

投资者问卷调查实证研究
耿中华*

该表涵盖四个方面,包括近百个问题,马克思以此来全面了解工人的劳动、生活和思想状况。20世纪以来,结构式的调查问卷被越来越多地使用,特别是在满足定量需求方面。目前,调查问卷通常被定义成由一系列与研究目标有关的问题组成的问题集(questionnaire)。这些问题对实现调查目的至关重要,以需要调查的内容为依据设计而成,形式上多采用各种类型的题目,

续表

项目	完全同意	同意	无所谓	不同意	完全不同意
很难看	1	2	3	4	5
采用好的材质	5	4	3	2	1
设计太差	1	2	3	4	5
颜色太少	5	4	3	2	1
款式太少	1	2	3	4	5

(二) 语义差异量表
 语义差异量表 (semantic differential scale) 是一种态度量表, 结构上是由两部分构成: 第一部

(三) 二分量表
 二分量表 (dichotomous scale) 是对于极端表述的调查, 通常是对两个相互排斥的项目进行回答。二分量表一般以“是”和“否”, 或者“同意”和“不同意”形式出现, 可以理解成是语义差异量表的一种极端简化。(见表 3)

表 3 二分量表示例

事项	是	否
1. 今天你是否忘记服药?	是	否
2. 过去一周你是否忘记服药?	是	否
3. 治疗期间, 当你感觉身体不适时, 你是否咨询过医生?	是	否

(四) 等级顺序量表
 等级顺序量表 (ordinal scale) 通常是用来对多个待了解者排序。等级顺序量表与其他量表最大的不同在于, 该量表是将全部待了解对象同时展示给被调查者, 并要求他们对这些对象排序或打分。例如, 表 4 所示就是请被调查者对 7 个汽车品牌进行排名, 每个品牌都会有相应的、

目的。

(二) 非量表问卷分析步骤

相对于量表问卷,非量表问卷用到的复杂统计模型和数据处理方法相对较少,但分析思考框架基本类似。统计方法上,一般是通过频数分

卷作答时长与投资者回答真实有效性之间的关系、作答激励对投资者问卷作答真实性的影响三方面。对于影响关系分析,由于因变量属于二元变量,可采用 Logistic 回归分析。为了尽量消除样本量不同对结果的影响,使对比效果更准确,本文通过选取 2019 年、2021 年的投资者知权、行权、维权现状调查数据做实证对比研究,探究影响关系,其中 2019 年有效样本 12,116,2021 年有效样本 12,464。

(四) 2019 年实证分析

本文涉及 10 个影响投资者准确做出维权意愿的自变量

表 5 2019 年变量两两之间的相关性

变量	答题时	地区	年龄	性别	教育背景	职业	金融从业人员	入市时间	入市资金	股票资金占家庭资产之比
皮尔逊相关性	-.029**	.000	-.181**	.078**	.059**	.063**	.228**	-.041**	-.007	-.012
Sig. (双尾)	.001	1.000	.000	.000	.000	.000	.000	.000	.453	.191
个案数	12, 116	12, 116	12, 116	12, 116	12, 116	12, 116	12, 116	12, 116	12, 116	12, 116

注.**在.01级别(双尾),相关性显著。

*在.05级别(双尾),相关性显著。

通过相关性分析发现,维权意愿的真实表达与投资者的答题用时、年龄、性别、教育背景、职业、是否属于金融从业人员、入市时间、股票资金占家庭资产之比有相关性,与投资者的地区、入市资金无相关性。

3. 多变量叠加分析
进行单变量分析后,本文通过 SPSS 的 Logistic 回归模型验证多变量叠加后的综合作用,探究多因素对投资者回答真实性的影响,结果见表 6。

表 6 2019 年全变量 Logistic 回归检验

变量	B	标准误差	瓦尔德	自由度	显著性	Exp(B)
答题用时	.000	.000	12.257	1	.000	1.000
地区	-.001	.002	.355	1	.551	.999
年龄	-.409	.027	226.664	1	.000	.664

但通过线下面访发现，投资者回答在“入市资金”这一类问题时，相对于单一地回答入市资金总额，往往不太愿意回答入市资金占家庭总资产的比例。其背后原因一方面可能是入市资金属于一个绝对数，在记忆和计算上更方便；另一方面可能是基于隐私的考虑，投资者希望保护自己的财富信息。为此，需要验证入市资金和股票资金占家庭总资产之比对最终回答结果的实际影响有多大区别，再结合投资者偏好做出最终保留谁的决断。下面先进行保留股票资金占家庭总资产之比计算，而后进行保留入市资金计算，结果见表7。

表7 2019年 Logi

部显著。

上述两个变量的计算结果显示,股票资金占家庭资产之比与入市资金占家庭资产之比这两个变量,各自都对投资者准确回答有显著影响,当两者合并出现时入市资金占家庭资产之比会更为显著。因此,可以推断,当调查对象是广泛性的大众投资者或入市资金占家庭资产之比敏感度不高时,基于问卷应该有的无差别接受性前提要求,应该选择入市资金占家庭资产之比这个绝对指标。当调查对象是小部分的特殊群体或调查本身是对某一类群体的专享调查时,应该优先选择股票资金占家庭总资产之比这个相对指标。

4. 2019 年三权调查数据实证结论

通过全变量回归检测,再结合相关性分析可以发现:

第一,投资者的年龄、性别、教育背景、是否金融从业、入市时间、入市资金、股票资金占家庭总资产之比会对投资者的维权意愿产生统计意义上的显著影响。

第二,投资者的地区分布不会对维权意愿产生统计意义上的显著影响,即不同地域的投资者受到损失后,维权意愿没有显著差异。

第三,投资者答题用时不会对投资者真实意愿的表达产生实际影响,不会对投资者真实意愿的表达产生统计意义上的显著影响。

个自变量综合影响后的结果,因此,对于7个变量之间是否有相关性以及关联性的强弱,不是本次实证分析重点关注的內容。本次重点关注投资者在受到损失后,7个单因素对维权选择是否有影响,7个因素的叠加影响使投资者在问卷上反映出的信息是否真实有效,以及推断投资者对问卷的可接受度。

本文通过 SPSS 软件,运用双变量相关性检验判断初始影响因素,结果见表 9。

表 9 2021 年双变量之间的相关性

变量		所用时间	省市	年龄	教育背景	入市时间	入市资金	风险承受等级
维权选择	皮尔逊相关性	-.128 **	.013	-.115 **	-.055 **	-.058 **	.073 **	-.071 **
	Sig. (双尾)	.000	.133	.000	.000	.000	.000	.000
	个案数	12,464	12,464	12,464	12,464	12,464	12,464	12,464

56

3 22

变量全部显著, 结果见表 11。

表 11 2021 年变量 Logistic 回归模型 (剔除地区因素)

变量 答题用时	B	标准误差	瓦尔德	自由度	显著性	Exp(B)
------------	---	------	-----	-----	-----	--------

职业因素却不会。

二是地域因素的确是众多调查中需要重要考虑的因素,但对于投资者受损后的维权意愿而言,投资者的地区分布不同不会对维权意愿产生统计意义上的显著影响。

标。当调查涉及的对象是小部分的特殊群体或调查本身是对某一类群体的专项调查时,应该优先选择股票资金占家庭总资产之比这个相对指标。这些措施可以在