过度医疗与转诊制：一个排队论下的博弈模型①

王文娟[1]，王季冬[2]
（1. 中央财经大学政府管理学院，北京 100081；2. 威斯康星大学麦迪逊分校经济系，麦迪逊 53706，美国）

摘要：建立了一个转诊系统中排队和博弈的集成模型，患者可能有轻微或严重型疾病，而医生可能将轻微型疾病诊断为严重型疾病，进行过度治疗. 本文探究政府如何通过配置医院规模和调节医疗服务价格影响医患博弈，使患者倾向于社区首诊、医生选择说实话. 研究表明，社区医院与三甲医院通过规模进行竞争，社区医院规模越大或三甲医院规模越小，诱导需求越能得到抑制. 轻微型疾病治疗方案的价格较低时，诱导需求以正概率存在；高于某阈值时，诱导需求消失. 提高严重型和轻微型疾病治疗方案的价差，能够抑制诱导需求，但不能完全消除，也不能改善患者福利；和三甲医院相比，社区医院在改善患者福利上效果更明显.

关键词：过度医疗；转诊系统；排队；诱导需求；医患博弈
中图分类号：F224.3；F280　文献标识码：A　文章编号：1007－9807（2019）02－0063－14

0　引　言

医疗服务是典型的信任商品（credence goods）：首先，服务卖方（医生）是决定买方（患者）需求的专家，而买方对自身需要何种服务几乎一无所知；其次，由于检验真实健康状态的成本极高[1]，作为代理人的医生的努力水平不仅不可观测，而且在交易完成后也无法验证. 上述因素为医生诱导患者购买大量不必要的检查和药物提供了机会，也就是过度医疗（overtreatment）. 除医疗服务外，维修、出租车、建筑等产品或服务同样具有信任商品的特征，卖方有机会诱导患者过度消费. 这种患者真实需要外的“需求”被称为供给方诱导需求（supplier－induced demand）. 诱导需求的存在不仅符合经济学直觉，而且被大量实证和实验研究验证[2－8]. 这种具有欺骗消费者性质的现象，被称作消费者欺诈（consumer fraud）或专家欺骗（fraudulent expert）[9].

Wolinsky[10]，Pitchik 和 Schotter[11] 以及 Emons[12] 等学者的工作奠定了信任品交易问题的博弈论基础. 其后的研究增强了模型的适用性，使该领域的研究趋于成熟. 它的一般假设如下：买方有昂贵型和廉价型两种服务需求，卖方相应有昂贵型和廉价型两种服务；卖方从昂贵型需求中能获得更多利润，因此有为廉价型需求买方提供昂贵型服务的动机；卖方的策略空间为是否欺骗买方，买方的策略空间为是否接受卖方提供的服务；均衡状态下买卖双方一般都采用混合策略——卖方以一定概率欺骗，而买方以一定概率接受服务. 此类研究的核心论题是：在什么样的市场机制下，卖方（专家）最有可能说实话[13－14]（play truthfully）.

同其他信任品相比，医疗服务还肩负社会责任. 过度医疗增加了患者的医疗费用，激化了医患矛盾. 为缓解“看病难，看病贵”问题，政府已从许多方面作出尝试. 目前世界范围内广泛采用的分级诊疗体系，让患者首先接受基本医疗服务，之后根据初步诊断结果再到相应机构接受专业治疗，是对医疗服务进行规范的有益探索：在欧洲有广

① 收稿日期：2018－02－08；修订日期：2018－07－21.
基金项目：国家自然科学基金资助项目（71473284；71603298）.
作者简介：王文娟（1965—），女，山西大同人，博士，教授. Email：wangwenjuan@cufe.edu.cn

为流行的全科医生制度(general practitioner,GP);在美国,健康维护组织(health maintenance organization,HMO)扮演着基层卫生组织的角色. 尽管它们只提供较为初级的医疗服务,却发挥着医疗系统"看门人"的关键职能(gatekeeping)[15-16],根据患者真实需要分配相应服务,有助于节约医疗资源,节省不必要的花费.

在这样的大背景下,分级诊疗体系下的决策分析成为国外公共卫生领域的重要课题. Garcia-Marinosod 等[17] 和 Allard 等[18] 学者围绕医保偿付制度和医生薪酬制度对全科医生的转诊决策进行研究,在他们的模型中,医生在能力和利他程度上有差异,而这些是医生的私人信息. 基于医生同管制者博弈的结构,设计出可以降低信息不对称的负面影响、引导全科医生最大化患者收益的合约. 而国内相关研究则比较欠缺. 陈妍等[19] 研究了在三甲医院和社区医院存在竞争的背景下,政府应如何进行补贴以有效利用医疗资源.

在医疗运作管理的建模中,排队模型是分析患者产生及就医选择的经典模型. 本刊之前的文章已经做了详尽的综述[20],此处不再赘述. 以往的文献一般用具有一类患者的 M/M/1 排队模型描述医疗系统,本文则应用一个具有轻微型和严重型两类患者的排队模型描述医疗系统,使模型更符合现实.

在以上几方面工作的基础上,本文建立了一个排队和博弈的集成模型,以研究政府应如何为医疗服务定价才能引导转诊系统良好运行,从而减少过度医疗. 同经典的信任品文献相比,本文从管制者的角度出发寻求提升社会福利的方案,为已有研究提供了一个新的思路. 在模型的假设上,国外同类研究并未考虑中国国情. 例如,在 Hilger[21] 的模型中,患者观察不到医生的成本; 在 Mimra 等[22] 的模型中,医生之间通过价格进行竞争,而在我国,医疗服务价格水平目前基本由政府划定并公示. 因此,本文调整了相关的假设.

已有的研究或是将医患关系看成是单纯的不完全信息序贯博弈,忽略了患者产生的动态性和随机性,或是把患者就医行为当作单纯的运筹问题进行优化,忽略了医疗服务的信任品性质. 因此,本文构建了一个将医疗服务的信任品属性同医疗服务需求的随机性相结合的模型. 本文旨在说明,除了医疗系统的人员组织和资源配置外,患者和医生的行为同样是影响系统效率的不容忽视的因素.

1 基本模型

1.1 模型假设

假设仅有两种疾病,轻微型 N 和严重型 S,在患病人群中的比例分别为 α 和 $1-\alpha$,$0<\alpha<1$. 病人的产生(两种类型病人的总和)是参数为 λ 的 Poisson 过程. 个体患病时能够立刻察觉到自己患病,但无法分辨疾病类型. 同时,有 N 型和 S 型两种疗法,价格分别为 p_N 和 p_S,价差 $p=p_S-p_N>0$. S 型疗法能治愈两种疾病,N 型疗法只能治愈 N 型疾病. N 型疗法和 S 型疗法的治疗时间服从均值分别为 $1/\mu_1$、$1/\mu_2$ 的负指数分布. $\mu_1>\mu_2$,代表着严重疾病的治疗需要比轻微疾病更长的周期. 为简化分析,假设 p_S 和 p_N 分别是医生实施相应治疗的个人净收入. 用 Y_A 表示医生的单位时间期望收益.

和陈妍等[19] 的研究相似,考虑由一家三甲医院和社区医院构成的转诊系统. s_A 和 s_C 分别代表三甲医院和社区医院的医生数. 和他们的假设不同的是,三甲医院和社区医院的定价相同,但功能不同. 两家医院都能对两种疾病进行诊断,都能实施 N 型疗法,但只有三甲医院能实施 S 型疗法. 在我国,医院的等级划分和医院的规模挂钩,而医生数是衡量医院规模的重要指标. 因此本文中医生数代理医院规模,自然有 $s_A>s_C$. 类似的,在后文,本文用下标 A 和 C 区分两家医院的参数.

病人发现自己生病,立即去医院就医,并且忽略患者到达医院所需的时间. 这样的假设保证了到达医院的患者是参数不变的稳定 Poisson 流. N 型患者和 S 型患者被治愈后,分别获得 V_N 和 V_S 的效用. 消费者心理的研究表明,等待往往会降低消费者从服务或产品中获得的效用,或者说消费者是延时敏感的(delay-sensitive). Naor[23] 首先考虑了顾客的延时敏感性. 其后,在很多研究企业定价、生产或库存策略的文献中,延时敏感(或等待敏感)成为重要的假设,如 Afèche 和 Mendelson[24],Mendelson 和 Whang[25],陈剑和张楠[26],周文慧等[27]. 在卫生经济学和医疗运作管理相关

文献中,该假设在之前的研究中已有体现,如寇宗来[28]在患者的预算中引入了麻烦成本,陈妍等[19]更直接地假设患者看病就医要承担等待成本. 本文假设患者延时成本在数值上等于期望排队时间 T_A、T_C. 并假设患者的效用是准线性的: 期望净效用等于从疾病治疗中获得的期望效用减去花费和期望延时成本. V_N 和 V_S 足够大,使得患者去看病总是获得正的净效用,因此不存在患者不去医院就医的情况. 在某些研究中,患者的策略空间是接受或者拒绝医生的治疗方案[29-31],而本文中,患者根据期望净效用做出如下选择: 直接去三甲医院就医或者先到社区医院进行首诊. 患者可以采用混合策略,以一定概率选择到社区医院首诊,以一定概率选择越过社区医院直接到三甲医院. 在排队系统到达稳态时,患者采用混合策略体现为一部分患者选择转诊,而一部分患者直接去三甲医院.

在我国,政府可以直接对三甲医院进行监管以确保其公益性,却难以对医生具体的医疗行为进行监管,这就为医生的投机行为创造了空间(如图 1 所示). 而特定的工资结构为医生实施过度医疗提供了动力. 除此之外,医院内部管理者和医生之间也可能存在利益上的矛盾或关联. 本文为了简化分析,假设过度医疗的直接原因是医生的投机行为. 在一些文献中,医生的利他性(altruism)是一个重要的私人信息. 如 Liu[32]将专家分为自私的和有良知的两种,本文并不作此区分,而认为医生普遍具有投机性,会通过欺骗 N 型患者最大化收益. 同样,允许医生使用混合策略. 在相关的研究中,易余胤等[33]假设博弈的参与者以一定概率采取机会主义行为,本文有相似的设定: 假设医生以概率 r 欺骗 N 型患者,将其诊断为 S 型患者,而在社区医院,由于不能开展 S 型疗法,无法通过欺骗 N 型患者获取更多利益,医生将如实报告 N 型患者的病情. 而 S 型患者,在社区医院接收诊断后,需要转诊到三甲医院接受治疗,对于转诊的患者,社区医院不再收取费用. 本文假设医生不会将 S 型患者诊断为 N 型患者,这一假设在文献中被称为责任(liability). 另外,本文只考虑向上转诊的情况.

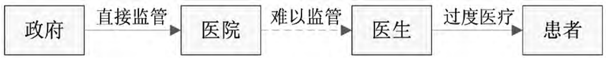

图 1　过度医疗的产生

Fig. 1 The emergence of overtreatment

患者到达医院后,以等概率选择每位医生. 这样,三甲医院和社区医院形成具有合作性质的并联 M/M/1 排队系统,曾银莲和李军[34]为此类排队系统提供了一个分析框架. 稍有不同的是,转诊的存在使得本文的队列又有了串联的性质,总体上如图 2 所示. 和社区医院不同,三甲医院需要服务两类患者,根据 De Smit[35]具有两类顾客的 M/M/1 队列和只有一类顾客的 M/H_2/1 队列等价. 其中 H_2 表示服务时间服从超指数分布. 在医疗实践中,诊断和治疗往往有明确的分界. 杜创[36]将诊断和治疗分开考虑,发现诊断和治疗具有不同的效应. 本文则不作此区分,把诊断和治疗统一囊括在服务时间中. 设三甲医院和社区医院的首诊到达率分别为 λ_A 和 λ_C. 假设 $\mu_2 > \lambda/s_A$ 和 $\mu_1 > \lambda/s_C$,这意味着医疗服务能力充足,排除了队长发散的情况. 根据排队理论计算出两家医院的期望排队时间:

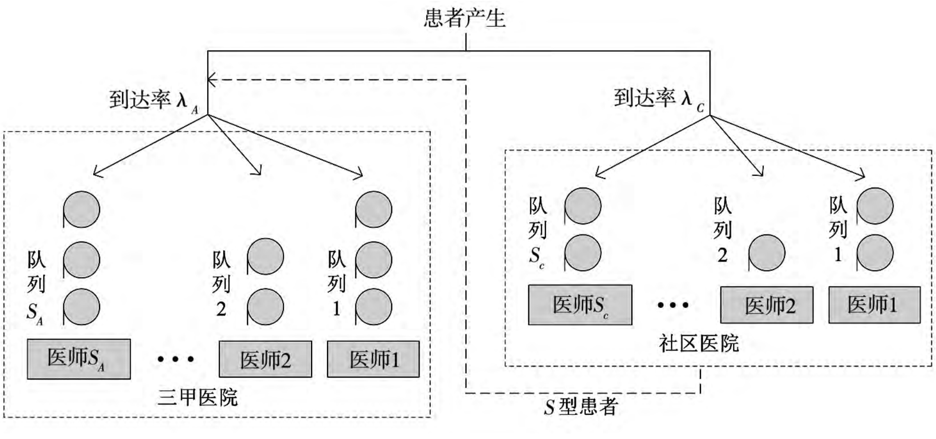

图 2　转诊系统示意图

Fig. 2 Illustration for the referral system

引理 1 患者在社区医院的期望排队时间 $T_C = \frac{s_C}{\mu_1 s_C - \alpha\lambda_C} - \frac{1}{\mu_1}$，在三甲医院的期望排队时间

$$T_A = \frac{(\lambda - \alpha\lambda_C)(p_1\mu_2^2 + p_2\mu_1^2)}{(\mu_1\mu_2 s_A - (\lambda - \alpha\lambda_C)(p_1\mu_2 + p_2\mu_1))\mu_1\mu_2}$$

其中 $p_1 = \frac{\alpha(\lambda - \lambda_C)}{\lambda - \alpha\lambda_C}$，$p_2 = \frac{(1-\alpha)\lambda}{\lambda - \alpha\lambda_C}$ 分别是三甲医院 N 型患者和 S 型患者的比例.

模型中，μ_1，μ_2 代表了医疗技术，短期内一般不可改变，视之为常量；α 和 λ 是疾病的特征，代表着不同类型的疾病，一般也认为不可变；s_A，s_C，p_N 和 p 则是政府有能力调节管控的变量. 医院和患者的决策时序如下：

第一阶段 自然赋予患者不同类型的疾病.

第二阶段 患者决定到社区医院或三甲医院，最大化期望效用. 对于 N 型患者，若选择到社区医院首诊，将直接接受 N 型疗法；若选择在三甲医院首诊，有 r 的概率被诊断为 S 型患者，接受 S 型疗法. 对于 S 型患者，不管在何处首诊，都将被诊断为 S 型患者，并到三甲医院接受 S 型疗法.

第三阶段 对于 N 型患者，三甲医院的医生做出欺骗或如实相告的决策.

由于患者源源不断地产生，这三个阶段相互影响. 在均衡时，医院的到达率不再改变，所有新产生的患者都采用相同的策略，医生欺骗患者的概率不再改变，并成为公共知识. 由排队论可知，若患者采用纯策略，那么 $\lambda_C = 0$ 或 $\lambda_C = \lambda$；若患者采用严格混合策略，$0 < \lambda_C < \lambda$. 医生和患者之间的决策实际上构成 Stackelberg 两阶段博弈，文章以下部分将采用逆向归纳法求解均衡.

1.2 患者选择行为

患者到社区医院进行首诊，有 α 的概率被直接治愈，有 $1-\alpha$ 的概率转诊. 因此选择社区医院进行首诊的期望净效用为

$$U_C = (1-\alpha)(V_S - p_S - T_C - T_A) + \alpha(V_N - p_N - T_C) \tag{1}$$

直接到三甲医院接受治疗，意味着当患者为 N 型患者时，有 r 的概率以 p_S 的价格治愈. 因此直接到三甲医院治疗的期望净效用为

$$U_A = \alpha(1-r)(V_N - p_N) + \alpha r(V_N - p_S) + (1-\alpha)(V_S - p_S) - T_A \tag{2}$$

后文将看到，医生的欺骗概率和两种疗法价差的乘积 rp 是极为重要的变量，令 $h = rp$，它代表了三甲医院的医生平均在每个 N 型患者身上攫取的剩余，对于每个患者来说，它是期望剩余损失. p 固定时，h 和 r 成正比. 由于 $r \in [0,1]$，$h \in [0,p]$. 为便于分析，不妨先令 h 在 $(0,+\infty)$ 变动，考察 h 对医院收益的影响，再研究 e 对 h 的限制对医生策略的影响. 下面给定 h，研究患者能作出的最优反应 (best response)，有如下结论成立:

引理 2 令

$$H = -\frac{(1-\alpha)\lambda}{\mu_2(\mu_2 s_A - (1-\alpha)\lambda)} + \frac{\lambda}{\mu_1(\mu_1 s_C - \alpha\lambda)}$$

给定 h，患者的最优反应 λ_C^m 唯一存在，这意味着 λ_C^m 是 h 的函数，记为 $\lambda_C^m = \lambda_C^m(h)$，有：

若 $H \leqslant 0$，那么 $\lambda_C^m = \lambda$，$\lambda_A^m = 0$；

若 $H > 0$，那么 $0 < \lambda_C^m < \lambda$，并且

$$\lambda_C^m(h) = \begin{cases} \frac{-\sigma_2 + \sqrt{\sigma_2^2 - 4\sigma_1\sigma_3}}{2\sigma_1}, & h \in [0,H) \\ \lambda, & h \in [H,+\infty) \end{cases}$$

在 $(0,H)$ 上连续可微，且 $\frac{\mathrm{d}\lambda_C^m}{\mathrm{d}h} > 0$.

其中

$$\sigma_1 = h\alpha^2\mu_1\mu_2^2 + \alpha(1-\alpha)\mu_2^2 \tag{3}$$

$$\sigma_2 = (\alpha h\mu_1 + 1)\mu_1\mu_2^2 s_A - \alpha h\mu_1^2\mu_2^2 s_C - (\alpha h\mu_1 + 1)((1-\alpha)\lambda\mu_1\mu_2 - \alpha\lambda\mu_2^2) + \alpha((1-\alpha)\lambda\mu_1^2 + \alpha\lambda\mu_2^2) + \alpha\mu_1\mu_2^2 s_C \tag{4}$$

$$\sigma_3 = \alpha\lambda h\mu_1^2\mu_2^2 s_C - (1-\alpha)\lambda\mu_1^3 s_C - h\mu_1^2\mu_2 s_C(\mu_1\mu_2 s_A - (1-\alpha)\lambda\mu_1) - \alpha\lambda\mu_1\mu_2^2 s_C \tag{5}$$

可见，当患者的期望剩余损失在某个阈值内时，患者将采取混合策略，以正概率选择三甲医院和社区医院，此时三甲医院和社区医院对于患者来说无差异. 在该阈值外，社区医院严格占优于三甲医院，患者将采取纯策略，去社区医院进行首诊. 后一种情况下，投机行为不复存在. 下面假设 $H > 0$，对前一种情况中医生的策略进行分析.

1.3 医生的策略

在了解到患者的最优反应 $\lambda_C^m(h)$ 后，三甲医院的医生求解如下最优化问题

$$\max\ Y_A = (1-\alpha)\lambda p_S + \alpha(\lambda - \lambda_C^m(h))(p_N + h)$$
$$\text{s.t.}\quad h \geqslant 0$$

医生的单位时间收益函数可以这样理解: 第一部分是所有 S 型患者带来的收益,由于 S 型患者只能在三甲医院接受治疗,所以这一部分收益是固定收益; 第二部分是 N 型患者带来的收益,$p_N + h$ 是从一位患者身上获得的收益,等于 N 型疗法的费用加上通过欺骗攫取的剩余.

为简化分析,暂时忽略了 $h \leqslant e$ 的约束,并假想医生可以通过调整 r 改变 h. 也就是说,Y_A 可以看作 h 的函数,这意味着,h 而非 r,才是影响患者福利和医生收益的关键变量. 当医生最大化其收益时,增大 e 尽管能减小医生欺骗患者的概率,却并不能减小过度医疗给患者带来的福利损失. 设 $h' = \arg\max_{h>0} Y_A$,r^* 为医生的最优策略,由于 r 在紧集 $[0,1]$ 变动,r^* 的存在性得以保证(但不一定唯一). 设 h^* 为博弈到达均衡时 h 的值($h^* = r^* p$,不同于 h'),$\lambda_C^* = \lambda_C^m(h^*)$ 为社区医院的均衡到达率. U_A^*,U_C^* 为均衡时患者的期望效用. 有如下结论:

定理 1 (1) $U_A^* = U_C^*$; (2) $0 < \lambda_A^* < \lambda$, $0 < \lambda_C^* < \lambda$; (3) $h', h^* < H, r^* \leqslant \min\{1, H/p\}$

可见,均衡时,医生会通过调整 r 引导患者采取混合策略,以至于患者选择不同医院带来期望净效用无差异. 结合引理 2,定理 1 意味着在均衡时,$H > 0$ 是患者采用混合策略的充要条件: 在 $H \leqslant 0$ 时,所有患者都将选择到社区医院进行首诊; 在 $H > 0$ 时,患者选择严格的混合策略(或者说,总有一部分患者越过社区医院直接到三甲医院). 也就是说可能存在如下的情形: 政府通过调节 s_A 和 s_C 控制 H 的值,实现完全的转诊,此时,社区医院最大程度地发挥了看门人的作用. 后文将证明,如果允许社区医院的规模超过三甲医院,这个想法可以实现.

当 $H > 0$ 时,医生通过调整 r,使得 $h^* \in [0, H)$,它蕴含了两方面信息: 第一,$\min\{1, H/p\}$ 是 r^* 的上界; 第二,h^* 和 h' 落在 $\lambda_C^m(h)$ 和 Y_A 的可微区域,可以通过讨论一阶条件进一步研究医生的策略. Y_A 对 h 一阶微分的形式十分复杂,定理 2 确保了它在 H 左邻域内的表现:

定理 2 $\exists h_1 \in (0, H)$, $\forall h \in [h_1, H)$, $\frac{\mathrm{d}Y_A}{\mathrm{d}h} < 0$.

而定理 3 及其推论阐述了 p_N 对医生策略的影响存在阈值效应,通过控制 N 型疗法的费用,可以确定一阶条件解的范围: 当 p_N 高于某个阈值时,医生不再欺骗患者,当 p_N 低于某个阈值时,医生以正概率欺骗患者,并且在价差足够高时,欺骗的概率和价差成反比.

定理 3 $\exists \underline{p_N} > 0$ 和 $h_2 \in (0, H)$,当 $p_N < \underline{p_N}$ 和 $h \leqslant h_2$ 时,$\frac{\mathrm{d}Y_A}{\mathrm{d}h} > 0$;

$\exists \overline{p_N} > 0$,当 $p_N > \overline{p_N}$ 时,$\forall h \in [0, H)$,$\frac{\mathrm{d}Y_A}{\mathrm{d}h} < 0$,以至于 $r^* = 0$

推论 1 当 $p_N < \underline{p_N}$ 时,$h', h^* \in (0, H)$,且 $\frac{\mathrm{d}Y_A}{\mathrm{d}h}\big|_{h=h'} = 0$,$r^* > 0$. 此时,若 $p \geqslant h'$,则 $h^* = h'$,$r^* = h^*/p$.

此前的结论探讨了均衡的存在性和范围. 由于目标函数表达形式复杂,只能揭示一阶导的局部性质. 通过对导数值符号的分析可知,当 $p_N < \underline{p_N}$ 时,一阶条件是使 Y_A 最大的必要条件. 一阶导的全局性质模糊,使得当前的假设并不能保证该条件的充分性,因为 $\frac{\mathrm{d}Y_A}{\mathrm{d}h} = 0$ 的解和博弈的均衡可能不唯一,为比较静态分析带来不便. 为此,本文对参数施加限制,给出一个使 $\frac{\mathrm{d}Y_A}{\mathrm{d}h}$ 单调的的充分条件:

定理 4 当有序数对 (s_A, s_C) 使得

$$\frac{\mu_2\alpha((1-\alpha)(\mu_1-\mu_2)\lambda + \mu_2^2 s_A)(\mu_1 s_C - \alpha\lambda)^3}{s_C((\alpha-1)\lambda\mu_1 - \mu_1 s_A \mu_2)^3} > \quad (6)$$

时,$\frac{\mathrm{d}^2 Y_A}{\mathrm{d}h^2} < 0$. 称该条件为单调性条件.

推论 2 假设单调性条件成立,并且 $p_N < \underline{p_N}$,那么 h' 和 h^* 皆存在且唯一. 且 $\frac{\mathrm{d}Y_A}{\mathrm{d}h}\big|_{h=h'} = 0 \Leftrightarrow h' = \arg\max_h Y_A$. 此时,考虑到 $h \leqslant p$,有 $r^* = \min\{h'/p, 1\}$,$h^* = \min\{h', p\}$.

但在 $s_A > s_C$ 的假设下,单调性条件未必成立. 因此下一部分放松该假定,讨论多个社区医院

的情况.

2 存在多家社区医院的情况

在基本模型中本文假设转诊系统中只有一家社区医院. 实际上,现实的转诊系统往往是一个三甲医院对接多个基层社区医院,为便于模型化,本部分做如下的定义: 称 R^{n+1} 中的向量 $x = (s_A, s_{C1}, \ldots, s_{Cn})$ 为一个转诊系统. s_A 代表三甲医院的医生数,s_{Ci} 代表第 i 家社区医院的医生数. 假设在同一个转诊系统内,患者在不同医院间辗转的时间可以忽略不计,以下定理保证了基本模型同样适用于存在多家社区医院的情形.

定理 5 对于两个转诊系统 $x = (s_A, s_{C1}, \ldots, s_{Cn})$ 和 $x' = (s_A', s_{C1}', \ldots, s_{Cn}')$,若它们满足 $s_A = s_A'$; $\sum s_{Ci} = \sum s_{Ci}'$,那么在 x 和 x' 中,患者选择三甲医院的概率、医生的策略、患者在社区医院和三甲医院的排队时间相同.

因此,当社区医院的医生总数固定时,社区医院的数量没有实质性的影响. 因此,可以把多个社区医院看作一家"大规模"社区医院. 在基本模型中,$s_A > s_C$; 当存在多家社区医院时,实际上是在研究 s_C 较大,甚至 $s_C \geqslant s_A$ 的情形. 实际上,放松 $s_A > s_C$ 的限制,可以得到更广泛的结论.

定理 6 允许 $s_C \geqslant s_A$. 当 s_C 足够大时: ①存在使 $H \leqslant 0$ 的 (s_A, s_C); ②存在使得 $H > 0$ 同时使单调性条件成立的 (s_A, s_C).

可见,政府可以通过增加社区医院的数量和规模,在整体上增加提供基层医疗服务的人员,引导所有患者选择先到社区医院进行初步诊疗. 只有基层医疗的力量足够强大时,才能起到"看门人"的作用. 放松 $s_A > s_C$ 的假定,让本文可以利用一阶导的单调性对转诊系统进行比较静态分析. 在基本模型其他结论的证明中,并未要求 $s_A > s_C$. 因此,在多家社区医院存在时,基本模型的其他结论仍然成立.

3 比较静态和数值算例

以医生的策略为支点,分析由 s_A, s_C, p_N, p 表示的政府行为对博弈双方的收益的影响. 考虑存在多家社区医院的情形(即允许 $s_C > s_A$),并假设单调性条件成立. 固定 p, h' 代表了医院的最优策略. 有:

定理 7 $\frac{\partial h'}{\partial s_A} > 0$, $\frac{\partial h'}{\partial s_C} < 0$, $\frac{\partial h'}{\partial p_N} < 0$.

定理 7 体现了社区医院和三甲医院之间对患者的竞争,而且竞争通过规模实现: 医生从患者攫取的剩余随三甲医院的规模递增,随社区医院的规模递减. 尽管社区医院不直接参与博弈,转诊作为患者的一种选择,对三甲医院的垄断起到了制约作用. h' 是选择三甲医院的 N 型患者遭受的损失,它反映患者个人的损失,但不等于患者整体的损失,因为当 h' 减小时,可能有更多的患者选择三甲医院,由此带来更大的福利损失.

对于单调性条件未必成立的情况,本文结合数值计算进行分析. 设定外生参数 $\lambda = 80$, $\mu_1 = 0.5$, $\mu_2 = 0.4$, $\alpha = 0.7$, $V_S = 100$, $V_N = 50$. 设定医院规模和服务定价的基准线为: $s_A = 800$, $s_C = 400$, $p_N = 0$, $p = 5$. 令 s_A 在 800 到 1 600 间变动,s_C 在 400 到 1 200 间变动,并考虑 p_N 取 0、2、4 以及 p 取 5、6、7 时各个指标的变动.

图 2 展示了三甲医院医生的欺骗概率,也就是实施过度医疗的概率随 s_A、s_C、p_N 和 p 的变动. 欺骗概率随三甲医院的规模递增,随社区医院的规模递减,随价差的增加而减小,并且随 p_N 的增大而减小,当 p_N 超过 2 时突变为零. 同时观察到,在当前的参数取值下,和三甲医院相比,社区医院规模的变动使欺骗概率发生的变化更迅速.

3.1 三甲医院医生的收益

记三甲医院医生的均衡收益为 Y_A^*.

推论 3 $\frac{\partial Y_A^*}{\partial s_A} > 0$, $\frac{\partial Y_A^*}{\partial s_C} < 0$,并且在固定价差 p 的情况下,$\frac{\partial Y_A^*}{\partial p_N} > 0$.

s_A, s_C 对 Y_A^* 的影响同对 h' 的影响方向相同,但 p_N 对 Y_A^* 产生了同对 h' 相反的影响. 倘若政府也关注医生的福利,把医生的收益考虑在社会总福利中,那么调整 s_A, s_C, p_N, p 实际上调整了总剩余在医生和患者之间进行分配. 图 3 展示了 Y_A^* 的变动情况: 三甲医院医生的收益随着价差的增加而增加,随小病疗法价格的增加而增加,随三甲医院的规模递增,随社区医院的规模递减.

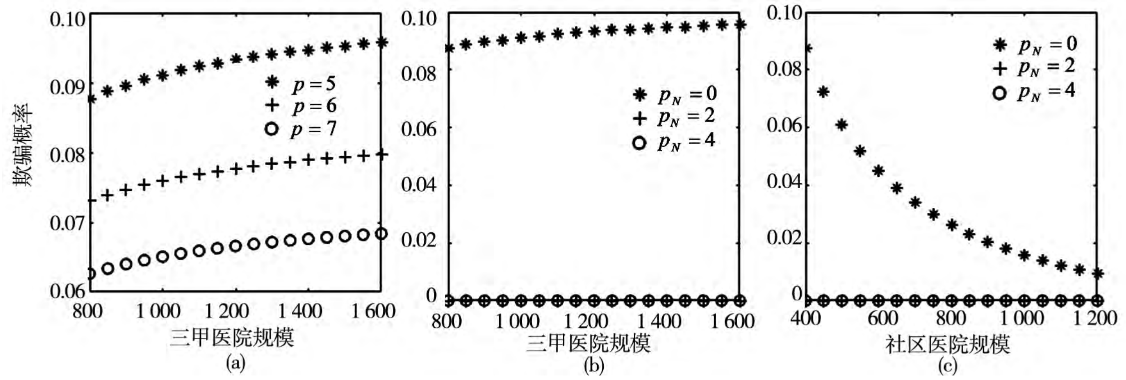

图 3　医生欺骗 N 型患者的概率随医院规模的变动

Fig. 3 Relation between cheating probabilities and hospitals' capabilities

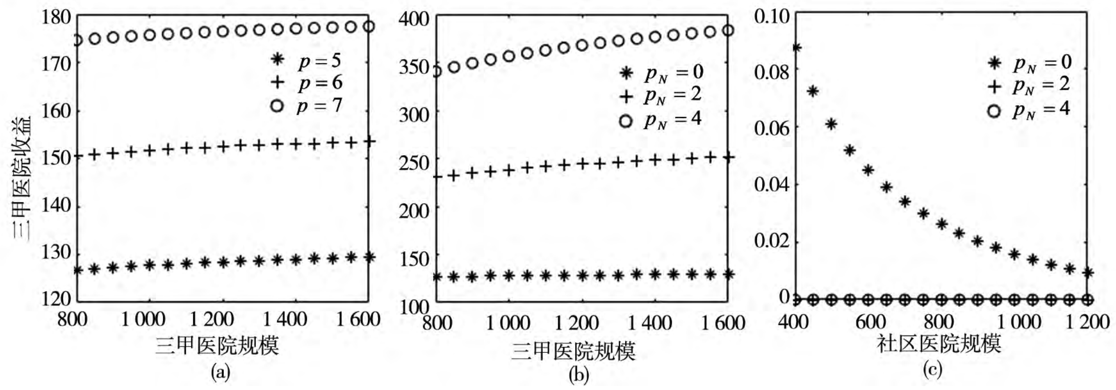

图 4　医生的收益随医院规模的变动

Fig. 4 Relation between physicians' payoff and hospitals' capabilities

3.2　患者的效用

由引理 2 和定理 1，不论政府如何调配医疗资源，也不论医生是否选择投机，均衡时患者的效用总为 U_C^*．也就是说，在三甲医院进行首诊带来的效用是衡量一个医疗系统患者满意度的标准．图 5 展示了患者效用随政府行为的变动：它关于两个医院的规模都呈递增的趋势，但随两种疗法的价格递减．

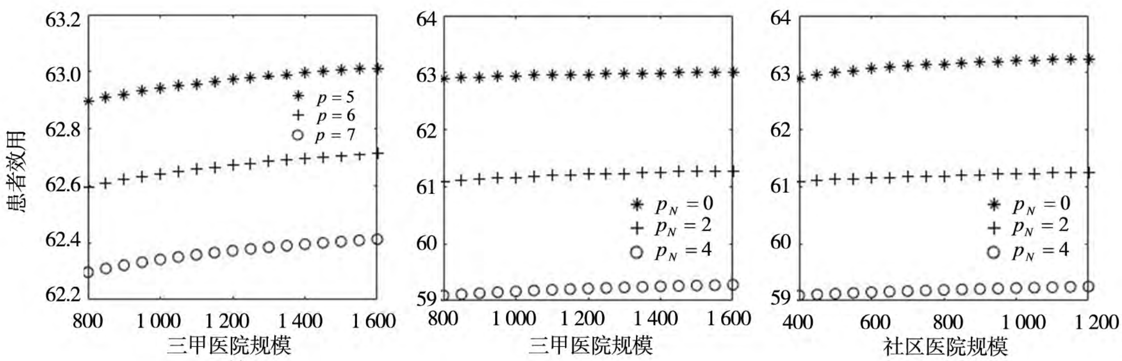

图 5　患者的效用随医院规模的变动

Fig. 5 Relation between patients' utility and hospitals' capabilities

3.3　排队时间

记转诊系统 x 的平均排队时间为 $T=((\lambda-\alpha\lambda_C)T_A+\lambda_C T_C)/\lambda$．图 5 展示了 T 的变动情况．平均排队时间随医院的规模递减．和模型的预测结果一致，改变价差没有改变排队时间．当轻微型疾病疗法的价格从 0 提高到超过消除过度医疗的

阈值时，排队时间减小并不再随着 p_N 的增大而改变. 价差对排队时间的影响随着三甲医院规模的增大而增大，但随着社区医院规模的增大而减小，并趋向于 0.

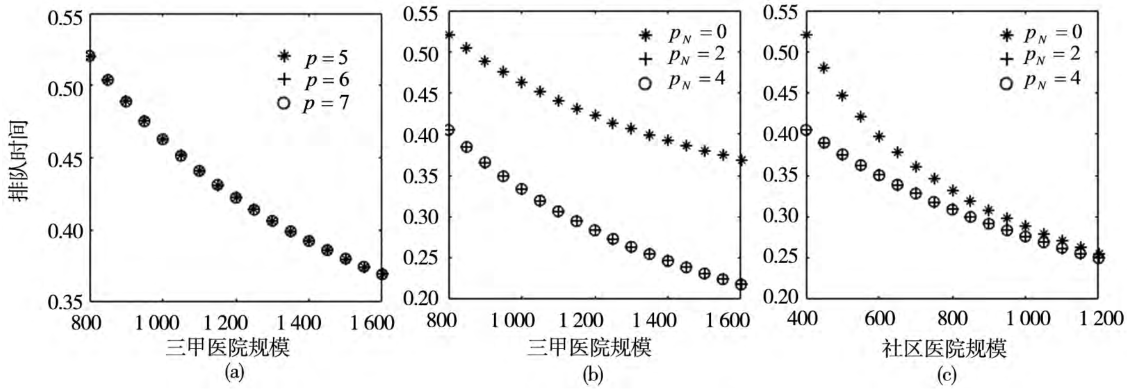

图 6 排队时间随医院规模的变动

Fig. 6 Relation between queuing time and hospitals' capabilities

3.4 福利损失

过度医疗降低了患者福利. 这种福利损失可以从两个方面考虑: 第一，患者多支付了医药费; 第二，患者原来选择的最优策略被扭曲，使排队时间增加. 因此，本文用三个指标衡量过度医疗带来的福利损失: 患者支付的超额医药费 $\Delta p = \alpha h(\lambda - \lambda_C^*)$; 增加的排队时间 $\Delta T = T|_{h=h^*} - T|_{h=0}$; 患者总体的效用损失 $\Delta U_C^* = U_C^*|_{h=0} - U_C^*|_{h=h^*}$. 从社会总福利的角度看，$\Delta T$ 是真正的无谓损失，Δp 只是从患者到医生的福利转移.

推论 4 $\dfrac{\partial \Delta p}{\partial p_N} < 0$

图 7 展示了福利损失随政府行为的变动. 第一，不管是何种福利损失，都不随 p 的改变而改变，这再一次印证了，从社会福利的角度看，价差是个无关紧要的变量; 第二，超额等待时间大于零，说明诱导需求的确同时加重了“看病难”问题，患者为了避免被欺骗遭受损失，调整策略，以至于增加了排队时间; 第三，随轻微型疾病疗法价格的上升，福利损失逐渐消失; 第四，增加三甲医院的规模，反而放大了患者的福利损失，也放大了总福利的损失——超额等待时间，但增大社区医院的规模提高了患者的福利，也降低了总福利的损失.

同时注意到，在图 3 和图 7 中，各指标随社区医院规模的变动都比随三甲医院规模的变动更迅速. 这意味着，对于政策制定者来说，调整社区医院的规模是比调整三甲医院规模更有效的方案.

4 结束语

本文就过度医疗问题，考虑由三甲医院和社区医院组成的转诊系统，把医患博弈嵌套在排队模型中，对其均衡进行了探讨. 医院的规模和医疗服务的定价是重要的政策变量，政府通过调节社区医院和三甲医院的规模、调节不同疾病疗法的价格影响医患博弈，从而缓解甚至消除过度医疗，提升患者福利和社会总福利.

本文揭示了“看病难”和“看病贵”相互耦合的经济学直觉，也解释了国内大医院不断扩张，“看病难”问题却始终没有得到解决的反直觉现象: 三甲医院的规模代表了其垄断实力. 扩大三甲医院的规模使三甲医院的医生更有能力欺骗患者，实施诱导需求，患者为了最大化自身效用，只好调整策略，追求次优(second best) ，最终增加无谓损失. 与三甲医院相比，社区医院的规模是更好的政策杠杆，扩大社区医院的规模，能够引导患者选择转诊，实现分级诊疗: 三甲医院治疗严重型疾病，社区医院治疗轻微型疾病; 同时，也能显著缓解过度医疗，降低排队时间. 调节轻微型疾病疗法的价格同样有助于消除过度医疗，但可能降低患者的效用. 而严重性疾病疗法和轻微型疾病疗法的价差无助于改善患者福利.

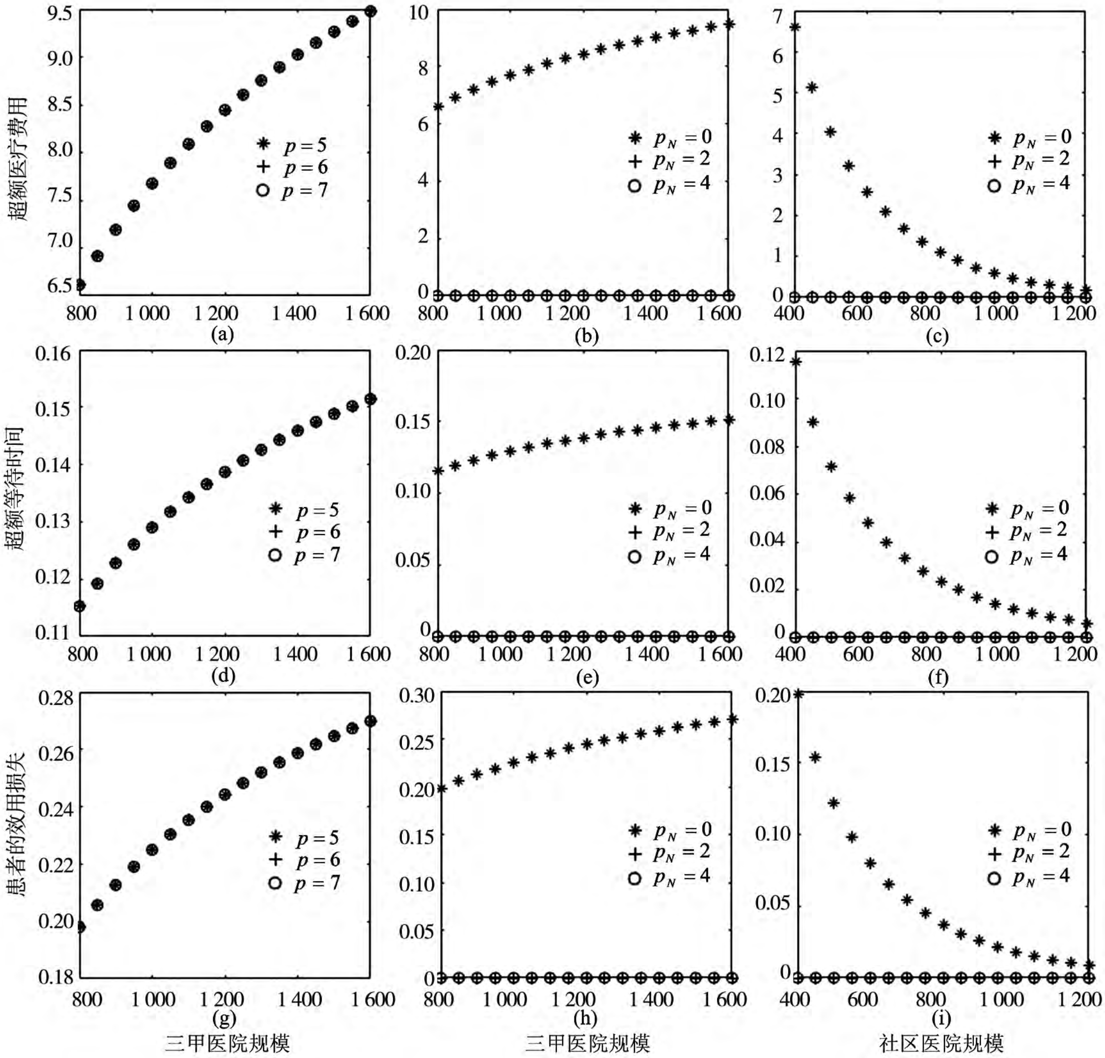

图 7　福利损失随医院规模的变动

Fig. 7 Relation between welfare loss and hospitals' capabilities

本文提供了转诊制下过度医疗现象的基本分析框架，有若干拓展方向: 考虑了同质性的患者，事实上，不同患者间对医疗服务的需求可能存在差异; 本文假定转诊系统的服务能力充足，这在某些地区可能不成立，因为我国的医疗资源分配极度不均，存在某些地区过剩，而某些地区短缺的情况; 在本文中，医疗费用完全由患者支付，而医保支付者在医疗系统中的地位举足轻重，很可能改变原有的均衡，后续的研究中，可以引入医保支付者，构建“供给者—需求者—支付者”的三方博弈.

参 考 文 献:

[1] Ching-To Albert Ma, Thomas G, McGuire. Optimal health insurance and provider payment [J]. American Economic Review, 1997, 87(4): 685 – 704.

[2] Labelle R, Stoddart G, Rice T. A re-examination of the meaning and importance of supplier-induced demand [J]. Journal of Health Economics, 1994, 13(3): 347 – 368.

[3] Balafoutas L, Kerschbamer R, Sutter M. Second-degree moral hazard in a real-world credence goods market [J]. The Economic Journal, 2017, 127(599): 1 – 18.

[4] Shigeoka H, Fushimi K. Supplier-induced demand for newborn treatment: Evidence from Japan [J]. Journal of Health Eco-

nomics, 2014, 35(5): 162 – 178.

[5] Bejarano H, Green E, Rassenti S. Payment scheme self-selection in the credence goods market: An experimental study [J]. Journal of Economic Behavior and Organization, 2017, 142: 396 – 403.

[6] Kerschbamer R, Sutter M. The economics of credence goods: A survey of recent lab and field experiments [J]. CESifo Economic Studies, 2017, 63(1): 1 – 23.

[7] Kerschbamer R, Slutter M, Dulleck U. How social preferences shape incentives in (experimental) markets for credence goods [J]. The Economic Journal, 2017, 127(600): 393 – 416.

[8] Kerschbamer R, Neururer D, Sutter M. Insurance coverage of customers induces dishonesty of sellers in markets for credence goods [J]. Proceedings of the National Academy of Sciences, 2016, 113(27): 7454 – 7458.

[9] Balafoutas L, Beck A, Kerschbamer R, et al. What drives taxi drivers? A field experiment on fraud in a market for credence goods [J]. Review of Economic Studies, 2013, 80(3): 876 – 891.

[10] Wolinsky A. Competition in a market for informed experts' services [J]. RAND Journal of Economics, 1993, 24(3): 380 – 398.

[11] Pitchik C, Schotter A. Honesty in a model of strategic information transmission [J]. American Economic Review, 1987, 77(5): 1032 – 1036.

[12] Emons W. Credence goods and fraudulent experts [J]. RAND Journal of Economics 1997, 28(1): 107 – 119.

[13] Dulleck U, Kerschbamer R. On doctors, mechanics, and computer specialists: The economics of credence goods [J]. Journal of Economic Literature, 2006, 44(1): 5 – 42.

[14] Frankel A, Schwarz M. Experts and their records [J]. Economic Inquiry, 2014, 52(1): 56 – 71.

[15] Malcomson J. Health service gatekeepers [J]. RAND Journal of Economics, 2004, 35(2): 401 – 421.

[16] Brekke K, Nuscheler R, Straume O. Gatekeeping in health care [J]. Journal of Health Economics, 2007, 26(1): 149 – 170.

[17] Garcia-Marinoso B, Jelovac I. GPs' payment contracts and their referral practice [J]. Journal of Health Economics, 2003, 22(4): 617 – 635.

[18] Allard M, Jelovac I, Léger P. Treatment and referral decisions under different physician payment mechanisms [J]. 2011, 30(5): 880 – 893.

[19] **陈 妍，周文慧，华中生，等. 面向延时敏感患者的转诊系统定价与能力规划 [J]. 管理科学学报**，2015，18(4)：73 – 83.

Chen Yan, Zhou Wenhui, Hua Zhongsheng, et al. Pricing and capacity planning of the referral system with delay-sensitive patients [J]. Journal of Management Sciences in China, 2015, 18(4): 73 – 83. (in Chinese)

[20] **杜少甫，谢金贵，刘作仪. 医疗运作管理：新兴研究热点及其进展 [J]. 管理科学学报**，2013，16(8)：1 – 19.

Du Shaofu, Xie Jingui, Liu Zuoyi. Progress and prospects in an emerging hot topic: Healthcare operations management [J]. Journal of Management Sciences in China, 2013, 16(8): 1 – 19. (in Chinese)

[21] Hilger N. Why don't people trust experts? [J]. Journal of Law and Economics, 2016, 59(2): 293 – 311.

[22] Mimra W, Rasch A, Waibel C. Price competition and reputation in credence goods markets [J]. Games and Economic Behavior, 2016, 83(11): 337 – 352.

[23] Naor P. The regulation of queue size by levying tolls [J]. Econometrica, 1969, 37(1): 15 – 24.

[24] Afèche P, Mendelson H. Pricing and priority auctions in queueing systems with a generalized delay cost structure [J]. Management Science, 2004, 50(7): 869 – 882.

[25] Mendelson H, Whang S. Optimal incentive-compatible priority pricing for the M/M/1 queue [J]. Operations Research, 1990, 38(5): 870 – 883.

[26] **陈 剑，张 楠. 针对等待敏感顾客的缺货补偿与库存策略研究 [J]. 管理科学学报**，2008，11(3)：53 – 62.

Chen Jian, Zhang Nan. Study on backorder incentives and inventory control policies with time-based customer behavior [J]. Journal of Management Sciences in China, 2008, 11(3): 53 – 62. (in Chinese)

[27] **周文慧，黄伟祥，吴永忠，等. 提高顾客等待满意度的两类排队管理策略 [J]. 管理科学学报**，2014，17(4)：1 – 10.

Zhou Wenhui, Huang Weixiang, Wu Yongzhong, et al. Two kinds of queuing management policy for improving customer satisfaction with waiting [J]. Journal of Management Sciences in China, 2014, 17(4): 1 – 10. (in Chinese)
[28] 寇宗来. "以药养医"与"看病难,看病贵" [J]. 世界经济, 2010, 33(1): 49 – 68.
Kou Zonglai. Funding hospitals through drug sales and the problem of inadequate and overly expensive medical services [J]. The Journal of World Economy, 2010, 33(1): 49 – 68. (in Chinese)
[29] Emons W. Credence goods monopolists [J]. International Journal of Industrial Organization, 2001, 19(3): 375 – 389.
[30] Alger I, Salanié F. A theory of fraud and overtreatment in experts markets [J]. Journal of Economics & Management Strategy, 2006, 15(4): 853 – 881.
[31] 黄 涛, 颜 涛. 医疗信任商品的信号博弈分析 [J]. 经济研究, 2009, (8): 125 – 134.
Huang Tao, Yan Tao. Analysis of signaling game for medical credence goods [J]. Economic Research Journal, 2009, (8): 125 – 134. (in Chinese)
[32] Liu T. Credence goods markets with conscientious and selfish experts [J]. International Economic Review, 2011, 52(1): 227 – 244.
[33] 易余胤, 肖条军, 盛昭瀚. 合作研发中机会主义行为的演化博弈分析 [J]. 管理科学学报, 2005, 8(4): 80 – 87.
Yi Yuyin, Xiao Tiaojun, Sheng Zhaohan. Evolutionary game analysis on opportunistic behavior in cooperative R&D market [J]. Journal of Management Sciences in China, 2005, 8(4): 80 – 87. (in Chinese)
[34] 曾银莲, 李 军. 存在外部市场的排队系统合作费用分配方法 [J]. 管理科学学报, 2017, 20(11): 88 – 99.
Zeng Yinlian, Li Jun. Cost allocations for cooperation among queueing systems with external market [J]. Journal of Management Sciences in China, 2017, 20(11): 88 – 99. (in Chinese).
[35] De Smit J. On the queue GI/M/s with customers of different types or the queue GI/H_m/s [J]. Advanced Applied Probability, 1983, 15(2): 392 – 419.
[36] 杜 创. 价格管制与过度医疗 [J]. 世界经济, 2013, 36(1): 116 – 140.
Du Chuang. Price control and overtreatment [J]. The Journal of World Economy, 2013, 36(1): 116 – 140. (in Chinese)

Overtreatment and referral systems: A game embedded in a queuing model

WANG Wen-juan[1], *WANG Ji-dong*[2]
1. School of Government, Central University of Finance and Economics, Beijing 100081, China;
2. Department of Economics, University of Wisconsin-Madison, Madison 53706, USA

Abstract: This paper develops an integrated model applying queueing theory and game theory for the referral system. Patients may have severe or moderate illness. Physicians could report moderate illness as severe illness and conduct overtreatment. This paper studies how a government could allocate the capacities of hospitals and adjust medical pricing to influence the physician-patient game so that the patients could get more primary care in community hospitals and the physicians could tell the truth. The study indicates that the community hospital competes with class 3-A hospitals via their capacities. Induced demand is restricted by larger community hospitals or smaller class 3-A hospitals. When the price of the treatment for moderate illnesses is low, induced demand occurs with positive probability; when that price is larger than some threshold, induced demand disappears. Raising the difference between the prices of treatments for severe and moderate illnesses could restrict induced demand, but could not eliminate it, nor could this action enhance patients' welfare. In enhancing patients' welfare, community hospitals are better than class 3-A hospitals.

Key words: overtreatment; queuing system; referral system; induced-demand; physician-patient game

附录:

引理 1 证明 考虑到去社区医院的 S 型患者要再去三甲医院,三甲医院的总体到达率为 $\lambda_A + (1-\alpha)\lambda_C = \lambda - \alpha\lambda_C$,每位医生接受到的患者的到达率为 $(\lambda - \alpha\lambda_C)/s_A$. 其中有 $p_1 = \dfrac{\alpha(\lambda - \lambda_C)}{\lambda - \alpha\lambda_C}$ 比例的患者是 N 型患者,$p_2 = \dfrac{(1-\alpha)\lambda}{\lambda - \alpha\lambda_C}$ 的患者是 S 型患者. 由 Pollaczek-Khintchine 引理

$$T_A = \frac{(\lambda - \alpha\lambda_C)(p_1\mu_2^2 + p_2\mu_1^2)}{(\mu_1\mu_2 s_A - (\lambda - \alpha\lambda_C)(p_1\mu_2 + p_2\mu_1))\mu_1\mu_2}$$

由于诊断时间为零,社区医院中只有 N 型患者影响排队时间,即影响排队时间的到达率为 $\alpha\lambda_C$,则每位医生接受到的患者的到达率为 $\alpha\lambda_C/s_C$. 由排队论的基本结论可得平均排队时间. 证毕.

引理 2 证明 患者的选择满足如下条件:

若 $U_A > U_C$,患者选择直接到三甲医院接受治疗,$\lambda_A^m = \lambda$,$\lambda_C^m = 0$;若 $U_A = U_C$,患者采用混合策略,λ_A^m,$\lambda_C^m \in (0,\lambda)$;若 $U_A < U_C$,患者选择到社区医院首诊,$\lambda_A^m = 0$,$\lambda_C^m = \lambda$.

构造 $G(h,\lambda_C) = U_A - U_C = -\alpha h - \alpha T_A + T_C$. 求导得

$$\frac{\partial G}{\partial h} = -\alpha < 0$$

$$\frac{\partial G}{\partial \lambda_C} = \frac{\alpha s_C}{(\mu_1 s_C - \alpha\lambda_C)^2} + \frac{\alpha^2((1-\alpha)(\mu_1 - \mu_2)\lambda + \mu_2^2 s_A)}{((\alpha-1)\lambda\mu_1 - \alpha\lambda\mu_2 + (\alpha\lambda_C + \mu_1 s_A)\mu_2)^2} > 0.$$

$\because \forall h > 0, G(h,0) < 0$,所以 $U_A > U_C$ 的情况不会出现.

令 $H = -\dfrac{(1-\alpha)\lambda}{\mu_2(\mu_2 s_A - \lambda + \alpha\lambda)} + \dfrac{\lambda}{\mu_1(\mu_1 s_C - \alpha\lambda)}$,那么 $G(H,\lambda) = 0$.

若 $H \leqslant 0$,那么 $\forall h > 0, G(h,\lambda) < 0$,所有患者选择到社区医院首诊.

若 $H > 0$,当 $h \geqslant H$ 时,$G(h,\lambda) < 0$,此时患者只选择到社区医院. 当 $h \in [0,H)$ 时,$G(h,\lambda) > 0$,又 $\because G(h,0) < 0$,所以存在唯一的 λ_C^m 使得 $G(h,\lambda_C^m) = 0$,此时患者采用混合策略,λ_A^m,$\lambda_C^m \in (0,\lambda)$. 解 $G(h,\lambda_C^m) = 0$,即得文中表达式. 由隐函数定理,得 $\dfrac{\mathrm{d}\lambda_C^m}{\mathrm{d}h} = -\dfrac{\partial G}{\partial h}\Big/\dfrac{\partial G}{\partial \lambda_C^m} < 0$ 证毕.

定理 1 证明 将 $\lambda_C^m(h)$ 带入到 Y_A 中,得到

$$Y_A = \begin{cases}(1-\alpha)\lambda p_S + \alpha(\lambda - \lambda_C^m(h))(p_N + h), & h \in [0,H) \\ (1-\alpha)\lambda p_S, & h \in [H,+\infty)\end{cases}$$

因为当 $h \in [0,H)$ 时,$\alpha(\lambda - \lambda_C^m(h))(p_N + h) > 0$,因此 $h \in [H,+\infty)$ 是劣策略,三甲医院的医生一定会调整 r 使得 $h \in [0,H)$. 因此均衡时,$U_A^* = U_C^*$,$0 < \lambda_A^* < \lambda$,$0 < \lambda_C^* < \lambda$. 证毕.

定理 2 证明

$$\frac{\mathrm{d}\lambda_C^m}{\mathrm{d}h} = -\frac{\partial G}{\partial h}\Big/\frac{\partial G}{\partial \lambda_C^m} = \left(\frac{s_C}{(\mu_1 s_C - \alpha\lambda_C^m)^2} + \frac{\alpha((1-\alpha)(\mu_1 - \mu_2)\lambda + \mu_2^2 s_A)}{((\alpha-1)\lambda\mu_1 - \alpha\lambda\mu_2 + (\alpha\lambda_C^m + \mu_1 s_A)\mu_2)^2}\right)^{-1}$$

$\dfrac{\mathrm{d}\lambda_C^m}{\mathrm{d}h}$ 有界,设 $0 < M_1 < \dfrac{\mathrm{d}\lambda_C^m}{\mathrm{d}h} < M_2$. 由 $\dfrac{\mathrm{d}Y_A}{\mathrm{d}h} = \alpha\left(-\dfrac{\mathrm{d}\lambda_C^m}{\mathrm{d}h}(h + p_N) + \lambda - \lambda_C^m\right)$,知 $\lim\limits_{h\to H-}\dfrac{\mathrm{d}Y_A}{\mathrm{d}h} = -\lim\limits_{h\to H-}\alpha\dfrac{\mathrm{d}\lambda_C^m}{\mathrm{d}h}(H + p_N) < -\alpha M_1(H + p_N)$,且 $\dfrac{\mathrm{d}Y_A}{\mathrm{d}h}$ 在 $[0,H)$ 连续,所以 $\exists h_1 \in (0,H)$,$\forall h \in [h_1,H)$,$\dfrac{\mathrm{d}Y_A}{\mathrm{d}h} < 0$. 证毕.

定理 3 证明　$\frac{\mathrm{d}Y_A}{\mathrm{d}h}=\alpha\left(-\frac{\mathrm{d}\lambda_C^m}{\mathrm{d}h}(h+p_N)+\lambda-\lambda_C^m\right)$，取 $\xi\in(0,H)$，当 $h\in[0,\xi)$ 时，$\lambda-\lambda_C^m$ 有正的下界，设 $\lambda-\lambda_C^m>M_3$，令 $\underline{p_N}=h_1=\frac{M_3}{2M_2}$，于是 $\forall p_N<\underline{p_N},\forall h\in(0,h_1]$：

$$\frac{\mathrm{d}Y_A}{\mathrm{d}h}>\alpha(-M_2(h+p_N)+M_3)>\alpha(-M_2(h_1+\underline{p_N})+M_3)=0.$$

取 $\overline{p_N}=\frac{\lambda}{M_1}$，当 $p_N>\overline{p_N}$ 时，$\forall h\in[0,H)$，$\frac{\mathrm{d}Y_A}{\mathrm{d}h}<\alpha(-M_1(h+p_N)+\lambda)<0$. 因此 $h^*=0$，$r^*=0$.　　证毕.

推论 1 证明　当 $p_N<\underline{p_N}$ 时，Y_A 在 $(0,h_2)$ 递增，所以 $r^*>0$，同时 Y_A 在 (h_1,H) 内递减，因此最大值在 $(0,H)$ 取到，且一定是驻点 h'. 若 $p\geqslant h'$ 则该驻点能被取到，从而 $h^*=h'$，$r^*=h^*/p$.　　证毕.

定理 4 证明　在定理 2 的证明中，注意到，$\frac{\mathrm{d}\lambda_C^m}{\mathrm{d}h}$ 是 λ_C^m 的函数. 记 $\frac{\mathrm{d}\lambda_C^m}{\mathrm{d}h}=g(\lambda_C^m)$. 根据链式法则 $\frac{\mathrm{d}^2Y_A}{\mathrm{d}h^2}=-\alpha g(\lambda_C^m)(g'(\lambda_C^m)(p_N+h)+2)$.

而 $g'(\lambda_C^m)=\dfrac{\left(\frac{2\mu_2\alpha^2((1-\alpha)(\mu_1-\mu_2)\lambda+\mu_2^2s_A)}{((\alpha-1)\lambda\mu_1-\alpha\lambda\mu_2+(\alpha\lambda_C^m+\mu_1s_A)\mu_2)^3}-\frac{2\alpha s_C}{(\mu_1s_C-\alpha\lambda_C^m)^3}\right)}{\left(\frac{s_C}{(\mu_1s_C-\alpha\lambda_C^m)^2}+\frac{\alpha((1-\alpha)(\mu_1-\mu_2)\lambda+\mu_2^2s_A)}{((\alpha-1)\lambda\mu_1-\alpha\lambda\mu_2+(\alpha\lambda_C^m+\mu_1s_A)\mu_2)^2}\right)^2}$，注意到 $\frac{2\mu_2\alpha^2((1-\alpha)(\mu_1-\mu_2)\lambda+\mu_2^2s_A)}{((\alpha-1)\lambda\mu_1-\alpha\lambda\mu_2+(\alpha\lambda_C^m+\mu_1s_A)\mu_2)^3}-\frac{2\alpha s_C}{(\mu_1s_C-\alpha\lambda_C^m)^3}$ 关于 λ_C^m 递减. 倘若 $\frac{2\mu_2\alpha^2((1-\alpha)(\mu_1-\mu_2)\lambda+\mu_2^2s_A)}{((\alpha-1)\lambda\mu_1-\mu_1s_A\mu_2)^3}-\frac{2\alpha s_C}{(\mu_1s_C-\alpha\lambda)^3}>0$，就有 $\forall h\in[0,H)$，$\frac{\mathrm{d}^2Y_A}{\mathrm{d}h^2}<0$.

证毕.

推论 2 证明　由定理 2，易知当 $p_N<\underline{p_N}$ 时，$\frac{\mathrm{d}Y_A}{\mathrm{d}h}>0$，$\forall h\in[0,h_2)$. 有零点定理，结合单调性条件，$\frac{\mathrm{d}Y_A}{\mathrm{d}r}=0$ 有唯一解. 若 $h'<p$，则 $r^*=h'/p$；若 $h'\geqslant p$，则 $r^*=1$.　　证毕.

定理 5 证明　记患者到第 i 家社区医院接受治疗的期望效用、医院的患者到达率和排队时间分别为 U_{Ci}、λ_{Ci} 和 T_{Ci}. 记 $\lambda_C=\sum\lambda_{Ci}$. 均衡时，有 $U_{Ci}=U_{Cj}$，这等价于 $T_{Ci}=T_{Cj}$，于是

$$\frac{1}{\mu_1-\alpha\lambda_{Ci}/s_{Ci}}-\frac{1}{\mu_1}=\frac{1}{\mu_1-\alpha\lambda_{Cj}/s_{Cj}}-\frac{1}{\mu_1}\Rightarrow\frac{\lambda_{Ci}}{s_{Ci}}=\frac{\lambda_{Cj}}{s_{Cj}}\Rightarrow\frac{\lambda_{Ci}}{s_{Ci}}=\frac{\sum\lambda_{Cj}}{\sum s_{Cj}}$$

$$\Rightarrow T_{Ci}=\frac{1}{\mu-\alpha\sum\lambda_{Cj}/\sum s_{Cj}}-\frac{1}{\mu}=\frac{1}{\mu-\alpha\lambda_C/\sum s_{Cj}}-\frac{1}{\mu},\forall i$$

同时，又有 $U_A=U_{Cj}$. 从前文的证明中了解到，给定 $\sum s_{Cj}$ 和 s_A，存在唯一的 λ_C 满足上述等式. 因此，若 $\sum s_{Cj}=\sum s_{Cj}'$，$s_A=s_A'$，就有 $\lambda_C=\lambda_C'$. 因此，x 和 x' 中患者的期望效用、选择三甲医院的概率、排队时间相等.　　证毕.

定理 6 证明

①只要 s_C 使得 $(1-\alpha)\mu_1(\mu_1s_C-\alpha\lambda)\geqslant\mu_2(\mu_2s_A-(1-\alpha)\lambda)$，就有 $H\leqslant0$.

②须证明满足如下约束的有序数对 (s_A,s_C) 组成的集合非空

$$\begin{cases}\dfrac{2\mu_2\alpha^2((1-\alpha)(\mu_1-\mu_2)\lambda+\mu_2^2 s_A)}{((\alpha-1)\lambda\mu_1-\mu_1 s_A\mu_2)^3}-\dfrac{2\alpha s_C}{(\mu_1 s_C-\alpha\lambda)^3}>0\\ (1-\alpha)\mu_1(\mu_1 s_C-\alpha\lambda)<\mu_2(\mu_2 s_A-(1-\alpha)\lambda)\\ \mu_2 s_A>\lambda\\ \mu_1 s_C>\lambda\end{cases}$$

在数值算例所选取的值正是满足条件的值，因此定理得证. 证毕.

定理 7 证明 由隐函数定理，$\frac{\partial h'}{\partial s_A}=-\frac{\partial^2 Y_A}{\partial h\partial s_A}/\frac{\partial^2 Y_A}{\partial h^2}$，$\frac{\partial^2 Y_A}{\partial h\partial s_A}=-\frac{\partial\lambda_C^m}{\partial s_A}[g'(\lambda_C)(p_N+h)+1]$. 在定理 4 的证明中看到，$g'(\lambda_C^m)>0$.

而 $\frac{\partial\lambda_C^m}{\partial s_A}=-\frac{\partial G}{\partial s_A}/\frac{\partial G}{\partial\lambda_C}$，$\frac{\partial G}{\partial s_A}>0$，$\frac{\partial G}{\partial\lambda_C}>0$. 所以 $\frac{\partial h'}{\partial s_A}>0$，同理 $\frac{\partial h'}{\partial s_C}<0$. 而 $\frac{\partial^2 Y_A}{\partial h\partial p_N}=-g(\lambda_C^m)<0$，所以 $\frac{\partial h'}{\partial p_N}<0$.

证毕.

推论 3 证明 由包络定理

$$\frac{\partial Y_A^*}{\partial s_A}=\frac{\partial Y_A}{\partial s_A}\Big|_{h=h'}=-\alpha\frac{\partial\lambda_C^m}{\partial s_A}\Big|_{h=h'}(p_N+h')>0$$

$$\frac{\partial Y_A^*}{\partial s_C}=\frac{\partial Y_A}{\partial s_C}\Big|_{h=h'}=-\alpha\frac{\partial\lambda_C^m}{\partial s_C}\Big|_{h=h'}(p_N+h')<0$$

$$\frac{\partial Y_A^*}{\partial p_N}=\frac{\partial Y_A}{\partial p_N}\Big|_{h=h'}=(1-\alpha)\lambda+\alpha(\lambda-\lambda_C^*)>0$$

证毕.

推论 4 证明 $\Delta p=\alpha(h+p_N)(\lambda-\lambda_C^*)-\alpha p_N(\lambda-\lambda_C^*)$. 由包络定理 $\alpha(h+p_N)(\lambda-\lambda_C^*)$ 对 p_N 的导数为 $\alpha(\lambda-\lambda_C^*)$，而 $\alpha p_N(\lambda-\lambda_C^*)$ 对 p_N 的导数为 $\alpha(\lambda-\lambda_C^*)-\alpha\frac{\partial\lambda_C^m}{\partial h}\Big|_{h=h1}\frac{\partial h'}{\partial p_N}p_N$. 因此 $\frac{\partial\Delta e}{\partial e_N}=\alpha\frac{\partial\lambda_C^m}{\partial h}\Big|_{h=h1}\frac{\partial h'}{\partial p_N}p_N<0$.

证毕.