两阶段基于信号博弈的声誉模型①

肖条军, 盛昭瀚
(南京大学管理科学与工程研究院,南京 210093)

摘要:建立了一个两时期基于信号博弈的声誉模型. 有两个参与人进行博弈,拥有私人信息的参与人叫发送者,另一个不拥有私人信息的叫接收者;该声誉模型研究 L 类(低能力类型) 发送者是否有动机在第一时期建立声誉. 证明显示,如果 L 类发送者在第一时期建立声誉,则他在第二时期的最优信号更大,在第一时期的效用更小,但是,他将在第二时期获得更高的效用.

关键词:声誉; 信号博弈; 算法
中图分类号:O225　　文献标识码:A　　文章编号:1007 - 9807(2003)01 - 0027 - 05

0 引　言

企业与企业、企业与消费者之间经常存在一种博弈关系. 当进行多阶段博弈时,声誉起很大的作用,上一阶段的声誉往往影响下一阶段及以后阶段的效用(利润). 现阶段良好的声誉往往意味着未来阶段有较高的效用,因此,某种类型的参与者有可能假装成另一种类型的参与者,建立声誉. 在博弈快结束的时候利用声誉获取更高的效用. 但是,建立声誉是需要成本的,这就涉及到两者的权衡问题. Kreps 等[1]将不完全信息引入重复博弈,建立了著名的 KMRW 声誉模型,解开了有限重复博弈的悖论. Barro[2]和 Vickers[3]使用 KMRW 声誉模型研究了政府是否选择制造通货膨胀的问题. Aoyagi[4]研究了动态 Stackelberg 博弈和声誉的关系. Abreu 和 Gul[5]建立了一个基于讨价还价理论的声誉模型,研究并突出了讨价还价的态度对讨价还价结果的影响. 侯光明等[6]研究了在多阶段动态博弈情况下隐蔽行为的隐性约束机制设计问题. Cole 和 Kehoe[7]将声誉模型推广到一个领域的声誉可以对多个领域产生影响的情形. 以 KMRW 模型为代表的传统声誉模型,一般考虑的是重复博弈,在单阶段进行静态博弈,并没有考虑非完全重复博弈——上一阶段的决策(非后验推断)影响下一阶段的效用函数,如果第 2 阶段参与者不采取积极的行动,这一影响又会消失;一般的声誉模型,在单阶段假设是同时进行博弈,没有考虑在单阶段进行动态博弈(如信号博弈)的情形. 但在现实生活中,这两种情形却较为普遍. 本文建立两阶段基于信号博弈(第一阶段进行信号博弈)的声誉模型,进一步推广 KMRW 声誉模型.

1 模型的基本描述

有两个参与人,进行两时期博弈. 在第一时期,参与人 1 拥有私人信息,有两种类型 $t = L, H$; L 可理解为低能力类型, H 可理解为高能力类型,在不同的具体模型中,有不同的含义. 在第一时期,参与人 2 不知道参与人 1 的类型. 参与人 1 先行动,发送信号 $\lambda_1 \geq 0$,发送信号前预测到参与人 2 将采取什么样的行动,参与人 2 接收到信号后,推断参与人 1 的类型,根据类型采取行动 $q_1 \geq 0$,也可以写出第 2 时期的相应记号.

在第一时期,参与人 1 的效用函数设为

$$U_1(t, \lambda_1, q_1) = \alpha q_1 - b_1(t)\lambda_1^2 - e_1\lambda_1 \quad (1)$$

其中: $\alpha > 0$,等式右侧第 1 项表示参与人 2 采取行

① 收稿日期:2001 - 03 - 13;修订日期:2002 - 08 - 23.
基金项目:国家自然科学基金资助项目(70171028, 70071049, 79830010).
作者简介:肖条军(1973 —),男,湖南隆回人,博士后.

动 q_1 时,参与人 1 的正效用,参与人 2 的行动越大,给参与人 1 带来的正效用越大;$e_1 > 0$, $b_1(t) > 0$, $b_1(L) > b_1(H) > 0$,第二、三项表示发送信号带来的负效用,发送信号越大,带来的负效用越大,L 类型发送信号的负效用大于 H 类型的负效用.

参与人 2 第一时期的效用函数设为

$$V_1(t, \lambda_1, q_1) = c(t)\lambda_1 q_1 - dq_1^2 \tag{2}$$

其中:$0 < c(L) < c(H)$,参与人 2 更偏好于 H 类型的参与人 1,等式右侧第 1 项表示参与人 2 采取行动 q_1 时的正效用;$d > 0$,第 2 项表示参与人 2 采取行动 q_1 时的负效用,边际负效用递增.

在第二时期,参与人 1 的效用函数设为

$$U_2(t, \lambda_1, \lambda_2, q_2) = \alpha q_2 - b_2(t)\lambda_2^2 - f\lambda_2 + e_2\lambda_1 \tag{3}$$

其中:$\alpha > 0$; $b_2(t) > 0$; $b_2(L) > b_2(H) > 0$; $f > 0$; $e_2 > 0$. 右侧最后一项表示第 1 时期发送的信号为第 2 时期带来的正效用,其他解释同上.

2 不考虑声誉时博弈模型的解

由于第一时期的信号大小影响参与人 1 第二时期的效用,因此,两参与者在进行第 1 时期博弈时,必须充分考虑到第二时期各自的最优决策,可以采用逆向归纳法来求解整个博弈.

2.1 第二时期博弈模型的解

由前述可知,在第一时期,两参与人进行信号博弈,信号博弈的均衡结果只包含唯一的分离均衡(参见下述定理 1),参与人 2 推断参与人 1 是 H 类型的后验推断或为 1,或为 0,并且在第二时期按照这一信仰进行行动,因此,第二时期的信息是完全信息. 此时,两参与人之间的博弈是 Stackelberg 博弈,需采用逆向归纳法求解[8].

解参与人 2 在 $i = 2$ 时的效用函数(2) 关于 q_2 的一阶条件,得其最优反应函数

$$q_2(t, \lambda_2) = \frac{1}{2} c(t) d^{-1} \lambda_2 \tag{4}$$

显然二阶条件是满足的,因此,式(4) 是参与人 2 的最优行动反应. 从式(4) 可以看出,参与人 2 的最优行动反应函数是参与人 1 的信号的增函数,即信号越大,最优行动越高.

将式(4) 代入式(3) 后,对 λ_2 求偏导得一阶条件

$$\frac{\partial U_2(t, \lambda_1, \lambda_2, q_2(t, \lambda_2))}{\partial \lambda_2} = \frac{1}{2}\alpha c(t) d^{-1} - f - 2b_2(t)\lambda_2 = 0 \tag{5}$$

很显然,二阶条件满足,因此,一阶条件式(5) 的解是第二时期参与人 1 的最优信号,解一阶条件式(5),可得最优信号为

$$\lambda_2^*(t) = \frac{1}{4}\alpha b_2^{-1}(t) c(t) d^{-1} - \frac{1}{2} b_2^{-1}(t) f \tag{6}$$

假设 1 $\frac{\partial U_2(L, \lambda_1, 0+0, q_2(L, 0+0))}{\partial \lambda_2} > 0$,即 L 类型参与人 1 发送一个小信号比不发送信号带来更高的效用(0 + 0 表示此偏导为在 $\lambda_2 = 0$ 时偏导的右极限,下同).

由于 $0 < c(L) < c(H)$ 以及其他的参数都是正的,所以,可以证明成立不等式

$$0 < \frac{\partial U_2(L, \lambda_1, 0+0, q_2(L, 0+0))}{\partial \lambda_2} < \frac{\partial U_2(H, \lambda_1, 0+0, q_2(H, 0+0))}{\partial \lambda_2}$$

即对 H 类型参与人 1 也成立类似于假设 1 中的不等式,因此,从式(6) 可知,$\lambda_2^*(t) > 0$. 进一步,从上面的描述可以推得 $\lambda_2^*(H) > \lambda_2^*(L)$,即在第二时期,$H$ 类型的参与人 1 发送的最优信号高于 L 类型发送的最优信号.

2.2 第一时期信号博弈的解

在第二时期,信息是不对称的,参与人 1 拥有信息优势,两者之间进行信号博弈. 由于参与人 1 第 1 时期的信号大小影响他第二时期的效用大小,因此,从长远角度考虑,他必须注意到这些影响,记时期贴现因子为 δ, $0 < \delta \leq 1$,则参与人 1 第一时期的决策目标是极大化他的两时期贴现效用

$$U(t, \lambda_1, q_1) = U_1(t, \lambda_1, q_1) + \delta U_2(t, \lambda_1, \lambda_2^*(t), q_2(t, \lambda_2^*(t))) \tag{7}$$

解参与人 2 在 $i = 1$ 时的效用函数(2) 关于 q_1 的一阶条件,得其最优反应函数

$$q_1(t, \lambda_1) = \frac{1}{2} c(t) d^{-1} \lambda_1 \tag{8}$$

很显然,二阶条件是满足的,因此,式(8) 是参与人 2 的最优行动反应函数.

假设 2 $0 < \delta \leq e_1 / e_2$,

$$\frac{\mathrm{d} U_1(L, 0+0, q_1(L, 0+0))}{\mathrm{d} \lambda_1} > -\delta e_2$$

可以证明,对于 H 类型参与人 1,也类似地成立假设 2. 解得式(8) 后,根据信号博弈均衡结果的唯一性定理[9,10] 可以证明定理 1.

定理 1　如果参与人1、2的效用函数分别由式(7)、(2) 确定,且假设 1、2 成立,则两参与人之间满足直观标准的精练贝叶斯均衡结果(ISGPBE) 存在且唯一,唯一的 ISGPBE 是分离均衡.

证明　当假设 1、2 成立时,只要验证效用函数式(7)、(2) 满足文献[9]中的唯一性定理的 6 个假设条件即可. 具体验证类似于文献[10]中有关定理的验证.

因此,可以根据文献[9]中的 ISGPBE 的算法来计算. 首先,计算在完全信息下的最优解. 将式(8) 代入式(7) 后,对 λ_1 求导,得一阶条件

$$\frac{\mathrm{d}U(t,\lambda_1,q_1(t,\lambda_1))}{\mathrm{d}\lambda_1} = \frac{1}{2}\alpha c(t)\,d^{-1} - 2b_1(t)\lambda_1 - e_1 + \delta e_2 = 0 \qquad (9)$$

二阶条件也显然满足,因此,一阶条件式(9) 的解是最优解,解之得

$$\lambda_1^*(t) = \frac{1}{2}b_1^{-1}(t)\left(\frac{1}{2}\alpha c(t)\,d^{-1} - e_1 + \delta e_2\right) \qquad (10)$$

由假设 2 也可以证明 $\lambda_1^*(H) > \lambda_1^*(L) > 0$. 下面假定假设 1 和假设 2 成立. 从上可知,在完全信息下,类型 t 的最优点为 $(\lambda_1^*(t), q_1(t,\lambda_1^*(t)))$. 在不完全信息下,$L$ 类参与人 1 对应的 ISGPBE 就是完全信息下的最优点 $(\lambda_1^*(L), q_1(L,\lambda_1^*(L)))$(此处省略后验推断). 过该点的 L 类参与人 1 的无差异曲线与曲线 $q_1(H,\lambda_1) = \frac{1}{2}c(H)\,d^{-1}\lambda_1$ 的交点方程为

$$U\left(L,\lambda_1,\frac{1}{2}c(H)\,d^{-1}\lambda_1\right) = U\left(L,\lambda_1^*(L),\frac{1}{2}c(L)\,d^{-1}\lambda_1^*(L)\right) \qquad (11)$$

解方程(11) 得最大根

$$\lambda_{1+}^* = \frac{1}{2}b_1^{-1}(L)\left[\frac{1}{2}\alpha c(H)\,d^{-1} - e_1 + \delta e_2 + \sqrt{\left(\frac{1}{2}\alpha c(H)\,d^{-1} - e_1 + \delta e_2\right)^2 - 4b_1(L)E}\right] \qquad (12)$$

其中

$$E = \left(\frac{1}{2}\alpha c(L)\,d^{-1} - e_1 + \delta e_2\right)\lambda_1^*(L) - b_1(L)\lambda_1^{*2}(L)$$

记 $\lambda_1^*(H) = \max\{\lambda_1^*(H), \lambda_{1+}^*\}$,在不完全信息下,省略后验推断,则 H 类参与人 1 对应的 ISGPBE 是 $\left(\lambda_1^*(H), \frac{1}{2}c(H)\,d^{-1}\lambda_1^*(H)\right)$.

3　声誉模型及其解

上面的模型假定 L 类参与人 1 不利用声誉,是诚实的. 但事实上,如果从利用声誉中可以获得利益,那么,L 类参与人 1 有动机在第一时期建立声誉,并且在第二时期利用声誉(H 类型参与人 1 是没有动机的).

当考虑声誉时,如果 L 类参与人 1 在第一时期建立声誉,并且在第二时期利用声誉,则 L 类参与人 1 在第一时期发送信号 $\lambda_1^*(H)$. 在第二时期参与人 2 将参与人 1 错认为是 H 类型的,将采取需求行动 $q_2(H,\lambda_2)$,这时,L 类参与人 1 的效用函数为

$$U_2(L,\lambda_1^*(H),\lambda_2,q_2(H,\lambda_2)) = \frac{1}{2}\alpha c(H)\,d^{-1}\lambda_2 - b_2(L)\lambda_2^2 - f\lambda_2 + e_2\lambda_1^*(H) \qquad (13)$$

求解方法同上,可以得到此时的最优信号为

$$\bar{\lambda}_2^*(L) = \frac{1}{4}\alpha b_2^{-1}(L)\,c(H)\,d^{-1} - \frac{1}{2}b_2^{-1}(L)\,f \qquad (14)$$

由于 $c(H) > c(L) > 0$,公式(6)、(14) 中的其他变量相同且大于 0,因此,$\bar{\lambda}_2^*(L) > \lambda_2^*(L) > 0$,即 L 类参与人 1 在第一时期建立声誉后,其第二时期的最优信号大于不考虑声誉时的最优信号. 于是有下面的命题 1.

命题 1　对于 L 类参与人 1,上一时期的声誉可以支撑下一时期更高的信号.

在第一时期,如果 L 类参与人 1 建立声誉,则他发送与 H 类型同样的信号,这时,他获得的效用为 $U_1(L,\lambda_1^*(H),q_1(H,\lambda^*(H)))$,于是可得两时期的贴现效用为

$$\bar{U}(L,\lambda_1^*(H),q_1(H,\lambda_1^*(H))) = U_1(L,\lambda_1^*(H),q_1(H,\lambda_1^*(H))) + \delta U_2(L,\lambda_1^*(H),\bar{\lambda}_2^*(L),$$

$$q_2(H,\bar{\lambda}_2^*(L)))\tag{15}$$

在第一时期,L 类参与人 1 不建立声誉时,他第一时期的最优效用为 $U_1(L,\lambda_1^*(L),q_1(L,\lambda_1^*(L)))$.

命题 2 如果 L 类参与人 1 在第一时期假装成 H 类参与人 1,且在第二时期利用声誉,则在均衡时,他第一时期的效用小于或等于不考虑声誉时他第一时期的效用,即成立不等式

$$U_1(L,\lambda_1^*(H),q_1(H,\lambda_1^*(H)))\le U_1(L,\lambda_1^*(L),q_1(L,\lambda_1^*(L)))$$

证明 根据信号博弈 ISGPBE 的求解过程可知

$$U(L,\lambda_1^*(H),q_1(H,\lambda_1^*(H)))\le U(L,\lambda_1^*(L),q_1(L,\lambda_1^*(L)))\tag{16}$$

根据式(7),要证明命题 2,只要证明下面的不等式(17)成立即可,

$$U_2(L,\lambda_1^*(L),\lambda_2^*(L),q_2(L,\lambda_2^*(L)))\le U_2(L,\lambda_1^*(H),\lambda_2^*(L),q_2(L,\lambda_2^*(L)))\tag{17}$$

由于 $e_2>0$,从表达式(3)可以看出,式(3)是 λ_1 的严格增函数,又从信号博弈 ISGPBE 的求解过程可知,$0<\lambda_1^*(L)<\lambda_1^*(H)$,因此,可以推得不等式(17)成立,从而命题 2 成立. 类似地可以证明命题 3.

命题 3 如果 L 类参与人 1 在第一时期假装成 H 类参与人 1,且在第二时期利用声誉,则在均衡时,他第二时期的效用大于他第一时期不假装成 H 类型参与人 1 时第 2 时期的效用,即成立.

$$U_2(L,\lambda_1^*(L),\lambda_2^*(L),q_2(L,\lambda_2^*(L)))<U_2(L,\lambda_1^*(H),\bar{\lambda}_2^*(L),q_2(H,\bar{\lambda}_2^*(L)))$$

证明 由 $\bar{\lambda}_2^*(L)$ 的含义可知下面的不等式成立

$$U_2(L,\lambda_1^*(H),\bar{\lambda}_2^*(L),q_2(H,\bar{\lambda}_2^*(L)))\ge U_2(L,\lambda_1^*(H),\lambda_2^*(L),q_2(H,\lambda_2^*(L)))\tag{18}$$

由于 $c(H)>c(L)>0,e_2>0,\lambda_1^*(H)>\lambda_1^*(L)$,由式(3)可以证明

$$U_2(L,\lambda_1^*(H),\lambda_2^*(L),q_2(H,\lambda_2^*(L)))>U_2(L,\lambda_1^*(L),\lambda_2^*(L),q_2(L,\lambda_2^*(L)))\tag{19}$$

综合式(18)和式(19)即可证明命题 3.

由命题 2、3 可知,如果 L 类参与人 1 在第一时期建立声誉且在第二时期利用声誉,他将在第一时期损失一部分效用,但在第二时期将获得更多的效用,理性的 L 类参与人 1 将权衡得失,如果所得大于所失,则在第一时期建立声誉,相反,则不建立声誉. 具体地说,只要比较 $\bar{U}(L,\tilde{\lambda}_1^*(H),q_1(H,\tilde{\lambda}_1^*(H)))$ 和 $U(L,\lambda_1^*(L),q_1(L,\lambda_1^*(L)))$,哪种情况下的贴现效用大,则选择哪种情况.

4 声誉模型的算例

上面建立了两时期的信号博弈模型,并且在此基础上,建立了基于信号博弈的声誉模型,求出了模型的解,下面举例说明如何求解.

有两个参与人,进行两时期考虑声誉的博弈,设各变量取值: $\alpha_1=10,b_1(L)=2,b_1(H)=1.8,c(L)=5,c(H)=6,d=2,e_1=3,\alpha_2=9,b_2(L)=2.2,b_2(H)=2,e_2=2.5,f=2,\delta=1$.

在计算过程中精确到 0.001,据本文公式计算可得

$$\frac{\partial U_2(L,\lambda_1,0+0,q_2(L,0+0))}{\partial\lambda_2}=9.25>0$$

即假设 1 成立;$\lambda_2^*(L)=2.102$,在第二时期,H 类参与人 1 发送信号 $\lambda_2^*(H)=2.875$,而参与人 2 采取行动 $q_2(H,\lambda_2^*(H))=4.313$;由于 $e_1/e_2=1.2>1=\delta$, $\frac{\mathrm{d}U_1(L,0+0,q_1(L,0+0))}{\mathrm{d}\lambda_1}=9.5>-2.5=-\delta e_2$,所以假设 2 成立,因此,可以进一步计算;$\lambda_1^*(L)=3.000,\lambda_1^*(H)=4.028,\lambda_{1+}^*=5.660$,所以,在第一时期,$H$ 类参与人 1 发送信号 $\lambda_1^*(H)=\max\{\lambda_1^*(H),\lambda_{1+}^*\}=5.660=\lambda_{1+}^*$,而参与人 2 采取行动 $q_1(H,\lambda_1^*(H))=8.490$;建立并利用声誉时,$L$ 类参与人 1 第二时期的最优信号 $\bar{\lambda}_2^*(L)=2.614$,比较建立并利用声誉与不建立声誉时的两时期总贴现效用大小有 $\bar{U}(L,\lambda_1^*(H),q_1(H,\lambda_1^*(H)))=33.027>27.723=$

$U(L, \lambda_1^*(L), q_1(L, \lambda_1^*(L)))$,于是参与人 1 选择建立声誉;$L$ 类参与人 1 第一、二时期发送的信号大小分别是 5.660,2.614,参与人 2 第一、二时期的行动分别为 8.490,$q_2(H, \bar{\lambda}_2^*(L)) = 3.921$.

5　结束语

本文建立了一个基于第一时期进行信号博弈的两时期声誉模型,在单时期本身进行动态博弈,与简单的重复博弈不同,这里第二时期的决策影响到第二时期的效用,以至影响到第二时期的决策. 本文也可以用来研究经济问题,但由于篇幅较大,只好精简成该种形式出现,具体经济中的声誉模型参见文献[10];信号博弈的其他应用研究参见文献[11,12]. 该模型也可以扩展到一个场地的声誉影响到多个场地的声誉的情形.

参 考 文 献:

[1] Kreps D, Milgrom P, Roberts J, *et al*. Rational cooperation in the finitely repeated prisoners' dilemma [J]. Journal of Economic Theory, 1982, 27: 245—252

[2] Barro R. Reputation in a model of monetary policy with incomplete information[J]. Journal of Monetary Economics, 1986, 17: 3—20

[3] Vickers J. Signaling in a model of monetary policy with incomplete information [J]. Oxford Economic Papers, 1986, 38: 443—455

[4] Aoyagi N. Reputation and dynamic Stackelberg leadership in infinitely repeated games [J]. Journal of Economic Theory, 1996, 71: 378—393

[5] Abreu D, Gul F. Bargaining and reputation [J]. Econometrica, 2000, 68(1): 85—117

[6] Hou Guangming, Jin Jun, Gan Renchu. The reputation model of multi-stage dynamic game [J]. Journal of Beijing Institute of Technology, 1999, 8(1): 1—6

[7] Cole H L, Kehoe P J. Models of sovereign debt: Partial versus general reputations [J]. International Economic Review, 1998, 39(1): 55—70

[8] 盛昭瀚. 主从递阶决策论 ——Stackelberg 问题[M]. 北京:科学出版社,1998

[9] Yu Gang, Sheng Zhaohan, Xiao Tiaojun. An effective algorithm for computing equilibrium outcome of a class of signaling games [J]. International Journal of Information Technology & Decision Making, 2002, 1(2): 209—228

[10] 肖条军. 纵向型企业集团的 R&D 及其经济增长的信号博弈分析[D]. 南京:东南大学,2000

[11] 肖条军,盛昭瀚. 纵向型企业集团 R&D 及其经济增长的博弈分析[J]. 管理科学学报,2002,5(4):1—6

[12] 肖条军,盛昭瀚. 多级企业集团 R&D 决策的信号博弈分析[J]. 系统工程学报,2002,17(1):93—96

Two-period reputation model based on signaling game

XIAO Tiao-jun, SHENG Zhao-han

Graduate School of Management Science & Engineering, Nanjing University, Nanjing 210093, China

Abstract: This paper sets up a two-period reputation model based on signaling game. There are two players in the game, the one is sender with private information, and the other is receiver without private information. The reputation model studies the type L senders to see whether they have motive to set up reputation in period one or not. We show that the optimal signal of the type L sender is larger in the period two if he sets up reputation in the period one. The utility of the type L sender is less in the period one if he sets up reputation in the period one, but he will get the higher utility in the period two.

Key words: reputation; signaling game; separating equilibrium