

知情交易与中国股市博彩溢价*

孔东民 代昀昊 李捷瑜

[摘要]本文发现中国股市存在博彩(投机)溢价,且无法为 Fama-French 三因子模型解释。尽管在组合构造期内,博彩型股票存在显著溢价,但组合的超额收益会迅速消失,并未给投资者(或投机者)带来持续的财富效应。基于知情交易概率测度,我们进一步发现中国股市的知情交易者驱动(或引发)了博彩型股票溢价;在组合构造后的月份,并没有明显的知情交易者存在,这意味着知情交易者在基于私人信息获利之后,那些随后进入市场的投资者(动量交易者)无法获得超额收益。

关键词: 博彩溢价 知情交易概率 三因子模型

JEL 分类号: G10 G12 G14

一、导 言

博彩(或投机)心理在人们的生活中根深蒂固^①,心理学研究对此也提供了充分的证据(Holliday and Fuller, 1975; Brenner, 1983; Walker, 1992)。就彩票而言,因为极小的中奖概率投资者在彩票上的支出成本远远超过预期回报(多数彩票的返还率只在 60%左右)。按照经典的教科书理论,仅仅通过期望收益和投资风险的权衡,理性的投资者不会参与彩票的购买。这显然与现实各种彩票盛行的事实相矛盾。因此,必然还有一些额外的因素吸引了投资者。Golec and Tamarkin(1998)认为即使是风险规避的投资者,也可能会牺牲一些收益来换取一个赢得巨额回报的机会,因为这种潜在的巨额回报给投资者提供了额外的效用。

格林斯潘(前美联储主席)在针对上世纪 90 年代末的互联网泡沫中公司股价过高以及市场炒作的现象,曾提到“博彩溢价”(lottery premium)的概念(Wessel, 1999)。他的这一说法主要指投资者在互联网狂热中对投机性更强的股票过多关注,从而把价格推高的现象。Statman(2002)也认为股市和博彩一样属于“负和游戏(negative-sum game)”,平均而言投资者是亏损的。对于人们为何会购买股票或者彩票的行为,Statman 从行为金融学的角度结合投机和赌博给出了一些逻辑上的探讨。

尽管投资者的博彩或投机行为的相关研究对于我们从理论和现实两方面理解股票市场运行、投资者行为和制度监管有重要意义,但受限于合适的股票分类方法或投资组合构建策略,一直未见针对“博彩溢价”的严肃学术研究。正如 Kumar(2009)所说的那样,第一个原因在于基于赌博动机的投资决策的证据很难收集,这是因为人们的博彩偏好以及资产组合的投资决策不容易被观察到。第二个原因在于准确界定何为博彩型股票也存在困难。

在本文中,我们借鉴 Kumar(2009)对于小投资者参与博彩型股票的研究,把市场个股划分为三个不同类型,即博彩型股票、非博彩型股票及其它类型。其中,博彩型股票能够对基于博彩动机

* 孔东民,华中科技大学经济学院,硕士生导师;代昀昊,华中科技大学经济学院,硕士研究生;李捷瑜,中山大学岭南学院,硕士生导师。作者感谢在华中科技大学经济学院举办的研讨会上各位提出意见的参与者。本文得到国家自然科学基金(70803013)的资助。

① 除非特别说明,否则本文将在大致相同的意义上使用赌博、博彩与投机这几个词。Borna and Lowry(1987)在专门讨论赌博与投机的文章中无差别地使用了这些说法。也可见 Han, Lee, and Liu(2009), Kumar(2009)等。

的投资进行较准确的定义。具体而言,Kumar(2009)从彩票的特征出发,根据股票的三个维度(价格、偏度和波动)排序,然后寻找出那些同时具有最低价格、最大偏度以及最高波动性的股票来代理具有博彩特征的股票。^①正是基于这样一种划分,我们可以对中国股市的投机型(或博彩型)股票的溢价进行研究。

在博彩型股票的分类基础上,我们进一步引入 Easley, Kiefer, O'Hara and Paperman(EKOP, 1996)的知情交易概率测度 *PIN*(probability of informed trading),考察知情交易在股市的投机行为中扮演了什么角色,以此考察博彩动机与投资者决策以及知情交易之间的关系。

本研究的主要贡献如下:第一,我们引入 Kumar(2009)的股票分类方法,发现中国股市存在显著的“博彩溢价”且无法被常见的因子模型解释(如 Fama-French 三因子模型),从而把格林斯潘(1999)所提到的“博彩溢价”这一说法,首次从更为严谨的角度给出证据;第二,结合 EKOP(1996)的知情交易概率测度,我们进一步考察了知情交易者在股市投机(博彩)中的行为及角色。最后,研究结果对中国股市的投机行为和投资决策提供了直接的借鉴。

全文的安排如下:第二部分讨论相关文献;第三部分构建研究假说;第四部分给出研究方法和数据说明;第五部分报告检验结果并分析;最后给出结论。

二、文献回顾

考虑到目前并未见基于博彩行为与知情交易二者相结合的研究,本文从这两个角度分别进行文献回顾。

我们参照 Kumar(2009)的做法,根据股票的三个维度(价格、偏度和波动),寻找出那些同时具有最低价格、最大偏度以及最高波动性的股票来代理具有博彩特征的股票。对于这三个博彩的维度而言,考虑到波动性向来是代理风险传统变量,因此博彩型或者投机型股票的风险较高的特征在一定程度上可以用波动来代理。关于此点,不再赘述。

对于股票价格,最近也有些研究证据显示投资者会在投资时把价格高低作为分类标准之一。例如 Baker, Greenwood, and Wurgler(2009)认为公司管理层有动力把股票分拆,通过价格的降低迎合那些给低价股更高估值的投资者,随后的实证验证了这一假说。Green and Hwang(2009a)同样发现投资者确实按照价格对股票进行分类,而且不同价格组合的收益之差无法被其它因子所解释。尽管 Kumar(2009)并未引用有关股价影响投资者分类的研究,但他的研究在直觉上提到这一点。

最近,逐渐有一些研究开始关注股票收益率的偏度问题。例如 Hong and Stein(2003)基于行为金融模型推导出收益率对于更高的交易量会呈现出负偏的特征。Zhang(2005)基于月度数据发现股票的偏度对于股票的收益具有很强的预测性。Green and Hwang(2009b)则发现对于 IPO 的股票而言,期望偏度更高的个股有着更高的首日回报率,而且他们进一步发现这一类股票往往更多地由机构投资者手中转移到个人投资者手中,这说明个人投资者更偏好具有博彩特征的股票。

整体上来说,关于股票博彩特征(价格分组、偏度预测性)的研究只是在近期才有所发展(主要得益于一些风格特征、分类标准、博彩特征的逐渐清晰)。就我们所知,有关的文献尚不多见,国内的相关研究则更是缺乏。

对于知情交易而言,EKOP(1996)主要是基于 Easley and O'Hara(1987, 1992)的序贯交易模型提出了估计知情交易概率 *PIN*的模型,用于衡量利用私有信息进行交易的概率。*PIN*反映了各类利用私有信息进行的交易(Jayaraman, 2008),*PIN*值越低,知情交易概率越低,信息对称程度越高。

^① 对于这样一种分类,Kumar 给出了详细的分析和阐释;投资组合的构建步骤则见后文。

Easley, Hvidkjaer, and O'Hara(2002)结合 Fama and French(1993)有关资产定价体系的三因素模型,按照 EKOP 提出的 *PIN*模型计算出知情交易概率,把它作为回归的第四个因子后,发现知情交易概率与价格正相关而且显著性很高,说明知情交易概率越高,所要求的风险补偿也越高,所以他们认为 *PIN*作为一种风险因子是有意义的,目前大量研究引用这一发展较为成熟的测度代理知情交易概率^①。自从 EKOP(1996)提出知情交易概率的估计方法后,*PIN*测度已经在文献中得到广泛运用(如 Easley, Kiefer and O'Hara, 1996; Easley, O'Hara and Paperman, 1998; Easley, Hvidkjaer and O'Hara, 2002; Easley and O'Hara, 2004)。

国内对知情交易概率的研究,主要有杨之曙、姚松瑶(2004)利用 *PIN* 测度发现交易活跃的股票的知识交易概率比不活跃的股票要低一些。私人信息对交易不活跃的股票影响更大。他们进一步发现交易活跃与交易不活跃股票之间的买卖价格差可以用 *PIN*来解释。刘元海、陈伟忠(2004)在序贯交易模型基础上对中国股市知情交易概率进行测定。王春峰、董向征和房振明(2005)将知情交易概率作为衡量信息不对称程度的一种直接指标,发现我国股市知情交易概率具有和流动性正相关,波动性负相关的特性。攀登、施东晖(2006)新建了一个关于知情交易概率的模型,发现知情交易确实会影响股价波动;知情交易者会被好的市场质量吸引,但随后又会降低市场质量;知情交易概率的日内行态为倒J型。王展翔(2007)发现我国股市具有较高的信息不对称,但由于流动性较高且知情交易者对坏消息的处理能力有限,所以总体上知情交易比率并不高。韩立岩、郑君彦和李东辉(2008)发现 *PIN*对价格变动存在影响,但对收益产生了负效应。张宗新(2008)也基于 *PIN*测度,发现在我国股权分置改革基本完成的背景下,上市公司很多重大事件的背后都有着私人信息交易行为,存在内幕交易的可能性很高。孔东民、申睿(2008)也利用 *PIN*测度发现中国股市信息效率较低股票有更高的知情交易。

整体上来说,国内基于 *PIN*的研究也发现了该测度可以在一定程度上代理知情交易行为,但目前大多数研究都基于金融市场微观结构的角度进行,我们则是利用知情交易与本文所发现的博彩溢价相结合,考察知情交易与市场投机的关系。

三、研究假说

在引言部分,我们讨论了彩票和人们的博彩心理。而在以流动性著称的股票市场,博彩(投机)更是屡见不鲜。Barberis and Huang(2008)构建了理论模型,他们基于心理学的累积前景理论(cumulative prospect theory)(Tversky and Kahneman, 1992),从博彩行为的角度研究股票市场中投资者基于风险的决策问题。事实上,仅仅通过我们的直觉感受,也很容易发现投机或赌博行为确实会明显影响到投资者在买卖股票中决策行为。例如,很多股市投资者会用较大概率的损失换取只有极小可能发生的高额收益,中国股市频繁发生的概念炒作、题材炒作等事件都体现了这一类特征,事后往往证明,所谓的热点概念或题材绝大部分都是虚假的。中国股市近 20 年的成长中,大部分阶段都伴随着较强投机与泡沫,Mei, Scheinkman, and Xiong(2009)在针对 A、B 股溢价的研究中也特别强调了中国股市的投机性。对于发达国家的股票市场而言,我们也可以很容易回溯到 17 世纪荷兰郁金香

^① 必须指出,学术界也有一些反对 *PIN*可以测度知情交易概率的观点。例如, Mohanram and Rajgopal(2009)认为,他们只能在 EKOP(1996)的样本期间内发现同样的证据,但在其它样本期间,结论就未必尽然。他们对 *PIN*和资产回报的关系持怀疑态度。Duarte, Jefferson and Young(2007)检查 *PIN*是否被信息不对称或者是流动性因素定价。他们首先建立一个模型,将 *PIN*分解成两部分,一部分与信息不对称相关,另一部分与流动性相关。在 Fama-MacBech 回归中发现与信息不对称相关的部分没有被定价,而与流动性相关的部分被定价。因此他们推断,只有与信息不对称无关的流动性因素解释了 *PIN*与预期回报横截面的关系。这样一来,*PIN*在代理知情交易概率方面就显得效力低下了。此外, Benos and Jochech(2007)也发现盈余公告之前的知情交易概率要比公告之后的更低,这似乎也反映了 *PIN*反映知情交易概率的原意相左。

香泡沫、18世纪英国南海泡沫和法国密西西比泡沫、日本20世纪90年代的泡沫以及眼前正在发生的全球金融风暴。

结合格林斯潘针对上世纪90年代末的互联网泡沫中公司股价过高以及市场炒作的现象所提到的“博彩溢价”(Wessel, 1999),针对这一问题,我们在每个月度构造组合,并通过 t -检验考察不同组合收益之差是否显著异于零。

在此基础上,我们将在后文进一步检查常见的一些因子模型是否可以解释这种溢价行为。直觉上,如果这种溢价来自于投资者的博彩心理,那么应该不会被常见的因子模型所解释。因此,我们根据几个常见的资产定价收益方程(常数收益、CAPM收益以及Fama and French三因子收益过程)考察回归的截距项是否显著。如果截距项可以显著地拒绝零假设,这说明我国存在显著的“博彩溢价”且无法被常见的因子模型所解释。Jensen(1968)在考察基金经理是否可以击败市场时,就用了这样的思路。据此,我们提出假说1:

假说1:中国股市存在显著的“博彩溢价”,且无法被常见的CPAM模型或者三因子模型所解释。

为了深入探讨投资者对“博彩溢价”的影响,我们进一步引入不同组合的知情交易测度加总变量。之所以引入知情交易,其原因在于我们想考察博彩型股票的在构建期的引致原因。一般而言,博彩型股票的投机性更强,直觉上投机性的来源在于概念炒作和一些公司信息发布,这在某种程度上与信息揭露时间相关,如果知情交易的确知道内幕消息,他们应该在这个期间体现出明显的增加。另一方面,我们也在直觉上认为知情交易可能在某些情况下对股价进行“操纵”,他们往往借助一些外生信息把股价推高,然后卖出。进一步,我们考察知情交易是否能在一定程度上引致下一期的投机行为,如果我们发现在时期 t 所构造的不同博彩组合在上一期($t-1$)知情交易程度之差可以预测时期 t 的组合收益之差,这也可以在一定程度上说明知情交易会早于股价的投机泡沫,从而(间接地)说明知情交易可能引致随后的投机泡沫的假说。据此,提出假说2:

假说2:知情交易可以显著影响“博彩溢价”,并在常见的因子模型之外存在显著的解释力。即滞后期的知情对未来的“博彩溢价”有显著解释力。

基于假说2的分析,如果知情交易确实显著影响当期“博彩溢价”,说明知情交易在博彩型股票中存在,那么知情交易者应该是基于更有效的信息判断(这里我们并不关心信息的来源渠道),从而在投资操作上应该比一般投资者更“聪明”,因此我们期望下一期的知情交易应该不会取决于本期的“博彩收益”。即知情交易的买卖行为与常见的正反馈操作策略(也称惯性交易)无关^①,即我们在构建组合以后的时间,应该看不到明显的知情交易。这同时也意味着,在随后的期间,因为知情交易者的退出,市场价格取决于噪音交易者,其收益应该在更大程度上体现出随机的或反转(因为噪音交易者的过度反应)特征,从而“博彩溢价”在随后的期间会消失甚至反转。据此,提出假说3:

假说3:一方面,“博彩溢价”在投资组合构造期以后会消失;另一方面“博彩溢价”与下一期的知情交易之间无关或者负相关。

^① 例如,Sentana and Wadhvani(1992)发现英国股市投资者存在明显正反馈形态;Bange(2000)采用美国个人投资者协会(American Association of Individual Investors, AAI)的调查数据则发现美国个人投资者存在明显的惯性交易行为。国内的攀登、施东晖和曹敏(2003)分析了国内的四个证券营业部投资账户数据,也发现中国股市超过三分之一的投资者(主要是个人投资者)在交易决策时采用了正反馈的趋势交易策略;孔东民(2008)也发现,中国股市的个体投资者不以股票的基本价值为参照,具有明显的“买涨杀跌”等惯性行为。

四、数据来源与研究设计

(一)数据来源说明

本文基于 A 股上市公司的月度数据,所有数据均来源于 CCER 数据库。其中,在计算知情交易概率测度时,我们利用高频分笔数据,其它变量的计算主要基于价格收益数据库和财务数据库。研究的样本期间为 2000 年 1 月至 2005 年 5 月,选取这段期间原因在于 2000 年之前的高频分笔数据难以获得,而 2005 年 5 月之后中国股市的股改事件,则会在较大程度上影响我们研究的主要变量。为保证数据的有效性并消除异常样本对研究结论的影响,本文在选择样本时按照以下原则处理数据:为消除 IPO 效应,IPO 上市当月不计入样本中;利用 Winsorized 的方法对 5%和 95%(或者 1%和 99%)的异常值进行整理。

(二)研究设计

1、知情交易概率测度 PIN 的计算

知情交易概率测度最早是由 Easley, Kiefer, O'Hara and Paperman (1996) 所提出,我们按照 Easley, Engle, O'Hara and Wu (2008) 的方法进行计算,该测度的基本原理是市场交易被认为由知情交易引起,并且服从泊松分布。在一系列假设下,可以根据订单的买卖方向和订单数量情况,从金融微观结构角度来推测私人信息引致交易的发生概率问题。具体而言, PIN 一般是按照下面的步骤进行参数估计,首先确定单位时间的似然函数:

$$L(\theta|B,S) = (1-\alpha)e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S!} + \alpha \delta e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B!} e^{-(\mu+\varepsilon_s)} \frac{(\mu+\varepsilon_s)^S}{S!} + \alpha(1-\delta) e^{-(\mu+\varepsilon_b)} \frac{(\mu+\varepsilon_b)^B}{B!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S!} \quad (1)$$

其中, B 和 S 分别表示单位时间内的买单和卖单数量,其余的 5 个参数均为待定变量,即 $\theta = (\alpha, \mu, \delta, \varepsilon_b, \varepsilon_s)$ 需要利用极大似然法进行估计。这里, α 是信息事件发生概率, δ 是坏消息的概率, μ 表示知情者提交订单的到达率, ε_b 表示非知情者提交买单到达率, ε_s 表示非知情者提交卖单到达率。在推断订单买卖方向的时候,我们采用 Lee and Ready (1991) 的方法进行确定,即如果当前交易价格大于前一买卖报价的中位数,则认为属于买方发起的交易,否则,则认定为卖方发起的交易。如果在此准则下无法加以判断,则再往前追溯一笔交易直至可以判断为止。

在假设每个交易日的消息互相独立的情况下,我们可以很容易给出一段时期内(设为 I)的似然函数,如下:

$$L(\theta|M) = \prod_{i=1}^I L(\theta|B_i, S_i) \quad (2)$$

显然,对两边取对数,在不改变单调性的情况下,可以表达为求和的形式,在此基础上,我们根据优化程序很容易计算出 $(\alpha, \mu, \delta, \varepsilon_b, \varepsilon_s)$ 。然后求得 PIN 为:

$$PIN = \frac{\alpha \mu}{\alpha \mu + \varepsilon_b + \varepsilon_s} \quad (3)$$

需要指出的是,在计算中我们发现中国股市的单向订单数量较大,这使得最大似然估计存在数值溢出问题,导致结果不稳定。为了解决这一问题,我们将中国股市的订单数量同比缩小 10 倍,这对 PIN 的结果没有影响,但是计算的精度和计算拟合的稳定性得到了极大增强。在最大似然估计中,每一只股票都选了 27 对组合(即 α, μ 和 ε 分别用 [0.25, 0.5, 0.75] 两两交叉)作为初始值,结果发现结果极为稳定,基本上不受初始值的影响。

2、“博彩溢价”投资组合的构建与检验

为了计算“博彩溢价”,我们首先参照 Kumar (2009) 的精神把所有的股票分为博彩型与非博彩型股票,具体的计算过程说明如下:

1) 在月度 t , 将市场上的所有股票按照期末收盘价、日度收益率的标准差与日度收益率偏度由

从小到大的顺序分别排序并平均分成 k 组。然后将 t 时期内同时出现在最低期末收盘价一组,最高日度收益率标准差一组,和最高日度收益率偏度一组的股票定义为博彩型股票。相应地,同时出现在最高期末收盘价一组,最低日度收益率标准差一组和最低日度收益率偏度一组的股票地应为非博彩型股票。考虑到我国沪深两市股票总数,因此我们选择 $k=3$ (或 5)来进行研究。

2)在得到每月的博彩型股票和非博彩型股票组合后,分别对两个组合计算加权平均收益率。其中权数为组合内各股票在当月末的 A 股流通市值与组合的总 A 股流通市值之比,组合内各股票收益率为其月收益率。

3)计算博彩型与非博彩型股票组合在当月的加权收益率差 $DiffRet_t$,并用 t -检验来检验 $DiffRet_t$ 是否在统计上显著异于零。

4)紧接组合形成期,分别计算 t 月博彩型与非博彩型组合在 $t+1$ 月和 $t+3$ 月,即未来一个月和未来三个月内的加权收益率差,得到 $DiffRet_{t+1}$ 和 $DiffRet_{t+3}$,并用 t -检验来检验 $DiffRet_{t+1}$ 和 $DiffRet_{t+3}$ 是否在统计上显著异于零。

基于上述同样步骤,我们按照 3 个月来划分时期对股票进行分组,得到博彩型与非博彩型股票组合在 $t-3$ 月到 $t-1$ 月,即 3 个月的组合构造期内的加权收益率差 $DiffRet_2$,并分别计算 $t-3$ 月到 $t-1$ 月内的博彩型与非博彩型组合在 t 月和 $t+2$ 月,即投资组合在未来一个月和未来三个月的加权收益率差,得到 $DiffRet_{2t}$ 和 $DiffRet_{t+2}$,用 t -检验来检验 $DiffRet_2$ 、 $DiffRet_{2t}$ 和 $DiffRet_{t+2}$ 是否在统计上显著异于零。

利用月度 PIN 数据,由同样的方法计算出博彩型与非博彩型股票组合在当月的加权 PIN 值之差 $DiffPIN_t$ 以及 t 月博彩型与非博彩型股票组合在 $t-1$ 月和 $t+1$ 月的加权 PIN 值之差,得到 $DiffPIN_{t-1}$ 和 $DiffPIN_{t+1}$ 。

3、实证模型

基于以上的组合和变量构建,我们提出实证模型来检验三个假说。针对假说 1,我们在每个月度构造组合,并通过 t -检验考察不同组合收益之差是否显著异于零。同时,我们采用如下三种收益生成过程进行检验,即常数收益、CAPM 收益以及 Fama and French(1993)的三因子收益:

$$\text{假说 1(a): } DiffRet_t = \alpha_t + \varepsilon_t \quad (4)$$

$$\text{假说 2(b): } DiffRet_t = \alpha_t + \beta_{IRM} RM_t + \varepsilon_t \quad (5)$$

$$\text{假说 3(c): } DiffRet_t = \alpha_t + \beta_{IRM} RM_t + \beta_{SMB} SMB_t + \beta_{HML} HML_t + \varepsilon_t \quad (6)$$

这里, $DiffRet_t$ 表示第 t 期的不同组合的收益之差; RM_t 表示第 t 期的市场收益率; SMB_t 和 HML_t 为 Fama and French(1993)三因子模型中的市值因子和 B/M 因子^①。如果常见的三因子模型无法完全解释“博彩溢价”,等式(4)-(6)的截距项将显著异于零。

针对假说 2,我们进一步在等式(4)-(6)的基础上引入不同组合的知情交易测度之差 $DiffPIN_t$ 。如果 $DiffPIN_t$ 的系数显著,这说明知情交易者确实影响到了这种“博彩溢价”。检验模型如下:

$$\text{假说 2: } DiffRet_t = \alpha_t + \beta_{PIN} DiffPIN_t \text{ (or: } DiffPIN_{t-1}) + \beta_{ConVar} ConVar_t + \varepsilon_t \quad (7)$$

这里, $ConVar$ 表示控制变量(即常见的一些因子)。我们期望在上述模型中, $DiffPIN_t$ or: $DiffPIN_{t-1}$ 变量在统计上显著,这意味着知情交易确实影响到了投机型溢价。同时,我们期望 $DiffPIN_{t-1}$ 的系数显著为正,从而意味着知情交易会影响到下一期的“博彩溢价”现象。需要指出,我们在文章中提到的“博彩溢价”,其背后都对应了投资者的投机和博彩行为,从而也隐含给出了关

^① 对于这两个因子而言,首先在以上年度末每股权益与收盘价的比值来度量 B/M 的高低;以当年度 6 月底流通市值来度量规模的大小,然后按照 Fama and French(1993),在每年 6 月底按流通市值的中位数样本分为大、小两组(B 和 S),按 B/M 分成高、中 and 低三组,即 $H(30\%)$ 、 $M(40\%)$ 和 $L(30\%)$;两者交叉分组形成 6 个组合,由此计算当年度 7 月到下年度 6 月每个组合价值加权月收益率。 HML 和 SMB 两个因子分别定义如下: $HML = (Ret_{SH} + Ret_{BL} - Ret_{SL} - Ret_{BL})/2$,表示高 B/M 与低 B/M 组合的收益差; $SMB = (Ret_{SL} + Ret_{SM} - Ret_{SH} - Ret_{BL} - Ret_{BM} - Ret_{BH})/2$,表示小规模公司和大规模公司组合的收益差。

于投资者行为的含义。

针对假说3,考察知情交易在下一期行为与本期“博彩溢价”的相关性问题,我们构造如下检验模型:

$$\text{假说 3: } \Delta \text{DiffPIN}_{t+1} (\text{or: } \text{DiffPIN}_{t+1}) = \alpha_t + \beta_{iRet} \text{DiffRet}_t + \beta_{iConVar} \text{ConVar}_t + \varepsilon_{it} \quad (8)$$

其中,表示博彩股票组合和非博彩股票组合在 $t+1$ 期与 t 期的知情交易概率加权均值差的差值, ConVar 表示外生控制变量 (RM , HML 和 SMB 等)。

五、实证结果分析

(一) 中国股市的投机溢价

这一部分主要针对假说1进行研究。表1提供了一些有关变量的描述性统计。可以看到,无论是基于一个月 (DiffRet) 还是基于三个月 (DiffRet2) 的频率,博彩型股票的收益率都高于非博彩型股票的收益率,但是知情交易与 DiffPIN 的情况则恰恰相反为负。我们将在后文对其进行更深入的分析。

表1 描述性统计

	DiffRet	DiffRet2	DiffPIN	RM	SMB	HML
均值	0.031	0.019	-0.029	-0.007	0.000	-0.001
标准差	0.062	0.079	0.038	0.062	0.031	0.012
最大值	0.163	0.171	0.026	0.159	0.100	0.020
最小值	-0.074	-0.114	-0.239	-0.134	-0.065	-0.045

注:本表报告了各变量的描述性统计。此处, DiffRet 是基于月度数据构造的博彩型股票与非博彩型股票组合的流通市值加权收益率之差,其中,在月度 t 对股票分类时,我们将市场上的所有股票按照期末收盘价、日度收益率的标准差与日度收益率偏度分别排序。然后将 t 时期内具有最低 k^{th} 的股票价格分位数,最高 k^{th} 标准差和偏度分位数性质的股票定义为博彩型股票。相应地,具有最高 k^{th} 的股票价格分位数,最低 k^{th} 标准差和偏度分位数性质的股票定义为非博彩型股票。考虑到我国沪深两市股票总数,我们选择 $k=3$ 。 DiffRet2 是基于三个月的频率构造组合,其它计算与 DiffRet 一致。 DiffPIN 为相应股票组合的知情交易概率加权均值之差, PIN 的计算与 Easley, Engle, O'Hara, and Wu (2008) 的方法一致,具体见正文。 RM 表示市场收益率, SMB 表示小规模公司组合和大规模公司组合的收益差, HML 表示高 B/M (账面值/市值) 与低 B/M 组合的收益差。

中国股市的“投机溢价”的基本检验结果图1a与图1b所示。可以看到,无论是基于一个月(图1a)还是基于三个月(图1b)的频率,博彩型股票组合都体现了更高的收益率。表2提供了进一步的统计检验,很容易看到,在组合构建期内,博彩型组合的与非博彩型组合的收益率之差 (DiffRet) 均显著拒绝其值为零的零假设,说明二者之差在统计上是显著的。这也给我们提供了中国股市存在“博彩溢价”的初步证据(假说1)。

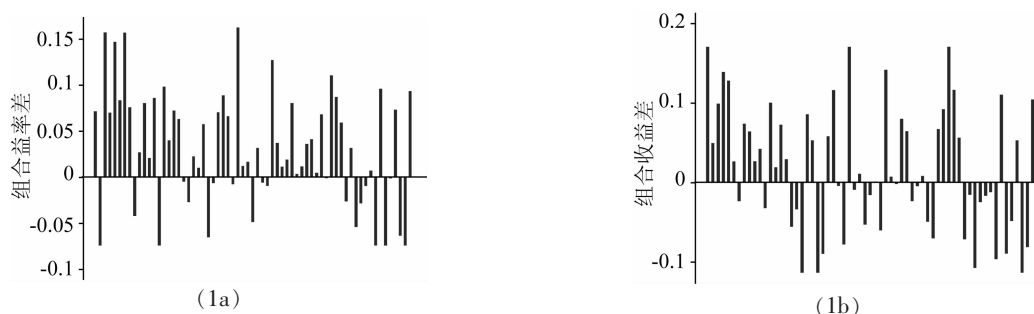


图1 (非)博彩型股票组合收益之差

注:本图描述了在博彩型股票和非博彩型股票在投资组合构建期内的收益之差。其中,图1a是基于月度频率构建组合,图1b是基于三个月的频率构建组合。组合的收益为流通市值加权收益率,具体的组合构建方法参见表1。

表 2 (非)博彩型股票组合收益之差的统计检验

$(H_0: DiffRet = 0)$	均值	$t-val$	$p-val$
$DiffRet$: 基于一个月	0.0312	4.087	0.001
$DiffRet2$: 基于三个月	0.0313	2.112	0.039

注:本表报告了(非)博彩型股票组合在投资组合构建期内的收益之差以及针对零假设 $H_0: DiffRet = 0$ 的统计检验结果。 $DiffRet$ 和 $DiffRet2$ 分别为基于月度频率和三个月频率构建的组合。组合的收益为流通市值加权收益率,具体的组合构建方法参见表 1。

不能被常见的因子所解释。

为了继续检验假说 1,即在“博彩溢价”显著的情況下,常见因子模型是否可以解释这种溢价?我们根据模型(4)–(6)进行回归估计。结果如表-3 所示。可以看到,在所有的三个回归中,模型的截距项都显著拒绝零假设,这说明“博彩溢价”不但存在,而且还

表 3 “博彩溢价”是否可以被不同因子模型解释?

因变量: $DiffRet_t$	常数回归	CAPM 回归	Fama-French 三因子
RM_t		0.630***	0.554***
		0.096	0.105
SMB_t			0.546***
			0.196
HML_t			0.274
			0.511
$Intercept$	0.031***	0.0353***	0.0350***
	0.008	0.006	0.006
Obs	65	65	65
$Adj-R^2$	0.000	0.405	0.482

注:本表根据等式(4)–(6)进行估计。因变量为 $DiffRet_t$, $DiffRet$ 为(非)博彩型股票组合在投资组合构建期内的收益之差。其中,组合的收益为流通市值加权收益率,具体的组合构建方法参见表 1。 RM_t 表示 t 期的市场收益率。 SMB_t 表示 t 期的小规模公司和大规模公司组合的收益差。 HML_t 表示 t 期的高 B/M (账面值/市值)与低 B/M 组合的收益差。 $Intercept$ 为截距。 Obs 为观测值的个数。 $Adj-R^2$ 为调整后的 R^2 。各变量对应的第一行为回归系数,第二行为异方差与时间序列稳健的标准误。***, ** 和 * 分别表示在 0.01, 0.05 和 0.10 的显著性水平上拒绝零假设。

(二)知情交易与投机溢价

表 4 根据模型(7)进行回归估计。回归结果为假说 2 提供了证据,即知情交易行为与“博彩溢价”存在显著关系。可以看到,在所有的三个回归中,无论是否加入不同的因子作为自变量,我们新引入的变量 $DiffPIN$ 的回归系数都显著地拒绝零假设,这说明在同期(组合构造器)，“博彩溢价”的确为知情交易所影响。因此,对于投机更强的博彩型股票而言,的确有知情交易会知道内幕消息,从而在这个期间体现出明显的知情交易行为,而且与组合的收益之差 $DiffRet$ 之间存在极为显著的正向关系。

尽管基于目前的证据,不太容易直接验证假说 2:知情交易是否在某些情况下对股价进行“操纵”,他们先借助信息把股价推高,然后卖出获利^①。但我们可以对这个问题进行一个初步的分析,表 4 考察是否上期的知情交易会引起下期的投机行为。从基于模型(7)的估计结果可以看到, $DiffPIN_{t-1}$ 的回归结果大都正显著,这说明,上一期的知情交易会引致下一期更大的投机泡沫,从而间接给出了证据。

^① 当然,在今后可以获得更为深入的数据情况下,我们可以考察是否知情交易首先引发了股价趋势,然后借助随后的公开的信息进一步获利而驱动市场的投机泡沫。

表4 知情交易与“博彩溢价”

因变量: $DiffRet$	常数回归			CAPM 回归			Fama-French 三因子		
	(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(3)
$DiffPIN_{t-1}$	0.170		-0.0893	0.367*		0.0233	0.368*		0.073
	0.263		0.296	0.202		0.212	0.191		0.207
$DiffPIN_t$		0.380*	0.415*		0.576***	0.567***		0.511***	0.482***
		0.201	0.232		0.146	0.167		0.144	0.166
RM_t				0.654***	0.689***	0.690***	0.575***	0.633***	0.633***
				0.0954	0.0879	0.0889	0.103	0.0986	0.0994
SMB_t							0.552***	0.415**	0.423**
							0.192	0.183	0.186
HML_t							0.231	0.346	0.333
							0.50	0.468	0.473
$Intercept$	0.0357***	0.0421***	0.0406***	0.0453***	0.0522***	0.0525***	0.0449***	0.0502***	0.0514***
	0.0104	0.0094	0.0106	0.008	0.0069	0.0077	0.0076	0.0068	0.0075
Obs	65	65	65	65	65	65	65	65	65
$Adj-R^2$	-0.009	0.039	0.025	0.417	0.509	0.501	0.479	0.543	0.537

注:本表根据等式(7)进行估计。因变量为 $DiffRet_t$, $DiffRet$ 为(非)博彩型股票组合在投资组合构建期内的收益之差。其中,组合的收益为流通市值加权收益率,具体的组合构建方法参见表1。 $DiffPIN$ 和 $DiffPIN_{t-1}$ 分别为相应股票组合在 t 期和 $t-1$ 期的知情交易概率加权均值之差, PIN 的计算与 Easley, Engle, O'Hara and Wu(2008)的方法一致,具体见正文。 RM_t 表示 t 期的市场收益率。 SMB_t 表示 t 期的小规模公司和大规模公司组合的收益差。 HML_t 表示 t 期的高 B/M (账面值/市值)与低 B/M 组合的收益差。 $Intercept$ 为截距。 Obs 为观测值的个数。 $Adj-R^2$ 为调整后的 R^2 。各变量对应的第一行为回归系数,第二行为异方差与时间序列稳健的标准误。***, ** 和 * 分别表示在 0.01, 0.05 和 0.10 的显著性水平上拒绝零假设。

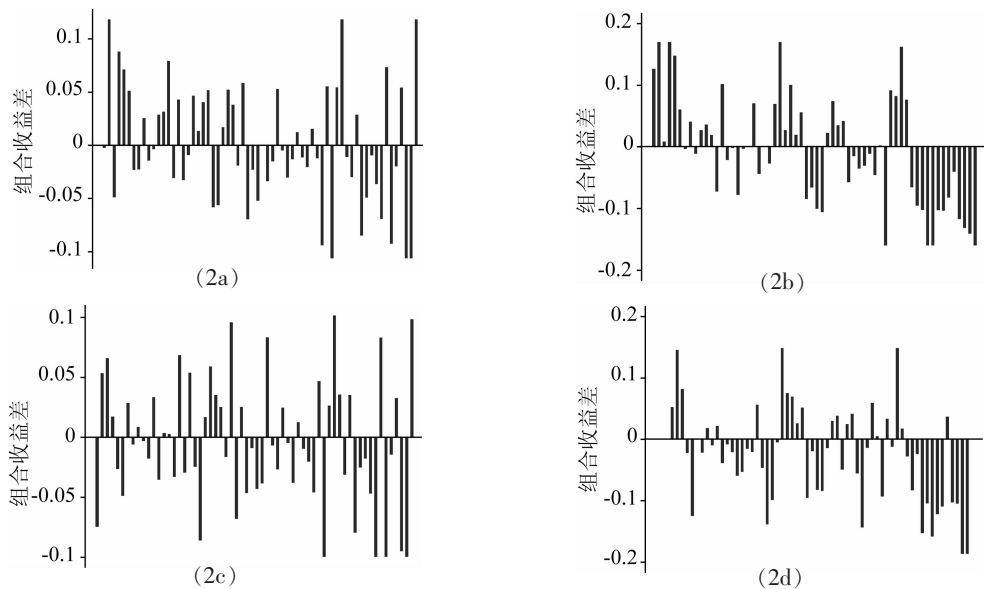


图2 (非)博彩型股票组合构建后期的收益之差

注:本图描述了博彩型股票和非博彩型股票在投资组合构建期内的收益之差。其中,图2a(图2b)是基于月度频率构建的组合在随后一个月(三个月)的组合收益之差;图2c(图2d)是基于三个月频率构建的组合在随后一个月(三个月)的组合收益之差。组合的收益为流通市值加权收益率,具体的组合构建方法参见表1。

(三)投机收益是否影响下一期的知情交易?

对于博彩型组合下期的收益的考察如图2所示。其中,在基于月频率构建组合的情况下,图

2a 和图 2b 分别报告了随后一个月和三个月的收益情况。可以看到,整体上,收益之差没有特定的形态体现,其均值在零附近波动。在基于三个月频率构建组合的情况下,图 2c 和图 2d 分别报告了随后一个月和三个月的收益情况,基本特征类似。表 5 提供了进一步的统计检验,很容易看到,在组合构建期的随后月度,博彩型组合的与非博彩型组合的收益率之差近似为零,均无法显著拒绝零假设,这说明博彩型股票的收益是短期的,并迅速回落,且有反转的倾向。尤其是对于基于三个月构建的组合而言,在随后的三个月平均跌了 2.5%,而且在统计上显著,这意味着博彩型股票存在明显的股价下跌现象。因此,那些随后进入市场的正反馈投资者(或动量交易者)无法获得超额收益,在一定情况下,甚至会出现明显的亏损。此处的结果意味着知情交易者在基于私人信息获利之后,那些随后进入市场的投资者(动量交易者)无法获得超额收益。

表 5 (非)博彩型股票组合收益在未来月度的统计检验

$(H_0: DiffRet = 0)$	均值	$t-val$	$p-val$
$DiffRet_{t+1}$	0.0002	0.0332	0.9736
$DiffRet_{t+3}$	-0.0067	-0.5988	0.5515
$DiffRet2_t$	-0.0047	-0.7122	0.4790
$DiffRet2_{t+2}$	-0.0254	-2.4979	0.0153

注:本表报告了(非)博彩型股票组合在投资组合构在未来一个月和三个月的益之差以及针对零假设 $H_0: DiffRet = 0$ 的统计检验结果。 $DiffRet$ 和 $DiffRet2$ 分别为基于月度频率和三个月频率构建的组合。组合的收益为流通市值加权收益率,具体的组合构建方法参见表 1。

根据前面的直觉,我们认为,如果知情交易确实存在,那么他们基于更有效的信息判断,在投资操作上不会取决于上一期的“博彩收益”。即知情交易的买卖行为与应该基于其私人信息,而不是历史的收益。或者说,知情交易在驱动了投机泡沫之后,会有更好的时机离开泡沫,从而在有可能出现收益反转的情况下(例如基于三个月频率构建的组合在未来三个月的收益显著为负),他们的行为应该与上一期知情交易没有明显关系。

表 6 给出了基于模型(8)的回归结果。可以看到,知情交易在大部分情况下与上一期的博彩型股票溢价为负相关或没有显著的相关性。在以 $\Delta DiffPIN_{t+1}$ 为因变量时,可以看到结果为负,且基本上拒绝零假设,这说明在第 t 期的博彩溢价越高,其随后的知情交易降低的越多。在以 $DiffPIN_{t+1}$ 为因变量时,可以看到结果基本上不显著,这说明在第 t 期的博彩溢价与随后的知情交易的水平值关系不大。这都支持了假说 3 的推断。当然,因为 $DiffPIN_{t+1}$ 只是水平值,根据表 4 的检验,如果其与 $DiffPIN_t$ 存在相关性,基于该变量的检验可能会存在一些内生性,因此我们更强调基于 $\Delta DiffPIN_{t+1}$ 的检验结果。

六、结 论

尽管博彩(或投机)心理在人们的生活中根深蒂固,而且格林斯潘(前美联储主席)在针对上世纪 90 年代末的互联网泡沫中公司股价过高以及市场炒作的现象,曾提到“博彩溢价”的概念,但一直未见到这方面严谨的学术研究。本文采用 Kumar(2009)的方法把股票分为博彩型和非博彩型,并发现了“博彩溢价”的存在,且无法为 Fama-French 三因子模型解释,这说明我国股市的确存在投机的溢价现象。而对我国博彩型股票的研究和与其他变量的相关分析,能够进一步地了解我国投资者的投资决策行为,也能为以后投资者在投资决策的选择提供一定的指导和参考价值。

我们发现尽管在组合构造期内,博彩型股票存在显著溢价,但组合的超额收益会迅速消失,并

表6 知情交易与上一期博彩收益

	因变量: $\Delta DiffPIN_{t+1}$			因变量: $DiffPIN_{t+1}$		
	常数回归	CAPM 回归	Fama-French 三因子	常数回归	CPAM 回归	Fama-French 三因子
$DiffRet_t$	-0.0514	-0.180*	-0.236**	0.0899	0.169**	0.105
	0.0733	0.0923	0.0984	0.0653	0.0839	0.0883
RM_t		0.200**	0.215**		-0.123	-0.13
		0.0913	0.0972		0.083	0.0873
SMB_t			0.215			0.286*
			0.16			0.144
HML_t			0.26			0.0271
			0.393			0.353
$Intercept$	-0.0126**	-0.00731	-0.0052	-0.0457***	-0.0489***	-0.0471***
	0.0050	0.0055	0.0056	0.0045	0.005	0.005
Obs	65	65	65	65	65	65
$Adj-R^2$	-0.01	0.05	0.059	0.029	0.062	0.123

注:本表根据等式(8)进行估计。因变量为 $\Delta DiffPIN_{t+1}$ 和 $DiffPIN_{t+1}$, $\Delta DiffPIN$ 表示博彩股票组合和非博彩股票组合在 $t+1$ 期与 t 期的知情交易概率加权均值差的差值; $DiffPIN_{t+1}$ 表示博彩股票组合和非博彩股票组合在 $t+1$ 期的知情交易概率加权均值之差。 PIN 的计算与Easley,Engle,O'Hara and Wu(2008)的方法一致,具体见正文。 $DiffRet_t$ 为(非)博彩型股票组合在投资组合构建期内的收益之差,具体计算方法参见表1。 RM_t 表示 t 期的市场收益率。 SMB_t 表示 t 期的小规模公司和大规模公司组合的收益差。 HML_t 表示 t 期的高 B/M (账面值/市值)与低 B/M 组合的收益差。 $Intercept$ 为截距。 Obs 为观测值的个数。 $Adj-R^2$ 为调整后的 R^2 。各变量对应的第一行为回归系数,第二行为异方差与时间序列稳健的标准误。***,**和*分别表示在0.01,0.05和0.10的显著性水平上拒绝零假设。

未给投资者带来持续的财富效应。为了深入考察知情交易是否驱动(或引发)了博彩型股票溢价,我们引入知情交易测度 PIN ,并发现在组合构造的月份,知情交易行为非常明显,但是在组合构造之后的月份,并没有明显的知情交易者存在。我们认为检验结果意味着知情交易者在基于私人信息获利之后,那些随后进入市场的正反馈交易者并无法根据历史收益获得超额收益。

参考文献

- 韩立岩、郑君彦、李东辉(2008):《沪市知情交易概率特征与风险定价能力》,《中国管理科学》,第1期。
- 孔东民(2008):《中国股市投资者的策略研究:基于一个噪音交易模型》,《管理学报》,第4期。
- 孔东民、申睿(2008):《 R^2 、异常收益与交易的知情概率成分》,《中大管理研究》,第3期,将发表。
- 刘元海、陈伟忠(2004):《中国股市知情交易概率的测定》,《同济大学学报》,第1期。
- 攀登、施东晖(2006):《知情交易概率的测度模型及其影响因素分析》,《管理世界》,第6期。
- 攀登、施东晖、曹敏(2003):《中国个人投资者采用股价趋势交易策略的经验研究》,《世界经济》,第11期。
- 王春峰、董向征、房振明(2005):《信息交易概率与中国股市价格行为关系的研究》,《系统工程》,第2期。
- 王展翔(2007):《关于我国股市知情交易水平及其性质的实证研究》,《当代财经》,第2期。
- 杨之曙、姚松瑶(2004):《沪市买卖价差和信息性交易实证研究》,《金融研究》,第4期。
- 张宗新(2008):《内幕交易行为测定:理论模型与实证分析》,《管理世界》,第4期。
- Baker, M., Greenwood, R. and Wurgler, J. (2009): "Catering Through Nominal Share Prices", *Journal of Finance*, 64, 2559-2590.
- Bange, M. (2000): "Do the Portfolios of Small Investors Reflect Positive Feedback Trading?" *Journal of Financial and Quantitative Analysis*, 35, 239-255.
- Barberis, N. and Huang, M. (2008): "Stocks as Lotteries: The Implications of Probability Weighting for Security Prices", *American Economic Review*, 98, 2066-2100.
- Benos, E. and Jochec, M. (2007): "Testing the PIN variable", Working paper.

- Borna, S. and Lowry, J. (1987): "Gambling and Speculation", *Journal of Business Ethics*, 6, 219–224.
- Brenner, R. (1983): *History—The Human Gamble*, Chicago: University of Chicago Press.
- Duarte, J. and Young, L. (2007): "Why is PIN Priced?" *Journal of Financial Economics*, 91, 119–138.
- Easley, D. and O'Hara, M. (1987): "Price, Trade Size, and Information in Securities Markets", *Journal of Financial Economics*, 19, 69–90.
- Easley, D. and O'Hara, M. (1992): "Time and the Process of Security Price Adjustment", *Journal of Finance*, 47, 577–605.
- Easley, D. and O'Hara, M. (2004): "Information and the Cost of Capital", *Journal of Finance*, 59, 1553–1583.
- Easley, D., Hvidkjaer, S. and O'Hara, M. (2002): "Is Information Risk a Determinant of Asset Returns?" *Journal of Finance*, 10, 2185–2221.
- Easley, D., Kiefer, N. and O'Hara, M. (1996): "Cream-Skimming or Profit-Sharing? The Curious Role of Purchased Order Flow", *Journal of Finance*, 51, 811–833.
- Easley, D., Kiefer, N. and O'Hara, M. (1997): "One Day in the Life of a Very Common Stock", *Review of Financial Studies*, 10, 805–835.
- Easley, D., Kiefer, N., O'Hara, M., and Paperman, J. (1996): "Liquidity, Information and Infrequently Traded Stocks", *Journal of Finance*, 51, 1405–1436.
- Easley, D., Kiefer, N., O'Hara, M., and Wu, L. (2008): "Varying Arrival Rates of Informed and Uninformed Trades", *Journal of Financial Econometric*, 4, 1–37.
- Fama, E. and French, K. (1993): "Common Risk Factors in the Returns on Stocks and Bonds", *Journal of Financial Economics*, 33, 3–56.
- Golec, J. and Tamarkin, M. (1998): "Bettors Love Skewness, not Risk, at the Horse Track", *Journal of Political Economy*, 106, 205–225.
- Green, C. and Hwang, B., (2009a): "Price-based Return Comovement", *Journal of Financial Economics*, 93, 37–50.
- Green, C. and Hwang, B., (2009b): "IPOs as Lotteries: Expected Skewness and First-Day Returns", Working paper.
- Han, B., Lee, Y. and Liu, Y. (2009): "Investor Trading Behavior and Performances: Evidence from Taiwan Stock Index Options", Working Paper.
- Holliday, J. and Fuller, P. (1975): *The Psychology of Gambling*, New York: Harper and Row Publishers.
- Hong, H. and Stein, J. (2003): "Difference of Opinion, Short-Sales Constraints, and Market Crashes", *Review of Financial Studies*, 16, 487–525.
- Kumar, A. (2009): "Who Gambles in the Stock Market?" *Journal of Finance*, 64, 1889–1933.
- Lee, C. and Ready, M. (1991): "Inferring Trade Direction from Intraday Data", *Journal of Finance*, 46, 733–746.
- Mei, J., Scheinkman, J., and Wei, X. (2009): "Speculative Trading and Stock Prices: Evidence from Chinese A-B Share Premia", *Annals of Economics and Finance*, 10, 225–255.
- Mohanram, P. and Rajgopal, S. (2009): "Is PIN Priced Risk?" *Journal of Accounting and Economics*, 47, 226–243.
- Sentana, E. and Wadhvani, S. (1992): "Feedback Traders and Stock Return Autocorrelations: Evidence from a Century of Daily Data", *Economic Journal*, 102, 415–425.
- Statman, M. (2002): "Lottery Players/Stock Traders", *Financial Analysts Journal*, 58, 14–21.
- Tversky, A. and Kahneman, D. (1992): "Advances in Prospect Theory: Cumulative Representation of Uncertainty", *Journal of Risk and Uncertainty*, 5, 297–323.
- Walker, M. (1992): *The Psychology of Gambling*, Oxford: Pergamon Press.
- Wessel, D. (1999): "Internet Mania is Like Lottery, Fed Chief Says", *Wall Street Journal*, January 29.
- Zhang, Y. (2005): "Individual Skewness and the Cross-section of Average Stock Returns", Working paper.

(责任编辑:程 炼)