

基于Lasso+分位数回归的上市保险公司市盈率微观影响因素分析

张春梅¹, 赵明清^{2*}, 姜丽慧²

¹青岛黄海学院大数据学院, 山东 青岛

²山东科技大学数学与系统科学学院, 山东 青岛

收稿日期: 2021年10月23日; 录用日期: 2021年11月13日; 发布日期: 2021年11月29日

摘要

市盈率是投资者选择保险业上市公司进行投资时首先要考虑的一个重要财务指标, 但关于上市保险公司市盈率影响因素的量化研究并不多见。本文利用2011~2019年我国保险业上市公司面板数据, 运用Lasso+分位数回归对我国上市保险公司市盈率的微观影响因素进行了分析, 结果表明: 每股净资产和资产净利率等7个财务指标对其有着显著正向影响, 而营业收入增长率和权益乘数对其有着显著负向影响; 在不同的分位数水平下, 各影响因素对其作用大小也有所不同。

关键词

上市保险公司, 市盈率, 面板数据, Lasso, 分位数回归

Analysis on the Microcosmic Affection Factors of the P/E Ratio of Listed Insurance Companies Based on Lasso+ Quantile Regression

Chunmei Zhang¹, Mingqing Zhao^{2*}, Lihui Jiang²

¹Big Data College, Qingdao Huanghai University, Qingdao Shandong

²College of Mathematics and System Science, Shandong University of Science and Technology, Qingdao Shandong

Received: Oct. 23rd, 2021; accepted: Nov. 13th, 2021; published: Nov. 29th, 2021

*通讯作者。

文章引用: 张春梅, 赵明清, 姜丽慧. 基于 Lasso+分位数回归的上市保险公司市盈率微观影响因素分析[J]. 应用数学进展, 2021, 10(11): 4024-4036. DOI: 10.12677/aam.2021.1011428

Abstract

The P/E ratio is an important financial index that investors must first consider when choosing a listed insurance company for investment. However, quantitative research on the factors affecting the P/E ratio of listed insurance companies is rare. Based on the panel data of China's listed insurance companies from 2011 to 2019, this article analyzes the microcosmic affection factors of the P/E ratio of Chinese listed insurance companies using the methods of Lasso+ quantile regression. The results show that: 7 financial indicators, such as net assets per share and net asset interest rate, have a significant positive impact on the P/E ratio; while the operating income growth rate and equity multiplier have a significant negative impact on the P/E ratio; the effect of each influencing factor on the P/E ratio varies due to different quantile levels.

Keywords

listed Insurance Companies, P/E Ratio, Panel Data, Lasso, Quantile Regression

Copyright © 2021 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

保险业是现代金融系统的三大支柱产业之一，其发展程度是衡量国家经济发展的重要指标。随着国民收入水平的快速提高，国内保险需求大幅增加。在国家政策的支持与激励下，保险业快速发展，但不论是在保险密度还是保险深度方面，同欧美发达国家相比仍处于较低的发展水平。为了进一步推动保险业的发展，2014年起，国务院关于保险服务业相关的各项政策在不断地发布和推进，使得保险业进入了一个崭新的发展阶段。2012年至2019年，我国已经成为世界上第二大保险国，年均保费收入增长率达18.67%。根据银保监会统信部数据显示，2019年1~10月，全国原保险保费收入累计37,058亿，比上年同期增长了12%。上述数据表明，目前中国保险市场发展迅速，未来的市场潜力庞大，保险行业是一个值得关注和投资的行业。伴随而来的是机构间的竞争愈加剧烈(见图1)和各地区的保险密度和保险深度有很大差别(见表1)。而从资本市场来看，我国保险行业的整体发展水平大致可以由国内的上市保险公司来判断，如果选择适当的企业价值评价方法，就能够正确地评价一个企业的价值，这有利于我们掌握保险行业的整体发展情况，准确了解保险公司的真实价值，同时对于投资者和管理人员来说，可以为股票投资和企业并购等提供决策依据。

当前，众多投资者通过分析市盈率来辨别投资风险和投资价值，以便于更好地选择合适的投资策略，同时政府金融监管部门也将市盈率作为重要参考资料来制定合理政策，引导资本市场健康和谐发展。市盈率是由股票价格比上年度每股收益所得出的(Benjamin Graham, 1934) [1]，在预测股票未来的获益能力中发挥一定的作用(Fairfield, 1994) [2]。国外学者通过对发达国家和发展中国家股票的市盈率影响因素进行分析，发现市盈率与每股收益、总投资回报率、GDP增长率、股利分配率等均存在显著的相关关系(White, 2000; Ramcharran H, 2002; Azam, 2010; Huang等, 2012) [3] [4] [5] [6]，同时预期成长性水平、股利支付水平、利率水平、市场风险程度和投资者信念等也都是影响市盈率的重要因素(Shamsuddin等, 2004; Zorn等, 2009; Jitmaneroj, 2017) [7] [8] [9]。国内学者以我国A股市场为研究对象，多角度分析市盈率，

发现流动比率、净资产收益率、流通股比率等均存在显著的相关关系(武一等, 2002; 徐筱凤等, 2005) [10] [11], 另外, 持股结构对上市公司市盈率也存在一定的影响, 即企业前几大股东持股比例、各董事持股比例的数值越大, 该企业的市盈率将越低(崔洁, 2013) [12]。同时根据上市公司市盈率影响因素, 通过建立多元回归模型可以发现, 不同因素对市盈率的影响大小不同, 其中净资产收益率和换手率对于市盈率的影响是最大的(宋光辉等, 2016; 陆雪丽, 2017; 韩荟雯, 2017; 林南玲等, 2018; 胡慧敏, 2019) [13] [14] [15] [16] [17]。

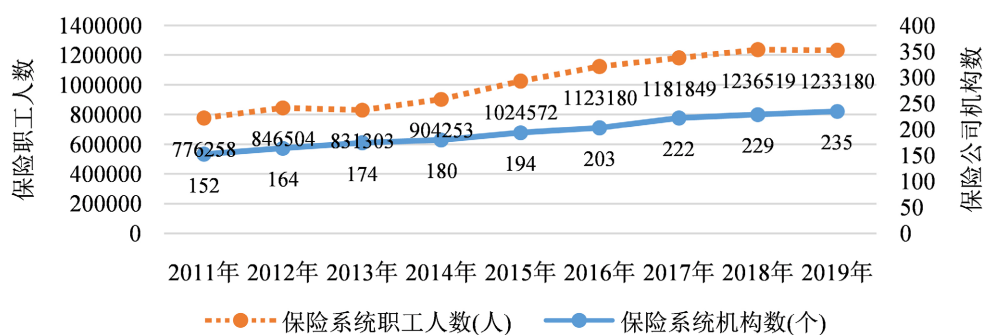


Figure 1. Trends in the number of insurance companies and the number of insurance employees in China
图 1. 我国保险公司机构数及保险职工人数变化趋势

Table 1. Insurance business situation of all over the country in 2019
表 1. 2019 年我国各地保险业务情况。

地区	保险密度(元/人)	保险深度(%)	地区	保险密度(元/人)	保险深度(%)
北京市	7472.24	4.59	广东省	2935.80	3.05
天津市	3160.19	3.49	广西壮族自治区	1093.14	2.51
河北省	2157.06	4.62	海南省	1760.72	3.07
山西省	1985.23	4.32	重庆市	2434.04	3.14
内蒙古自治区	2345.17	3.45	四川省	2109.13	3.76
辽宁省	1612.45	3.00	贵州省	1108.94	2.37
吉林省	2082.85	4.83	云南省	1235.29	2.55
黑龙江省	2098.35	5.84	西藏自治区	830.49	1.65
上海市	5615.45	3.56	陕西省	2200.53	3.27
江苏省	3844.67	3.10	甘肃省	1401.53	4.22
浙江省	3176.60	2.88	青海省	1336.94	2.70
安徽省	1785.83	3.01	宁夏回族自治区	2346.76	4.27
福建省	1950.91	1.80	新疆维吾尔自治区	2188.43	3.93
江西省	1466.87	2.74	大连市	4126.45	4.21
山东省	2258.68	3.18	宁波市	2882.46	2.47
河南省	2103.75	3.71	厦门市	4003.98	3.02
湖北省	2390.72	3.08	青岛市	5126.42	4.15
湖南省	1648.64	2.85	深圳市	8484.79	3.95

资料来源: 国家统计局官网。

总结目前的国内外研究文献可以发现,关于我国上市公司市盈率影响因素的研究,不管是在理论基础还是在实证分析方面都很少。在研究方法上,学者们从上市公司方面分别假设、检验并分析影响市盈率的因素,大部分以先验的财务理论知识为基础,选择少量的指标进行最小二乘回归或者其它具体金融模型的分析,来判断这些指标对市盈率是否有显著影响。由于不同学者的理解不尽相同,所以考虑的指标、使用的模型都很不一样,目前尚未得出一致性的结论,而且这些研究都只考虑了少量财务指标,可能会忽略对市盈率产生重要影响因素。同时,我们发现没有研究保险行业上市公司市盈率影响因素的文献。我国的上市保险公司日臻完善,吸引了越来越多的投资者,他们在选择企业进行投资时,首先考虑的一个指标也是市盈率,因此研究上市保险公司市盈率的影响因素就显得十分迫切和重要。

本文运用 Lasso+分位数回归模型分析我国上市保险公司市盈率的微观影响因素,既能筛除掉那些对市盈率影响甚微的指标,防止这些指标将重要因素湮灭,又能帮助我们考察市盈率影响因素处于不同分位数下的效应变化情况。

2. 理论模型

2.1. 面板数据的 Lasso 变量选择

Lasso 是 Tibshirani 在 1996 年提出的一种最常用于变量选择的方法,它是以压缩系数为基础来进行变量选择的[18]。此方法在常规的最小二乘法的基础上添加了限制条件,这个限制条件便是将系数的绝对值求和之后小于一个常数,通过这样一个限制条件的压缩,模型估计中一些回归系数的数值变为零,这方便我们进行变量的选择从而对模型进行更精准的描述。Lasso 不仅具有最佳子集选择层面方便阐述说明的优势,而且具有类似于岭回归的稳固性优势。

Lasso 等价的数学描述如下(Robert Tibshirani, 1996):

$$\arg \min \left\{ \sum_{i=1}^n \left(y_i - \alpha - \sum_{j=1}^p \beta_j x_{ij} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (1)$$

其中: n 为样本量, p 为指标个数; x_{ij} 为第 i 个样品的第 j 个指标变值, y_i 为第 i 个样品的因变量值; $\alpha, \beta_1, \dots, \beta_p$ 为待估参数; $\lambda \sum_{j=1}^p |\beta_j|$ 为惩罚项, λ 为非负惩罚参数,其值越大,删除的变量越多,反之删除的变量就越少。

易知,个体固定效应 Lasso 等价的数学描述为(李子强, 2014) [19]

$$\arg \min \left\{ \sum_{i=1}^N \sum_{t=1}^T \left(y_{it} - \alpha_i - \sum_{j=1}^p \beta_j x_{ijt} \right)^2 + \lambda \sum_{j=1}^p |\beta_j| \right\} \quad (2)$$

其中: N 为个体数, T 为时期数; x_{ijt} 为 t 时刻第 i 个个体第 j 个指标变值, y_{it} 为 t 时刻第 i 个个体的因变量值。

面板数据混合 Lasso 是其特例。

2.2. 面板数据的分位数回归

针对解释变量,分位数回归可以得到不同的分位数水平下的系数估计,显示出自变量随分位数变化其影响程度的变化,挖掘更丰富的信息,分析结果也更加全面深入。

假设随机变量 Y 的分布函数为 $F(y) = \text{prob}(Y \leq y)$, 则 Y 的 τ 分位数为使 $F(y) = \tau$ 的最小的 y 值,即

$$Q(\tau) = \inf \{ y : F(y) \geq \tau \}, \quad 0 < \tau < 1 \quad (3)$$

假定回归方程的形式为

$$y_i = \beta x_i + \varepsilon_i + \alpha, \quad i = 1, 2, \dots, n \quad (4)$$

其中: y_i 为被解释变量; $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})'$ 为解释变量; $\beta = (\beta_1, \beta_2, \dots, \beta_p)$ 为各解释变量的回归系数; ε_i 为随机扰动项; α 为截距项。分位数 τ 下的截距项和回归系数记为 α_τ 和 β_τ , 其估计通过最小化下式得出(Koenker R, 1978) [20]:

$$\arg \min_{\alpha_\tau, \beta_\tau} \left\{ \sum_{y_i \geq \beta_\tau x_i + \alpha_\tau} \tau |y_i - \beta_\tau x_i - \alpha_\tau| + \sum_{y_i < \beta_\tau x_i + \alpha_\tau} (1 - \tau) |y_i - \beta_\tau x_i - \alpha_\tau| \right\} \quad (5)$$

易知, 固定效应分位数回归估计为(罗幼喜, 2010) [21]

$$\arg \min_{\alpha_{it}, \beta_{it}} \left\{ \sum_{y_{it} \geq \beta_{it} x_{it} + \alpha_{it}} \tau |y_{it} - \beta_{it} x_{it} - \alpha_{it}| + \sum_{y_{it} < \beta_{it} x_{it} + \alpha_{it}} (1 - \tau) |y_{it} - \beta_{it} x_{it} - \alpha_{it}| \right\} \quad (6)$$

其中, $x_{it} = (x_{it1}, x_{it2}, \dots, x_{itp})'$ 。若面板数据模型为混合估计模型, 则 $\alpha_{it} = \alpha_\tau$ 。

3. 数据、结果与分析

3.1. 数据来源

截止到 2020 年 1 月, 我国保险行业上市公司仅有中国人寿、中国太保、新华保险、中国人保、中国平安、西水股份、天茂集团和中国太平共 8 家上市公司。为获得较多的样本量, 同时考虑数据的完整性和连续性, 本文选用 2011~2019 年我国保险业上市公司的时间序列和截面两个维度上的季度 A 股数据来构建模型, 避免了市场交易差异的影响, 根据市盈率的定义, 剔除了市盈率是负数以及有其它数据缺失的样本。本文所使用的数据来自于锐思数据库以及各公司披露的财务报表。

3.2. 指标选取

在微观层面, 影响我国上市保险公司市盈率的因素繁多, 为更加系统地进行分析, 本文根据指标选取的科学性、全面性、数据可获取等原则共选取了 30 个财务指标, 涵盖了盈利能力、偿债能力、成长发展能力和营运能力四个方面。本文拟定的各个指标及其含义(见表 2)。

Table 2. Influencing factors index of the P/E ratio of listed insurance companies and the descriptive statistics analysis of the index

表 2. 上市保险公司市盈率影响因素指标及其描述性统计分析

指标类型	符号	指标名称	观测值	均值	标准差	最小值	最大值
	PE	市盈率(%)	128	21.59	10.52	9.865	70.24
	AvgROE	净资产收益率(平均)(%)	128	8.973	4.821	0.828	21.64
	ROA	资产净利率(%)	128	0.876	0.447	0.110	2.063
	Netprfrit	销售净利率(%)	128	6.532	3.330	1.321	16.54
盈利能力	Opeprfrit	营业利润率(%)	128	8.283	4.294	1.802	19.07
	NAPS	每股净资产(元/股)	128	9.566	13.62	-8.470	111.5
	EPS	每股收益(元/股)	128	1.532	1.301	0.080	7.290
	OpeCPSgrrt	每股经营活动现金流量增长率(%)	128	374.0	3316	-201.3	37400

Continued

	Dbequrt	产权比率(%)	128	994.6	319.7	591.1	1761
	Dbtanequrt	有形净值债务率(%)	128	1134	522.2	600.8	2628
	Casrtsale	销售现金比率(%)	128	27.22	22.94	-15.69	128.2
偿债能力	Equmul	权益乘数(%)	128	10.14	2.299	6.809	14.59
	Dbastrt	资产负债率(%)	128	89.63	2.297	85.31	93.15
	OpeCFPS	每股经营活动现金流量(元/股)	128	6.372	6.505	-2.340	35.48
	CFPSgrrt	每股净现金流量(元/股)	128	0.907	4.019	-7.915	23.86
	EPSgrrt	每股收益增长率(%)	128	26.93	56.26	-80	223.1
	Opeincmgrt	营业收入增长率(%)	128	12.67	11.65	-19.31	48.25
	Totprfgrt	利润总额增长率(%)	128	30.89	61.23	-79.18	296.3
	Netprfgrt	净利润增长率(%)	128	29.57	55.33	-81.30	245.3
	Netassgrt	净资产增长率(%)	126	16.58	34.82	-8.232	376.7
成长发展能力	Totassgrt	总资产增长率(%)	128	0.148	0.069	0.035	0.316
	NAPSgrt	每股净资产相对年初增长率(%)	128	16.18	6.526	6.780	35.08
	EquYTDgrt	股东权益相对年初增长率(%)	128	11.51	33.58	-8.232	376.7
	TotastYTDgrt	资产总计相对年初增长率(%)	128	9.851	9.537	-1.989	95.06
	Susgrt	可持续增长率(%)	128	9.145	5.842	0.852	38.36
	PB	市净率(%)	128	2.132	0.441	1.387	3.467
	Equrat	股东权益周转率(次)	128	1.541	0.736	0.543	5.767
	Totassrat	总资产周转率(次)	126	14.95	14.36	1.587	101.9
营运能力	ToloperevPS	每股营业总收入(元/股)	128	23.39	12.36	4.668	53.44
	OpeprfPS	每股营业利润(元/股)	128	2.107	1.876	0.104	8.935
	Fixassrt	固定资产比率(%)	128	0.802	0.190	0.443	1.285

由表 2 可以看出, 2011~2019 年, Dbtanequrt 的波动比较大, 意味着企业对于负债的偿还能力在不同年份和不同省份之间存在明显差异; ROA、Totassgrt、PB、Equrat、Fixassrt 的波动很小, 体现了保险业上市公司的投资价值比较高, 能吸引日益增多的投资者进行投资。

3.3. 基于混合估计模型的分析

我们以 30 个影响因素指标为解释变量, 上市保险公司市盈率为被解释变量, 在数据经过标准化处理后, 建立面板数据的混合估计模型如下:

$$PE_{it} = \alpha + \beta_1 AvgROE_{it} + \beta_2 ROA_{it} + \dots + \beta_{30} Fixassrt_{it} + \mu_{it} \quad (i = 1, 2, \dots, 8; t = 1, 2, \dots, 19) \quad (7)$$

其中: i 表示个体, t 表示时间; μ_{it} 表示随机扰动项。用 Stata 15 软件进行参数估计, 结果(见表 3)。

Table 3. Results of mixed model estimates for the analysis of factors affecting the P/E ratio of listed insurance companies
表 3. 上市保险公司市盈率影响因素分析的混合模型估计结果

指标	估计系数	指标	估计系数
AvgROE	-2.709 (-1.26)	Opeincmgrrt	-0.158** (-2.52)
ROA	15.19 (1.30)	Totprfgrrt	0.0646* (1.87)
Netprfrit	1.084 (0.82)	Netprfgrrt	-0.0240 (-0.15)
Opeprfrit	-0.689 (-0.49)	Netassgrrt	-0.0983 (-0.89)
NAPS	1.162*** (2.68)	Totassgrrt	0.0717 (1.02)
EPS	3.526 (0.86)	NAPSGrrt	0.197*** (3.29)
OpeCPSgrrt	0.000256*** (3.17)	EquYTDgrrt	0.0389 (0.27)
Dbequrt	0.00390 (0.13)	TotastYTDgrrt	-0.0360 (-0.48)
Dbtanequrt	0.00450 (0.48)	Susgrrt	0.842** (2.18)
Casrtsale	-0.155** (-2.24)	PB	7.703*** (4.95)
Equmul	-1.642 (-0.58)	Equrat	-3.492 (-0.26)
Dbastrt	1.947 (0.92)	Totassrat	179.5** (2.57)
OpeCFPS	0.0434 (0.17)	ToloperevPS	-0.347 (-0.84)
CFPSgrrt	0.264 (1.05)	OpeprfPS	5.302* (1.75)
EPSgrrt	0.0852 (0.49)	Fixassrt	-1.540 (-0.19)
	_cons		-138.2 (-0.81)
	<i>N</i>		126
	r^2		0.719
	F		91.14

t statistics in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

由表 3 可以看出, r^2 为 0.719 较高, 拟合效果相对较好。混合估计模型中共有 9 个影响因素指标至少在 10% 的水平上显著。其中 NAPS 的系数为 1.162, 且在 1% 的水平上是显著为正的, 说明每股净资产对于市盈率的影响作用是正向的, 即每股净资产越高, 投资者越喜欢, 未来的前景越高, 相应的股票就越有可能被追捧, 从而推高股价, 使得市盈率也会相应的提高; 同理, OpeCPSgrrt、Totprfgrrt、NAPSGrrt、Susgrrt、PB、Totassrat、OpeprfPS 的系数分别为 0.000256、0.0646、0.197、0.842、7.703、179.5、5.302, 且至少在 10% 的水平上显著为正, 说明其对市盈率均存在显著的正向促进作用; Casrtsale 和 Opeincmgrrt 的回归系数分别为 -0.155 和 -0.158, 且都是在 5% 的水平上显著为负, 说明其对市盈率起到显著的负向促进作用。

传统的最小二乘回归, 不仅忽略掉了各解释变量之间的相互作用, 而且特别容易受到一些异常值的影响, 从而出现估计系数变得特别大或特别小这样不准确的情况, 因此仅仅通过上述分析结果, 不能够对市盈率与影响因素指标的相关性给予十足的把握。所以, 考虑到财务指标之间存在的共线性, 本文首先利用 Lasso 对影响因素指标进行筛选, 仅留下对市盈率影响比较显著的指标, 筛选掉那些影响甚微的指标, 防止这些指标将重要因素湮灭, 然后再进一步分析保留财务指标同市盈率之间的关系。

3.4. 基于 Lasso 的影响因素指标筛选

本文利用 Lasso 对上市保险公司市盈率的影响因素指标进行筛选(见表 4)。该表显示, 随着 λ 由大到小, Lasso 筛选出越来越多的指标进入模型: 当 $\lambda > 1602.10303$ 时, 由于惩罚力度过大, 所有指标的系数均为 0; 当 $\lambda = 1602.10303$ 时, 常数项首先进入模型; 当 $\lambda = 1459.77661$ 时, Netprfirt 进入模型; Dbequrt 进入时的 λ 最小, 为 0.482926。

Table 4. Variables entering with the change of λ in Lasso solution (part)

表 4. Lasso 求解中变量随 λ 变化进入情况(部分)

Knot	ID	Lambda	s	L1-Norm	EBIC	R-sq	Action
1	1	1602.10303	1	0.00000	598.07719	0.0000	Added_cons.
2	2	1459.77661	2	0.17120	596.69733	0.0629	Added Netprfirt.
3	4	1211.93219	3	0.46387	589.91991	0.1586	Added Opeprfirt.
4	7	916.78226	4	0.77409	581.50882	0.2543	Added EPSgrtt.
5	8	835.33785	5	1.25171	581.08720	0.2959	Added PB.
6	10	693.51216	6	2.50815	574.17705	0.3685	Added Netprfgrtt.
7	12	575.76598	7	3.56800	570.06558	0.4209	Added NAPS.
8	13	524.61651	9	4.09205	577.78987	0.4473	Added AvgROEOpeincmgrtt.
9	20	273.53549	10	6.66487	560.31867	0.5441	Added Equumul.
10	25	171.78826	12	8.04409	567.13254	0.5681	Added ROA Totprfgrtt.
11	26	156.52706	13	8.85142	572.16823	0.5741	Added NAPSgrtt.
12	28	129.95152	14	13.09804	575.59543	0.5853	Added Totassrat
13	29	118.40698	13	20.10604	565.39135	0.5964	Added OpeCPSgrtt, Removed AvgROEEPSgrtt.
14	30	107.88803	14	25.54462	570.35820	0.6022	Added TotastYTDgrtt.
15	33	81.61334	13	39.05414	559.62804	0.6144	Removed Netprfirt
16	34	74.36304	15	42.87931	572.21923	0.6175	Added CasrtsaleFixassrt.
17	35	67.75683	16	47.72412	577.66259	0.6216	Added Netassgrtt.
18	36	61.73750	15	52.94314	569.25364	0.6264	Removed Opeprfirt.
19	37	56.25291	16	56.84856	574.44479	0.6312	Added OpeprfPS.
20	39	46.70215	15	65.27153	563.51412	0.6430	Removed Netprfgrtt.
21	44	29.33031	16	77.81016	565.40300	0.6567	Added ToloperevPS.
22	46	24.35054	18	88.02066	576.87170	0.6625	Added EPS CFPSgrtt.
23	47	22.18731	21	93.55523	596.03731	0.6658	Added AvgROEEPSgrttTotassgrtt.
24	48	20.21625	22	101.43152	600.78738	0.6712	Added Susgrtt.
25	51	15.29285	23	125.85317	601.17801	0.6875	Added Dbtanequrt.

Continued

26	53	12.69639	24	138.77238	604.50590	0.6960	Added Dbastrt.
27	54	11.56848	25	144.81316	609.84432	0.6995	Added Equrat.
28	56	9.60436	26	157.79019	614.35039	0.7049	Added Netprfrt.
29	59	7.26534	25	173.83514	605.31051	0.7101	Removed TotastYTDgrtt.
30	60	6.61991	26	179.45200	611.55826	0.7114	Added EquYTDgrtt.
31	64	4.56285	27	197.71461	616.87176	0.7148	Added Opeprfrt.
32	66	3.78816	28	203.47225	623.18613	0.7159	Added TotsatYTDgrtt.
33	67	3.45163	29	205.91728	629.79284	0.7163	Added OpeCFPS.
34	68	0.48293	30	208.39511	636.59183	0.7172	Added Dbequrt

整个求解过程的路径(见图 2), 显示了对于不同影响因素指标回归系数的变化过程[22]。其中, 当 $\lambda = 0$ 时(图 2 中最左边), 不存在惩罚项, 故此时 Lasso 等价于最小二乘估计; 而当 λ 很大时(图 2 中最右边), 由于惩罚力度过大, 所有指标的系数均为 0。

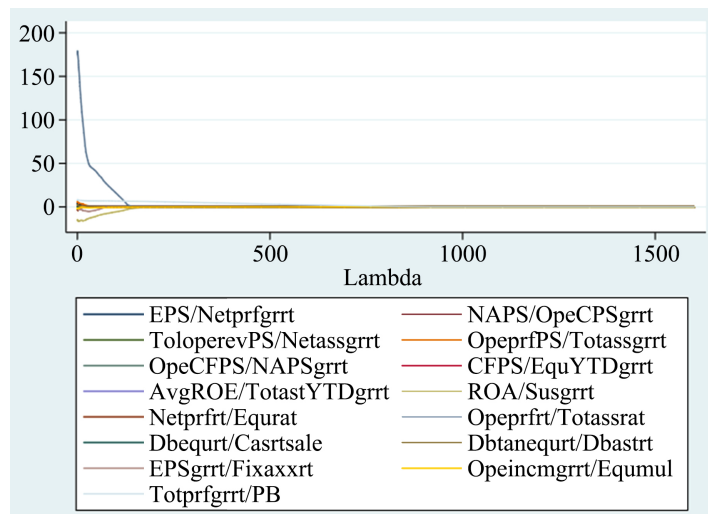


Figure 2. Lasso model solution path
图 2. Lasso 模型求解路径

Lasso 中的 λ , 如果设置的太大, 那么最后的模型参数将均趋于 0, 造成拟合不足; 如果设置的太小, 又会造成拟合过度[23] [24]。所以, λ 的取值一般需要通过交叉验证(梁子超, 2020) [25]来判断。本文采用 10 折交叉验证选择的最优 $\lambda = 98.304$, 对应的 Lasso 估计结果如表 5 所示, 第 2 列即为 Lasso 所估计的影响因素指标系数。其中, 除常数项外, ROA、Netprfrt、Opeprfrt、NAPS 等 13 个财务指标的系数为非零, 而剩下未出现的 17 个财务指标的系数是等于 0 的, 通过 Lasso 被筛选出去了。通过对比 Lasso 和混合估计的回归结果可以发现, 一些指标的回归系数变化很大, 比如 Totassrat 的回归系数由混合估计的 59.0737216 变成了 Lasso 估计的 16.7222840, 由此说明混合估计回归受各指标之间多重共线性的影响, 出现估计系数变得特别大或特别小的状况, 结果变的不准确。而 Lasso 通过财务指标的筛选在很大程度上克服了多重共线性的影响, 得出的结果更加准确(见表 5)。

Table 5. Lasso regression estimation results ($\lambda = 98.304$)
表 5. Lasso 回归估计结果($\lambda = 98.304$)

Selected	Lasso	Post-est OLS
ROA	5.7721727	13.7798443
Netprfirt	-0.0977533	0.5432623
Opeprfirt	-0.1694693	0.0731514
NAPS	0.3397298	0.4028520
OpeCPSgrrt	0.0001426	0.0002984
Equmul	-0.4167192	-0.7047170
Opeincmgrrt	-0.1056529	-0.1589642
Totprfgrrt	0.0157520	0.0294309
Netprfgrrt	-0.0255769	-0.0081671
NAPSGrrt	0.0297548	0.0513699
TotsatYTDgrrt	0.0126671	0.0752998
PB	6.7905728	7.6370377
Totassrat	16.7222840	59.0737216
_cons	23.5678945	19.9539254

3.5. 基于 Lasso 的分位数回归

在利用 Lasso 进行影响因素指标选择的基础上, 我们采用面板数据的分位数回归分析我国上市保险公司市盈率的微观影响因素, 并通过回归结果的显著性进一步找出具有重要影响的财务指标。

本文把经 Lasso 筛选出的指标分别代入 0.25、0.5 和 0.75 分位数回归方程(见表 6)。

0.25 分位点处对应的是市盈率相对较低的水平。从分位数回归结果可以看出, ROA、NAPS、OpeCPSgrrt、PB 的系数均在 0.01 水平下显著, Totassrat 的系数在 0.05 水平下显著, NAPSGrrt 的系数在 0.1 水平下显著。同时我们可以观察到, Totassrat、PB 和 ROA 这三个指标的回归系数较大, 这是因为这三个指标的数据在取值范围上相对较小。在 0.25 分位点处, NAPS 在 0.01 水平下显著, 其系数为 0.359, 说明其对市盈率有显著的正向影响, 而且在当其他条件不变的情况下, 该指标增加一个单位, 市盈率将平均提高 0.359 个单位。而 OpeCPSgrrt 这个指标的回归系数的绝对数值较小, 虽对市盈率有显著正向影响, 但系数相对较小, 表明这个指标对市盈率的影响不是很大。

处于 0.5 分位点处的上市保险公司, 代表了所有公司市盈率的平均水平。从分位数回归结果看, ROA、NAPS、PB 这三个指标的系数均在 0.01 水平下显著, 并且对上市保险公司市盈率有正向影响, 其回归系数分别为 6.556、0.366 和 6.363; Opeincmgrrt 和 Equmul 分别在 0.01 和 0.05 显著水平下对上市保险公司市盈率有负向影响, 其回归系数分别为-0.123 和-0.602; OpeCPSgrrt 和 NAPSGrrt 对市盈率有正向影响, 且在 0.05 的显著水平下显著。在 0.25 分位点处, Totassrat 这一指标的回归系数为 23.54, 但在该分位点处该指标的影响不显著。同时, 与较低分位点处相比, 0.5 分位点处 Opeincmgrrt 和 Equmul 的回归系数绝对值明显变大。

0.75 分位点处对应的是市盈率相对较高的水平。ROA、NAPS、PB 这三个指标依旧对上市保险公司的市盈率有正向影响; Equmul 和 Opeincmgrrt 这两个指标对上市保险公司的市盈率有负向影响。不过,

随着分位点的增高, OpeCPSgrtt 和 NAPSgrtt 这两个指标对保险公司市盈率的影响不再显著。

表 6 显示, 随着分位数的增加, ROA 的回归系数不断增大, 且在至少 0.05 的显著性水平下显著, 说明随着 PE 的增大, ROA 对 PE 的影响越来越大; NAPS 是反映上市保险公司盈利能力的指标, NAPS 的分位数回归系数呈现先增加后减少的趋势, 由 0.359 到 0.366, 再到 0.357, 对上市保险公司市盈率的影响呈倒“U”型, 由此说明, NAPS 对 PE 的条件分布两端的影响小于其中间部分的影响, 也就是说增加 NAPS 对低 PE 和高 PE 的影响都比较大, 但影响最大的是中间部分; 随着分位数的增加, OpeCPSgrtt 的分位数回归系数呈现递减趋势, 由 0.000431 到 0.000344 再到不显著, 由此说明随着 PE 的增大, OpeCPSgrtt 对 PE 的影响是越来越弱的; Equmul 是反映公司偿债能力的指标, 当 PE 值处于中高水平时, Equmul 对市盈率有显著的负向促进作用; 当 PE 值较小时, Opeincmgrtt 对 PE 的影响结果是不显著的, 但随着 PE 值的不断增大, Opeincmgrtt 对 PE 有着负向的促进作用, 该作用至少在 0.05 的水平上显著。Opeincmgrtt 能够体现一个企业在销售所得方面的收益增长状况, 当该企业的营业收入增长率为正值时, 说明这个企业的收益在不断提高, 相应地股票收益率就会越大, 市盈率就会变小, 所以营业收入增长率同市盈率存在负相关关系; 当 PE 值处于中低水平时, NAPSgrtt 对 PE 有着正向的促进作用, 该作用至少在 0.05 的水平上显著; PB 是反映公司成长发展能力的指标, PB 的分位数回归系数呈现先增后减的趋势, 而且在 0.75 分位点处的回归系数是要小于在 0.25 分位点处的回归系数, 说明 PB 在中间时, 对 PE 的影响是最大的, 在尾段时, 对 PE 的影响是最小的; Totassrat 这一指标主要反映公司的营运能力, 当 PE 值处于较低分位数水平时, Totassrat 对 PE 有显著的正向促进作用。

Table 6. Quantitative regression results of influencing factors after Lasso variable selection
表 6. Lasso 变量选择后的影响因素分位数回归结果

	$\tau = 0.25$ PE	$\tau = 0.5$ PE	$\tau = 0.75$ PE
ROA	5.224*** (3.37)	6.556*** (3.69)	7.932** (2.49)
Netprfrit	-0.0427 (-0.26)	-0.1562 (-0.84)	-0.1557 (-0.47)
Opeprfrit	-0.103 (-0.41)	-0.196 (-0.19)	-0.251 (-0.34)
NAPS	0.359*** (3.66)	0.366*** (3.25)	0.357* (1.77)
OpeCPSgrtt	0.000431*** (3.52)	0.000344** (2.45)	0.000267 (1.06)
Equmul	-0.309 (-1.36)	-0.602** (-2.30)	-0.876* (-1.86)
Opeincmgrtt	-0.0565 (-1.41)	-0.123*** (-2.68)	-0.171** (-2.07)
Totprfgrtt	0.0133 (0.68)	0.00519 (0.23)	0.0575 (1.43)
Netprfgrt	-0.0154 (-0.69)	-0.0193 (-0.76)	0.0360 (0.79)
NAPSgrtt	0.0802* (1.82)	0.110** (2.17)	0.0638 (0.70)
TotastYTDgrtt	0.0244 (0.48)	0.0804 (1.38)	0.109 (1.04)
PB	6.161*** (6.29)	6.363*** (5.66)	5.312*** (2.63)
Totassrat	23.54** (2.41)	12.91 (1.15)	23.47 (1.17)
_cons	15.05*** (4.61)	23.08*** (6.17)	31.34*** (4.66)
<i>N</i>	126	126	126
<i>r</i> ²	0.8626	0.8664	0.8752

t statistics in parentheses; * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4. 研究结论

本文针对保险业上市公司市盈率微观影响因素,从财务状况的四个方面建立指标体系后,先是实证分析了应用面板数据混合模型的不足,接着进一步利用 Lasso 方法对财务指标进行筛选,在指标筛选的基础上将分位数回归用到了保险业上市公司市盈率微观影响因素分析,结论是:每股净资产、资产净利率、市净率、每股经营活动现金流量增长率、总资产周转率、每股净资产相对年初增长率对上市保险公司市盈率有着显著正向影响,而营业收入增长率和权益乘数对上市保险公司市盈率有着显著负向影响;当市盈率在不同的分位数水平下,各影响因素对其作用大小也有所不同。本文的研究结论有利于我们对上市保险公司市盈率的更深刻理解,并有利于我们更准确把握其走势,这对上市保险公司经营决策和投资者的投资决策都具有重要指导意义。

本文将 Lasso 和分位数回归有机结合起来分析上市保险公司市盈率影响因素是一有益尝试,该方法可用于其它类似研究。

基金项目

国家自然科学基金青年基金项目(61502280); 山东科技大学研究生导师指导能力提升计划项目(KDYC17018); 2019 年山东省高等学校青创人才引育计划建设团队项目。

参考文献

- [1] (美)本杰明·格雷厄姆, (美)戴维·多德. 证券分析[M]. 巴曙松, 陈剑, 等, 译. 成都: 四川人民出版社, 2019.
- [2] Fairfield, P.M. (1994) P/E, P/B and the Present Value of Future Dividends. *Financial Analysts Journal*, **50**, 23-31. <https://doi.org/10.2469/faj.v50.n4.23>
- [3] White, C.B. (2000) What P/E Will the U.S. Stock Market Support? *Financial Analysts Journal*, **56**, 30-38. <https://doi.org/10.2469/faj.v56.n6.2401>
- [4] Ramcharran, H. (2002) An Empirical Analysis of the Determinants of the P/E Ratio in Emerging Markets. *Emerging Markets Review*, **3**, 165-178. [https://doi.org/10.1016/S1566-0141\(02\)00004-3](https://doi.org/10.1016/S1566-0141(02)00004-3)
- [5] Azam, M. (2010) Factors Influencing the Price-Earnings Multiples and Stock Values in the Karachi Stock Exchange. *Interdisciplinary Journal of Contemporary Research in Business*, **9**, 1280-1284.
- [6] Huang, A.G. and Wirjanto, T.S. (2012) Is China's P/E Ratio Too Low? Examining the Role of Earnings Volatility. *Pacific-Basin Finance Journal*, **20**, 41-61. <https://doi.org/10.1016/j.pacfin.2011.07.003>
- [7] Shamsuddin, A.F.M. and Hillier, J.R. (2004) Fundamental Determinants of the Australian Price-Earnings Multiple. *Pacific-Basin Finance Journal*, **12**, 565-576. <https://doi.org/10.1016/j.pacfin.2004.02.001>
- [8] Zorn, T.S., Dudney, D.M. and Jirasakuldech, B. (2009) P/E Changes: Some New Results. *Journal of Forecasting*, **28**, 358-370. <https://doi.org/10.1002/for.1097>
- [9] Jitmaneroj, B. (2017) Does Investor Sentiment Affect Price-Earnings Ratios? *Studies in Economics and Finance*, **34**, 183-193. <https://doi.org/10.1108/SEF-09-2015-0229>
- [10] 武一, 陈述云, 袁媛. 总体市盈率与公司业绩流通股本间的实证分析[J]. 西南民族大学学报(人文社科版), 2002, 23(6): 148-151.
- [11] 徐筱凤, 李寿喜. 中国企业市盈率: 理论分析与经验证据[J]. 世界经济文汇, 2005(4): 172-178.
- [12] 崔洁, 罗鄂湘. 上市公司股权结构、激励与市盈率关系的研究[J]. 西安财经学院学报, 2013, 26(2): 16-21.
- [13] 宋光辉, 孙影, 董艳. A 股上市公司市盈率水平的驱动因素研究[J]. 财会月刊, 2016(5): 102-105.
- [14] 路雪丽, 叶陈刚. 创业板上市公司市盈率影响因素分析[J]. 国际商务财会, 2017(8): 61-65+74.
- [15] 韩荟雯. 沪深 A 股市盈率及其影响因子研究[D]: [硕士学位论文]. 湘潭: 湘潭大学, 2017.
- [16] 林南玲, 巫彦衡, 林巧玲, 曾凡. 上市公司市盈率影响因素的实证分析[J]. 开发性金融研究, 2018(3): 37-49.
- [17] 胡慧敏. 创业板上市公司市盈率影响因素的实证研究[J]. 价值工程, 2019, 38(31): 190-194.
- [18] Tibshirani, R. (1996) Regression Shrinkage and Selection via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, **58**, 267-288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>

- [19] 李子强, 田茂再, 罗幼喜. 面板数据的自适应 Lasso 分位回归方法研究[J]. 统计与信息论坛, 2014, 29(7): 3-10.
- [20] Koenker, R. and Bassett, G. (1978) Regression Quantiles. *Econometrica*, **46**, 33-50. <https://doi.org/10.2307/1913643>
- [21] 罗幼喜, 田茂再. 面板数据的分位回归方法及其模拟研究[J]. 统计研究, 2010, 27(10): 81-87.
- [22] 李扬, 曾宪斌. 面板数据模型的惩罚似然变量选择方法研究[J]. 统计研究, 2014, 31(3): 83-89.
- [23] 白仲林. 面板数据的计量经济分析[M]. 天津: 南开大学出版社, 2008.
- [24] 范永东. 模型选择中的交叉验证方法综述[D]: [硕士学位论文]. 太原: 山西大学, 2013.
- [25] 梁子超, 李智炜, 赖铿, 林卓琛, 李铁钢, 张晋昕. 10 折交叉验证用于预测模型泛化能力评价及其 R 软件实现[J]. 中国医院统计, 2020, 27(4): 289-292.