第11卷第3期 管理科学学报 Vol. 11 No. 3
2008第6月 JOURNAL OF MANAGEMENT SCIENCES IN CHINA Jun. 2008

针对等待敏感顾客的缺货补偿与库存策略研究[①]

陈 剑[1,2], 张 楠[1,2]

(1. 教育部人文社会科学重点研究基地 清华大学现代管理研究中心, 北京 100084;
2. 清华大学经济管理学院, 北京 100084)

摘要: 随着企业竞争的激烈化,顾客对于等待时间越来越敏感,产品对于他的效用不仅与价格有关,而且也与获得该产品的等待时间相关.考虑了等待敏感型顾客的选择行为,针对一个生产与需求过程都存在随机性的企业,研究了差异定价及相应的库存策略.当企业持有库存时,顾客不需要等待;而当企业发生缺货时,如果顾客购买,就需要等待企业生产;企业需针对这两种情况进行差异定价.然而差异定价策略可能导致顾客采取投机的行为.文章首先推导了不存在投机行为时企业的最优策略,分析了企业实行差异定价策略的条件;进一步,还给出了防止投机行为的条件,以及存在投机行为时企业的最优策略.

关键词: 顾客选择行为; 等待成本; 延迟交付; 差异定价; 库存控制

中图分类号: F224.3; F253.4 **文献标识码**: A **文章编号**: 1007-9807(2008)03-0053-10

0 引 言

顾客对于等待是敏感的,这已成为不争的事实.价格不再是企业竞争的唯一要素,顾客越来越关注企业的响应速度.对于顾客来说,等待时间首先意味着沉没成本(sunk cost),其次还有焦虑和未满足的急切心理.Becker认为只要企业用固定的生产能力服务随时到来的顾客,顾客就需要付出两个成本:产品的价格(显性成本)和等待成本(隐性成本)[1]. Maister以及Taylor等研究了顾客等待时的心理,结果表明等待时间将影响顾客对于服务的评价[2,3]. Png与Reitman对马萨诸塞州四个城市的加油站做了统计分析,结果表明等待时间与需求是负相关的[4].用等待因子(impatience factor)表示顾客的等待敏感性,等待因子越大,顾客对等待时间越敏感.把顾客从到达至拿到产品期间因等待而付出的成本叫做等待成本(delay cost),其中包括时间的机会成本以及等待所导致的其它物质上和精神上的成本.

考虑到顾客会将等待成本作为购买产品或服务的总成本的一部分来进行决策,当发生缺货时,由于需要等待,对于等待比较敏感的顾客就会离开,从而使企业的利润受到损失.对此,企业一方面可以从供应管理的角度出发,压缩提前期[5]或实行快速响应策略(quick response strategy)[6];另一方面还可以从需求管理的角度出发,利用定价策略来引导需求,根据不同等待时间制定不同的价格(如:当缺货发生时给与价格折扣,以补偿顾客等待的成本),以此取得更高的利润.

与本文相关的研究首先涉及到缺货时顾客被部分保留的模型. Montgomery等和Park等假设缺货时有固定比例的需求会再订货,即缺货保留率是一定的,这些文章优化了企业的库存策略[7,8]. Abad, Papachristos与Skouri以及San José等则假设缺货保留率与等待时间有关,不再是固定数值,但其仍然是外生变量[9~11].只有Decroix与Arreola-Risa考虑了在缺货时给予顾客补偿的方

① 收稿日期: 2007-05-28. 修订日期: 2008-04-18.
基金项目: 国家自然科学基金资助项目(70621061; 70518002).
作者简介: 陈 剑(1962—), 男, 福建人, 博士, 教授, 博士生导师. Email: chenj@sem.tsinghua.edu.cn.

法,试图控制缺货保留率[12].

目前运作管理领域对顾客基于时间选择行为的研究集中在服务(队列)定价领域,这些文献假设顾客对等待时间的敏感性不同,企业根据顾客的等待敏感性对优先级或者服务台进行差异定价.根据采用的价格机制不同,可以将这些文献分为两类. Naor, Mendelson 与 Wang, Bradford 等采用了直接定价的方法[13~15]. Lui, Afeche 与 Mendelson, Kittsteiner 与 Moldovanu 等研究了优先级拍卖机制[16~18]. Chen 和 Zhang 对这部分文献进行了详细的综述[19].可以看到,由于这部分文献针对的是服务队列,所以只研究了企业的定价策略,不涉及库存策略.

与以往缺货部分保留模型的研究不同,本文除了库存策略外,还分析了顾客的选择行为,研究了缺货时的差异定价策略,从而达到对缺货保留率的控制.与考虑顾客基于时间选择行为的文献相比,本文将研究范围从服务业扩展到制造业,除了优化差异定价策略外,同时还考虑了企业的库存策略.此外,由于在现实中,差异定价可能导致顾客采取投机的行为,我们还给出了防止投机行为的条件.文章采用了 Mendelson 与 Wang[14] 给出的顾客效用函数形式,并将其转化为成本函数形式,但没有假定等待成本是线性的,而采用更一般的假设,即效用是关于等待时间的凸函数.

主要符号如下.

S—— 企业库存水平上限,决策变量;

p_1—— 企业持有库存时产品的价格,决策变量;

p_2—— 企业缺货时产品的价格,决策变量;

p_0—— 产品的市场价格;

β—— 缺货时企业给与顾客的价格补偿,$\beta = p_0 - p_2$;

λ—— 顾客的最大期望到达率;

λ_1—— 企业持有库存时顾客的期望到达率;

λ_2—— 企业缺货时顾客的期望到达率;

θ—— 顾客的等待因子,服从$[\theta_{\min},\theta_{\max}]$上的分布 $\Phi(\cdot)$;

W—— 缺货时顾客的等待时间;

$C_d(\cdot,\cdot)$—— 等待成本,关于θ和W的函数;

c_v—— 单位可变成本;

c_p—— 单位需求丢失损失;

c_h—— 单位时间单位库存的持有成本.

1 模型描述

企业生产并出售一种产品,产品的价值及数量不随着时间发生改变.假设产品生产时间服从负指数分布,平均生产时间为 $1/\mu$,不失一般性,假设 $\mu = 1$. 企业采取连续检查库存的 $(S-1,S)$ 生产策略,即当库存数量到达 S 时停止生产,库存量低于 S 时生产.企业位于自由竞争的市场,此产品的市场价格为 p_0,即顾客总能够以价格 p_0 购买到类似的产品.当持有库存时,企业定价为 p_1.当发生缺货时,企业定价为 p_2,此时如果顾客选择等待,企业将按照先到先服务的方式供货.我们对企业的定价作如下的规定:1) 在有库存的时候,企业只以价格 p_1 销售产品;2) 缺货时只以缺货价格 p_2 销售产品;3) 企业只公布产品当前的价格.在这种定价方式下,不存在投机行为,即顾客等待至价格较低的时段再购买的行为.在第 4 节中,将放宽这个假设,考虑顾客的投机行为.

假设每个顾客一次购买一单位产品,其到达过程为 Poisson 过程,与企业的生产过程独立.顾客的最大期望到达率为 λ,实际到达率则与顾客的选择行为相关.顾客到达后首先观察到产品的价格,然后做出决策:在此企业处购买,或者离开以价格 p_0 到别的企业购买.顾客做出选择的依据是最大化自己的效用,包括购买产品获得的价值(或者顾客对于产品的估价)以及付出的成本(包括产品的价格与等待成本).显然,对于相同的产品,无论顾客在哪里购买该产品,其获得的价值(或对产品的估价)是相同的,因此顾客只需要考虑每种选择下付出的成本即可.当顾客选择离开时,其成本为 p_0;当顾客选择购买时,如果企业持有库存,顾客的成本为 p_1;如果发生缺货,我们假定顾客看不见缺货时排队的情况,因为排队队列(订单)是企业的私有信息,往往不公布,所以此时顾客需要计算期望成本 $E(C_2(\theta))$

$$E(C_2(\theta)) = p_2 + E(C_d(\theta,W)) \tag{1}$$

其中,W 是顾客从到达至拿到产品的等待时间,$C_d(\theta,W)$ 是等待成本.

等待因子 θ 表示顾客对等待的敏感性,等待因子越大,敏感性越高;它是顾客的私有信息,但

经过市场调查，企业可以得到其分布函数，假设顾客的等待因子独立同分布，在$[\theta_{\min},\theta_{\max}]$上服从分布$\Phi(\cdot)$.

假设 1　$\Phi(\cdot)$是连续二次可微的递增凹函数. 等待成本$C_d(\theta,W)$是关于θ和W的二次连续可微的递增凸函数，并且是关于(θ,W)的上模(super-modular)函数.

该假设表示，等待时间W越长或者等待因子θ越大，则等待成本越高；随着等待因子θ(或者W)的增加，边际等待成本递增；并且等待时间增加一单位时，顾客的等待因子越大，其等待成本增加的越多.

假设 2　$(p_0 - c_v)\lambda > c_h$.

该假设保证企业销售产品有利可图.

2　模型求解

模型的求解可以划分为三个步骤：1) 为了得到企业的利润函数，求解系统达到均衡状态时的指标，包括企业持有库存的概率、平均库存持有量等. 2) 在给定的库存和价格策略下分析顾客行为. 3) 将顾客行为代入企业的利润函数，求解最优的策略.

2.1　利润函数

引入$\{X(t),t\geqslant 0\}$来表示库存(排队等待的顾客)数量. 当$X(t)>0$时，其表示库存数量；否则$|X(t)|$，表示排队等候的顾客的数量. 由于我们假设生产时间和顾客到达间隔为负指数分布，$\{X(t),t\geqslant 0\}$显然是一个马尔可夫过程，其能够达到稳态的充要条件是$\lambda_2<1$. 定义$\rho_1 := 1/\lambda_1$，根据状态转移平衡方程，得到以下的指标.

1) 企业持有库存的概率$p_+ = p(X(t)>0)$.

$$P_+(S,\lambda_1,\lambda_2) = \frac{K(S,\lambda_1)}{K(S,\lambda_1)+L(\lambda_2)} \tag{2}$$

式中

$$L(\lambda_2) = \frac{1}{1-\lambda_2}$$

$$K(S,\lambda_1) = S\cdot 1(\rho_1 = 1) + \rho_1\frac{1-\rho_1^S}{1-\rho_1}\cdot 1(\rho_1\neq 1) \tag{3}$$

其中：$1(\cdot)$为示性函数，当ρ_1不同时$K(S,\lambda_1)$形式不同.

2) 单位时间平均库存持有量$E(I)$.

$$E(I) = G(S,\lambda_1)P_+(S,\lambda_1,\lambda_2) \tag{4}$$

其中，$G(S,\lambda_1)$是企业在有货状态下的平均库存，具体形式如下

$$G(S,\lambda_1) = 0\cdot 1(S=0) + \frac{S+1}{2}\cdot 1(S\geqslant 1,\rho_1 = 1) + \left(\frac{1}{1-\rho_1} - \frac{S\rho_1^S}{1-\rho_1^S}\right)\cdot 1(S\geqslant 1,\rho_1\neq 1) \tag{5}$$

得到上述参数之后，我们可以将企业的目标函数表示为

$$\begin{aligned}\pi(S,p_1,p_2) &= P_+((p_1-c_v)\lambda_1 - c_p(\lambda-\lambda_1)) - c_hE(I) + (1-P_+)((p_2-c_v)\lambda_2 - c_p(\lambda-\lambda_2)) \qquad (6)\\ &= P_+((p_1+c_p-c_v)\lambda_1 - c_hG) - c_p\lambda + (1-P_+)(p_2+c_p-c_v)\lambda_2\end{aligned}$$

显然，顾客的行为受到企业定价策略的影响，即：不同情况下顾客的期望到达率λ_1和λ_2与企业采取的价格策略相关. 在下面一节中，将分析价格策略与顾客选择行为的关系.

2.2　顾客选择行为

顾客到达时有两种选择，离开或者购买. 其期望成本如表 1 所示.

表 1　顾客各类选择的期望成本

Table 1 The customer's expected costs under different choices

选择行为	有库存状态	缺货状态
离开	p_0	p_0
购买	p_1	$E(C_2) = p_2 + E(C_d(\theta,W))$

如果企业制定价格$p_1>p_0$，那么在有库存状态下，直接购买是顾客的劣策略，所有的顾客都会离开. 在这种情况下，企业一旦持有库存，则永远处于没有需求的状态，这对企业是不利的，所以企业制定的价格一定满足$p_1\leqslant p_0$，此时所有到达的顾客都会购买产品. 而只要满足$p_1\leqslant p_0$，p_1越高，企业利润越高，所以最优定价$p_1^* = p_0$，此时期望到达率$\lambda_1^* = \lambda$.

如果顾客到达时企业处于缺货状态，那么顾客将首先估计等待时间. 已知某时刻缺货，由于生产时间与需求到达间隔都为负指数分布，所以顾客从到达至拿到产品的等待时间服从负指数

分布，

$$W(\lambda_2) \sim \exp(1/(1-\lambda_2)) \tag{7}$$

性质1 缺货时顾客的等待时间 $W(\lambda_2)$ 是关于 λ_2 的强随机递增凸函数.

证明 容易证明 $W(\lambda_2)$ 与 $\exp(1)/(1-\lambda_2)$ 同分布，而 $1/(1-\lambda_2)$ 是关于 λ_2 的递增凸函数，根据Shanthikumar与Yao(1991)对强随机递增凸性的定义[20]直接得到性质1. 证毕.

此时顾客的选择行为取决于自己的等待因子，等待因子较小的顾客会留下排队等候拿到产品，而对等待较为敏感，即等待因子较大的顾客，则会离开并以价格 p_0 在别处购买. 选择付出价格 p_2 并排队等候的顾客必然满足 $E(C_2) \leqslant p_0$，即，

$$p_2 + E(C_d(\theta, W)) \leqslant p_0 \tag{8}$$

定义 θ_β 表示缺货时愿意等待并购买产品的顾客的最高等待因子，即

$$\theta_\beta := \theta_{\max} \wedge \arg\{\theta: p_2 + E(C_d(\theta, W(\lambda_2))) = p_0\} \tag{9}$$

由于等待因子低于 θ_β 的顾客将购买产品，而其他顾客会离开，所以此时期望到达率为

$$\lambda_2 = \lambda\Phi(\theta_\beta) \tag{10}$$

显然，缺货时的期望需求 λ_2 是潜在需求 λ 的一部分，$\Phi(\theta_\beta)$ 相当于顾客保留率. 根据公式(9)，(θ_β) 是 λ_2 的函数，所以公式(10)的两边都包含 λ_2，是一个平衡方程，必须求解此方程才能将 λ_2 表示为价格 p_2 的函数. 为方便起见，将价格 p_2 表示为 λ_2 的函数，即得到缺货时的价格函数，这与得到需求函数是等价的. 为此定义

$$D(y) = \Phi^{-1}(y/\lambda), y \in [0, \lambda] \tag{11}$$

性质2 在 $[0,\lambda]$ 上，$D(y)$ 是关于 y 的递增凸函数，并且 $D(0) = \theta_{\min}$，$D(\lambda) = \theta_{\max}$.

性质2可以由假设1中 $\Phi(\cdot)$ 的性质直接得到. 根据公式(10)和(11)，$\theta_\beta = D(\lambda_2)$，所以 $D(y)$ 表示的是当期望到达率为 y 时，选择留下等待的顾客中的最高的等待因子. 再根据公式(9)，可以将缺货时的价格 p_2 表示为 λ_2 的函数，

$$p_2(\lambda_2) = p_0 - E(C_d(D(\lambda_2), W(\lambda_2))), \quad \lambda_2 \in [0, \lambda \wedge 1^-] \tag{12}$$

注意到公式(12)中，λ_2 的取值范围是 $[0, \lambda \wedge 1^-]$，这是为了保证系统可以达到稳态. 根据此价格函数，在 λ_2 的取值范围 $[0, \lambda \wedge 1^-]$ 内，λ_2 与 p_2 是一一对应的，可以将求解缺货时的价格 p_2 转化为求解缺货时的期望到达率 λ_2.

引理1 缺货时的价格函数 $p_2(\lambda_2)$ 是关于 λ_2 的递减凹函数.

证明 根据假设1、性质1和性质2，由Yao[21]得到，$C_d(D(\lambda_2), W(\lambda_2))$ 是关于 λ_2 的强随机递增凸函数. 根据Shanthikumar和Yao[20]，强随机递增凸性强于随机递增凸性，所以 $C_d(D(\lambda_2), W(\lambda_2))$ 是关于 λ_2 的随机递增凸函数. 根据Shaked与Shanthikumar对随机递增凸性的定义[22]，$E(C_d(D(\lambda_2), W(\lambda_2)))$ 是关于 λ_2 的递增凸函数. 所以 $p_2(\lambda_2)$ 是关于 λ_2 的递减凹函数. 证毕.

为了简化 $p_2(\lambda_2)$ 的形式，定义 $\beta(\lambda_2)$ 表示发生缺货时企业给与顾客的价格补偿，

$$\beta(\lambda_2) = E(C_d(D(\lambda_2), W(\lambda_2))), \quad \lambda_2 \in [0, \lambda \wedge 1^-] \tag{13}$$

从公式(13)可以看出，缺货时的补偿函数等于某个顾客的期望等待成本，这个顾客就是所有留下的顾客中等待因子最大的那个. 由于补偿与等待成本相同，此顾客无论选择留下还是离开，期望成本都是 p_0. 而其他留下的顾客，由于等待因子比较小，所以期望等待成本小于企业给予的补偿，选择留下将得到更高的效用.

假设3 $p_0 > \beta(0)$，$p_\pi \geqslant \beta(0)$.

该假设要求产品的市场价格大于只保留一个顾客时的补偿，否则为了在缺货时保留顾客，企业必须定价为负值；同时销售产品的利润与防止需求丢失损失之和不小于保留一个顾客的成本，否则给予缺货补偿将降低企业的利润. 由于 $p_1^* = p_0$，$\lambda_1^* = \lambda$，根据 $\beta(\lambda_2)$ 的定义，将公式(6)中企业单位时间的平均利润转化为库存水平 S 和缺货时的期望到达率 λ_2 的函数，即

$$\pi(S, \lambda_2) = P_+(S, \lambda, \lambda_2)(p_\pi\lambda - c_h G(S, \lambda)) - c_p\lambda + (1 - P_+(S, \lambda, \lambda_2))(p_\pi - \beta(\lambda_2))\lambda_2 \tag{14}$$

2.3 最优策略

由于持有库存时的最优价格 $p_1^* = p_0$，只需优化库存水平 S 和缺货价格 p_2 即可. 根据引理1及假设2与3，可以得到如下的定理.

定理1 当不存在投机行为时，企业的最优策略是唯一的. 如果参数不满足条件(15)，那么

最优的 $S^* = S^{c1}, p_2^* = p_0 - \beta(0)$. 如果参数满足条件(15)，则 $\lambda_2^{nc} > 0$，此时，(a) 如果 $\lambda_2^{nc} \leqslant \lambda$，则 $S^* = S^{nc}, p_2^* = p_0 - \beta(\lambda_2^{nc})$；(b) 如果 $\lambda_2^{nc} > \lambda$，则 $S^* = S^{c2}, p_2^* = p_0 - \beta(\lambda)$.

$$\hat{\pi}^* := \max_S \pi(S,0) < p_\pi - \beta(0) \tag{15}$$

$$S^{c1} = \inf\{S : \frac{d\pi(S,0)}{ds} = 0\} \tag{16}$$

$$\lambda_2^{nc} = \arg\{\lambda_2 : \frac{\partial\pi(S^{nc},\lambda_2)}{\partial\lambda_2} = 0\},$$

$$S^{nc} = \inf\{S : \frac{d\pi(S,\lambda_2^{nc})}{dS} = 0\} \tag{17}$$

$$S^{c2} = \inf\{S : \frac{d\pi(S,\lambda)}{dS} = 0\} \tag{18}$$

定理 1 的证明见附录. 分析定理 1 中的条件(15)，如果其满足，那么 $\lambda_2^* = \lambda_2^{nc} \wedge \lambda > 0$，否则 $\lambda_2^* = 0$，即缺货时没有顾客愿意等待(需求全部丢失)，这种情形相当于企业不提供缺货补偿，即不实行差异定价，所以条件(15) 可以看作企业采取差异定价策略的条件. 公式(15) 的左边是企业不提供缺货补偿时单位时间的最优利润 $\hat{\pi}^*$，注意到我们假设企业单位时间生产率 $\mu = 1$，所以 $\hat{\pi}^*$ 也可以表示不提供缺货补偿时单位产品的平均利润. 公式(15) 的右边 $p_\pi - \beta(0)$ 表示企业实行缺货补偿时保留第一个顾客的单位利润. 公式(15) 说明，企业是否采用差异定价，取决于出现缺货时保留第一个顾客的单位利润 $p_\pi - \beta(0)$ 与不提供缺货补偿时企业单位产品的最优利润 $\hat{\pi}^*$ 的关系. 如果前者大于后者，那么企业将采取差异定价策略，即在缺货时给予顾客一定的补偿，使一部分顾客选择留下等待；否则企业不会实行差异定价，此时缺货时的需求全部丢失.

推论 1 p_π 越大，单位持有成本 c_h 越大，顾客最小等待因子 $\theta_{\min}$ 越小，企业越倾向于采取差异定价策略，即在缺货时给予顾客价格补偿.

只需分析 p_π、c_h 和 $\theta_{\min}$ 变化时公式(15) 左右两边的变化量即可证明推论 1. 根据推论 1，差异定价策略适用的产品具有边际利润较高、需求丢失损失较高、持有成本较大的特点. 对于这类产品，企业需要在缺货时给予顾客一定的补偿来保留部分顾客，这一方面可以在缺货时增加销售收入，一方面可以降低过高的库存水平. 而如果顾客的最小等待因子较小，保留顾客变得较为容易，企业也会实行差异定价，以在缺货时保留顾客.

接下来分析最优策略的性质. 由于利润函数形式复杂，根据定理 1 难以直接得到最优策略的显示表达形式，但容易证明最优策略具有如下的性质.

推论 2 p_π 越大，最优库存水平 S^* 越大，最优缺货补偿 $\beta(\lambda_2^*)$ 越小；单位库存成本 c_h 越大，S^* 越小，$\beta(\lambda_2^*)$ 越大；当 θ 以随机序增加时，S^* 是非减的.

根据推论 2，p_π 越大，即产品边际利润越大、需求丢失损失越大，企业越倾向于在缺货时保留顾客，此时企业应该制定较高的库存上限，以防止缺货；而一旦发生缺货，企业将给予较高的价格补偿，以保留更多的顾客. 而单位持有成本越大，企业越不倾向于保持过高的库存水平，所以应该制定较小的库存上限；而当发生缺货时，应给予较高的补偿，以留住更多的顾客. 当 θ 以随机序增加时，相当于顾客整体对于等待变得更加敏感，此时企业不希望发生缺货，所以最优库存水平上升.

前面我们将库存水平 S 作为连续变量进行分析，而现实中 S 必须是整数. 如果根据定理 1 求得的 S^* 不是整数，那么最优的库存水平应该为最接近 S^* 的整数. 记 S^* 向下取整为 $\lfloor S^* \rfloor$，向上取整为 $\lceil S^* \rceil$，企业需要比较库存水平为 $\lfloor S^* \rfloor$ 和 $\lceil S^* \rceil$ 时利润的大小，取能够带来较高利润的库存水平，并制定相应的缺货时的价格 p_2^*. 如果两者带来的利润相同，那么按照企业的思维，取较小的 $\lfloor S^* \rfloor$. 考虑整数约束后，推论 2 中最优策略的递增(递减) 性质需要修改为非减(非增) 性质.

3 存在投机行为的模型

第 2 节的模型假设企业只公布产品当前的价格，所以不存在顾客的投机行为. 本节则假设有库存和缺货时的价格是事先公布的，顾客可能根据价差采取投机行为. 考虑到问题的复杂性，这里我们只考虑等待成本为线性的情形. 如果两类价格是事先公布的，那么顾客到达时有三种选择：(1)

离开;(2) 购买;(3) 等待至下一状态以下一状态的价格购买.

顾客只有在有利可图的情况下才会投机,但是要等待至下一状态发生才能实施,其期望等待时间一定大于直接购买的期望等待时间,所以只有当下一状态的价格较低时,顾客才会投机. 对于企业来说,以较低的价格在未来销售,其收益小于当时以较高的价格卖出产品,所以企业不希望出现顾客投机现象. 为了最大化自己的期望利润,企业不会设置有货时的价格低于缺货时的价格,所以缺货时不存在投机行为. 但为了补偿顾客的等待,企业可能会设置缺货时的价格低于有货时的价格,所以当企业持有库存时,某些顾客可能会采取投机行为. 此时假如只有一个顾客选择投机,记其期望等待时间为 w_s,期望成本 $E(C_s) = p_2 + \theta w_s$.

3.1 防止投机行为的条件

防止投机行为只需保证顾客投机时的期望成本不低于不投机时的期望成本即可,即 $E(C_s) \geqslant p_1$. 为了得到 $E(C_s)$ 的形式,首先必须得到 w_s. w_s 是顾客到达时企业持有库存,但其等待至缺货时才购买,并进一步等待拿到货物的期望时间,可以分为两部分,一部分是顾客从到达(此时企业有货) 等待至缺货,即库存为 0 的时间. 当库存降为 0 时,产品价格降为 p_2,由于只有此顾客投机,一旦企业库存降为 0,顾客就会购买并等待企业生产;另一部分是等待企业生产的时间. 由于此顾客是缺货发生时最先到达的顾客,所以等待生产的期望时间为 $1/\mu = 1$. 显然,w_s 的第一部分与企业最高库存水平 S 有关,所以首先给定 S,求解 w_s.

引理2 给定库存水平 S,只有一个顾客投机时,其期望等待时间 $w_s(S)$ 为公式(19) 所示,其中 $\rho := 1/\lambda$, $1(\cdot)$,为示性函数.

$$w_s(S) = \frac{1}{1-\lambda}\left(\begin{matrix}\dfrac{2\rho^{S+2}}{1-\rho^2}\left(\dfrac{S}{K(S,\lambda)}-1\right)\\ -G(S,\lambda)\end{matrix}\right)\times 1(S>0,\lambda\neq 1) + \frac{(S+1)(4S-1)}{12}\times 1(S>0,\lambda=1)+1 \tag{19}$$

引理 2 的证明见附录. 为防止投机行为,使顾客行为达到 Nash 均衡的充分必要条件是:对于所有顾客,即 $\forall\theta\in[\theta_{\min},\theta_{\max}]$,满足 $p_2+\theta w_s(S)\geqslant p_1$. 而由于每个顾客投机时期望等待时间 $w_s(S)$ 是相同的,所以如果等待因子最小的顾客不选择投机,那么就没有顾客会投机. 这样我们就得到了达到 Nash 均衡的充要条件,如引理 3 所示.

引理3 避免顾客的投机行为,使顾客达到 Nash 均衡的充要条件是,S、p_1 和 p_2 满足,

$$p_2 + \theta_{\min} w_s(S) \geqslant p_1 \tag{20}$$

3.2 最优策略

无投机行为模型中 $p_1^* = p_0$,所以如果无投机行为模型的最优值满足 $p_2^* + \theta_{\min} w_s(S^*) \geqslant p_0$,那么根据公式(20),该最优值满足防止投机的约束,也是投机行为模型中的最优解. 但是如果最优解不满足 $p_2^* + \theta_{\min} w_s(S^*) \geqslant p_0$,那么就需要重新求解. 投机行为模型中的最优值记为 $\tilde{S}_1^*$, $\tilde{p}_1^*$ 和 $\tilde{p}_2^*$. 给定 p_2 与 S, p_1 必须同时满足约束(20) 和 $p_1 \leqslant p_0$,而由于企业必然制定最优价格 $\tilde{p}_1^*$ 为其约束的上限,所以企业有两种策略.

策略 1 $\tilde{p}_1^* = p_0$

此时根据约束(20),$\tilde{p}_2^*$ 与 $\tilde{S}^*$ 必须满足 $\theta_{\min} w_s(\tilde{S}^*) + \tilde{p}_2^* \geqslant p_0$,其中 $\tilde{p}_2^* = p_0 - \beta(\tilde{\lambda}_2^*)$,$\tilde{\lambda}_2^*$ 和 $\tilde{S}^*$ 为模型(21) 的最优解,

$$\begin{aligned}\max_{S,\lambda_2}\tilde{\pi}(S,\lambda_2) =& P_+(S,\lambda,\lambda_2)(p_\pi\lambda - c_h G(S,\lambda)) + \\ & (1-P_+(S,\lambda,\lambda_2))(p_\pi-\beta(\lambda_2))\lambda_2 - c_p\lambda \\ \text{s.t.}\quad & \theta_{\min} w_s(S) \geqslant \beta(\lambda_2)\end{aligned} \tag{21}$$

策略 2 $\tilde{p}_1^* = \tilde{p}_2^* + \theta_{\min} w_s(\tilde{S}^*)$

此时约束(20) 满足,但由于 $\tilde{p}_1^* \leqslant p_0$,所以 $\tilde{p}_2^*$ 与 $\tilde{S}^*$ 必须满足,$\theta_{\min} w_s(\tilde{S}^*) + \tilde{p}_2^* \leqslant p_0$ 其中,$\tilde{p}_2^* = p_0 - \beta(\tilde{\lambda}_2^*)$,$\tilde{\lambda}_2^*$ 和 $\tilde{S}^*$ 为模型(22) 的最优解,

$$\begin{aligned}\max_{S,\lambda_2}\tilde{\pi}(S,\lambda_2) =& P_+(S,\lambda,\lambda_2)\times \\ & \left(\begin{matrix}(p_\pi-\beta(\lambda_2)+\theta_{\min}w_s(S))\lambda \\ -c_h G(S,\lambda)\end{matrix}\right) + (1- \\ & P_+(S,\lambda,\lambda_2))(P_\pi-\beta(\lambda_2))\lambda_2 - c_p\lambda \\ \text{s.t.}\quad & \theta_{\min} w_s(S) \leqslant \beta(\lambda_2)\end{aligned} \tag{22}$$

策略 1 与策略 2 比较

如果无投机行为模型中的最优值不满足约束(20),那么根据定理 1 及 $w_s(S)$ 的递增性容易证明策略 1 的最优值一定满足 $\theta_{\min} w_s(\tilde{S}^*) =$

$\beta(\tilde{\lambda}_2^*)$. 而 $\theta_{\min}w_s(\tilde{S}^*) = \beta(\tilde{\lambda}_2^*)$ 满足策略2的约束条件 $\theta_{\min}w_s(S) \leqslant \beta(\lambda_2)$，并且此时策略2中有库存时的最优价格 $\tilde{p}_1^* = p_0 - \beta(\tilde{\lambda}_2^*) + \theta_{\min}w_s(\tilde{S}^*) = p_0$，所以策略1的最优值位于策略2的可行域中，策略2优于策略1. 由此得到定理2.

定理2　如果无投机行为模型的最优解满足 $p_2^* + \theta_{\min}w_s(S^*) \geqslant p_0$，那么 $p_1^* = p_0$、S^* 与 p_2^* 也是投机行为模型的最优解. 否则，投机行为模型的最优解为，$\tilde{p}_1^* = p_0 - \beta(\tilde{\lambda}_2^*) + \theta_{\min}w_s(\tilde{S}^*)$，$\tilde{p}_2^* = p_0 - \beta(\tilde{\lambda}_2^*)$，其中 $\tilde{\lambda}_2^*$ 与 $\tilde{S}^*$ 为模型(22)的最优解.

3.3　数值分析有库存时的最优价格

在第3节中，证明有库存时的最优价格 $p_1^* = p_0$. 而在存在投机行为模型中，根据定理2，企业可能会制定 $\tilde{p}_1^* \leqslant p_0$，以防止顾客投机. 下面我们采用数值分析方法简单分析当单位持有成本 c_h 和顾客的最小等待因子 $\theta_{\min}$ 变化时，存在投机行为模型中企业有库存时最优价格的变化情况. 数值实验参数设置为 $\lambda = 1.4$，$p_0 = 500$，$c_v = 60$，$c_p = 60$，θ 在 $[\theta_{\min}, \theta_{\min} + 50]$ 上服从均匀分布，$C_d(\theta, W) = \theta W$.

图1中，固定 $\theta_{\min} = 20$；图2中，固定 $c_h = 100$. 分析图1和图2，当单位持有成本 c_h 较小和顾客最小等待因子 $\theta_{\min}$ 较大时，$\tilde{p}_1^* = p_0$. 比较 $\theta_{\min}w_s(\tilde{S}^*)$ 与 $\beta(\tilde{\lambda}_2^*)$，可以发现此时 $\theta_{\min}w_s(\tilde{S}^*) \geqslant \beta(\tilde{\lambda}_2^*)$，根据引理3可知，此时不存在投机行为. 而当单位持有成本 c_h 较大和顾客最小等待因子 $\theta_{\min}$ 较小时，有库存时的最优价格低于市场价格，即 $\tilde{p}_1^* < p_0$，这是为了满足引理3的Nash均衡约束. 分析图1，根据推论2，单位持有成本 c_h 越大，企业越倾向于制定较低的库存水平和较低的缺货价格，根据引理3约束条件(20)，这样越可能造成投机，所以单位持有成本较大时，企业必须降低有库存时的价格，以防止顾客采取投机的行为. 分析图2，当顾客的最小等待因子 $\theta_{\min}$ 较小时，根据推论2，企业的最优库存水平较低，在这种情况下，从有库存到缺货的时间相对较短，此时容易发生投机行为，所以企业需要制定较低的有库存时的价格.

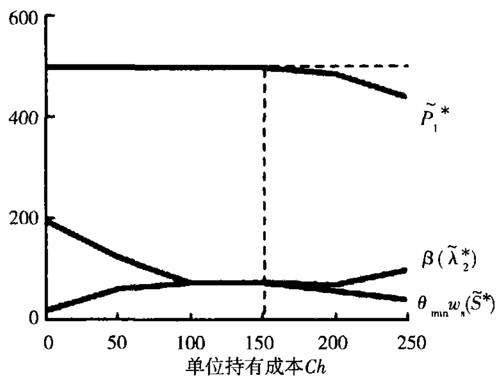

图1　最优有货价格关于单位持有成本的变化

Fig. 1 The optimal in-stock price as a function of c_h

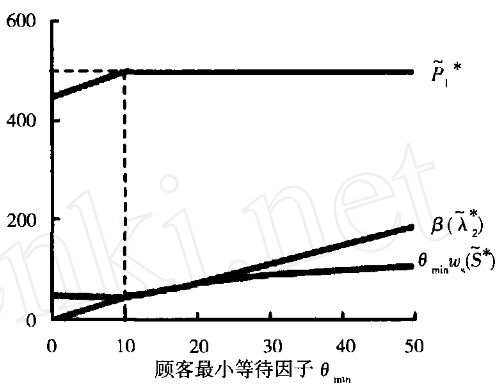

图2　最优有货价格关于最小等待因子的变化

Fig. 2 The optimal in-stock price as a function of $\theta_{\min}$

比较无投机行为模型与存在投机行为模型下有库存时的最优价格. 由于存在市场价格，只要有库存时的价格不高于市场价格，所有的顾客都会购买产品. 所以在无投机行为模型中，有库存时的最优价格为市场价格. 但是当存在投机行为时，为了防止顾客等待至缺货再购买的行为，当单位持有成本较高时或者顾客的最小等待因子较小时，企业应制定低于市场价格的有货价格，这虽然不能增加有货时的需求，但是可以放宽发生投机行为的条件，从而增加缺货时的需求，提高总的利润. 所以企业在制定价格策略时，不能只考虑当前的需求，还应该综合考虑当前需求和未来缺货需求之间的关系.

4　结　论

本文针对等待敏感型顾客，研究了自由竞争市场上企业的缺货补偿与库存策略. 文章的主要结论包括：(1) 不考虑投机行为时，企业应该制定

有库存时的价格为市场价格，缺货时给予的价格补偿应该等于留下等待的顾客中等待因子最大的顾客的等待成本；等待因子较小的顾客选择留下，等待因子较大的顾客将离开.（2）不考虑投机行为时，产品单位利润较高、需求丢失的损失较高时，企业应该制定较高的库存水平，同时在缺货时给与较高的价格补偿；而产品持有成本较高时，企业应该制定较低的库存水平，并且在缺货时制定较低的价格.（3）不考虑投机行为时，产品的市场价格越高、单位可变成本越低、需求丢失成本越高、单位持有成本越高，企业越倾向于采用缺货差异定价策略；反之，企业不应该采取差异定价策略.（4）考虑投机行为时，当单位持有成本较大或者顾客的最小等待因子较小时，企业应该制定低于市场价格的有货价格，以放宽防止投机行为的条件.

文章可以从以下几个方面进行扩展. 最直接的扩展就是考虑更多的库存模型. 本文考虑的企业采取的是 $(S-1,S)$ 库存策略. 而在现实中，很多企业会采取 (Q,r) 和 (s,S) 策略，生产时间负指数分布的假设也会放宽. 在这些库存模型中，如何制定相应的库存策略，并在缺货时制定合理的价格，是值得探讨的问题. 此外，本文假设企业的库存和排队状况是不可见的，顾客到达时需要估计等待时间进行决策. 而在有的企业中，排队状况是可见的，或者企业会公布排队状况，那么顾客将根据到达时的排队情况做出决策，此时的需求函数将与本文不同，这必将影响企业的最优策略. 所以企业可以通过控制等待信息的公开程度来优化利润. Guo 与 Zipkin 初步研究了等待信息的公开程度对企业利润的影响[23]，文章针对的是 M/M/1 排队系统，并且只探讨了定价策略. 而将此问题引入到更为一般的排队系统，并进一步考虑定价与库存结合策略，还有非常多的工作可以做.

附录

定理 1 的证明 定理 1 中求解的策略是 S 和 p_2，但由于 p_2 与 λ_2 在 λ_2 的取值域内存在一一对应关系，所以只需优化 S 与 λ_2，即求解模型(14)即可.

1）给定 S，证明函数 $\pi(S,\lambda_2)$ 是关于 λ_2 的单峰函数. 我们考虑单峰函数的一种特殊形式，即当 $\pi(S,\lambda_2)$ 关于 λ_2 的一阶偏导小于等于 0 时，二阶偏导小于 0. 将一阶偏导小于等于 0 的条件带入二阶偏导，根据引理 1，容易得到此结论. 此时，如果 $\lambda_2=0$ 时利润 $\pi(S,\lambda_2)$ 关于 λ_2 的偏导小于等于 0，那么企业的利润随着 λ_2 的增加而一直减小，最佳的 $\lambda_2^*(S)=0$；否则，一阶偏导开始时大于 0，在 $\lambda_2=1^-$ 处小于 0，$\lambda_2^*(S)=\lambda_2^{nc}(S)\wedge\lambda$，其中 $\lambda_2^{nc}(S)=\arg\{\lambda_2:\partial\pi(S,\lambda_2)/\partial\lambda_2=0\}$.

2）证明 $\pi(S,\lambda_2^*(S))$ 是先增加后减少的单峰函数.（a）假设 $\lambda_2^*(S)=\lambda_2^{nc}(S)$，与第(1)部分证明过程相同，容易证明当利润 $\pi(\lambda_2^{nc}(S),S)$ 关于 S 的一阶导数等于 0 时，二阶导数小于 0，即 $\pi(S,\lambda_2^{nc}(S))$ 是关于 S 的单峰函数.（b）假设 $\lambda_2^*(S)=\lambda$ 或 $\lambda_2^*(S)=0$，此时由第(2)部分的证明容易得到 $\pi(S,0)$ 和 $\pi(S,\lambda)$ 也是单峰函数.（c）由于 $\pi(S,\lambda_2^{nc}(S))$、$\pi(S,0)$ 和 $\pi(S,\lambda)$ 都是单峰函数，$\pi(S,\lambda_2^{nc}(S))\geqslant\pi(S,0)$，$\pi(S,\lambda_2^{nc}(S))\geqslant\pi(S,\lambda)$，所以 $\pi(S,\lambda_2^*(S))$ 也是单峰函数. 根据假设 2 和假设 3，此单峰函数在 $S=0$ 处的导数大于 0，在 $S\to\infty$ 处的导数小于 0，所以必然与 0 有交点，而且一旦导数下降为 0，将一直小于等于 0，即利润一直下降，所以最优值应取在一阶条件为 0 的最小处.

3）分析 $\lambda_2^*=0$ 的条件. 容易证明只有当 $\lambda_2^{nc}(S)=0$ 时，$\pi(S,\lambda_2^{nc}(S))$ 与 $\pi(S,0)$ 才有交点，所以 $\lambda_2^*=0$ 的充要条件是，$\pi(S,\lambda_2^{nc}(S))$ 与 $\pi(S,0)$ 有交点，即

$$\exists S,\text{s. t.}\frac{\partial\pi(S,\lambda_2)}{\partial\lambda_2}\Big|_{\lambda_2=0}=-\frac{K(S,\lambda)}{(K(S,\lambda)+1)^2}(p_\pi\lambda-c_hG(S,\lambda))+\frac{1}{K(S,\lambda)+1}(p_\pi-\beta(0))=0$$

上式等价于，$\exists S$ 使得

$$\frac{K(S,\lambda)}{K(S,\lambda)+1}(p_\pi\lambda-c_hG(S,\lambda))=\pi(S,0)=p_\pi-\beta(0)$$

而根据第(2)部分的证明，$\pi(S,0)$ 是先增后减的单峰函数，所以 $\lambda_2^*=0$ 等价于

$$\max_S\pi(S,0)=\hat{\pi}^*\geqslant p_\pi-\beta(0)$$

证毕.

引理2的证明　首先求解w_s的第一部分w_s^1，如果$s=0$,那么$w_s^1=0$;否则,当系统有i个库存时,库存数量回到0的期望步数$N(i)$满足下面的关系

$$\begin{cases} N(0)=0 \\ N(1)=1+\dfrac{\lambda}{\mu+\lambda}N(0)+\dfrac{\mu}{\mu+\lambda}N(2) \\ \cdots \\ N(i)=1+\dfrac{\lambda}{\mu+\lambda}N(i-1)+\dfrac{\mu}{\mu+\lambda}N(i+1) \\ \cdots \\ N(S)=1+N(S-1) \end{cases}$$

根据上式,可以求得

$$N(i)=\left(\frac{1+\rho}{1-\rho}i-\frac{2(1-\rho^i)}{(1-\rho)^2}\rho^{S-i+2}\right)\cdot 1(\lambda\neq 1)$$

$$+(2iS-i^2)\cdot 1(\lambda=1)$$

而每步发生的时间长度为$1/(\mu+\lambda)$,所以顾客到达时系统中有i个库存时,顾客等待至企业库存为0的期望时间$w_s^1(X(t)=i)$为

$$w_s^1(X(t)=i)=\frac{1}{1+\lambda}N(i)$$

所以

$$w_s^1=\frac{1}{1+\lambda}\sum_{i=1}^{S}N(i)P(X(t)=i\mid X(t)>0)$$

$$=\frac{1}{1-\lambda}\begin{pmatrix}\dfrac{2\rho^{S+2}}{1-\rho^2}\left(\dfrac{S}{K(S,\lambda)}-1\right)\\ -G(S,\lambda)\end{pmatrix}\cdot 1(\lambda\neq 1)$$

$$+\frac{(S+1)(4S-1)}{12}\cdot 1(\lambda=1)$$

而$w_s(S)=w_s^1+1$,得证.　　证毕.

参 考 文 献:

[1]Becker G S. A theory of the allocation of time[J]. Economic Journal, 1965, 75(299): 493—517.

[2]Maister D. Psychology of Waiting Lines[M]. Harvard Business School Cases. Boston: Harvard University, 1984. 71—78.

[3]Taylor S. Waiting for service: The relationship between delays and evaluations of service[J]. Journal of Marketing, 1994, 58(2): 56—69.

[4]Png I P L, Reitman D. Service time competition[J]. RAND Journal of Economics, 1994, 25(4): 619—634.

[5]宋华明, 马士华. 二阶段供应链中提前期压缩的影响与协调[J]. 管理科学学报, 2007, 10(1): 46—53.
Song Huaming, Ma Shihua. Effect and coordination of lead-time compression in two-echelon supply chain[J]. Journal of Management Sciences in China, 2007, 10(1): 46—53. (in Chinese)

[6]鲁其辉, 朱道立, 林正华. 带有快速反应策略供应链系统的补偿策略研究[J]. 管理科学学报, 2004, 7(4): 14—23.
Lu Qihui, Zhu Daoli, Lin Zhenghua. Return policies in supply chain system with quick response strategy[J]. Journal of Management Sciences in China, 2004, 7(4): 14—23. (in Chinese)

[7]Montgomery D C, Bazaraa M S, Keswani A K. Inventory models with a mixture of backorders and lost sales[J]. Naval Research Logistics Quarterly, 1973, 20(2): 253—255.

[8]Park K S. Inventory model with partial backorders[J]. International Journal of Systems Sciences, 1982, 13: 1313—1317.

[9]Abad P L. Optimal price and order size for a reseller under partial backordering[J]. Computers and Operations Research, 2001, 28(1): 53—65.

[10]Papachristos S, Skouri K. An inventory model with deteriorating items, quantity discount, pricing and time-dependent partial backlogging[J], International Journal of Production Economics, 2003, 83: 247—256.

[11]San José L A, Sicilia J, García-Laguna J. An economic lot-size model with partial backlogging hinging on waiting time and shortage period[J]. Applied Mathematical Modeling, 2007, 31: 2149—2159.

[12]Decroix G A, Arreola-Risa A. On offering economic incentives to backorder[J]. IIE Transactions, 1998, 30: 715—721.

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

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

[15]Bradford R M. Pricing, routing and incentive compatibility in multi-server queues[J]. European Journal of Operational Research, 1996, 89(2): 226—236.

[16] Lui F. An equilibrium queuing model of bribery[J]. Journal of Political Economy, 1985, 93(4): 760—781.

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

[18] Kittsteiner T, Moldovanu B. Priority auctions and queue disciplines that depend on processing time[J]. Management Science, 2005, 51(2): 236—248.

[19] Chen J, Zhang N. Customer Incentives in Time-based Environment[M]. Service Enterprise Integration. NY: Springer, 2006. 103—129.

[20] Shanthikumar J G, Yao D D. Strong stochastic convexity: Closure properties and applications[J]. Journal of Applied Probability, 1991, 28(1): 131—145.

[21] Yao D D. Stochastic Modeling and Analysis of Manufacturing Systems[M]. New York: Springer-Verlag, 1994.

[22] Shaked M, Shanthikumar J G. Stochastic convexity and its applications[J]. Advances in Applied Probability, 1988, 20: 427—446.

[23] Guo P, Zipkin P. Analysis and comparison of queues with different levels of delay information[J]. Management Science, 2007, 53: 962—970.

Study on backorder incentives and inventory control policies with time-based customer-choice behavior

CHEN Jian[1,2], *ZHANG Nan*[1,2]

1. Research Center for Contemporary Management, Key Research Institute of Humanities and Social Sciences at Universities, Tsinghua University, Beijing 100084, China;
2. School of Economics and Management, Tsinghua University, Beijing 100084, China

Abstract: We study the differential pricing and inventory control policies with time-based customer-choice behavior in a continuous-review setting, where production and demand arrival processes are stochastic. The customer's utility is decreasing in both of the price and the waiting time. To maximize his steady-state average profit, the firm determines a base-stock level, an in-stock price and an out-stock price. However, the differential pricing mechanism may cause speculation. We first present the optimal pricing mechanism and base-stock level when speculation does not exist. Then we derive the conditions under which the firm should adopt the differential pricing strategy. Finally, we provide the conditions to prevent speculation and derive the optimal policy.

Key words: customer-choice behavior; delay cost; backorder; differential pricing; inventory control