成员选择付出完全努力给A和B带来的收益也值得A和B付出完全努力。因此，如果在A和B之前行动的成员也付出完全努力，则AB两人都会付出完全努力。¹⁸考虑到共同知识假定，那些在A和B之前行动的成员也会预料到这点，因此他们也会付出完全努力。注意，在这种推理下，A和B的行动次序没有必要直接相邻。

现在再回到A和B两人同时行动博弈的情况（但所有其他成员序贯行动）。很显然，假如A和B在整个博弈的终点同时行动，则他们俩人都会选择偷懒。我们可以把他们两人的同时行动次序从博弈的终点向前移到第 $(n-m_A^*)$ 个和第 $(n-m_B^*)$ 个位置中较前的一个。这时A和B之间仍存在着一个有意思的协调博弈（本文不详细展开这方面的分析）。然而，正如任意一个

¹⁷ 可以将这里的AB博弈推广到混合博弈的情形。在混合策略博弈中，每个参与者会考虑对手策略的不确定性。在本文我们不进行这方面的分析，而只把注意力集中在纯策略博弈和纯策略均衡上。

¹⁸ 如果在A和B之前行动的成员中有任何一人不付出完全努力，则因为确保规则在A和B之后行动的成员也不会付出完全努力，因而A和B两人也不会付出完全努力。

两个参与者的协调博弈一样,这里存在着两个纳什均衡结果,因此整个博弈的均衡结局也是不确定的。

3. 例3的情形

与例2不同的是,在例3中除了A和B之外的 $(n-2)$ 个成员承诺付出完全努力当且仅当至少有 $(n-1)$ 个成员付出完全努力。这里A和B两个成员仍然完全自私自利,参与者对博弈有完全信息并且成员的承诺是共同知识。很显然,在这个例子中,如果成员同时行动则所有成员都偷懒(当所有成员同时行动时策略行为是不可能产生的)¹⁹。

考虑序贯行动情形,我们首先让A和B在博弈的起点同时行动而所有其他成员在他们之后序贯行动,这时AB之间进行的博弈是我们熟悉的斗鸡博弈:

		成员 B	
		完全努力	偷懒
成员 A	完全努力	5, 5	10, 4
	偷懒	4, 10	2, 2

上面所给出的收益仅仅只是用来说明如下的关系:假如A(B)偷懒,则对B(A)来说最好是付出完全努力。反过来说,假如A(B)付出完全努力,则对B(A)来说最好是偷懒。B(A)自己偷懒而A(B)付出完全努力对B(A)来说要比A(B)偷懒而B(A)付出完全努力好。如B(A)付出完全努力,A(B)付出完全努力对B(A)来说要比A(B)偷懒好。最后,对A和B来说,所有成员都付出完全努力要比所有成员都偷懒好。这里有两个纳什均衡,(偷懒,完全努力)和(完全努力,偷懒)。我们不清楚具体哪一个均衡会占上风。

然而,AB博弈均衡的不确定性不会影响整个博弈的均衡,两个AB博弈均衡都能确保剩下的 $(n-2)$ 个成员付出完全努力。然而,由于A和B之间是同时行动的,事实上两个纳什均衡都不令人信服。A和B中的任何一个都希望对方付出完全努力(这样会导致其他 $(n-2)$ 个成员也付出完全努力)而自己偷懒。困难就在于如何使得对手知道他选择的是偷懒。因为具有同样偏好的对手当然也会用相同的方式进行推理。因此若AB两人同时行动,两人都很有可能冒险去偷懒而寄希望于对手付出完全努力。如果出现这种情况的

¹⁹ 如果假定只有A是完全自私自利的并且只要有至少 $(n-1)$ 个成员付出完全努力则其他所有成员承诺付出完全努力,在成员同时行动的情况下,则A会认为自己的选择并不影响其他成员的选择,因此A总是偷懒(偷懒是A的占优策略)。均衡中,除A之外的所有 $(n-1)$ 个成员或者全部偷懒或者全部付出完全努力。所以整个博弈的一个均衡是A偷懒而其他成员付出完全努力,另一个均衡是所有成员都偷懒。

话，结局则是（偷懒，偷懒），尽管这个均衡不是一个纳什均衡。为了便于后文的参照，我们把（偷懒，偷懒）称为 AB 博弈的“貌似合理的结局”。与整个博弈相应的“貌似合理的结局”是所有成员都偷懒。^{20 21}

然而当 A 和 B 序贯行动时，并不存在上述的均衡与“貌似合理结局”之间的分歧。显然，（A 或 B）谁第一个行动都会选择偷懒，而第二个行动者会付出完全努力。因此整个博弈的均衡结局是除了一个成员（A 或 B）外的所有其他（ $n-1$ ）个成员都付出完全努力。

同前面一样，如果 A 和 B 在整个博弈的终点行动，则上述均衡结局会发生改变。在该情形下，不论 A 或 B 中的哪一个首先行动，A 和 B 两人都会选择偷懒。这时整个博弈的均衡结局是全体偷懒。这里我们也可以确定出 A 和 B 各自在整个博弈行动顺序中的截止点，在这些截止点 A 和 B 两人之间进行的 AB 博弈会导致两人中的一人付出完全努力。这里截止点的确定方式同例 2。

（二）均衡结局和其他博弈情形

总结前面三个例子我们可以看到：均衡的结局取决于成员的承诺以及博弈的行动顺序（也就是说整个博弈是同时行动还是序贯行动，以及序贯行动下 AB 之间进行的是同时行动 AB 博弈还是序贯行动 AB 博弈）。当所有成员同时行动时，三个例子中的均衡结局都是全体偷懒；当成员序贯行动时（在例 2 和例 3 中不论 A 和 B 是同时行动还是序贯行动），博弈的均衡结局很大程度上取决于 A（例 1）或 A 和 B（例 2 和例 3）在整个博弈行动顺序中的位置。

当这些完全自私自利的成员在整个博弈中的行动顺序位于某一截止点之前时，他们会采取策略行为。否则，他们肯定会偷懒，这时三个例子中整个博弈的结局均是全体偷懒。当他们的行动位置在上述截止点之前时（例如，如果他们在整个博弈的始点行动时），在例 1 整个博弈的结局是全体付出完全努力。在例 2，若 AB 之间序贯行动则整个博弈的结局仍是全体付出完全努力；然而如果 AB 两人同时行动并在两人之间进行一个协调 AB 博弈的话，则整个博弈的结局既可能是全体偷懒也可能是全体付出完全努力。在例 3，若

²⁰ 这里的论证仅仅是直觉意义上的。当把上述博弈建模成混合策略博弈时，我们可以使有关论证更加严密。在某些参数下（例如上面双矩阵所给出的收益），可能存在一个混合策略均衡，在该混合策略均衡里，两个参与者都对两个纯策略赋以一定的正概率从而采用混合策略。取决于这个均衡的参数（每个参与者采用每个纯策略的具体概率），博弈的事后结局可以是（偷懒，偷懒）。当（偷懒，偷懒）发生的概率超过 0.5 时，我们可以称它为博弈的“貌似合理的结局。”

²¹ 此处的讨论同时也说明为什么在现在的情形下把 AB 建模成同时行动比建模成序贯行动更为恰当。就该博弈而言，谁最先行动就有先动优势。AB 两人都希望自己首先行动并偷懒。即使 AB 两人之间存在着事先的沟通（例如达成一个双方都付出完全努力的协议），他们每人都会被诱使去首先行动并获取先动优势。正因为如此，假定 A 和 B 之间同时行动显得更加合理。它与例 2 的纯粹协调 AB 博弈形成了鲜明的对照。在例 2 的纯粹协调 AB 博弈里似乎并不存在充分的理由去假定 AB 两人同时行动。事实上，这里假定他们两人序贯行动会更加恰当。

AB序贯行动,则在均衡中除了一个成员(A或B)外,所有($n-1$)个成员会付出完全努力;若AB同时行动并在两人之间进行斗鸡博弈,这时(偷懒,完全努力)和(完全努力,偷懒)都是可能的AB博弈纳什均衡,此时整个博弈的均衡结局再次是除了一个成员外其他成员都付出完全努力。然而,我们已经对这一情况下的纳什均衡解在AB博弈中的适用性提出了质疑,并且认为在这种情形下“貌似合理的结局”似乎应该是AB两人都偷懒,这样的话,整个博弈的“貌似合理的结局”是全体偷懒。

这三个例子从某种意义上说是三个初始博弈类型,对其中每一个的分析都可以得到进一步的延伸。例如,就例1来说,我们可以允许其余($n-1$)个成员采用更弱形式的承诺(包含更弱的限制条件)。至于例2,我们可以进一步增加完全自私自利成员的人数,在这种情形下,如果完全自私自利的成员同时行动,则协调博弈应该在三个或更多的成员之间进行而不是原来的两个成员之间进行。至于例3,我们可以或者进一步增加完全自私自利成员的人数或者进一步弱化成员承诺的限制条件,这里我们不准备对这些可能的情形作更深入的研究。

(三) 承诺的不完全信息和其他方面的考虑

通过上面的例子我们可以看到,尽管一些成员承诺遵守某些行为规则,而且这些规则在某些情况下会压倒完全自私自利的考虑,但是在不少结局中所有成员却像完全自私自利的成员一样选择偷懒。即使考虑到了策略行为,这种结果依然没有改变。下面我们对其他两个相关因素作进一步评论,它们也可能是帮助导致中国集体成员全体偷懒的原因。

首先,信息的不完全性。在上述分析中,我们假定了每个成员的承诺是共同知识,这在很大程度上是正确的。在中国,一个农业集体通常也就是一个村庄社区,由于集体成员在一生中的大部分时间是作为邻居或同事生活在一起,所以一集体成员对每个成员的承诺会有一些的了解。然而,这种了解永远也不可能是完全无缺的。一成员不时地会对其他成员的真实承诺产生怀疑,特别当周围所有的人看起来好像都在仅仅追逐个人利益时,悲观情绪就会逐渐产生。而通常在一个集体中,所有成员看上去也好像确实都只在追逐他们的个人利益。所以,根据拇指法则,只要没有足够的证据支持其他成员也遵守一定承诺的结论,他们就都会把所有其他成员看作是完全自私自利的。这种对每个成员真实承诺的谨慎的看法反过来会导致所有成员都只追逐个人利益,尽管他们中的许多人所遵循的价值观与专一追逐个人利益的价值观会相去甚远。

另一个同样妨碍集体成员进行合作的因素是一集体中完全自私自利成员的数量。如果一集体中有相当数量的自私自利成员存在,则其中任何一个成员都很难将自己行动所产生的影响与其他完全自私自利成员行动所产生的影

响区分开来（假定成员承诺是共同知识）。如果遵守承诺的成员对不合作做出的回应是同样也不合作，一个完全自私自利的成员可以认为这不是因自己而是因其他自私自利成员的不合作而引起的。在这种情形下，每个完全自私自利的成员总会认为自己不合作是有利的，因为他相信即使自己合作，其他自私自利成员也不会合作，从而导致所有成员都偷懒。²²

如果一个完全自私自利的成员总是认为其他一些完全自私自利的成员会偷懒并因此导致所有成员都偷懒，则该成员就会认为自己的行为选择与整个博弈的结局毫无关系，策略行为因此失去了存在的基础，因为我们知道策略行为存在的基础是完全自私自利的成员可以看到自身的行动与其他成员的行动之间存在着某种联系。

五、承诺和个人利益：为什么公社内部合作失败了

如同我们所熟知的，原指望成员之间展开的合作在中国公社内部绝大部分都失败了。是什么原因导致了合作的失败呢？我们在这一部分对一些学者给出的解释进行评论和比较，并表明：要对合作失败做出恰当的解释，应该把集体成员看成是遵守着某些承诺的；并且我们可以通过对集体成员所通常具有的一些承诺结构的分析来解释合作失败的原因。

（一）建立在个人利益基础上的解释

我们首先从 Lin（1990）的解释开始，该解释曾引起极大的争论。Lin 观察到在 1958 年下半年公社化运动之前，中国集体还是取得成功的，这表现在全要素生产率的增加上。然而，公社化后的中国农业先在 1959—1962 年的大饥荒中陷入了深重的危机，然后在 1962—1978 年间陷入了长期的停滞状态（以全要素生产率来衡量）。研究中国农业的不少学者基本上同意了 Lin 对有关证据的解读，虽然也有人对其某些细节提出了质疑。Lin 认为，中国公社内部（以及通常在任何农业环境下）对成员努力的观察和监督即使未接近不可能也是非常困难的，并以此作为其分析的起点。他认为：

……一个农业合作社要取得成功就不可避免地依赖于合作社成员之间默契达成的一个自我约束的协议。然而，只有当合作社成员在其他成员不遵守协议时有权退出合作社，这种自我执行的协议才能维持。在合作化的初级阶段，退出权普遍得到了很好的尊重。因此这时自我执行的协议在大多数集体中还能维持并带来了整个农业生产绩效的提高。（Lin, 1990, P. 1249）

²² 上面的例 2 是一个很好的例子。例 2 中若 AB 同时行动则他们两人之间进行 AB 博弈的一个均衡结局是两个人都偷懒，从而导致全体成员都偷懒（假如所有其他成员随后序贯行动）。这时 A 和 B 两人都可认为不是自己而是对手的行为引发了所有成员都偷懒，并且他们两人都可认为即使他自己不偷懒对手也会偷懒并因此导致所有成员都偷懒。

Lin 上面所说的自我执行协议与他关于成员的退出权假说有关, 这里的退出权是指, 当集体绩效不好并且成员单干的收益要大于其留在集体中所能获得的收益时, 该成员可以退出集体(集体成员在合作化运动初期拥有这种退出权)。一个集体的绩效依赖于成员的勤劳与合作。如果“勤快”的成员退出集体, 则集体的绩效会变差, 从而使得更多的成员退出集体, 最终导致一个集体完全解散。然而, 由于规模经济的原因, 集体的成功最终是符合成员的个人利益的。集体的解散会对包括那些选择不合作的成员在内的所有成员都起到一个惩罚作用。当那些“懒惰”的成员意识到他的偷懒最终会损害到自身利益时, 他们就会对自己的偷懒行为进行克制。只要这个退出机制能够发挥作用的话, 就会对“懒惰”成员的偷懒行为起到一定的抑制作用。按照 Lin 的观点, 这解释了集体在合作化初期所取得的成功, 当时成员拥有退出不成功集体的权利; 并且它也解释了公社化运动(1958年下半年)后集体所面临的失败, 因为此时集体成员不再拥有退出权。

有学者就 Lin 对相关时段历史的阐释提出了异议(例如 Kung, 1993)。但这些问题不是本文所要关心的。就本文的分析而言, 有必要指出的是 Lin 把成员分为两类: “勤快的”和“懒惰的”, 或“有所顾忌的 (scrupulous)”和“无所顾忌的 (unscrupulous)”(Lin, 1993)。然而, 应该注意到这些区分完全是从成员的个人偏好和对个人利益追求的角度做出的。尽管像“有所顾忌的”和“无所顾忌的”这样的术语已包含着人们的道德品质问题, Lin 在他的分析中实际上并没有考虑到道德因素。Lin 对成员类型存在的差异只是从他们在“时间偏好, 个人能力以及其他禀赋”上存在的差别来解释的 (Lin, 1990)。

Lin 所假定的惩罚偷懒的自我执行机制(该机制取决于成员是否拥有退出权)的逻辑基础遭到了 Dong 和 Dow (1993) 的质疑。他们提出了不同的防止偷懒的惩罚机制, 这就是“报复性的偷懒”, 每个成员答应在无人偷懒时每期都努力工作, 但威胁只要有证据表明在过去曾有人偷懒则在随后的每期中都会进行报复性偷懒(无限期重复博弈)。由于直接监督一个成员是否偷懒是非常困难的, 为简单起见可假定这种监督是完全不可能的。如果本期产量低于本期所有成员付出完全努力而应该得到的产量(假定不存在生产不确定性), 则表明有其他成员选择了偷懒。假定成员不拥有退出权, 在一定的成员偏好和生产技术假定下(见前第三部分), Dong 和 Dow 证明, 在下面两种情形下, 不允许成员退出集体实际上会增加对集体成员偷懒的威慑力: (1) 当成员间对闲暇的偏好差异并不很大并且规模经济并不显著时; (2) 当集体足够大并且集体是在一个个体农作能够独立存在的环境下运作时。当成员各不相同, 这些条件能确保对所有 i 有 $U_i^P > U_i^N$, 而这是不具有退出权时报复性偷懒对成员偷懒有更大威慑力的必要条件。²³

²³ 回忆第三部分, U_i^P 和 U_i^N 分别是成员 i 进行个体农作和作为集体(全体成员偷懒)的一员时一期收益所产生的效用。

要明白这是为什么并不困难。若 $U_i^P > U_i^N$ ，则不具有退出权时的报复性偷懒会比具有退出权而从事个体农作对成员 i 具有更大的威慑力。若要使报复性偷懒对所有成员的偷懒行为都具有更大的威慑力，则对所有成员该不等式都须成立。如果成员间对闲暇的偏好存在较大差异的话，则有可能上述不等式只对部分成员成立。同样地，如果规模经济非常显著的话，则也有可能上述不等式（至少对部分成员）不成立。至于 Dong 和 Dow 所指出的第二种情形，我们只需注意当一个集体足够大时则每个成员的努力贡献会变得相当小（参见第三部分方程 5），这会对 U_i^N 起下降的作用（在极限 U_i^N 等于 0，在这种情形下不等式退化为只要求个体农作的收益为正）。

Dong 和 Dow 的报复性偷懒策略存在可置信方面的问题，在下文我们会对此进行讨论。此处我们姑且认为它是一个可置信策略，那么，要使 Lin 的退出权假说成立，我们一方面要求对部分成员有 $U_i^P > U_i^N$ （以便这些成员在他人偷懒时确实会选择退出）；另一方面要求对其他成员有 $U_i^P < U_i^N$ （以便不存在退出权意味着对他们偷懒行为的威慑力会下降）。这也正是 Lin（1993）后来承认的。²⁴我们需要强调的是，如果成员之间不存在这种程度的差异（例如假定成员完全相同），则 Lin 的退出权假说就完全站不住脚。

（二）建立在承诺基础上的解释

Lin 的退出权假说及其所引发的争论有助于强调成员类型具有差异性。Lin 把这种差异性归因于成员在“时间偏好，个人能力以及其他禀赋”方面的差别。此处我们并不试图完全驳倒对成员差异性原因或者说对成员差异性性质的这种解释（根据前文第二部分所指出的我们也知道要完全做到这点是不可能的）。但我们可以利用前文所定义的承诺概念提供一个不同的解释。

在第二部分我们已考虑了使承诺成为人类行为决定因素的理论基础。我们现在需要做的是检验那些证明中国公社成员存在承诺差异性的证据。要在这里进行一次详细的检验是不可能的，我们只引用当时的一些政策制定者、实践者和学者的一些重要的言论来加以说明。

我们首先引用刘少奇的话。刘是 20 世纪 50 年代中期到 60 年代中期中国的国家主席，也是合作化运动初期（按照 Lin 的观点这时的合作化取得了成功）随后进行的公社化和随后所发生的农业危机直至危机复苏这整个期间中国的最高层决策者之一。在 1962 年的一次会议上，来自某地方的干部向刘汇报说，新生的人民公社所面临的一个主要问题就是每个集体受到少数捣蛋鬼的拖累。针对这点，刘回答说公社所面临的困难是因为“捣蛋鬼”并没有排

²⁴ Lin（1993）提出了一种“ j 型”成员。对这些成员有 $U_j^N > U_j^C > U_j^P$ 。这个规定比确保退出权具有更大威慑力所必需的条件要强。注意，报复性偷懒触发策略（trigger strategy）在这种类型的成员存在时会使得全体偷懒成为一个确定的结果。

除在集体之外,而在合作化运动的初期基本上是将捣蛋鬼排除在集体之外的,但这一做法在后来并没有得到坚持。用刘自己的话说:

初级社时,我们采用的是自愿合作的方式。如果你不是好人,我不会让你加入。现在好人和坏人都加入到其中。在每个生产队中都存在着个别捣蛋鬼。如果我们不采取措施约束他们的话,这些捣蛋鬼会对每个成员都会产生不良影响。(《中国农业合作史资料》,1998, No. 1)

我们也许可以把刘的对好人和坏人的二分法解释为由于成员之间存在着“时间偏好、个人能力和其他禀赋”上的差异。然而,这种解释看起来并不十分令人信服。而如果我们把这个二分法解释为成员承诺上的差异似乎更为贴切,即成员中的一些人专一地追逐个人利益而另一些人则遵守某些承诺。注意,促使刘作上述讲话的原因是地方干部对当时形势的估计,这对我们应该是很有启发意义的。²⁵

下面我们再引用陈永贵的话。陈在20世纪60年代中期前一直担任大寨大队的党支部书记(该大队在20世纪60年代中期到70年代末成为了全中国农村集体学习的典范)。在一篇描述大寨大队“劳动核算”经验(衡量成员劳动投入量的方法)的文章里,陈(1966)提到了两类完全不同的人——狡猾人和老实人。按照陈的说法,这两种不同类型的成员在面对同样的环境时所做出的反应是截然不同的。例如:在计件工分制时,狡猾人会不惜以牺牲工作的质量为代价,把他们的精力放在尽可能多地完成工作的数量上。老实人却对工作的质量予以足够的重视,结果他们完成的工作数量和工分都比狡猾人要少。由于检验工作的质量非常困难(Lin, 1990),因此,一集体往往仅仅按照成员完成的工作数量而不考虑工作的质量去核算成员的劳动。结果是,狡猾人在报酬上获得了更多的好处,而老实人则由于重视了工作的质量反而得到更少的报酬。

尽管陈试图把成员行为的这种差异归因于他们政治觉悟不同(那个时代的标志),但狡猾的和老实的这两个术语意味着成员在道德品质上存在着巨大的差异,并且这两个术语看起来也与我们对成员做出的“完全自私自利的”和“遵守承诺的”区分非常吻合。把成员行为的差异仅仅归因于他们在“时间偏好,个人能力和其他禀赋”上的差异是很难令人信服的。

最后我们引用Unger的文章。Unger与他的合作者一起写了一部关于广东省陈村所经历的激烈的合作化阶段、公社化阶段和改革开放阶段的记录性著作,该著作也许是众多关于中国农村的记录性报告中最为精确、最为详细和最具权威性的报告之一(Chan et al., 1984)。

²⁵ 刘的讲话在其他方面也有启发性。它密切地反映了Lin的退出权假说。然而,刘的论题是集体阻止捣蛋鬼加入的权力,而不是成员从集体退出的权利。参见Data和Li(1999)。

可以观察到，首先，尽管陈村的村民总是关注着他们家庭的利益，但他们并非专一地追逐家庭利益。他们知道为提高农业生产率所付出的共同努力会提高所有人的生活水准。并且他们中的大多数人甚至愿意支持实行大寨式的更加平均主义的分配方法——只要能力更差的邻居们福利的增加最终不是以牺牲他们的家庭利益为代价。

困难在于如何协调家庭和集体的利益。生产队总是会遇到“搭便车”的问题。一个人的所有邻居不会都把集体目标与他们自己更狭隘的个人利益摆在同等地位，因此，他们之中也不是所有人必定会承担公平的工作份额。事情如果是这样的话，除了傻瓜之外还会有谁为了他们的“搭便车”邻居们竭尽全力地工作呢？(Unger, 1985, P137；斜体字为本文作者所加。)

(三) 建立在承诺基础上的解释和建立在个人利益基础上的解释的对比

前面我们对建立在承诺基础上的和建立在个人利益基础上的有关人民公社内部合作失败原因的解释进行了评论。接下来我们从几个重要的方面对这两个不同的解释进一步进行比较。我们认为承认承诺也是人类行为的决定因素的解释更加令人信服。

(1) 报复性偷懒触发策略的可置信性。上文在谈到 Dong 和 Dow 的报复性偷懒触发策略时，我们并没有强调指出该策略在可置信性上存在着问题。然而，正如 Dong 和 Dow 自己指出的，该策略只是在子博弈完美意义上是可置信的，也就是说如果偷懒已经发生并被发现（通过本期产量低于预期产量而发现），并且所有其他成员在下期以及随后的所有期确实采取报复性偷懒时，则任何成员执行触发策略都符合自身的利益。²⁶然而，该策略并不满足更严格的可置信性检验——抗重新谈判检验。如果有人已经偷懒并按照触发策略要求其他成员在随后所有期采取报复性偷懒，这时那些未偷懒的成员肯定会进行以下推理：“过去的就让它过去吧”，也就是说放弃惩罚。因为这里的惩罚除了惩罚偷懒者外，还会惩罚未偷懒者。因此，未偷懒的成员不可能采取这种报复性惩罚，特别是如果这样做会导致全体成员面临大饥荒的威胁；而在当时的资源和技术条件下，如果全体成员都偷懒，公社面临大饥荒不是一件不可能的事。即便全面的大饥荒不可能真的发生，只要重新谈判带来的收益大于采取报复性偷懒带来的收益，那些未偷懒成员仍然有可能重新谈判而拒绝采用报复性偷懒。令 j 为某一期的偷懒成员，所有其他成员（所有 $i, i \neq j$ ）会推断如果他们不进行报复则随后各期都有效用：

$$U_i^{C-j} = (1/n) \left[\theta e_j^S + \theta \sum_{k \neq j} e_k^* - F \right] - v_i(e_i^*),$$

$$i, k = 1, 2, \dots, n; i, k \neq j.$$

²⁶ 正如第三部分所指出的，在前述的关于技术和成员偏好的假定下，对于所有其他成员努力选择的所有可能组合，一个成员偷懒时的收益比他付出完全努力时的收益始终大一个固定的数量。当所有其他成员偷懒（或以偷懒进行报复）时，一成员最好的选择也是偷懒（或以偷懒进行报复），该结论在更一般化的条件下也成立。

然而如果他们报复,他们每人每期只能得到 U_i^N 。显然,我们预期对所有 $i \neq j$ 有 $U_i^{C-j} > U_i^N$ 。所以所有这些成员不会采用报复性偷懒,报复性偷懒触发策略是不可置信的。²⁷

正是因为每个未偷懒成员也像偷懒成员一样仅仅关注他的个人利益(他在决定是否采取报复性偷懒时完全是从自身利益角度出发的),最终使得报复性偷懒触发策略不可置信。除非成员遵守某种“超越”个人利益的行为规则,否则每个未偷懒成员在随后每期都采取偷懒来加以报复是不可置信的行为。

事实上,我们只需对 Dong 和 Dow 的努力付出重复博弈略加修正就能使之成为一个努力付出重复确保博弈。设有完全自私自利成员 i 。除该成员外,所有其他成员都承诺只要所有成员在每期合作,他们也会在每期合作,但是只要有人在任一期不合作,则所有这些成员中的任何一人在随后各期都不合作。要得到全部合作的结局,只要求对成员 i 来说,下面不等式成立:

$$U_i^C(1 - \delta_i) \geq U_i^S + \delta_i U_i^N(1 - \delta_i), \quad (8)$$

其中 δ_i 是成员 i 的主观贴现因子。不等式的左边是无限期合作所产生的收益,不等式右边是成员 i 得自他在第一期里偷懒而所有其他成员都努力但在以后各期所有成员都偷懒的总收益。可以把该重复确保博弈看作是在本文第四部分所谈到的一期确保博弈的推广。²⁸

(2) 无限期重复博弈。即使忽略触发策略的非抗重新谈判性,要使这个策略取得预想的结果,这里的一期博弈必须重复进行无限次。因为正如大家所熟知的,任何有限重复博弈都可以从博弈的最后阶段进行反向求解,这样每个成员在博弈的最后一期,倒数第二期……直至博弈的任意一期都选择他的一期占优策略——偷懒。当人类寿命是有限的并且博弈所给定的环境会随时间而发生改变时,博弈重复进行无限次的假定严格地说是非现实的,尽管无限期博弈确是一个可带来在有限期情形下不可能得到的合作结局的分析工具。然而,它要求一期博弈重复进行无限次才能得到合作的结果,这一点与其说表明了合作的可能性,倒不如说充分地说明了完全自私自利的参与者很难会

²⁷ 以同样的方式,我们可以得出,在某些条件下, Lin 的退出策略虽然满足了子博弈完美但并不满足抗重新谈判性。我们再次令 j 为某期的偷懒成员。如果所有其他成员因此而退出集体但成员 i 没有退出,则只有成员 i 和 j 两人留在集体里。假定 j 在随后各期继续偷懒(但程度不同,参见方程 5)并假定 i 继续付出完全努力。成员 i 在随后各期的收益是 $(1/2)[\alpha(e_i^S + e_i^*) - F] - v_i(e_i^*)$ 。只要 F 足够小,则这个收益小于 U_i^P 。注意这个条件是与个体农作能够独立存在的条件相一致的。所以,我们可以得出结论退出策略满足子博弈完美。但要它具有抗重新谈判性,则要求 $U_i^P > U_i^{C-j}$ 成立。然而这是不可能成立的。 U_i^{C-j} 与 U_i^C 仅仅相差很小,而由于规模经济 U_i^C 明显大于 U_i^P 。由于我们这里主要关心的是报复性偷懒策略的可置信性而不是退出策略的可置信性,所以本文不对此作深入讨论。

²⁸ 注意,假若不存在确保规则,在报复性偷懒触发策略下(假定该策略可置信),要得到全体成员合作,则要求不等式(8)对每个成员 $i = 1, 2, \dots, n$ 都成立。

选择合作，除非他们的行为也被某些非个人利益的因素所支配。

本文第四部分的分析表明，一旦我们考虑承诺，就可以得到所有成员都肯定不合作的结局，也可以变为全体成员都合作的结局。注意，该结论既不依赖于无限期重复博弈，也不依赖于一期博弈是否得到任何次数的重复。

(3) 公社内部合作失败的原因。本文第四部分的分析表明，公社内部的合作失败可以归因于集体成员具有确保型承诺的某些特殊的结构。同时，在考虑到承诺的情况下，成员的行动顺序（如果成员序贯行动的话）也对博弈结果有重大影响。按照承诺结构和成员行动顺序的不同，我们或者可以得到全体成员偷懒的结局，或者可以得到全体成员付出完全努力的结局；还有可能是部分成员偷懒而其他成员付出完全努力的结局（建立在个人利益基础上的解释并没有考虑到这种可能性）。

在基于个人利益解释的框架里，全体偷懒或全体付出完全努力都可能是无限重复博弈的一个结局，最终会得到哪一个结局取决于成员的主观贴现因子和他们选择的对偷懒的惩罚方式。定义每个成员 i 的主观贴现因子 δ_i 的两个临界值，它们满足：

$$\begin{aligned}\delta_i^1 &\equiv (U_i^S - U_i^C) / (U_i^S - U_i^P), \quad i = 1, \dots, m; \\ \delta_i^2 &\equiv (U_i^S - U_i^C) / (U_i^S - U_i^N), \quad i = 1, \dots, m.\end{aligned}$$

设对所有 i 有 $U_i^P > U_i^N$ ，则对所有 i 有 $\delta_i^1 > \delta_i^2$ 。在该情形下，我们可以得到：

(i) 若对所有 i 有 $\delta_i \geq \delta_i^1$ ，则不管对偷懒的惩罚是采用退出集体还是采用报复性偷懒的方式，我们得到全体成员合作的结果；(ii) 若对所有 i 有 $\delta_i^1 > \delta_i \geq \delta_i^2$ ，则对偷懒采用报复性偷懒的惩罚方式（也就是说不拥有退出权）仍有可能得到全体合作的结果，但对偷懒采用退出的惩罚方式不可能得到全体合作的结果；(iii) 若对所有 i 有 $\delta_i < \delta_i^2$ ，则不管采用何种方式惩罚偷懒，结果都是全体偷懒。²⁹我们再次强调以上结论假定了报复性偷懒策略是可置信的。

我们已经指出，Lin 的退出权假说若要成立，则集体中应有部分成员有 $U_i^P > U_i^N$ （第一组成员），而其他成员有 $U_i^P < U_i^N$ （第二组成员）。对于第一组成员（也只对该组成员），上面的结论依然成立。但对第二组成员，上面关于三种不同情形的定义就不成立了。³⁰在同时有这两种不同类型的成员存在时，

²⁹ 这些结论和不等式(8)是建立在同样的推理上的，即一个完全自私自利的成员在任意一期都选择合作当且仅当他得自全部成员永远合作的总收益 $U_i^C / (1 - \delta_i)$ 超过或至少不少于该成员在一期偷懒并因此导致其他成员永远的报复性偷懒时该成员的收益 $U_i^S + \delta_i U_i^N / (1 - \delta_i)$ 。如果要使得所有成员都合作，则需对所有 i 要求上述收益比较成立。这些结论是对 Dong 和 Dow (1993) 在假定成员相同的情况下所做出的结论的推广。

³⁰ 当然我们只需在定义上述三种不同情形的不等式中交换 δ_i^1 和 δ_i^2 的位置，也可对第二组成员定义与上述三种情形完全类似的三种不同情形。

如果第一组中的任何成员偷懒,第二组中没有任何成员会从集体退出。另一方面,只要第二组中有成员偷懒,第一组中的成员都会退出集体。因此 Lin 的退出权假说只有当我们考虑潜在的偷懒者是来自第二组成员时才成立。³¹

我们通过对基于个人利益的解释和基于承诺的解释进行对比可以看出,前者把注意力集中在成员的主观贴现因子上,而后者则把重点放在成员承诺的结构上。在这两种解释下,根据不同的博弈情形,均衡结局都可以是全体成员偷懒,也可以是全体成员付出完全努力。一部分成员努力工作而另一部分成员偷懒的中间情形可以在基于承诺的解释里得到,但 Lin (1990, 1993) 和 Dong 和 Dow (1993) 都没有对这一情形进行研究。这种结果的多样性实际上恰好和公社的现实相吻合。尽管中国大多数农村集体在取消退出权后表现不佳,(主要原因当然是困扰着成员的激励问题),但这个时期也出现了一些成功的集体或者至少集体的绩效并不是清一色的差。这些绩效上的差异也许是因为不同集体里的成员有不同的贴现因子,但也可能是由于不同集体的成员有不同的承诺结构。通过本文的例子可以看出,后者可能是更恰当的解释。

六、结 语

本文介绍了 Sen 关于承诺的思想以及承诺在决定人类选择时所起的作用,特别是 Sen 原来在两人博弈框架内阐述的确保规则在这里得到了进一步的强调并被推广到 n 个人博弈的情形。我们对确保型承诺的不同形式进行了分类。我们所关心的问题本质上是经典的集体行动问题,即在实行集体目标(一个对社会和个人都更好的结局)与追逐狭隘的个人利益之间存在着的根本冲突。对这个问题的标准描述是囚徒困境博弈。

我们分析了集体行动问题的一个特殊表现形式,即中国公社内部的合作问题。一个集体要取得成功,则集体的每个成员都应该努力工作。然而,在一定的关于成员偏好和生产技术的假定下,不管其他成员如何行动,偷懒都符合每个成员的个人利益。也就是说,偷懒是每个成员的占优策略。

中国集体化农业的不平凡历史以及集体化成功或失败的原因近来在学术界成为争论的一个主题。Lin (1990, 1993) 以及 Dong 和 Dow (1993) 都在无限期博弈背景下假定了成员会对偷懒采取惩罚机制。成员在任何一期的偷懒都会引发所有其他成员在随后各期进行报复。在 Lin 的例子中,惩罚采取“勤快”的成员退出集体并最终导致集体解体的方式。不用说这种形式的惩罚只有成员拥有退出集体的权利时才是可行的。在 Dong 和 Dow 的例子中,惩罚采用报复性的偷懒方式,它在成员不具有退出权时取代了退出作为惩罚的方式。文献中的大部分争论集中到这两个惩罚机制的相对威慑力上。由于中国农村集体成员在 1958 年公社化运动后不再拥有退出权,所以上述争论具有

³¹ 注意,退出威胁也存在不可置信的问题。参见注 27。

历史相关性。

虽然文献中争论的注意力主要集中在两个惩罚体系的相对威慑力上，它们也突出了中国农村集体的一个重要方面，即成员类型存在着显著的差异。Lin 试图完全从成员在“时间偏好，个人能力以及其他禀赋”方面的差别来解释上述差异性。这种解释与建立在专一的个人利益基础上的人类选择理论是完全一致的。在这篇论文中，我们把建立在这一理论基础上的集体成员努力水平选择模型与包含承诺的集体成员努力水平选择模型进行了对比。我们发现能使 Lin 的退出权以及 Dong 和 Dow 的报复性偷懒惩罚机制成立的假定条件是非常苛刻的，并且这些条件带有很大的限制性。与此形成对照的是，包含承诺的模型避免了做出同样的具有极大限制性的苛刻假定，并且它所提供的对成员努力水平选择的解释和理解通常更加令人信服。

除了为中国公社合作失败的原因提供一个更加令人信服的解释外，本文的一个更大的目的是把承诺和道德因素（或更一般意义上的非个人利益因素）重新引入到当前关于中国经济改革和发展的讨论中。中国以市场为导向的经济改革不可避免地要给合法地追逐个人利益以应有的地位。由于由市场驱动和市场管理的现代经济的一个根本支柱是追逐个人利益，所以这样做既是正确的也是恰当的，但它不是惟一的支柱。我们要不断重复地强调，更加发达和更加成功的市场经济实际上是建立在一个融个人利益因素和非个人利益因素在一起的因素集合体之上的，正是这样的集合体才使得那些市场体系成功地运作着。

然而，为了正确理解一个融个人利益因素和非个人利益因素在一起的因素集合体可怎样对中国经济发展做出贡献，我们首先需要对这些因素是如何决定个人行为和个人决策的问题有一个深入的了解。希望本文对特定的人民公社内部合作问题的研究以及对 Sen 关于承诺和确保原则观点的介绍，起到抛砖引玉的作用。

参 考 文 献

- [1] Chan, A., R. Madsen and J. Unger, *Chen Village: The Recent History of a Peasant Community in Mao's China*. Berkeley: University of California Press, 1984.
- [2] Chen, Yonggui, "Adhere to Socialist Labour Management", *People's Daily (Renmin Ribao)*, 22 March 1966 (in Chinese). English translation in *Survey of China's Mainland Press*, 12 April 1966, 3—26.
- [3] Data, B. and M. Liu, "Work Incentives in a Model of Collective with and without Universal Membership", Discussion paper series, No. 99/35, Department of Economics, University of York, 1999.
- [4] Dong, X. and G. Dow, "Does Free Exit Reduce Shirking in Production Teams?" *Journal of Comparative Economics*, 1993, 17(2), 472—484.
- [5] Fan, G., "Amoral Economics", *Dushu*, 1998, 6, 50—55 (in Chinese).
- [6] Kung, J. K., "Transaction Costs and Peasants' Choice of Institutions: Did the Right to Exit Really Solve the Free-rider Problem in Chinese Collective Agriculture?" *Journal of Comparative Economics*, 1993, 17(2), 484—503.

- [7] Lin , J. Y. , “ Collectivization and China ’s Agricultural Crisis in 1959—1961 ” , *Journal of Political Economy* , 1990 , 98(6) , 1228—1252.
- [8] Lin , J. Y. , “ Exit Rights , Exit Costs , and Shirking in Agricultural Cooperatives : A Reply ” , *Journal of Comparative Economics* , 1993 , 17(2) , 504—520.
- [9] Liu , M. , “ Exit Right , Retaliatory Shirking , and the Agricultural Crisis in China ” , *Journal of Comparative Economics* , 1993 , 17(2) , 540—559.
- [10] Sen , A. K. , “ Labour Allocation in a Cooperative Enterprise ” , *Review of Economic Studies* , 1966 , 33 , 361—371.
- [11] Sen , A. K. , “ A Game-Theoretic Analysis of Theories of Collectivism in Allocation ” , in *Growth and Choice* , T. Majumdar ed. London : Oxford University Press , 1969.
- [12] Sen , A. K. , “ Choice , Orderings and Morality ” , in *Practical Reason* , S. Korner ed. , Oxford : Blackwell , 1974.
- [13] Sen , A. K. , “ Rational Fools : A Critique of the Behavioural Foundations of Economic Theory ” , *Philosophy and Public Affairs* , 1977 , 6 , 317—344.
- [14] Unger , J. , “ Remuneration , Ideology , and Personal Interest in a Chinese Village , 1960—1980s ” , in *Chinese Rural Development : The Great Transformation* , W. L. Parish ed. , New York : M. E. Sharpe , 1985.
- [15] 《中国农业合作史资料》 , 1988 年 , 第 1 期。

“ I ’ ll Cooperate Provided You Cooperate ” : An Application of Sen ’ s Assurance Rule to the Cooperative Behavior on the Chinese Commune

MINGQUAN LIU

(*Hopkins-Nanjing Center , Nanjing , China*)

Abstract In social interaction a person ’s behaviour may be bounded by a variety of “ commitment ” in that she may sometimes suspend her relentless pursuit of selfinterest , contrary to what a “ completely individualistic ” player does in the classic Prisoners ’ Dilemma game. Commitment can have diverse “ shades ”. In a situation requiring cooperation between two agents , one may agree to cooperate when *assured* that the other does also(Sen , 1969 , 1974). Where cooperation is required between more than two agents($n > 2$) , one may agree to cooperate if one other , two others , ... , or all($n - 1$) others cooperate. Finally an agent may agree to unilaterally and unconditionally cooperate. My paper aims to analyze different shades of commitment by collective members on China ’s communes (1958—1978). Depending on configurations of members ’ commitments , different outcomes of cooperation become possible and were sometimes observed. This interpretation of the nature of the cooperation problem on communes is contrasted with two models (Lin , 1990 and 1993 ; Dong and Dow 1993) which only admit selfinterest as a determinant of collective members ’ effort choice.

JEL Classification A13 , C72 , P32