

团队生产方式下医疗服务支付方式 与医患间最优委托权配置*

刘自敏 张昕竹

【摘要】 医疗服务结果是医患共同努力程度的表现。通过比较实现社会福利最优的团队合作方式以及医患效用最大化的非合作博弈方式的委托权配置,表明在不存在监督能力时采用后付费方式必然会带来供方诱导需求,在不存在监督能力时采用预付费方式会带来逆向选择问题;当引入监督能力后,预付费方式在一定条件下能够实现委托权的最优配置,并同时提高社会福利与患者的效用。

【关键词】 团队生产 医疗支付方式 委托权 供方诱导需求

【中图分类号】 F062.5 **【文献标识码】** A **【文章编号】** 1000-2952(2012)05-0052-10

一、引言

在传统的信息经济学定义中,关于委托人与代理人的界定是非常简单的。如不掌握私人信息的一方被称为委托人(Principal),掌握私人信息的一方是代理人(Agent)(张维迎,1995^①;陈钊,2009^②);主动缔约方称为委托人,受邀缔约方称为代理人(Stadler&Castrillo,1997^③)。这在诸多的经济学现象中是适用的,如公司与销售人员关于销售合同的制定、投资者与经理关于经理人员收入结构的合同制定等。

但在卫生经济学领域,尤其是在医患关系间,按照是否掌握私人信息,将病人称为委托人,医生称为代理人,却存在着与标准的委托—代理理论不符合的诸多现象与结果。主要表现为以下三个方面:

(1) 在标准的委托—代理理论中,研究假定是委托人风险中性、代理人风险规避,但在医患关系中,很明显,作为委托人的病人是风

险规避的,且其风险规避度很大;而作为代理人的医生则可能是风险规避或风险中性的。

(2) 张维迎(1996)^④认为,委托人之所以是委托人是由于他必须对代理人的行为后果承担责任。但现实却是作为委托人的病人极度害怕对代理人行为承担责任,因为代理人的行为后果与委托人的身体健康直接相关。

(3) 在传统的委托—代理理论中,委托人不能

* 基金项目:国家自然科学基金项目“利用非线性定价促进能源节约的基础理论与实证研究”(批准号71173236)。

① 张维迎:《博弈论与信息经济学》,上海人民出版社1995年版,第20页。

② 陈钊:《信息与激励经济学(第二版)》,上海人民出版社2009年版,第108页。

③ Inés Macho-Stadler, J. David Pérez-Castrillo, Optimal Auditing with Heterogeneous Income Sources, International Economic Review, 1997, 38(4): 951-968.

④ 张维迎:《企业的企业家:契约理论》,上海人民出版社1996年版,第56页。

直接参与生产，他只是通过委托权来获得代理人生产的收益，但是在大量的生产过程中，团队生产是一种普遍形式，在医疗服务过程中，医疗服务的良好结果是医生和患者共同努力才能达到的。如果医患有一方不配合，那么要达到良好的治疗效果是不可能的。

我们首先对委托权进行界定。剩余索取权是委托权的核心内容。Grossman&Hart (1986)^① 将委托权定义为控制权，Knight (1921)^② 认为，委托权是剩余索取权和控制权的统一。根据张维迎 (1995)、Harris&Raviv (1991)^③ 以及 Dewatripont &Tirole (1994)^④ 的定义，亦将委托权定义为剩余索取权和控制权的统一。Fama&Jensen (1983)^⑤ 将剩余索取权定义为对在契约中事前不能明确规定的那部分索取收益的权利，以及在契约中事前不能明确规定的那部分控制权的权力。而控制权则为委托人要求代理人干代理人不愿意自愿干某些事情的权力。张维迎 (1996) 认为，委托权安排之所以重要，是因为它能够影响代理人的行动。Coase (1937)^⑥ 指出，雇主—雇员（委托—代理）关系的显著特征就是雇主可以指示雇员去干什么。从这点看，医患关系中，医生具有很强的影响病人的权力，如专业诊断、过度医疗、诱导需求等，而病人却与之相反，几乎不具备任何与医生谈判的能力，毫无控制权可言（在存在第三方支付机构，如医疗保险机构或政府时，该情况可能会改变）。从这个角度来看，医生也更像委托人，病人更像代理人。

根据张维迎 (1995, 1996) 的模型，团队成员在生产上的相对重要性和监督上的相对有效性是决定最优委托权安排的决定因素。根据这个结论，在医患关系中，医生的行为明显是重要的且难于监督（尤其是不存在第三方监督时）。我们将结合医疗服务市场的具体情况，在此模型基础上分析医患间的委托权配置，并试图回答在医患关系中，最优的配置是谁作为委托人？谁作为代理人？或委托权应该如何医患之间进行合理的分配？哪些因素影响医患间委托权的配置？

二、不存在监督能力下的最优委托权

我们首先分析不存在监督能力时的最优委

托权配置。委托权中的一个重要部分即控制权，指委托人要求代理人做出比没有监督时更大的工作努力的权力。但对单个病人来说，即使他拥有委托权，由于医疗服务的特殊性和专业性，他也很难去监督医生的努力。因此，不存在监督能力时，委托权退化为仅包括剩余索取权。

我们考虑在不同收费方式下最优委托权是否有变化。收费方式大类可分为预付制与后付制以及两者的混合制，预付制包括总额预付 (Global Budget)、按病种收费 (DRG)、按人头付费 (Capitation)；后付制包括按服务项目付费 (Fee-For-Service)、按服务单元付费 (Service Unit)。我们分别选取后付费中最常见的按项目付费方式，预付费中最常见的按人头付费方式来分析后付费和预付费模式下在最优委托权分配中的差异。

1. 后付费下的最优委托权配置

后付制指医疗费用实际发生后，病人以实际发生的医疗费用为基础向医疗机构或医生支付医疗费用的方式。我国长期以来采用的就是这种支付制度，最典型的的就是按项目付费模式。

Alchian&Demsetz (1972)^⑦ 认为，在团队生产中，产出是所有成员共同努力的结果。在他们的模型中，团队成员均为代理人，考察的是代理人团队中的成员个人努力和团队产出之间的关系。但在本文中，由于委托权的配置，

① Sanford J. Grossman and Oliver D. Hart, The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration, *Journal of Political Economy*, 1986, 94 (4): 691-719.

② F. H. Knight, *Risk, Uncertainty and Profit*, Boston: Houghton Mifflin, 1921.

③ Milton Harris, Artur Raviv, The Theory of Capital Structure, *The Journal of Finance*, 1991, 46 (1): 297-355.

④ Mathias Dewatripont, Jean Tirole, A Theory of Debt and Equity: Diversity of Securities and Manager-Shareholder Congruence, *The Quarterly Journal of Economics*, 1994, 109 (4): 1027-1054.

⑤ Eugene F. Fama, Michael C. Jensen, Agency Problems and Residual Claims, *Journal of Law and Economics*, 1983, 26 (2): 327-349.

⑥ R. H. Coase, *The Nature of the Firm*, *Economica*, 1937.

⑦ Alchian A. A., Demsetz H., *Production, Information Costs and Economics of Organization*, *American Economic Review*, 1972, 62: 777-795.

医生与患者均有可能称为委托人或代理人，因此我们考察医生与患者组成的团队。在医生与患者组成的团队中，产出（治愈疾病或者提供健康）是由医生与患者共同努力完成的。假定一个团队由医生（Doctor）与患者（Patient）两个成员组成，为集中讨论激励问题，假定成员都是风险中性的，^① 医生的冯·诺依曼-摩根斯坦因（von Neumann-Morgenstern）效用函数为： $U_D(y_D, e_D) = y_D - C_D(e_D) = y_D - \frac{e_D^2}{2}$ ， y_D 是医生收取的医疗费用， e_D 是医生的努力程度， $C_D(e_D) = \frac{1}{2}e_D^2$ 满足 $C'(e) > 0, C''(e) > 0$ 。在按项目付费且不存在医疗保险的情况下，病人的效用函数为 $U_P(y_P, e_P) = Y - y_D - \frac{1}{2}e_P^2$ 。 e_P 是病人的努力程度，如病人配合医生治疗的付出等。

令 Y 为医生与病人联合努力的产出（如治愈或健康的价值），其生产函数为： $Y = f(e_D, e_P) + \varepsilon = e_D e_P^{1-\alpha} + \varepsilon$ ，此处 ε 是均值为 0 的随机变量，期望产出为 $E(Y) = E(f(e_D, e_P) + \varepsilon) = e_D e_P^{1-\alpha}$ ，柯布-道格拉斯函数^②反映了医患共同努力提高病人健康的特征，此生产函数刻画出团队工作的特征，即每个成员努力的边际生产率 $\frac{\partial Y}{\partial e_i}$ 都是另一个成员努力（投入）的递增函数。将产出对努力的弹性系数 α 与 $1-\alpha$ 解释为医患在生产中各自的重要性（ $0 < \alpha \leq 1$ ）。医疗服务一般可分为预防性医疗服务（preventive care），包括常规体检、疫苗接种、健康教育与健康档案等；治疗性医疗服务（curative care），包括门诊就医、急诊、住院医疗等。显然，在一般的疾病治疗过程中， $\alpha > \frac{1}{2}$ ，在某些预防医学实践或传染病诊治过程中， α 可能小于 $\frac{1}{2}$ 。

产出的分配采取如下线性形式：

$$y_D = w_D + \beta(Y - w_D - w_P) = (1-\beta)w_D + \beta(Y - w_P)$$

$$y_P = Y - y_D = w_P + (1-\beta)(Y - w_D - w_P) = \beta w_P + (1-\beta)(Y - w_D)$$

w_D 是医生获得的最低固定工资， w_P 是病人

获得的最低努力治疗的货币形式。 β 与 $1-\beta$ 是医生与病人的收入分配比例。在完全的按服务付费模式中， $w_D = w_P = 0$ ，但更现实的情况是 $w_D = w_P > 0$ 。 w_D 是医生由于必须付出的最低努力 \bar{e}_D （如仅满足按时坐班）获得的固定工资， w_P 是病人获得的最低努力治疗的货币形式。

我们将 β 等价于委托权安排。 $\beta = 1$ 即医生是委托人，病人是代理人； $\beta = 0$ 即病人是委托人，医生是代理人； $0 < \beta < 1$ 即医生与病人互为委托人和代理人。最优的 β 即为最优的委托权安排。与张维迎（1995，1996）不同的是，我们的模型中由于存在最低努力 \bar{e}_D 及 \bar{e}_P ，即使某一方未获得委托权，产出 Y 仍然为正，即将委托权完全分配给某一方也可能是最优选择，从而可以将 $\beta = 0$ 与 $\beta = 1$ 以及 $0 < \beta < 1$ 时的不同状态进行比较。

假定固定工资 w_D 与 w_P 不影响努力水平 e_i ，最优的 β 是最大化总福利 $W(\beta)$ 。 $W(\beta) = U_D(y_D, e_D) + U_P(y_P, e_P) = EY - C_D(e_D) - C_P(e_P)$ 。委托权安排的最优化问题为：

$$\max W(\beta) = e_D e_P^{1-\alpha} - \frac{1}{2}e_D^2 - \frac{1}{2}e_P^2 \quad (1)$$

$$s. t. (1) e_D \in \arg \max ((1-\beta)w_D +$$

$$\beta(e_D e_P^{1-\alpha} - w_P) - \frac{1}{2}e_D^2) \quad (2)$$

$$s. t. e_D \geq \bar{e}_D$$

$$(2) e_P \in \arg \max (\beta w_P + (1-\beta)$$

$$(e_D e_P^{1-\alpha} - w_D) - \frac{1}{2}e_P^2) \quad (3)$$

$$s. t. e_P \geq \bar{e}_P$$

(1) 模型求解

两个约束条件要实现各自的最大化效用，此时实现的是医患的非合作博弈解，即在给定契约下，医患分别确定自己的最佳努力水平，最大化各自的效用水平。

求解 (2)、(3) 式，其纳什均衡为：

^① 张维迎（1995）认为，风险规避度与相对重要性和监督的相对有效性相互作用于委托权安排，但即使传统意义上的委托人比代理人更害怕风险，如果委托人的相对重要性足够大，对他监督非常困难，则让委托人拥有委托权仍然优于让代理人拥有委托权。

^② 可以使用更一般的 CES（不变替代弹性）函数，但结论不会改变。

$$e_D^{NE} = \max \{ (\beta\alpha)^{\frac{1-\alpha}{2}} [(1-\beta)(1-\alpha)]^{\frac{1-\alpha}{2}}, \bar{e}_D \} \quad (4)$$

$$e_P^{NE} = \max \{ (\beta\alpha)^{\frac{\alpha}{2}} [(1-\beta)(1-\alpha)]^{\frac{\alpha}{2}}, \bar{e}_P \} \quad (5)$$

我们重点讨论各自努力水平超过最低努力时的情况^①，将(4)、(5)式对 β 求导，得医生的努力 e_D^{NE} 在 $\beta = \frac{1+\alpha}{2}$ 处达到最大值，病人的努力 e_P^{NE} 在 $\beta = \frac{\alpha}{2}$ （即 $1-\beta = \frac{1+(1-\alpha)}{2}$ ）处达到最大值。即每个成员纳什均衡努力达到的剩余控制权份额只与该成员的相对重要性正相关。容易得到，二阶条件满足。

同时，医生效用最大化的最优委托权 $\beta_D = \frac{1+\alpha}{2}$ 以及病人效用最大化的最优委托权 $\beta_P = \frac{\alpha}{2}$ 之差为 $\beta_D - \beta_P = \frac{1}{2}$ ，病人出让给医生的委托权份额总比医生期望获得的最优份额要少。如果不能有效地对医生进行监督，超过的部分就成为产生供方（医方）诱导需求的剩余控制权收益。当然，很多时候医生并不能获得全部的供方（医方）诱导需求的剩余控制权收益，这取决于双方的谈判能力（如现实中病人也可能存在监督能力），因此，现实的医方的剩余控制权 $\beta_P < \beta_{Real} < \beta_D$ 。

而求解(1)式中社会福利函数 $W(\beta)$ 的最大值所表达的是一个团队合作问题，即决定一个契约 $\{e_D, e_P, \beta\}$ ，使其最大化社会福利，其解为：

$$\beta^* = \frac{(\alpha+\alpha^2) - \sqrt{(\alpha+\alpha^2)(2-\alpha)(1-\alpha)}}{2(2\alpha-1)}, \alpha \neq \frac{1}{2} \text{ 时}; \quad (6)$$

$$\beta^* = \frac{1}{2}, \alpha = \frac{1}{2} \text{ 时}$$

由(6)式，可得 $\frac{\partial \beta^*}{\partial \alpha} > 0$ 。因此，当病人患有自己不熟悉的疾病（ $\alpha > \frac{1}{2}$ ）或患有严重的大病重病时（ $\alpha \rightarrow 1$ ），医生所能掌握的委托权非常大（ $\beta^* \rightarrow 1$ ），其现实表现即为剩余索取权非常大，这就能够解释诱导需求在一些复杂疾病中发生的概率和程度更大，而在一些小病或者预防

性治疗中（ $\alpha < \frac{1}{2}$ ），诱导需求的量则要小很多。

(2) 供方诱导需求的必然性

命题1：当不存在监督能力时，在后付费制度下，追求收益最大化的医生必然导致供方诱导需求。

从前面可知，医生的努力 e_D^{NE} 在 $\beta_D = \frac{1+\alpha}{2}$ 处达到医生效用的最大值，而社会总福利最大时的最优 β 为 β^* ，要使医生的努力刚好能够实现社会总福利最大，需要 $\beta_D = \beta^*$ ，即：

$$\frac{1+\alpha}{2} = \frac{(\alpha+\alpha^2) - \sqrt{(\alpha+\alpha^2)(2-\alpha)(1-\alpha)}}{2(2\alpha-1)},$$

该式当且仅当 $\alpha = 1$ 时成立。

此时 $\beta_D = \beta^* = 1$ ，医生拥有所有的剩余索取权，即医生实行了完全价格歧视，榨取了病人所有的消费者剩余。病人仅仅获取了最低的保留效用，福利无任何改进。

而病人的努力 e_P^{NE} 在 $\beta_P = \frac{\alpha}{2}$ 达到病人效用的最大值，而社会总福利最大时的最优 β 为 β^* ，要使病人的努力刚好能够实现社会总福利最大，需要 $\beta_P = \beta^*$ ，即：

$$\frac{\alpha}{2} = \frac{(\alpha+\alpha^2) - \sqrt{(\alpha+\alpha^2)(2-\alpha)(1-\alpha)}}{2(2\alpha-1)},$$

该式当且仅当 $\alpha = 0$ 时成立，与现实不符（ $0 < \alpha \leq 1$ ）。

因此，在按项目付费模式下，当 $0 < \alpha < 1$ 时， $\beta_D^{NE} - \beta^* > 0$ 成立。因为 $\frac{\partial \beta^*}{\partial \alpha} > 0$ ，可以看到，此时为了最大化个人效用，医生掌握了过多的委托权，付出了过多的努力，从而带来供方诱导需求。因为 $\frac{\partial U_D}{\partial \beta} = e_D e_P^{1-\alpha} - W_D - W_P > 0$ ，所以诱导需求给医生增加效用，增加的效用为 $U_D(\beta_D^{NE}) - U_D(\beta^*)$ 。但此时带来社会福利的损失，社会福利的损失为 $W(\beta_D^{NE}) - W(\beta^*)$ ，因为 $\frac{\partial U_P}{\partial \beta} = -(e_D e_P^{1-\alpha} - W_D - W_P) < 0$ ，因此 $W(\beta_D^{NE}) - W(\beta^*) = [U_D(\beta_D^{NE}) + U_P(\beta_D^{NE})] -$

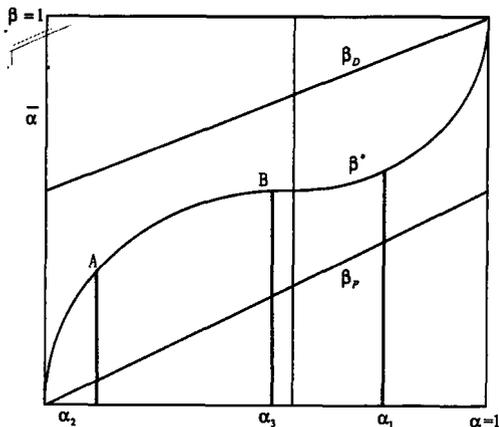
① 我们会在本部分第2节讨论采用最低努力时的情况，为避免重复，此处这种情况略去。

$[U_D(\beta^*) + U_P(\beta^*)] = [U_D(\beta_D^E) - U_D(\beta^*)] + [U_P(\beta_D^E) - U_P(\beta^*)] < U_D(\beta_D^E) - U_D(\beta^*)$ ，即病人的福利被医生侵占。

同时，当 $0 < \alpha < 1$ 时， $\beta_D^E - \beta^* < 0$ ，此时病人努力不足，其原因是由于信息不对称，病人不知道如何施展努力，只能被动接受医生建议。病人努力不足给自己带来的效用损失为： $U(\beta_D^E) - U(\beta^*)$ ，同时带来社会福利损失为 $W(\beta_D^E) - W(\beta^*)$ ，且 $W(\beta_D^E) - W(\beta^*) = [U_D(\beta_D^E) + U_P(\beta_D^E)] - [U_D(\beta^*) + U_P(\beta^*)] = [U_D(\beta_D^E) - U_D(\beta^*)] + [U_P(\beta_D^E) - U_P(\beta^*)] > U(\beta_D^E) - U(\beta^*)$ 。即病人的福利仍然被医生侵占。

在图 1 可以看到三种状态 $\beta^* = \frac{(\alpha + \alpha^2) - \sqrt{(\alpha + \alpha^2)(2 - \alpha)(1 - \alpha)}}{2(2\alpha - 1)}$ ； $\beta_D = \frac{1 + \alpha}{2}$ ； $\beta_P = \frac{\alpha}{2}$ 下的努力水平与委托权的关系。

图 1 实现不同目标的最优努力水平与委托权关系图



对 $\beta_D - \beta^*$ (或 $\beta_P - \beta^*$) 求导，可得 α_1 点。从图 1 中可以看到， $\beta_P < \beta^* < \beta_D$ ，即 $\beta^* < \beta_D$ ， $1 - \beta^* < 1 - \beta_P$ ，使得医患某一方期望收入达到最大的剩余索取权大于使得社会福利最优的剩余索取权分配，即非合作博弈解是劣于团队合作解的。

当疾病为一般性治疗疾病时 ($\alpha \geq \frac{1}{2}$)，在 α_1 点处，医生的努力 e_D^E 与社会总福利最大时的最优的 β^* 之间差距最大，此时诱导需求最强，带来的福利损失最大。当疾病为预防性疾病或传染病时 ($\alpha < \frac{1}{2}$)，从图 1 中可以看出，由于存在着固定工资 w_D ，即医生至少要承担部分重要性

(责任) $\bar{\alpha} > 0$ 的限制，因此若 $\bar{\alpha} < \alpha_2$ 时，最优的诱导需求努力点为 $\bar{\alpha}$ 处。若 $\bar{\alpha} > \alpha_2$ 时，医生不能在 $[0, \alpha_2]$ 内寻求最大的诱导需求量。此时医生转向 $[\bar{\alpha}, \frac{1}{2})$ 内寻找最优的诱导需求点，通过 A 点且斜率与 β^* 切线相同 (为 $\frac{1}{2}$) 的直线与 β^* 曲线交与 B 点。可以看到，此时医生寻求的最优诱导需求点位于 $[\alpha_3, \frac{1}{2})$ ，且接近于 $\frac{1}{2}$ ，即医生总是强调自己在治疗过程中的重要性。同时也可以看到， $\alpha < \frac{1}{2}$ 时的诱导需求量小于 $\alpha > \frac{1}{2}$ 时，这也与现实情况相符合，一般来说，小病治疗过程中医生开“大处方”的难度比急病重病要难得多。

此时医生诱导需求所获得的效用为 $U(\beta_D) - U(\beta^*)$ ，我们可以采用一个无量纲的相对指标——诱导需求度来测量。

诱导需求度 (Supplier-Induced Demand Degree) = $\frac{U_D(\beta_D) - U_D(\beta^*)}{U_D(\beta^*)} = \frac{U_D(\beta_D)}{U_D(\beta^*)} - 1$

(3) 算例

我们使用具体数据进行算例分析，在后付费的情况下，我们需要设定的数据为医患各自的重要性 α ，同时由于基本固定工资 (收益) W_D, W_P 的存在，需要对其进行假定。我们分别分析在 $W_D = W_P = 0.05$ ， $\alpha = 0.3$ 、 $\alpha = 0.4$ 、 $\alpha = 0.7$ 及 $\alpha = 0.8$ 时算例的分析结果，见表 1 和表 2。

表 1 $\alpha < \frac{1}{2}$ 时后付费下的福利及效用比较

α	$\alpha = 0.3$			$\alpha = 0.4$		
	$\beta^* =$	$\beta_D =$	$\beta_P =$	$\beta^* =$	$\beta_D =$	$\beta_P =$
β	0.364	0.65	0.15	0.433	0.7	0.2
e_D^E	0.179	0.211	0.111	0.212	0.245	0.137
e_P^E	0.361	0.237	0.404	0.297	0.197	0.335
$W(\beta)$	0.211	0.179	0.187	0.193	0.165	0.169
Y	0.292	0.229	0.274	0.260	0.215	0.234
U_D	0.104	0.111	0.070	0.097	0.100	0.068
U_P	0.107	0.067	0.117	0.096	0.065	0.101
诱导需求度	$\frac{U_D(\beta_D) - U_D(\beta^*)}{U_D(\beta^*)} = 7.11\%$			$\frac{U_D(\beta_D) - U_D(\beta^*)}{U_D(\beta^*)} = 3.72\%$		

表 2 $\alpha > \frac{1}{2}$ 时后付费下的福利及效用比较

α	$\alpha = 0.7$			$\alpha = 0.8$		
	$\beta^* =$	$\beta_D =$	$\beta_P =$	$\beta^* =$	$\beta_D =$	$\beta_P =$
β	0.636	0.85	0.35	0.710	0.9	0.4
e_D^E	0.361	0.404	0.237	0.452	0.503	0.290
e_P^E	0.179	0.111	0.211	0.145	0.084	0.178
$W(\beta)$	0.211	0.187	0.179	0.247	0.222	0.205
Y	0.292	0.274	0.229	0.360	0.352	0.263
U_D	0.107	0.117	0.067	0.132	0.150	0.073
U_P	0.104	0.070	0.111	0.115	0.072	0.132
诱导需求 求度	$\frac{U_D(\beta_D) - U_D(\beta^*)}{U_D(\beta^*)} = 8.77\%$			$\frac{U_D(\beta_D) - U_D(\beta^*)}{U_D(\beta^*)} = 13.3\%$		

从以上不同的医患重要性 α 的算例结构可以看出，当 $\alpha < \frac{1}{2}$ 时，医生的诱导需求能力相对较弱，诱导需求度较低，而 $\alpha > \frac{1}{2}$ 时，医生的重要性上升后，医生的诱导需求度大大提高，而大量的疾病治疗正属于这种情况。因此，如果缺乏有效的监管手段，很难阻止医生诱导需求的发生。有学者根据中国健康与营养调查 (CHNS) 数据实证计算出，我国医疗服务市场中达成的医疗服务价格相对于公正的基准价格要高出 26% 左右。即相当于计算出的诱导需求度为 26%。可见当前我国医疗服务市场的供方诱导需求程度已经极度严重。

(4) 社会福利与患者效用同时优化的两难

那么能否实现产出最大且实现社会福利的最优呢？答案是肯定的。 $E(Y) = E(f(e_D, e_P) + \epsilon) = e_D e_P^{1-\alpha}$ ，将纳什均衡下的努力水平带入其中，得：

$$Y = f(e_D, e_P) = e_D e_P^{1-\alpha} = (\beta\alpha)^{\alpha} [(1-\beta)(1-\alpha)]^{1-\alpha}$$

对 α 求导，得 $\alpha + \beta = 1$

对 β 求导，得 $\alpha = \beta$

因此， $\alpha = \beta = \frac{1}{2}$ 时，可以实现产出 Y 最大

化，又实现社会福利的最优，但却不能实现医生与患者双方效用的最大化。因为此时并没有考虑医患双方的（努力）成本，很明显，为了追求产出最大与社会福利最优，此时过度的成

本被投入了。这在某种程度上解释了世界各国医疗费用不断攀高的原因，以美国为例，存在因过度追求治疗效果与延长患者生命而大量采用新的治疗技术以及孤儿药的研发等情况。据 1995 年对 50 名经济学家进行的一项调查，81% 的经济学家同意这样一个观点：“过去 30 年，医疗卫生事业占 GDP 份额不断上升的首要原因是医疗技术的变化” (Iglehart, 1999)。^① 保险方也认为，近年来保险费用每年都在增长，其中 40% 缘于新技术的使用 (Shenkin, 1994)。^② 2009 年，美国的医疗费用占 GDP 的比例达到了 17.6%，约为 25000 亿美元。

并且，在现实中要达到医患双方成员的相对重要性 α 、委托权的安排 β 同时为 $\frac{1}{2}$ ，是需要做出诸多努力的，比如引入第三方支付机构（医疗保险），增加患者的谈判能力和重要性，同时增加病人的努力程度；改变医生的收入模式，减少医生的过度努力。

那么能否同时实现个人效用最大与社会福利最优呢？在按服务收费的模式下，答案是否定的，前面我们已经证明，因为三者的委托权最优无一重合。根据 Holmstrom (1982) 的结论，通过打破预算平衡的方法可以实现个人收益最大与社会福利最优，方法是把原来的医患收入约束 $\sum_{i=D,P} y_i = Y$ 改为 $\sum_{i=D,P} y_i < Y$ ，并将收入设置为在医患双方在 β^* 时的努力 e_i^* ($i = D, P$) 下，收入为 $W(\beta_i^*)$ ($i = D, P$)，偏离这一努力则为 0，但此时实行的已经不是按服务收费的模式了。

2. 预付费下的最优委托权配置

预付制指在医疗费用发生之前，按一定的标准将医疗费用预先支付给医疗机构或医生。支付标准在一定时期内是固定的，一段时期后按实际情况的变化再做出相应的调整。按人头付费模式是一种典型的预付费支付方式。

① J. K. Iglehart, Expenditures, New England Journal of Medicine, 1999, 340: 70-76.

② Peter S. Shenkin, D. Quentin McDonald, Cluster Analysis of Molecular Conformations, Journal of Computational Chemistry, 1994, 15 (8): 899-916.

在按人头付费的模式下,医生与病人的支付函数发生了变化,医生的效用函数为 $U_D(y_D, e_D) = \bar{y}_D - C_D(e_D) = \bar{y}_D - \frac{1}{2}e_D^2$, 病人的效用函数为 $U_P(y_P, e_P) = Y - \bar{y}_D - \frac{1}{2}e_P^2$ 。 \bar{y}_D 为按人头付费的且与医生努力无直接关系的支出水平。生产函数仍然为: $Y = f(e_D, e_P) + \varepsilon = e_D^\alpha e_P^{1-\alpha} + \varepsilon$ 。产出的分配为:

$$y_D = w_D + \bar{y}_D = (w_D + \bar{y}_D) +$$

$$0 * (Y - w_D - w_P - \bar{y}_D);$$

$$y_P = w_P + (Y - w_D - w_P - \bar{y}_D)$$

可以看到,医生与病人的收入与剩余索取权 β 无关 ($\beta=0$),而与按人头支付的标准 \bar{y}_D 有关系。从整个医院或行业角度来看, \bar{y}_D 受到 β 的间接影响,但对医患这个小团队来说, \bar{y}_D 是外生决定的。

命题2:当不存在监督能力时,在预付费制度下,追求收益最大化的医生必然导致逆向选择,追求最低努力。

最优化问题为:

$$\max W(\beta) = e_D^\alpha e_P^{1-\alpha} - \frac{1}{2}e_D^2 - \frac{1}{2}e_P^2$$

$$s. t. (1) e_D \in \arg \max (w_D + \bar{y}_D - \frac{1}{2}e_D^2)$$

$$s. t. e_D \geq \bar{e}_D$$

$$(2) e_P \in \arg \max (e_D^\alpha e_P^{1-\alpha} - w_D - w_P - \bar{y}_D - \frac{1}{2}e_P^2)$$

$$s. t. e_P \geq \bar{e}_P$$

解上式,得到纳什均衡下的努力水平为:

$$e_D^* = \bar{e}_D;$$

$$e_P^* = \max \{ (1-\alpha)^{\frac{1}{1-\alpha}} \bar{e}_P^\alpha, \bar{e}_P \}$$

因此,在按人头付费模式下,委托权完全属于病人所有,此时医生的努力总是维持在最低水平,因为他无法分享到直接努力的成果,此时医生是明显的代理人角色,此时几乎不会发生道德风险问题,因此不会发生诱导需求,反而会发生医疗服务供给不足,因为付出的努力越小,医生的效用越大,此时存在严重的逆

向选择问题,医生会选择付出最少努力能够治愈而支付人头费的患者。

在按人头付费时,病人获取所有的剩余控制权,但由于医生的努力不足,在医生的重要性 $\alpha > 1/2$ 时,医生不努力带来的产出 Y 很小,病人明显无法获取更多的剩余控制权收益,在这种情况下,只有加强对医生的监督,从而增加医生的努力程度,病人作为委托人才可能获取更多的剩余控制权收益,即委托权收益。

3. 小结

通过比较后付费(按项目付费)与预付费(按人头付费)两种不同的支付方式,可以得出,在不存在监督能力的情况下,预付费必然导致追求个人效用最大化的理性的医生采取诱导需求行为,这导致了医生的努力程度超过了社会福利最大化时的最优努力程度,医生的效用超过了社会福利最大化时医生的效用,他通过诱导行为获得了更多的效用。并且,由于缺乏有效的监督手段,患者缺乏合适的努力途径,患者个人的努力不足,患者的个人效用低于社会福利最大化时患者的效用。在后付费制度下,产出 Y 最大化、社会福利 $W(\beta)$ 最优,医生和患者效用 $U_D(y_D, e_D)$ 、 $U_P(y_P, e_P)$ 最大化无法同时达到。

当采取按人头付费模式时,由于委托权完全归患者所有,出现与之相反的情况,不存在诱导需求但医生努力不足,导致产出(即治疗效果或健康)过低。这需要通过引入对医生的监督,从而提高医生的努力水平。

三、存在监督能力下的最优委托权

由于在不存在监督能力时采用按人头付费的方式会存在严重的逆向选择问题,因此,我们考虑当医患双方都存在监督能力时,最优委托权的安排。

1. 模型求解

令 e_i^s 表示医患在没有监督时的自我努力, e_i^b 为被监督时的努力,医生监督病人我们是很容易能够理解的,比如通过命令、专家建议等方式监督病人的(配合)努力。而近些年来,管理型医疗(Managed Care)机构得到发展,

这些组织有能力采取措施监督医生的行为，而加入这些组织的病人同样也就可以监督医生。

假设监督能力采取如下线性形式：

$$e_D^b = \mu e_P^s; \\ e_P^b = \rho e_D^s$$

μ 是病人监督医生的能力， ρ 是医生监督病人的能力，由于监督的不完全性，因此 $0 < \mu < 1, 0 < \rho < 1$ ，医患各自的反应函数为：

$$e_D = \max \{ (\beta \alpha)^{\frac{1}{2-\alpha}} e_P^{\frac{1-\alpha}{2-\alpha}}, \mu e_P^s \}; \\ e_P = \max \{ (1-\beta)(1-\alpha)^{\frac{1}{1-\alpha}} e_D^{\frac{\alpha}{1-\alpha}}, \rho e_D^s \}$$

应用 Fama&Jensen (1983) 与张维迎 (1995) 的证明，可知存在一个 $\beta_P \in (0, 1)$ 以及 $\beta_D \in (0, 1)$ ，且 $\beta_P < \beta_D$ ，使得 $\beta \in (0, \beta_P)$ 时，病人（通过第三方机构）监督医生，此时医患各自的努力水平为： $e_D = \mu e_P^s, e_P = e_P^s$ ；如果 $\beta \in (\beta_P, 1)$ ，医生监督患者，此时医患各自的努力水平为： $e_D = e_D^s, e_P = \rho e_D^s$ ；如果 $\beta \in (\beta_D, \beta_P)$ ，医患双方各自提供自我努力即可。

可求得：

$$\beta_P = \frac{(1-\alpha)\mu^{2(1+\alpha)}}{\alpha + (1-\alpha)\mu^{2(1+\alpha)}}; \\ \beta_D = \frac{(1-\alpha)}{\alpha \rho^{2(2-\alpha)} + (1-\alpha)}$$

很明显，当 $\beta=1$ 时，医生有积极性监督病人，因为此时他获得所有的剩余控制权，通过诱导需求使得病人仅能获取自己的保留效用，此时类似于当前我国按项目付费的方式。与之相反，在按人头付费模式下 ($\beta=0$)，医生没有积极性监督病人，但病人有积极性监督医生。同时，张维迎 (1995) 指出，如果监督足够有效（即 μ 与 ρ 足够大）， β^* 就劣于 $\beta=0$ 与 $\beta=1$ 。因此，我们比较 $\beta=0$ 与 $\beta=1$ 时的福利大小。

$$W(\beta=0) = (1-\alpha)\mu^{2\alpha} [1 - \frac{1}{2}(1-\alpha)(1+\mu^2)]$$

$$W(\beta=1) = \alpha \rho^{2(1-\alpha)} [1 - \frac{1}{2}\alpha(1+\rho^2)]$$

2. 社会福利与患者效用同时优化的条件

由于我们已经证明在不存在监督时，后付费制度下医生存在过度努力，病人努力不足，因此，在存在监督能力情况下，如果 $\mu > 0$ ，虽然会增加病人的努力程度，但同时也会继续

增加医生的努力，但由于 $\mu < 1$ ，即病人增加的努力量要大于医生的，这样将可能减少医生与病人间的努力差异程度。我们将证明存在监督能力时，采用按人头付费方式将可能实现委托权的优化配置。

命题 3：当存在监督能力且 $e_P^s \leq \frac{\mu^2}{1+\mu^2}$ 时，社会福利与患者效用能够实现同时优化。

按人头付费下，需要证明 $W(\beta=0) > W(\beta=1)$ 时，这样才有效。当 $\alpha \leq \frac{1}{2}$ ，且 $\rho < \mu$ 时能够保证该式一定成立。即医生的重要性相对较小（如在普通疾病或预防疾病时），且医生监督病人的能力小于病人监督医生的能力。这也是按人头付费为什么在基本医疗服务中大量使用的原因。当然，即使 $\alpha > \frac{1}{2}$ （但不能太大），如果 $\mu \geq \rho$ ，按人头付费模式仍然是实现社会福利最大化的最优模式，但 $\alpha > \frac{1}{2}$ 时要实现病人的强监督能力是非常困难的。或者 α 在足够小的情况下（某些轻微疾病）， μ 略微小于 ρ 也可能保证按人头付费模式仍然是实现社会福利最大化的最优模式。

当 $\beta=0$ 时，医生的效用为：

$$U_D(y_D, e_D) = \bar{y}_D + w_D - \frac{1}{2}e_D^2 = \\ \bar{y}_D + w_D - \frac{1}{2}(1-\alpha)^2 \mu^{2+2\alpha} \quad (7)$$

由 (7) 式得出医生追求效用最大化的条件 $\frac{\partial U_D}{\partial \alpha} = 0$ ，得 $\alpha = 1 - \frac{1}{\ln \mu}$ ，可以看出，在医生的重要性 α 增强的情况下，需要患者的监督能力 μ 也增强，而这正是医生难于监管之处。单位医生努力带来的产出增加，即医生的工作效率更高，由于医患各自努力的增加，随着产出 Y 的增加，病人的效用也随之增加。因此，通过加强患者的监督能力能够很好地实现社会福利最大化与病人效用最大化的同方向变化。

由于 $\beta=0$ ，此时医生是不愿意自愿付出努力的，即 $e_D = e_D^b = \mu e_P^s, e_P = e_P^s, e_D^s = 0$ 。

病人的效用为：

$$U_P = e_D^b e_P^{1-\alpha} - w_D - y_D - \frac{1}{2}e_P^2 = [(\mu e_P^s)^\alpha$$

$$\begin{aligned} & (e_p^*)^{1-\alpha} - w_D - \bar{y}_D - \frac{1}{2}]^2 \\ & = (\mu^\alpha e_p^* - \frac{1}{2})^2 - w_D - \bar{y}_D \end{aligned} \quad (8)$$

由(8)式得出,当 $e_p^* \leq \mu^\alpha$ 时,病人的效用都是处于增长状态。当 $e_p^* = \mu^\alpha$ 时,病人的效用最大,此时病人的努力 $e_D = e_p^* = \mu e_p^* = \mu^{\alpha+1}$ 。

社会福利函数为:

$$\begin{aligned} W &= U_P + U_D = e_D e_p^{1-\alpha} - \frac{1}{2} e_D^2 - \frac{1}{2} e_p^2 = [(\mu e_p^*)^\alpha \\ & (e_p^*)^{1-\alpha} - \frac{1}{2} (\mu e_p^*)^2 - \frac{1}{2}]^2 \\ & = [\mu^\alpha e_p^* - \frac{1}{2} (\mu e_p^*)^2 - \frac{1}{2}]^2 \end{aligned} \quad (9)$$

由(9)式得出,当 $e_p^* \leq \frac{\mu^\alpha}{1+\mu^2}$ 时,社会福利都是处于增长状态。且在 $e_p^* = \frac{\mu^\alpha}{1+\mu^2}$ 时,达到最大。

表3 存在监督能力下的福利及效用比较

	$\mu = 0.6,$ $\alpha = 0.3,$ $\bar{y}_D = 0.08$	$\mu = 0.6,$ $\alpha = 0.3,$ $\bar{y}_D = 0.0649$	$\mu = 0.7,$ $\alpha = 0.3,$ $\bar{y}_D = 0.1$	$\mu = 0.7,$ $\alpha = 0.3,$ $\bar{y}_D = 0.0969$	$\mu = 0.7,$ $\alpha = 0.4,$ $\bar{y}_D = 0.1$	$\mu = 0.7,$ $\alpha = 0.4,$ $\bar{y}_D = 0.0633$
e_P	0.858	0.858	0.899	0.899	0.867	0.867
e_D	0.515	0.515	0.629	0.629	0.607	0.607
Y	0.736	0.736	0.807	0.807	0.752	0.752
U_D	0.015	0.000	0.003	0.000	0.034	0.000
U_P	0.370	0.385	0.368	0.371	0.285	0.318
W	0.385	0.385	0.371	0.371	0.318	0.318

4. 现实中的优化策略

我们可以通过两种途径来加强患者的监督能力 μ ,一种是增强患者自身存在的监督能力,另一种则是第三方如医疗保险机构或政府通过病人施加给医生的监督能力。从而实现社会福利函数与医生效用的同方向增长。

患者存在监督能力的一个例子是慢性病患者的所谓“久病成医”,在长期与某个固定医院或医生的重复博弈中,多次的博弈带来的信息交互增加了病人关于这方面疾病的专门知识,医患间的信息差异越来越小。但是要想通

综合以上两式可知,当 $e_p^* \leq \frac{\mu^\alpha}{1+\mu^2}$ 时,病

人的效用函数与社会福利函数同时处于增长状态。此时病人通过付出努力,一方面直接增加了产出,另一方面监督了医生付出努力增加产出,实现了社会福利与患者效用的同时优化。

3. 算例

我们使用具体数据进行算例分析,在按人头付费且存在监督的情况下,我们需要设定的数据为病人的监督能力 μ ,医生的重要性 α ,按人头付费的额度 \bar{y}_D ,基本固定工资(收益) w_D, w_P 。我们分别分析在 $\mu = 0.6, W_D = W_P = 0.05, \alpha = 0.3, \alpha = 0.4$; $\bar{y}_D = 0.08, \bar{y}_D = 0.1$ 时算例的分析结果。由于当 $\alpha > \frac{1}{2}$ 时,需要满足的 $\mu \geq \rho$ 条件并不符合现实,因此我们仅考虑 $\alpha \leq \frac{1}{2}$ 的情况。

过病人的个体努力使得 $\rho < \mu$ 成立是非常困难的。

因此,我们有必要引入第三方机构来加强患者的监督能力。如通过健康维护组织(HMOs),它是一种集筹资与医疗保健服务于一体的新的预付保健组织形式。其成员只需按月或年交纳一笔固定的费用,就可以免费得到门诊、住院和预防等综合性卫生保健服务。目前在全美大约有600多个HMOs,覆盖人口4000万以上,已经成为美国医疗保健制度的一个重要组成部分。

四、结论及展望

通过建立医患间的团队生产模型，并对预付费及后付费两大类支付方式下最优委托权配置的比较分析可知，在病人无监督能力下，通过后付费方式支付医疗费用，必然会带来供方诱导需求行为，因此，这对于医疗费用控制是一个不利选择；在无监督能力下，通过预付费方式支付医疗费用，必然会带来医生消极工作，产生逆向选择行为。因此，有必要引入医患双方尤其是病人的监督能力，这样既能提高医生的工作努力程度，又能很好地控制医疗费用。

同时，我们可以看到，在医疗费用日趋高涨的背景下，控制医疗费用与激励医生努力工作是需要同时考虑的两大难题，要同时实现这两个目标，可行的医疗支付方式调整方向为：

(1) 不同类型的医疗服务需要采取不同形式的支付方式。通过理论及实践已经证明按人头付费方式在基本医疗中发挥了巨大作用，但是，在当 $\alpha \geq \frac{1}{2}$ ，且 $\rho > \mu$ 时，即一些大病重病，按人头付费并不是最佳选择。因此，我们必须针对不同类型的医疗服务采取不同的支付方式，如引入按照疾病的诊断种类定额付费等。

(2) 总体趋势是支付方式从后付制向预付

制演进。由于医疗保险支付制度的改革和完善，以及第三方支付机构的引入，增强了患者方的谈判能力。并且，对第三方机构来说，防范医生不努力的逆向选择行为比防范医生过度努力的道德风险行为，相对会容易一些。因此，纵观世界各国的医疗支付制度改革，大趋势都是从后付制向预付制演变。

(3) 支付方式由单一支付方式向混合支付方式发展。由于疾病的复杂性表现出来的医患双方的不确定性、医疗费用高速增长、供方诱导需求、道德风险、逆向选择等诸多问题，单纯采取某种支付方式很难同时实现几个目标的优化与改进，而各种支付方式都有其优点和缺点。考虑通过混合支付方式来实现不同支付方式的优缺点互补，不失为一个医疗费用支付的探索方向。

本文作者：刘自敏是中国社会科学院研究生院数量经济与技术经济系 2010 级博士研究生，西南大学经济管理学院讲师；张昕竹是江西财经大学产业经济研究院特聘教授，中国社会科学院规制与竞争研究中心主任研究员，博士生导师，经济学博士

责任编辑：王姣娜

Payment Pattern in Healthcare and Optimal Principalship Assignment between Physicians and Patients under Teamwork

Liu Zimin Zhang Xinzhu

Abstract: The outcome of healthcare service is the effect of both physicians and patients. This paper compares the principalship assignment between cooperative game for social welfare maximum and non-cooperative game for the physician-patient utilities maximum. The conclusion is that Post-Paid-System will lead to supplier-induced demand and FPS will lead to adverse select without supervision. After introducing supervision, Pre-Paid-System could achieve the optimal principalship assignment under some conditions, and improve social welfare and patient's utility.

Key words: teamwork; payment pattern in healthcare; principalship; supplier-induced demand