

以数据挖掘方法评估 投资者风险承受能力的应用

❶ 华西证券有限责任公司证券研究所投资者、投资管理研究小组*

一、前言

作为现代金融服务的基本原则和基本要求，投资者适当性制度是规范投资者和金融机构之间权利义务关系的重要制度，已为主要资本市场采纳。究其含义，是指金融机构在向其客户或潜在客户提供金融产品或服务时，应当确保其所提供的金融产品或服务与特定客户的财务状况、投资目标、知识和经验以及风险承受能力等相匹配¹。

适当性工作主要包括两个核心内容：了解客户与了解产品。了解客户是适当性工作的关键，也是适当性工作开展的基础。金融机构只有全面了解客户的信息，才能根据每一个客户的综合情况，向其提供适当的金融产品和服务，否则做出的判断是不准确、甚至是错误的，可能导致客户投资损失、金融机构的声誉和经济利益受损²。本文从证券公司了解客户工作中遇到的问题着手，

探讨了解客户的关键指标及其评价模型，发展出一套以客户账户数据为基础评估客户风险承受能力的实用方案。

二、证券公司了解客户工作中遇到的问题

根据中国证券业协会发布的《证券公司投资者适当性制度指引》规定，证券公司向客户销售金融产品或提供金融服务时需要了解客户的身份信息、财务状况、投资知识、投资经验、投资目标、风险偏好等相关信息。这些规定在执行中仍面临一些挑战：

（一）复杂性问题

需要了解的客户信息众多，缺乏结构、标准。对于为什么了解这些信息、怎样了解这些信息、了解这些信息之后怎样得出结论以及得出什么样的结论等问题容易认识不清，执行难度大。

* 小组成员：郝绎、王翠、王正、吴光胜、鄢黎

¹ 参见《投资者适当性制度是一项投资者保护制度》，网址：http://www.sac.net.cn/hyfw/hydt/201203/t20120326_43471.html

² 参见《五方面理解证券公司投资者适当性制度》，网址：http://www.cs.com.cn/xwzx/15/20120412_2/01/04/201204/t20120412_3313709.html

（二）成本效益问题

我国个人投资者众多，客户信息获取、分类、追踪、更改等工作需要证券公司大量技术与人力的投入，但是短期效益并不明显。因此，不少证券公司以被动满足各项监管及合规为主，多为一次性地对投资者风险承受能力做出评估，后续跟踪服务明显不足。此外，“了解客户”目前并未真正对证券公司带来业绩提升，导致投资顾问或产品销售者对此不够重视，整个工作流于形式，甚至存在为赚取佣金或达成业绩向投资者提供不合适的产品或服务的情况。

（三）信息可得性问题

现在获得的客户信息以开户信息、问卷测评和访谈为主，部分问题涉及客户隐私，投资者参与意愿不高，在有效配合方面存在保留，部分信息的可得性存在问题。

为了解决以上问题，更好地开展适当性管理工作，证券公司有必要研究一种简便、高效并且可大规模运用的方法，确保投资者基本的适当性得到满足。

三、用风险承受能力作为了解客户的关键指标

为了解决以上问题，我们尝试在众多信息中确定对客户适当性判断最为关键的因素，并舍弃相对次要的信息，把复杂的问题简单化。这种方法虽然不能全面反映投资者各方面的特征，但却能保证投资者适当性得到基本的满足。

通过研究和梳理，我们发现证券公司收集的众多客户信息可以帮助我们解决了解客户三个层面的问题（如图1所示）：（1）投资者是否有资格进入市场进行投资；（2）投资者在投资过程中能承受怎样的风险；（3）投资者的投资意愿及偏好。

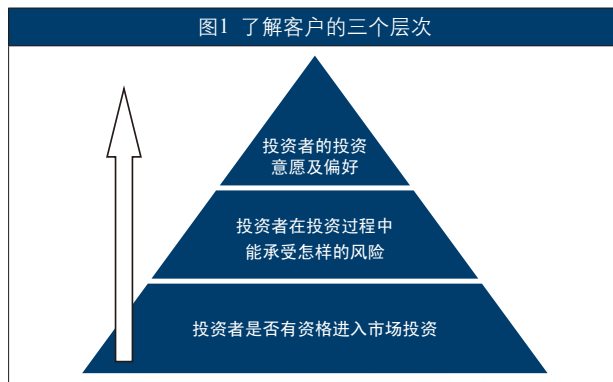
这三个层面的问题层层递进：

首先，证券公司需要判断投资者是否有资格进入市场投资，证券公司应按照交易所相关规定

严格落实准入工作。这方面的信息获得及判断比较容易做到，也是金融机构的法定义务。

在此基础上，证券公司需要了解投资者在投资过程中能够承受多大的风险。风险承受能力相对而言比较客观，是投资者能不能购买某类金融产品或服务的决定性因素。如果购买的金融产品或服务与投资者的风险承受不相匹配，投资者可能遭受不可接受的损失乃至彻底退出证券市场。基于风险承受能力的适当性管理多属于行业自律范畴，是投资者适当性的主要内容。

最后，证券公司在保证投资者于适当风险下进行投资的基础上，可以根据投资者的投资意愿和偏好提供更个性化的服务，以此来提高投资者在投资过程中的舒适度以及对证券公司服务的满意度。基于投资者偏好的适当性多属于证券公司自律范畴，既需要证券公司更多的投入，也能给证券公司带来差异化的竞争优势。



由上述分析可以发现，投资者的风险承受能力关系到其能否在市场中生存，是证券公司了解客户的核心因素。证券公司必须在满足投资者风险承受能力适当性匹配的基础上才能为其提供进一步的服务。证券公司可以运用该指标作为了解客户的核心指标，以此给投资者匹配风险承受适当的金融产品或服务，满足投资者基础的适当性需求，把复杂的问题简单化。

四、风险承受能力评价模型的构建

（一）风险承受能力的影响因素

投资者的风险承受能力受多方面因素影响，

不同的人由于家庭财力、投资时机、个人投资取向等因素的不同，风险承受能力各不相同；同一个人也可能因时间、市场环境及其他因素的变化，而表现出不同的风险承受能力。因此，风险承受能力不是一个单一的指标，而是投资者几方面情况的综合反映。

国外研究发现，投资者的风险承受能力包含两方面的内容³，一方面是风险容量，即投资者可以承担多大的损失；另一方面是风险感知，即投资者可以认知和控制多大的风险。这两方面分别从客观和主观上揭示了投资者承受风险的能力，对投资者的风险承受能力起到决定性的影响。在实际运用中，我们一般用投资者的财务能力表征其风险容量，用投资经验和投资能力表征其风险感知。

1. 财务能力

财务能力是投资者拥有的财富、收支状况、财务负担的综合体现，是衡量风险容量最直接、最关键的指标。财务能力越强，投资者能承担的损失就越大，风险承受能力也就越强。具体来说，财务能力主要包含以下几个因素：

(1) 当前拥有的财富多少，这一因素决定了投资者当前能承担的最大损失。

(2) 收入情况，包括收入的高低、收入的稳定性、工作保障等，主要影响了投资者从损失中恢复的能力。

(3) 财务负担，包括日常开支、大项支出、子女抚养与父母赡养投入等。财务负担越重，投资者能承受的投资损失就越小，从损失中恢复的能力就越弱，风险承受能力也就越弱。

2. 投资经验和投资能力

投资经验和投资能力包括投资者对投资标的的了解程度、对市场的认知、信息获取能力、分析能力等，它决定了投资者是否具备参与的条件（可以做出相应的投资决策并理解投资风险），

以及是否具有规避和控制风险的能力。证券市场的一个重要特点是影响因素众多、不确定性强，而投资决策能力是在这一市场成功的关键。投资能力差的投资者不能很好地判断市场，承受市场波动的能力也较差；投资能力强的投资者能很好地了解市场，理性应对市场变化，控制风险、规避风险的能力就越强。

除了以上两点，针对适当性管理工作，我们添加了年龄作为影响投资者风险承受能力的平衡因素。这是因为：（1）在各国适当性管理中，年龄都是一个重要因素，如美国和日本金融监管机构就对老年投资者的适当性做出了特别规定。

（2）年龄与生活状态有较高的相关性，可以评估投资者的收入状况和财务负担。（3）年龄的大小可以评估投资者用于积累财富和投资的时间跨度。年龄越小，用来积累财富的时间跨度就越长，也就拥有越强的从损失中恢复的能力。

（二）风险承受能力的评价方法

确定风险承受能力作为了解客户的核心指标后，需要明确风险承受能力的评价方法。表1比较了几种常见方法，其中：访谈的结果准确度最高，可以覆盖风险能力和风险意愿，但是投入成本极高、对相关咨询服务人员的要求也很高，分析的难度大，而且跟踪评价难度极大；问卷调查是现在风险承受能力评价的主要方法，侧重于对风险偏好和意愿的评价，这种方式可以快速获取评价结果，但是评价结果精确度低、客户的参与意愿不高，跟踪评价的难度也较大；数据分析方法在评价风险能力上更有优势，不仅投入少、便捷性高，而且易于持续跟踪分析，但是这一方法受数据覆盖面、数据质量等因素影响。鉴于证券公司数据仓库中不断增加的海量客户注册信息及交易数据，我们认为数据分析是目前较为理想的风险承受能力的评价方法。

³ Michael Kitces . The Three Faces Of Risk Tolerance [J].: Inside Information , 2009 (8) .

表1 风险承受能力主要评估方法比较

	优点	缺点
访谈	◎ 针对性强, 评估结果准确度极高	◎ 投入人力成本极高, 效率低 ◎ 对咨询服务人员要求高 ◎ 分析难度大 ◎ 跟踪评价难度极大
问卷调查	◎ 适用范围广 ◎ 可快速获取结果	◎ 评价结果准确度低, 无法精确分析 ◎ 客户参与意愿不高 ◎ 跟踪评价难度大
数据分析	◎ 投入成本低, 评价结果相对准确 ◎ 便捷性高, 方便持续性跟踪评价 ◎ 灵活性高, 易于根据业务变化做调整	◎ 对数据要求高, 评价精确度受数据质量影响较大

(三) 风险承受能力评价模型的表征指标体系

基于财务能力、投资经验和投资能力、年龄这三方面因素, 我们构建的投资者风险承受能力的分类指标体系如表2所示:

表2 风险承受能力评价模型指标体系

一级指标	二级指标	主要观测点
财务能力	账户平均资产	过去三年的平均资产
	账户最高资产	过去三年中的最高单日资产
投资经验和投资能力	开户年限	根据开户时间计算出的年限
	交易次数	过去三年的有效交易次数
	胜率	过去三年的交易胜率
年龄	年龄	根据注册信息计算出的年龄

1. 财务能力的表征指标

我们采用投资者的账户平均资产和账户最高资产两个指标来表征其财务能力, 理由如下: 长期而言(我们的样本时期是三年), 账户平均资产通常是客户在扣除日常支出之后的资产, 综合考虑了财富、收入以及负担; 股票账户的资产应是客户能承受较大风险的一部分资产, 尽管客户还可能还有其他资产, 但用股票账户里的资产评估客户的风险承受能力通常不会高估, 符合审慎的原则。

账户的平均资产可以用来表征投资者账户资产在一段时间内的基本情况。但是仅用平均资产一般会低估客户的实际可投资资产。在通常情况下, 投资者的可投资资产会在证券资产和其他金融资产间转换。例如, 有些投资者会不定期向证

券账户转入现金, 还有些投资者喜欢在股市交易清淡时把资金转出股市, 等市场好转再转回来。所以账户平均资产只反映了账户的基本情况而忽略了账户资金的波动情况。为此, 我们用账户最高资产作为调整指标。一段时间内的账户最高资产代表着投资者可投资资产的峰值, 也代表着投资者可投资资产的潜力。比如说某个投资者在三年内账户的平均资产为20万, 账户的最高资产为100万, 我们就认为投资者的可投资资产很可能是大于20万。选用这个指标可以在账户平均资产的基础上进一步准确估计其资产情况。

在以上指标选择的基础上, 为了避免不同地域经济发展水平的影响, 我们还需要根据不同地域的收入水平对结果进行修正。

2. 投资经验和投资能力的表征指标

投资者的投资经验和投资能力都是比较主观的概念, 是模型中比较难以量化的部分。对于这一指标, 我们通过投资者开户年限、交易次数、胜率的综合评价使结果尽可能准确。

开户年限和有效交易次数主要用于衡量投资者的投资经验。一般来说, 开户年限越长, 其经历的市场波动次数、范围也就越多, 投资经验也就越丰富。同理, 有效交易次数⁴越多, 其参与市场的机会、尝试过的投资策略也就越多, 投资经验也就越丰富。需要注意的是, 对于有效交易次数占比较低的客户⁵, 这一指标并不能很好反应投资者的交易能力, 计算其交易成功率的意义也不大, 所以对这类客户的风险承受能力并不能通过此评价方法来评定。

交易胜率主要从投资绩效的角度评估投资者的投资能力, 一般而言, 长期投资绩效好的投资者投资能力较强。即使有些投资者开户时间较长, 有效交易次数也较多, 但是交易胜率很低,

⁴ 有效交易次数指过去一段时间中, 买入、卖出可以匹配的股票支数, 即客户所有交易过的股票中, 可以完成买卖匹配(包括分红配股)的股票数量。

⁵ 我们定义为交易匹配率在60%以下的客户。

我们也不能认为其具有较好的投资能力。

3. 年龄的表征指标

年龄的表征指标比较简单，可以直接采用投资者的生理年龄表征。

(四) 风险承受能力评价相关指标的度量及标准化处理

三个维度的指标量纲都不同，我们对每类指标采取分档的方式进行标准化。其中，财务能力分为五档，投资经验和投资能力分为三档，年龄分为四档。对每个客户相关指标进行度量后，都会经过处理在这三个维度上分别有对应的分档值，然后通过分类处理获得最终结果。

1. 财务能力

我们用投资者账户资产表征其财务能力。这一指标用投资者过去三年的账户平均资产和最高单日资产的加权和来表征，计算公式为：

$$\text{账户资产} = \omega_1 * \text{平均资产} + \omega_2 * \text{最高单日资产}$$

我们采用了阶梯权重的形式⁶（见表3），权重值由平均资产与最高单日资产的差异程度决定，比如投资者账户三年内的最高单日资产是平均资产5倍以上时，由于两者差异较大，表明其最高单日资产可能是比较异常的状况，所以赋予账户平均资产70%的权重、单日最高资产30%的权重。

账户资产是模型中最重要最客观的指标，所以我们尽可能区分得更细一些，分为5万以下、5-15万、15-50万、50-200万、200万以上五档⁷。

表3 资产计算表

最高单日资产/平均资产	≤2	2< <5	≥5
ω_1	0.3	0.5	0.7
ω_2	0.7	0.5	0.3

2. 投资经验和投资能力

投资经验和投资能力是比较主观的概念，很难用数据直接量化。对于这一指标的度量，我们

通过分类评估法把投资者的投资经验和投资能力确定为三档：将开户年限、有效交易次数和交易胜率都分为三档，并对因此产生的二十七种组合的每一种情况进行评估，分别给出每一种组合的投资经验和投资能力的分档值，示例如表4。

(1) 开户年限分为1年内、1-2年及2年以上三档。

(2) 有效交易次数分为10次以内，10-20次，20次以上三档。

(3) 胜率分为低、中、高三档。由于胜率还受市场状况的影响，所以我们采用相对评价法，将所有投资者的胜率由高到低按2:3:5的比例划分，分别对应高中低三档。其中，胜率是指过去三年内，投资者的完整交易行为中盈利次数的占比。

表4 投资经验和投资能力评价示例

开户年限	有效交易次数	胜率	投资能力和投资经验分档
2年以上	10次以内		低
	10-20次	低	低
	10-20次	中或高	中
	20次以上	低	低
	20次以上	中	中
	20次以上	高	高

3. 年龄

借助成熟并在理财中广泛应用的生命周期模型，我们将年龄划分成30岁以下、30-50岁、50-65岁及65岁以上四个年龄段（如图2所示）：

(1) 30岁以前：通常单身或刚组建家庭，一般收入较低、积蓄较少。支出以日常消费、继续教育、交际娱乐支出为主，是为下一阶段进行知识储备、资金储备的时期。随着工作年限的增长、工作经历的丰富，预期收入将增长。

(2) 30-50岁：收入随工作经历的增长而增加，但子女教育、赡养长辈等会使经济负担加重，同时工作压力加大、竞争激烈，人的健康状

⁶ 对于三年最大资产低于10万或一年最大资产小于1万的客户，其风险评估的价值不大，因而可忽略这部分客户，或者直接定义为低风险承受能力的客户

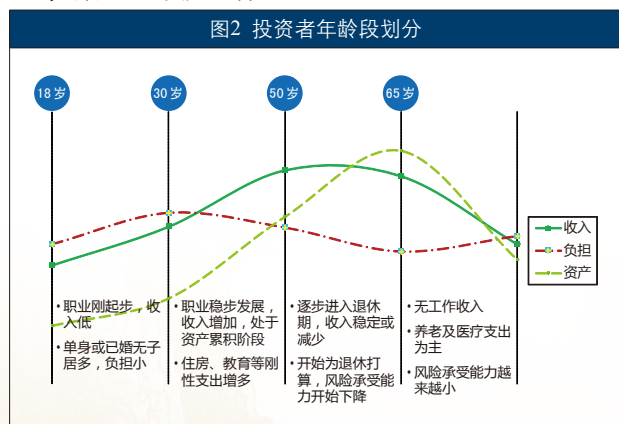
⁷ 资产阈值的设定可根据实际情况动态调整

况可能有所下降，这个阶段的医疗保险费用可能开始增加，风险承受能力相对单身期有所降低，出现购房、购车等大项支出。

(3) 50-65岁：退休前期，这个时期家庭稳定，子女经济逐渐独立，家庭负担降低。但随着步入老年，身体机能开始逐步降低。这个阶段的支出以日常消费、医疗保障和积累退休金为主。由于临近退休，风险承受能力进一步降低。

(4) 65岁以上：退休期，收入基本靠养老金，支出以日常消费、医疗保健和休闲娱乐为主。退休期的风险承受能力很低，对资金安全的要求远大于对收益的要求。

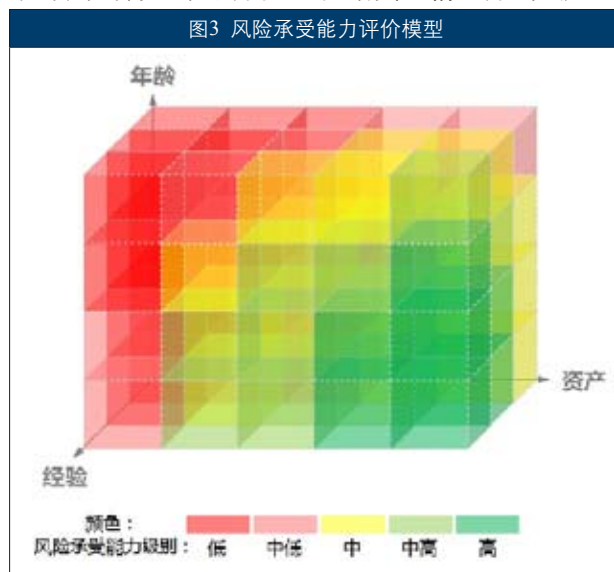
一般来说，年龄段的划分与年龄评级一致。但是，由于年龄在风险承受能力的评估中还隐含了对财务负担的衡量，因而我们还需要借助客户注册信息中婚姻状况、家庭人口等数据对年龄评估结果做个性化的修正，如对某28岁已婚且育有子女的投资者，其年龄评级应与大多数30-50岁这一年年龄段的评级一样。



(五) 分类法与专家评判法构建模型

风险承受能力的判断是一个很复杂的问题，目前多数金融机构采用问卷打分、加总，然后按总分将客户的风险承受能力分为三档或五档，这种方法在分数加总、分档的过程中存在大量的信息损失。为了减少这种信息损失，我们用分而治之的手段，采用组合分类的方法构建风险承受能力分类模型（如图3所示）：在表征财务能力、投

资经验和投资能力及年龄这三个指标的基础上，把投资者分为 $5 \times 4 \times 3 = 60$ 个细分类型，并运用专家评判法将60个细分类型根据实际情况分为5级。



如前文所说，风险承受能力是投资者综合情况的反应，很难用定量化的方法给出结果。专家评判法很好的解决了这个问题，用定性的方法给出定量化的结果。多位专家凭借各自的经验，按照细分的投资者情况对其风险承受能力给出合理的估计。

例如，投资者的细分类型是资产5万元以下，年龄30岁以下，投资经验和投资能力高，专家通过讨论认为，尽管这类投资者处于可以承受较高风险的生命周期阶段，并且有很高的投资经验和能力，但是由于资产过低，其实际的风险承受能力应该偏低，属于中低级别。

另一个例子，投资者的细分类型是资产50-200万，年龄65岁以上，投资经验和投资能力低，专家通过讨论认为，这类投资者拥有一定的资金量，本来可以承担较高的风险，但是鉴于年龄过高并且投资经验和能力不高，所以其实际的风险承受能力偏低，属于中低级别。具体示例见表5。

(六) 评价模型的有效性验证

为检验模型的有效性，我们选择了我司S、C两个营业部2009年3月30日至2012年3月30日的抽样数据来验证。

表5 专家打分示例

资产	年龄	经验	风险承受能力
15-50万	30岁以下	低	中低
		中	中高
		高	中高
	30-50岁	低	中低
		中	中
		高	中高
	50-65岁	低	中低
		中	中
		高	中高
	65岁以上	低	低
		中	中低
		高	中

1. 数据需求及预处理

由于需要数据来确定胜率、平均单次盈利、资产的阈值，需要获得的字段包括：客户id、开户时间、是否连续三年未交易（开户年限三年以上）、出生日期、过去三年的平均资产、过去三年的最高单日资产、过去一年的平均资产、过去一年的最高单日资产、过去三年可匹配股票数、可匹配股票中的盈利股票数。

为保证数据有效性，在获取数据前需要对样本进行如下预处理：

(1) 剔除数据期内未发生交易的休眠客户；

(2) 删除起始日期中交易标的不是A股的持仓数据；

(3) 定义起始日（20090330）所有持仓股票为当日买入，定义终止日（20120330）所有持仓股票为当日卖出；

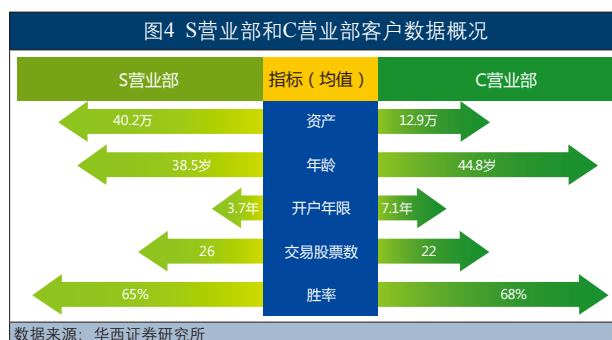
(4) 定义托管转入为当日买入，定义托管转出为当日卖出，并将此买卖记录添加进入股票交易表；

(5) 如当天某一只股票有一次以上的买卖记录，将当天的买卖记录进行合并，统一定义为买一次或者卖一次；

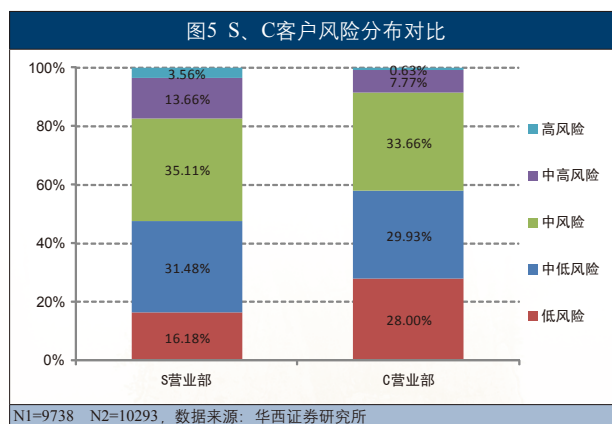
(6) 删除打新股记录（有卖无买）。

2. 数据概况及风险分布

初步的统计发现，S营业部客户更年轻，开户年限较短，但资产量较大，交易比较频繁，高风险客户更多。S营业部和C营业部客户基本情况对比如下：



在提取了基础数据并计算出两家营业部投资者风险承受能力的分布数据后，可以发现：风险承受能力评级为高的客户在两个营业部都只占很小比例，评级为低和中低的客户在两个营业部均达到50%或以上；资产均值较低的C营业部风险承受能力评级为低的客户比例更高。

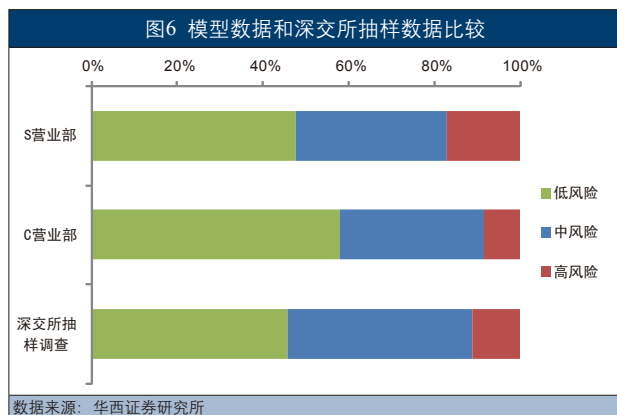


3. 风险承受能力分布状况的验证

深圳证券交易所发布的《中国股市个人投资者状态调查报告》显示，我国个人投资者的风险能力不容乐观，高达45.9%的个人投资者风险承受能力很弱，42.9%的个人投资者风险承受能力一般，只有11.2%的个人投资者风险承受能力较强⁸。

⁸ 深圳证券交易所综合研究所. 中国股市个人投资者状态调查报告[R].深圳: 2002.

因为调研数据多把个人投资者的风险承受能力分为三类，所以我们把模型中的五类风险承受能力评级做了相应的合并处理：中高风险和高风险合并为“高风险承受能力”、中低风险和低风险合并为“低风险承受能力”，中风险不变。分布情况的比较如图6所示，模型计算出的风险承受能力分布情况和调查报告的分布情况相类似，均表现出我国个人投资者中风险承受能力为中低和低的占大多数、风险承受能力高的占少数这一特点。通过与统计数据的对比，我们可以说模型具有一定的有效性。



五、风险承受能力评价模型研究的创新之处

相比以往的评价方法，我们认为，风险承受能力评价模型研究的创新之处有如下几点：

1. 明确了解客户三个层面的问题：（1）投资者是否有资格进入市场进行投资；（2）投资者在投资过程中能承受怎样的风险；（3）投资者的投资意愿及偏好。确定了风险承受能力为投资者适当性的核心属性。

2. 以财务能力、投资经验和能力这两个概念来对应风险承受能力的风险容量和风险感知这两个属性。

3. 建立了财务能力、投资经验和能力的指标表征体系，结合年龄这一平衡和综合因素，以分档、分类加专家判断的方式，提高了评估结果的合理性和准确度，有助于投资顾问缩小目标客户群体，也有助于在业务中对适当性进行更全面评估。

4. 系统定期自动分析投资者风险承受能力变化情况，降低了投资顾问的跟踪服务压力。SAC

参考文献：

- [1] 中证协发[2012]248号，证券公司投资者适当性制度指引，2012年12月。
- [2] 巴塞尔联合论坛，向零售客户销售适合的金融产品及服务报告（Customer suitability in the retail sale of financial products and services），2008年4月。
- [3] Robert H. Mundheim. Professional Responsibilities of Broker-Dealers: The Suitability Doctrine[J]. Duke Law Journal, 1965(3): 445-480.
- [4] Jaimie Sung, Sherman Hanna. Factors Related To Risk Tolerance[J]. Financial counseling and planning, 1996(7): 11-19.
- [5] Invesco Ltd. Rethinking Risk: Risk Tolerance = Ability + Emotion. 网址：<http://www.invescoaim.com/pdf/PPRRRT-BRO-1.pdf>.
- [6] Michael Kitecs. The Three Faces Of Risk Tolerance[J]. Inside Information, 2009(8): 7-10.
- [7] 杨巍，刘正瑶. 我国证券市场投资人分类管理制度研究——以商主体法定为视角[J]. 投资研究，2011(6): 16-19.
- [8] 赵德武，马永强等. 决策能力、风险偏好与风险资本[J]. 会计研究，2004(4): 52-58.
- [9] 马莉莉，李泉. 中国投资者的风险偏好[J]. 统计研究，2011(8): 63-72.
- [10] 深圳证券交易所综合研究所. 中国股市个人投资者状态调查报告[R]. 深圳：2002.
- [11] 谢磊. 我国股票投资者投资风险研究[D]. 长沙：中南大学，2006.