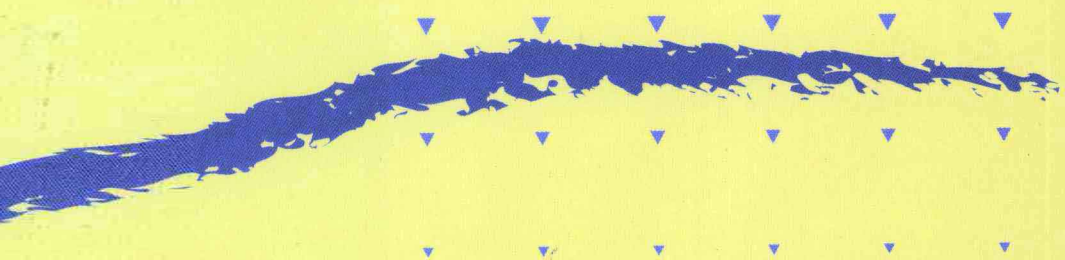


孙绍荣/等著

TOUZIZHE XINGWEI YANJIU

投资者行为研究



 复旦大学出版社

孙绍荣 方 勇 李 红
刘春霞 刘 丹 崔晓丽 / 著

TOUZIZHE XINGWEI YANJIU

投资者行为研究

复旦大学出版社

图书在版编目(CIP)数据

投资者行为研究 / 孙绍荣等著. —上海: 复旦大学出版社, 2009. 4

ISBN 978-7-309-06603-6

I. 投… II. 孙… III. 投资—经济行为—研究 IV. F830.59

中国版本图书馆 CIP 数据核字 (2009) 第 057391 号

投资者行为研究

孙绍荣 等著

出版发行 **复旦大学出版社** 上海市国权路 579 号 邮编: 200433
86-21-65642857(门市零售)
86-21-65100562(团体订购) 86-21-65109143(外埠邮购)
fupnet@fudanpress.com <http://www.fudanpress.com>

责任编辑 盛寿云

出品人 贺圣遂

印	刷	句容市排印厂
开	本	890×1240 1/32
印	张	8.125
字	数	219 千
版	次	2009 年 4 月第一版第一次印刷

书	号	ISBN 978-7-309-06603-6 / F·1481
定	价	20.00 元

如有印装质量问题, 请向复旦大学出版社发行部调换。

版权所有 侵权必究

前 言

本书是国家自然科学基金项目(70471066)的研究成果,也是上海市重点学科“系统管理”(S30504)的研究内容。在上述两个项目的研究工作中,我们研究团队付出了相当长时间、相当多的人力、物力。将之出版,为的是供有志于研究行为管理的同行共享和参考,节省精力和提高效率,也使上述项目研究的投入能够发挥更大的作用。

投资者行为规律的研究,涉及金融市场、房地产市场,对国家经济的管理与调控具有重要的意义。特别是在当前国际金融危机的严重形势下,这方面的研究更有相当重要的意义。

我国当前在经济管理特别是投资品市场管理中所遇到的许多问题,都与投资者行为有关,比如房地产市场的调控,涉及对投资者行为的管理;维护股市的稳定,涉及对机构及个人的投资者的投资行为的管理。为此,就要深入地研究投资者行为的规律,建立投资者行为研究的理论体系,为国家有关政策的出台、有关制度的设计,提供科学的理论基础。

当然,决定人们行为的因素是很复杂的。所以,投资者行为管理并不是一件简单的事,在许多情况下常常失败,各种管理措施常常达不到目标。但是,这并不能说明投资者行为管理没有意义,投资者行为管理研究没有价值。反之,恰恰说明我们对投资者行为管理的理论的研究还很不够、对投资者行为管理规律的掌握还很肤浅,还需要大力深入研究。

正如管子所说的,“天不变其常,地不易其则,春秋冬夏,不更其节”。投资者行为管理也与其他社会现象一样,有其内在的规律。作

为管理学研究,我们应当勇敢地面对投资者行为管理的复杂性,努力地探讨投资者行为的规律,只有这样,管理理论才能够起到有效地指导实践的作用。

本书各章节的作者分别为:孙绍荣负责第一、二章,方勇负责第三、四、五、六章,李红负责第七、八、九章,刘春霞负责第十、十一、十二章,刘丹负责第十三章,崔晓丽负责第十四章的整合。全书由孙绍荣、崔晓丽同志统稿。

本书的出版受到了复旦大学出版社的大力支持,在此谨表示感谢。

孙绍荣

2009年3月8日

目 录

第一章 投资行为概率的数学模型	1
第一节 引言	1
第二节 基本概念的定义	3
第三节 输入变量的预处理	5
第四节 行为概率模型	7
第五节 案例——风险投资行为模型	10
第二章 行为金融的创新进展	16
第一节 主流金融理论的困境与行为金融理论	16
第二节 金融行为控制方面的新进展	19
第三章 金融市场有效性研究	26
第一节 概述	26
第二节 有效市场假说	27
第三节 金融市场中的异象	31
第四节 并非有效的市场	34
第五节 国际股票市场运行效率的实证分析	37
第六节 小结	47

第四章 异质预期、噪声交易与价格波动	49
第一节 概述	49
第二节 DSSW 噪声交易模型	50
第三节 包含三类异质投资者的噪声交易模型	54
第四节 上升市场中机构投资者的投资行为实证分析	66
第五节 小结	77
第五章 中国证券投资者的预期模式研究	78
第一节 概述	78
第二节 中国证券投资者的预期模式实证分析	79
第三节 情绪指数与未来市场收益及波动率的关系实证 分析	88
第四节 投资者的非理性预期与金融危机	95
第五节 小结	96
第六章 噪声交易证券市场的投资策略模型	98
第一节 概述	98
第二节 基于投资者行为参数的股票指数广义回归神经网络 预测模型	99
第三节 基于过度反应的长期策略模型	109
第四节 小结	114
第七章 行为扩散动力学模型	116
第一节 单一行为扩散动力学模型	117

第二节	三峡游行为扩散实例	121
第三节	竞争性行为扩散动力学模型	125
第四节	小结	131
第八章	基于复杂网络的行为扩散模型	132
第一节	引言	132
第二节	复杂网络研究简介	134
第三节	建立模型	137
第四节	基础网络	140
第五节	试验结果与分析	142
第六节	小结	145
第九章	行为扩散研究的信息视角	147
第一节	含模仿偏好的信息层叠模型	147
第二节	股市中的从众行为与信息传播	153
第十章	投资者行为与计算机模拟方法	160
第一节	Agent 模拟方法	161
第二节	CA 模拟方法	162
第三节	两种模拟方法的异同	164
第四节	现有模拟方法的局限	165
第十一章	具有元胞结构的智能 Agent	167
第一节	模型的提出	167
第二节	模型的构建	167

第三节	模型的计算机实现	171
第十二章	从模拟结果看投资者行为	178
第一节	不同规模下的投资者行为	178
第二节	不同类型比例下的投资者行为	188
第三节	不同行为比例下的投资者行为	193
第四节	投资者行为规律总结	199
第十三章	投资者行为与股票市场波动及实证研究	202
第一节	投资者行为与股票市场波动	202
第二节	股票市场波动实证研究	203
第三节	股票市场波动规律总结	223
第十四章	总结与建议	233
参考文献		241

第一章 投资行为概率的 数学模型

第一节 引 言

无论是组织管理还是社会管理,常常需要用一定的制度或者政策来规范人们的行为。而某种制度或者某种政策出台前,管理者常常需要事先估计其可能的效果,即对人们行为的影响程度的大小,由此来对制度或政策进行成本效益分析,以决定这些制度或者政策是否需要出台。这就需要建立行为概率的数学模型,这种模型描述的是在人们对某种行为的回报的普遍预期下,人们选择该行为的概率是多少,在群体中,也就是选择该行为的比例有多少。

目前的模拟和描述行为的数学方法主要有:元胞自动机(cellular automata, CA)方法^[1],主要模拟个体之间的相互影响^{[2][3][4]};智能主体(Agent)技术,把群体中的大量个体看作是众多决策主体,主要模拟决策过程中各个 Agent 之间的相互影响,最后输出的是一个群体的决策结果^[5];随机激励学习(reinforcement learning, RL)方法,模拟个体决策的学习过程;遗传算法,模拟个体特征的演化与变异过程;定性模拟算法(QSIM 算法),模拟按不同特征的个体组成的小组等等^[6]。这些方法的共同特征,是构造出一个人造的并假设遵守一定规则的“虚拟人”,每个人都有一些离散的简单化的不同“状态”,并且“虚拟人”相互之间存在着一定的影响规则,把群体看作是由这样的“虚拟人”构成的,然后给出一定的初始条件,计算该群体最终演化的结果。其中,定性模拟算法(QSIM 算法)是把群体中的个体分成 n 个组,假设每个组

有不同的特征^[7],实质上这些“组”还是一种个体,在这一点上与其他的行为模拟算法并无本质的区别。

实际上,在人所组成的群体中,个体种类相当复杂多样,在群体中个体比较多的情况下,个体的种类过渡可以看作是连续的,很难通过简单地划分几个类型就能准确地描述。此外,个体的状态也是相当复杂的和连续的,并不是几个简单的状态就能代表的。就拿对人们的行为决策的过程的模拟来说,一般的行为模拟方法都把入看作是完全理性的,即对利益的判断是“无偏的”,但实际中的人的行为决策过程非理性成分相当多,仅仅就非理性的认知偏差来说,就有许多种情况,从而导致人们的行为决策常常偏离理性最优。

所以,这些基于离散的有限类型的有限状态的个体基础上的对行为的微观模拟,很难与实际相符。

此外,这些模拟方法大多以群体的演化的个体一致性为模拟的最终结果,即在个体相互影响的规则下,群体中的所有个体的状态趋于一致的某个状态。这显然与实际的群体中广泛存在的个体之间的差异性相矛盾。比如,从理性判断来看,吸毒行为的负回报是非常明显的:个人财产遭受损失,国家法律打击,因此多数社会成员并不参与吸毒行为。但是,无论打击多严厉,社会上却总是存在一定比例的吸毒人群。

这些模拟方法的局限性,可能也是当前在行为模拟方面理论研究较多但实际应用很少的一个重要原因。从各国的情况看,一些国家层面上的大规模政策的出台,还没有运用行为模拟来进行实验和预测。

本文的思路与上述方法不同,模拟的是在一定的回报预期下,群体中各有关行为的比例或者概率。这种模拟是建立在数据的宏观统计基础之上的,这种宏观统计包含了群体中的个体类型分布、理性判断、认知偏差等各种因素的综合影响。

社会管理的实践告诉我们,虽然个体之间的影响比较大,但群体之中的个体的行为并不能完全统一,有关的各种行为几乎都在一

定程度上存在着(即使一些看起来非常不合理的行为)。因此,对于制度设计和政策措施制定等管理工作来说,行为概率模型更有参考意义。

第二节 基本概念的定义

一、行为集 A_i 及行为 a_i

设行为主体的行为集为 A_i , 其中一共有 N 个行为, 行为主体必选择且只选择其中一个行为 a_i (即行为选择的唯一性), 并且对于这 N 个行为不存在选择时间的先后问题, 只能在同一时间点选择其中一个行为。

行为集中的元素各个“行为”之间的划分要求满足行为选择的唯一性。如果没有唯一性的要求, “各个行为”之间的划分就可以是任意的。比如, 对于一个投资行为集来说, 既可以把“买股票”和“买债券”定义为两个不同的行为, 也可以把“买 50% 的债券又买 50% 的股票”与“买 70% 的债券又买 30% 的股票”定义为两个不同的行为(注意行为集中的行为可以是连续的, 也可以是离散的)。在本文中, 行为集中的“各个行为”的定义要满足行为选择的唯一性, 即行为主体只能并且必须选择一个行为。显然, “买 x 万元的股票并且买 y 万元的债券 ($x, y \in [0, m]$, m 为相应的投资者的最大可用资金数量)”的行为划分是满足行为选择的唯一性的, 但如果把“买股票”和“买债券”分别定义为两个不同的行为, 则这种行为划分不满足唯一性, 因为投资者有可能选择既买股票又买债券, 即可以同时选择两个行为。

二、期望回报 $u(a_i)$

期望回报是行为概率模型的输入变量。

在众多个体所构成的群体中, 某个行为的期望回报是群体中的所有个体对该行为的回报的估计值的平均值, 而不是某个个体对行为回

报的估计值。这是因为在群体中,各个个体对同一行为的回报的估计往往有很大差别。在这种定义下,即使某行为的期望回报是负的,则选择该行为的概率仍然可能大于0,即仍然有个体选择该行为。比如2006年8月万科认购权证的行权,认购权证的价格为0.001元,行权价为3.64元,而万科股票的价格是6.90元左右。也就是大家基本上都知道,持有万科权证的投资者如果要行权,意味着要以每股6.90元左右的价格买入万科股票,然后再以每股3.64元的价格卖给万科公司,每股亏损3元多。这是一个明显负回报的行为,显然按常理是不应该有人选择“行权”这一行为的。但实际上却仍然有一部分投资者选择“行权”,根据交易所记录,8月29日,投资者行权3笔,行权5801份;30日,行权2笔,行权3100份;31日,行权4笔,行权2601份。仅在三个行权日里,投资者就对万科权证错误行权9笔,累计损失约3.75万元。

三、行为概率 $p(a_i)$

行为概率是模型的输出变量。

对于作为个体的行为主体来说,行为概率只是行为的不确定性的描述,即在他的行为集中选择各个行为的可能性就是行为概率。对于群体来说,则行为概率可以看作是行为个体选择该行为的个体数量比例。比如对于一个有100个人的群体,如果其中有10个人选择了行为 a_i , 则行为 a_i 的概率 $p(a_i) = 0.1$ 。

在群体中,如果人们对各个行为的期望回报的评价是一致的,则群体中的全部个体都会选择同一个行为,这样就不存在行为概率问题。但实际上,人们对同一行为的期望回报的估计常常有很大差别。这是因为有如下两个因素会影响人们对各行为的期望回报评价的一致性。

一个因素是认知偏差。由于一些认知的心理规律的原因,人们往往不能准确而客观地判断期望回报,而是存在一些规律性的偏差,即认知偏差。不同心理特征的人,知识偏差的强弱也不同。这就造成对

同一行为的期望回报,不同的人的判断结果往往具有较大的差别。这里须说明的是,认知偏差既有群体性的,也有个体性的。群体性的共有的认知偏差已经反映为期望回报中(因为期望回报是通过统计调查得到的)。只有个体性的认知偏差,才能对人们对期望回报评价的一致性产生影响。

另一个因素是路径状态造成的成本差异,简称路径成本差异。人们以往的行为选择,形成了行为路径。就选择当前的同一个行为来说,行为路径不同的人的成本是不同的。比如文学家写小说一般会比科学家写小说的效率高,但文学家却很难搞科学研究。实际上,人们的知识与经验、财富、地位、人际关系等都会形成对某些特定行为的优势,如果不利用这些优势而去选择不适合的行为,就会增加很多成本。

由于上述两个因素,导致人们对行为的集中的行为的期望回报的估计出现差别,从而在选择各种行为时出现不同的比例(而不是一致地都选择某个行为和都不选择另一些行为),即行为概率。

第三节 输入变量的预处理

一、整理行为集

在解决具体问题时,面对的初始行为集有时不符合行为选择的唯一性的要求,这时就须整理行为集,重新划分各个行为,使行为主体对行为的选择具有唯一性。

二、获取期望回报 $u(a_i)$

由于期望回报 $u(a_i)$ 是群体中各个个体对行为 a_i 的回报的估计的平均值,所以,必须用调查统计的方法获得期望回报。

期望回报的取值方式是一个重要问题,可以采用两种方法来估计期望回报。

一是收益率估计法,这种方法比较适于投资等一些经济行为。如

果行为集中的行为是投资采煤和投资发电,并且规定只能投资一个行业,则可以根据当时的情况分析和统计投资采煤和投资发电的行业平均收益率。

许多行为的期望回报由于没有经济的收益率概念,所以不能用收益率法估计。但是,这里人们对行为的价值仍然是有判断的,比如行为的重要性的、成绩的大小或者损失的严重程度等。这时,就可以采用语言变量法来估计期望回报。

语言变量法是把人们的主观感觉与相应的语言变量对应起来,而语言变量又与相应的分值对应,从而把各个个体的主观感觉转变成分数,最后把群体中各个个体的分数平均起来,就得到群体的期望回报。本文建议把语言变量分为 11 个等级,具体描述形式与对应分数见表 1-1。

表 1-1 语言变量与分值转换表

语言变量	对应分数
收益无穷大	∞
收益大	1
收益较大	0.75
收益中等	0.5
收益较小	0.25
无收益无损失	0
损失较小	-0.25
损失中	-0.5
损失较大	-0.75
损失大	-1
损失无穷大	$-\infty$

如果比较两种方法的特点,收益率法貌似客观,但实际上这种方法其实并不准确。这是因为,人们对行为的期望回报的估计,实际上

受到许多具体因素的影响。只有在估计期望回报时考虑这些因素,所得到的结果才可靠,即人们是根据自己的主观感觉而不是客观数据去选择行为的。

影响期望回报估计值的第一个因素,是人们的地位差异所引起的边际效用的差别。不同地位的人对同样收益率,所感觉到的重要性并不一定相同,这是因为边际效用的原因:100%的收益率对于经济贫穷的人来说可以是相当大的收益,但对于经济上已经富足的人来说,并不算什么重大的成绩。这个因素,收益率法是反映不出来的,但语言变量法却能够很好地反映。

第二个因素是认知偏差。认知偏差是人们在期望回报的认知过程中存在的一些规律性的偏差,这些偏差导致人们对期望回报的估计会在一定程度上偏离客观值。由于所估计的是当前尚未实现的期望回报,所以认知偏差的影响往往比当前客观现实的估计还要明显一些。而经济收益率法由于是客观数据,反而不能反映人们的认知特征。

第三个影响因素是人们的价值观。人们的追求常常具有很大的差异,有人追求经济,有人追求地位,有人追求奉献,有人只顾索取。这个因素,收益率法也反映不出来,只有语言变量法才能够反映出来。

此外,由于收益率法只适用于评估经济行为,因而应用范围比较有限,而语言变量法则适用面很广泛。

但是,收益率法的最大优点,就是从客观的统计数据中直接获取期望回报值,比较节省成本。而语言变量法要专门通过大量的调查才能得到期望回报值,成本比较高。

第四节 行为概率模型

一、模型的基本形式

设行为主体的行为集为 A_i , 其中一共有 N 个行为,该行为集满

足行为选择的唯一性,行为主体必选择且只选择其中一个行为 a_i , 则行为概率模型如下:

$$\left. \begin{aligned} p(a_1) &= \frac{e^{\beta_{10} + \beta_{11} u(a_1) - \beta_{12} u(a_2) - \dots - \beta_{1N} u(a_N)}}{1 + e^{\beta_{10} + \beta_{11} u(a_1) - \beta_{12} u(a_2) - \dots - \beta_{1N} u(a_N)}} \\ &\vdots \\ p(a_i) &= \frac{e^{\beta_{i0} - \beta_{i1} u(a_1) - \beta_{i2} u(a_2) - \dots + \beta_{ii} u(a_i) - \dots - \beta_{iN} u(a_N)}}{1 + e^{\beta_{i0} - \beta_{i1} u(a_1) - \beta_{i2} u(a_2) - \dots + \beta_{ii} u(a_i) - \dots - \beta_{iN} u(a_N)}} \\ &\vdots \\ p(a_{N-1}) &= \frac{e^{\beta_{(N-1)0} - \beta_{(N-1)1} u(a_1) - \beta_{(N-1)2} u(a_2) - \dots + \beta_{(N-1)(N-1)} u(a_{N-1}) - \beta_{(N-1)N} u(a_N)}}{1 + e^{\beta_{(N-1)0} - \beta_{(N-1)1} u(a_1) - \beta_{(N-1)2} u(a_2) - \dots + \beta_{(N-1)(N-1)} u(a_{N-1}) - \beta_{(N-1)N} u(a_N)}} \\ p(a_N) &= 1 - p(a_1) - \dots - p(a_i) - \dots - p(a_{N-1}) \end{aligned} \right\} \quad (1-1)$$

模型中,各符号的意义如前面概念部分所定义,其中的 β_{ij} 为回归系数,其取值区间为 $\beta_{i0} \in (-\infty, \infty)$, 当 $i \in [1, 2, \dots, N]$; $\beta_{ij} > 0$, 当 $i \in [1, 2, \dots, N]$, $j \in [1, 2, \dots, N]$ 。

对于 β_{ij} (当 $i \in [1, 2, \dots, N]$, $j \in [1, 2, \dots, N]$) 来说,在 $\beta_{ij} > 0$ 的情况下,只有 β_{ii} 前面的符号是正的,而其他 β_{ij} 前面的符号都是负的,这表示 $p(a_i)$ 随着 $u(a_i)$ 的提高而增加,随着 $u(a_j)$ 的提高而减少 ($j \neq i$)。这实际上反映了同一行为集中的各行为之间的资源制约关系:由于行为资源有限,所以只有同一行为的期望回报对该行为概率有正面影响,而其他行为的期望回报对其的影响是负的,即如果某个行为的期望回报提高,则该行为的概率也会变大,但如果其他行为的期望回报提高,则该行为的概率会变小。

从模型可以看出,在其他行为的期望回报不变的情况下:

$$\left. \begin{aligned} \lim_{u(a_i) \rightarrow \infty} p(a_i) &= 1 \\ \lim_{u(a_i) \rightarrow -\infty} p(a_i) &= 0 \end{aligned} \right\}$$

可见 $p(a_i)$ 的值域的确符合概率的要求。