

doi:10.19920/j.cnki.jmsc.2025.02.011

优胜劣汰还是逆向选择^①

——基于上市公司质量与股价表现关联的研究

靳馥境¹, 姜富伟², 唐国豪^{3, 4*}

(1. 北京交通大学经济管理学院, 北京 100044; 2. 厦门大学宏观经济研究中心/经济学院/
王亚南经济研究院, 厦门 361000; 3. 湖南大学金融与统计学院, 长沙 410006;
4. 湖南大学高性能分布式账本与数字金融教育部重点实验室, 长沙 410082)

摘要: 构建优胜劣汰的高质量资本市场是金融服务实体经济的关键。本文使用 90 多个财务特征指标等高维大数据和多种机器学习与计量方法提取基本面信息, 对上市公司质量进行评价并构建基本面质量综合指数, 研究基本面质量与股价表现的关联关系。研究发现, 上市公司基本面质量对股价表现具有显著为正的预测能力。其中, 由偏最小二乘法构建的基本面质量指数对股票横截面收益的预测能力最强, 年化收益近 38%, 且 CAPM、三因子和五因子模型对此不能解释。本文还从行为金融和经济周期等视角探索公司基本面质量对股票价格的影响机制, 发现市场情绪、有限套利、公司投资等决策和经济周期都有助于理解上市公司质量溢价现象。本文研究表明, 我国股票市场定价效率已稳步提高, 进入了“优胜劣汰”和价值投资阶段。

关键词: 基本面质量; 股票收益; 文本情绪; 有限套利; 大数据与机器学习

中图分类号: F832.5 **文献标识码:** A **文章编号:** 1007-9807(2025)02-0154-17

0 引 言

作为实体经济的晴雨表, 资本市场在我国金融运行中具有牵一发而动全身的作用, 是支持创新创业、推动供给侧改革和实现经济高质量发展的有效工具。相较于英美发达国家百余年证券发展历史, 我国股市以 30 余年的时间, 总市值居于世界第二, 取得喜人发展势头。但我国股票市场仍存在散户多于机构投资者、信息操纵多于信息披露、暴涨暴跌多于平稳慢牛等特征, 被质疑缺乏长期投资价值和支撑实体经济的能力。因此, 如何促进我国股市高质量发展, 避免劣股驱逐良股, 是目前我国资本市场改革最亟待解决的问题。习近平总书记也强调, 要深化资本市场供给侧改革, 提高

上市公司质量。证监会从强化监管方面, 聚焦提高上市公司质量, 营造良好的投资环境, 贯彻落实国务院《关于进一步提高上市公司质量的意见》。总之, 我国股票市场上公司是优胜劣汰还是逆向选择的结果, 不仅是投资者的关注点, 也是监管层关心的焦点问题。

本文使用 90 多个财务指标等高维数据和多种计量方法提取基本面信息, 对上市公司进行质量评价并构建基本面质量综合指数, 研究上市公司质量指数与股票表现的相关性, 以此分析我国股票市场对面质量的定价效率和发展阶段, 检验“优胜劣汰”还是“逆向选择”两大备择假设。本文还从行为金融和经济周期等视角, 探索公司质量对股票价格的影响机制, 进一步深入理解质

① 收稿日期: 2020-12-15; 修订日期: 2022-07-07.

基金项目: 国家自然科学基金资助项目(72072193; 71872195; 72003062; 72302013); 教育部人文社科基金资助项目(24YJC790159); 国家社科基金资助重大项目(22&ZD063); 中国博士后科学基金资助项目(2022M710344).

通讯作者: 唐国豪(1989—), 男, 湖南长沙人, 博士, 副教授. Email: ghtang@hnu.edu.cn

量溢价现象. 具体研究内容如下.

第一, 本文使用高维大数据和偏最小二乘法等现代机器学习与金融科技手段, 从大量公司特征数据中高效提取公司基本面信息, 评价并测度上市公司质量, 从而构建上市公司基本面质量综合指数. Asness 等^[1]发现, 价值投资者愿意为具有某些“高质量”特征的股票支付更高的价格, 这些公司特征即为公司质量的体现. 自上世纪七八十年代起, 公司规模、资产收益率、投资、研发等反应上市公司质量的特征指标层出不穷^[2, 3]. 面临基本面信息“爆炸式”增长, 投资者陷入高维陷阱, 导致投资决策困难, 因此如何降维构建公司基本面综合质量测度迫在眉睫. 本文首先在整合现有文献的基础上, 从高登 (Gordon) 股票估值理论模型出发, 联系价值投资实践, 从盈利能力、成长能力、营运策略、潜在价值和安全性五个方面收集 94 个特征指标; 然后参考 Jiang 等^[4], 采用主成分分析法 (principal component analysis, PCA)、珉玛麦克白回归法 (Fama-MacBeth regression, FM)、预测组合法 (forecast combination, FC) 和偏最小二乘法 (partial least squares, PLS), 以及缩放主成分分析法 (scaled principal component analysis, SP-CA)^[5] 和标准化法 (normalization approach, NA)^[1] 等多种机器学习与计量方法, 对高维基本面指标集进行信息提取和数据降维, 构建上市公司基本面综合质量指数.

第二, 分析研究了上市公司基本面综合质量指数与股票价格表现的关系, 以此判断我国资本市场供给侧改革是否已经起效, 我国股市是否已经进入注重公司质量和价值投资的时代, 表现为“优胜劣汰”; 还是我国股市仍处于低质量粗放发展阶段, 散户的投机行为扭曲资产价格并导致“逆向选择”问题. 应用单变量投资组合分析法, 本文发现基于偏最小二乘法 (PLS) 构建的质量指数 Q^{PLS} 具有最强且相对稳定的预测能力, 上市公司综合质量指数和股价表现呈现显著为正的相关性, 因此我国股市已进入高质量发展“优胜劣汰”阶段.

第三, 从行为金融和经济周期两个视角探索公司质量对股票价格的影响机制, 进一步深入理

解上市公司基本面质量溢价现象. 具体说, 本文分别从市场情绪、有限套利、公司决策和宏观金融、经济周期四个方面, 检验驱动机制. 结果表明: 市场情绪高涨、套利摩擦大、经济下行时期, 公司综合质量的预测能力更强, 但质量溢价不受证券、利率和汇率市场周期的影响; 而且优质公司更倾向于披露可读性强的年报、不盲目扩张资产规模, 降低信息不对称程度, 形成利好消息从而推高股价. 综上, 市场情绪、有限套利、公司决策和宏观经济周期, 都有利于解释公司基本面综合质量与股价表现的正相关关系.

第四, 通过对比 Q^{PLS} 与胡熠等^[6]、尹力博等^[7]、Asness 等^[1] 构建的综合质量测度, 及扩大基本面指标集后分别应用 PLS 和 NA 构建的综合质量测度, 均证明基于偏最小二乘法构建的公司综合质量指标预测能力最优. 同时, 基于 Q^{PLS} 的质量投资策略显著优于市场指数投资策略, 且该策略在极端市场条件下均可获得显著的正收益. 因此, 我国股票市场中存在显著且稳健的质量溢价.

本文的创新点如下. 第一, 结合高维大数据和偏最小二乘法等现代机器学习与金融科技手段, 为上市公司质量评价和测度以及基本面驱动的价值投资研究提供了新思路. 以往公司质量研究多集中于某一个或某一类指标, 例如盈利^[8]、分红^[9]、特质风险^[10]、市场风险^[11]等. 胡熠等^[6]、尹力博等^[7]、Asness 等^[1]、Jiang 等^[4] 分别研究了 8 个、14 个、16 个和 75 个上市公司特征. 而本文从高维数据出发, 在 90 多个上市公司特征基础上, 应用 PLS 等计量方法进行信息提取, 为股票收益预测提供了一个全新的、可以综合测度公司基本面质量的指标.

与本文构建质量指数思想最相近的研究为 Asness 等^[1]、胡熠等^[6]和尹力博等^[7]. 这些研究使用的数据较小、方法简单, 一般是从毛利率、beta 值等不足 20 个指标中, 采用标准化法提取复合指标. 本文与这些研究不同之处在于: 1) 数据集涵盖面更广、指标数量更多, 包括盈利能力、营运策略等五个方面, 多达 94 个基本面指标; 2) 使用多种高维数据降维方法进行信息提取, 最终选择优于传统标准化法的偏最小二乘法, 从公司特征

中筛选出对股票收益有预测作用的信息,使得公司基本面综合质量指数的测度精度和预测能力更强,同时相较于非线性模型,也能更好的理解指标的构建过程。

其次,本文为中国股票市场质量投资和价值投资的有效性补充了实证证据。研究发现,基本面综合质量与股票预期收益呈正相关关系,表明我国股票市场存在质量溢价。换言之,质优股表现优于垃圾股,证实基于公司综合质量的价值投资策略可行。第三,本文注重对价值投资的有效性进行深入理解与分析,创新性的引入非结构化的文本数据开展经济机制检验。在市场情绪方面,目前大多研究以 Baker 等^[12]构建情绪的方法为基础,检验投资者情绪与股票市场异象的关系^[13, 14],而开户数等是由情绪触发的结果变量,测度过于间接。本文以媒体情感倾向性可以描述投资者情绪为前提^[15],从基于新闻的个股情绪中构建市场情绪指标,来检验相关驱动机制。同时,采用从公司年报中提取的可读性指标,进一步理解信息不对称与质量溢价的关系。最后,从行为金融和宏观视角,分别检验了有限套利、公司决策、市场情绪、金融周期和经济周期,与公司质量溢价的关系,较全面的分析驱动上市公司质量溢价的经济理论机制。

1 数据与样本

1.1 数据来源及样本筛选

本文使用的财务数据、股票月度及日度收益、Fama 和 French 五因子^[16]、中国的无风险收益率等,均来自国泰安数据库(CSMAR);适用中国市场的三因子^[17]数据来源于 Robert F. Stambaugh 的个人主页^②;个股投资者情绪来自数行者科技公司(DataGo)的报刊新闻量化舆情数据库;年报可读性数据来源于爱文本(AI texts)研发的年报文本数据库。样本包括上海证券交易所和深圳证券交易所的全部 A 股,区间为 2002 年 1 月—2017 年 12 月。

数据筛选过程如下。首先,在构建投资组合

前,去除上季度变更为特殊转让或特别处理状态的股票,这些股票受到金融抑制并且缺乏市场流动性;第二,根据中国证券监督管理委员会行业分类,去除金融行业的上市公司;第三,剔除账面价值为负的上市公司,因为这类公司往往财务吃紧。第四,去除流通市值为负的公司。

1.2 基本面指标分类及选择

根据高登的固定股利增长模型(式 1),可知股价(*Price*)与股利(*Dividend*)、预期回报(*Ret*)和股利增长(*Growth*)息息相关

$$Price = \frac{Dividend}{Ret - Growth} \quad (1)$$

仿照 Fama 等^[16],将式(1)的左右两边分别除以账面价值(*BV*),再将分子乘以利润(*Profit*)再除以利润,整理过程如下式

$$\begin{aligned} \frac{Price}{BV} &= \frac{\frac{Dividend}{BV}}{Ret - Growth} = \frac{\frac{Profit}{BV} \times \frac{Dividend}{Profit}}{Ret - Growth} \\ &= \frac{Profitability \times Payout_ratio}{Ret - Growth} \quad (2) \end{aligned}$$

式(2)说明,股票预期收益率与公司盈利能力(*Profitability*)、成长能力(*Growth*)、营运策略(*Payout_ratio*)和价值(*Price/BV*)紧密相关。另外,上述估值模型要求股利稳定增长,所以基本面信息分类也将考虑公司的稳定性或安全性。直观上,公司盈利能力强,营运策略、成长性和安全性良好,潜在价值高,则公司质量较好,预期收益较高。指标分类说明如下,具体指标库构建见表 1。

1) 盈利能力,即公司赚钱的能力。

2) 成长能力,即公司成长潜力或发展潜力,主要是指盈利能力前后两期的变动。

3) 营运策略,包括公司分红和投融资策略、运营效率和盈余管理。

4) 潜在价值,主要指反应公司股价与基本面关系的相关指标。

5) 安全性,包括公司面临的系统性风险、非系统性风险和偿债能力。

② <http://finance.wharton.upenn.edu/~stambaug/>.

表 1 基本面指标分类及符号化^③

Table 1 Classification and notation of fundamental indicators

指标 1	符号 1	指标 2	符号 2	指标 3	符号 3
表 A：盈利能力(19 个)					
永久经济盈余	CEPS	现金流资产比	CFOA	现金资产比	CTA
息税前收益	EBIT	每股盈余	EPS	企业收益率	EY
F 评分	F	现金流构盈利因子	FOE	边际毛利	GM
毛利率 1	GP	投资收益占比	IP	净利润	NPOP
金融投资收益	RFI	净运营资产收益	RNA	资产收益率	ROA
股权收益率	ROE	投资型资产收益率	ROIC	应税所得与账面资产比	TBI
经营利润率	ROO				
表 B：成长能力(24 个)					
资本开销增长率	CAPXG	纳税变化额	CHTX	销售增长率	GS
可持续增长率	SG	纳税增长	TG	其他盈利指标变动	G_X ^④
表 C：营运策略(27 个)					
应计收入	ACC	公司年龄	AGE	资产换手率	ATO
公司投资	CI	现金及其等价物对固定资产净值比	CAC	资本换手率	CTO
财务杠杆	DA	股东权益变化	dBE	固定资产与存货变化率	dPIA
股利支付率	DPR	投资资产比	IA	每股无形资产	IBV
投资资本比	IC	存货变化率	IVC	存货增长率	IVG
每股净资产	NAPS	研发投入与销售收入之比	RDS	净运营资本	NOA
房地产投资量	RESTA	公司公允价值变动与股本比	RSGL	百分比应计收入	PACC
销售收入比	SAREV	销量现金比	SC	销量存货比	SI
公司规模	SIZE	销量增长与存货增长差	SMI	有形资产与营业收入比	TA
表 D：潜在价值(18 个)					
资产市值比	AM	账面市值比	BM	现金流负债比	CFD
现金流股价比	CFP	长期债务股本比	DLME	红利价格比	DP
市盈率	EP	基于现金流的估值因子	FM	公允价值价格比	FVAD
最大日收益	MAXRET	现金流量价格比	NCF	营运现金流价格比	OCFP
延迟股价	PRCDEL	研发支出与上市公司流通市值之比	RDM	反转	REV1
人民币交易量	RVOL	销量价格比	SP	股票换手率	TURN
表 E：安全性(6 个)					
系统性风险系数	BETA	系统性风险系数平方	BETASQ	非流动性	ILLIQ
异质性收益波动率	IVOL	人民币交易量波动率	S_RVOL	Z 评分	Z

1.3 投资者情绪代理变量

在新闻媒体情绪可以度量投资者情绪的前提下^[15]，本文选择基于报刊新闻个股情绪均值测度市场情绪。DataGo 数据库涵盖各大报纸媒体上与 A 股上市公司相关的所有新闻，兼顾新上市、退市公司及公司更名等多种情况，对广告等噪音进行清理，同时甄别“中关村”等较易混淆的上市公司新闻，保障了本文分析样本干净有效。该数据库以

句为单位切割新闻，将句子中的词汇褒贬性、词向量统计值、句子结尾标点符号类型等特征，输入至有监督的机器学习模型中，来判别句子表达的是积极、中性还是消极情绪；在此基础上，赋予句子不同位置差异化的权重，通过句子情绪的加权平均得到新闻的情感倾向性。该指标有别于以往对全文整体简单进行积极、中性、消极的离散性划分，构建了一个介于 -1 ~ 1 之间连续的新闻情绪

③ 基本面指标的文献来源可向作者索要。

④ G_X 表示任意盈利能力指标 X 较上一期的数值增量，例如资产收益率 ROA 的变动表示为 G_ROA。

测度,数值越大表示更积极的新闻.构建市场情绪指标方法如下:首先将新闻倾向性与股票代码合并,与股票 i 相关的任意一篇新闻 j 的情绪测度 $Senti_{i,j,t}$;然后将每个月 t 中每篇新闻的情绪 $Senti_{i,j,t}$ 取算术平均值,得到股票 i 在 t 月的个股情绪测度 $Senti_{i,t}$;最后将每个月的个股情绪测度取均值则为基于新闻的市场情绪测度 $Senti_t$.

1.4 有限套利代理变量

有限套利指标如下:第一,异质性波动($IVOL$).越高的异质性波动率表明公司股票估值的不确定性越强,套利风险越大.以当月此前 12 个月为一个样本期,即最多 250 个交易日的样本,将日度个股收益率对市场收益率进行回归,得到的残差的标准差即为个股异质性波动率;第二,非流动性指标($ILLIQ$)^[18].非流动性越高(或流动性越低)表明价格对每个订单流的影响越大,导致投资者的套利成本越高.将日度收益的绝对值除以日度人民币计价的交易量,计算此前 12 个月上述比值均值的自然对数,为非流动性指标;第三,交易额($RVOL$).交易额越低,套利者受限程度越高,股票错误定价水平越高.以每日交易量与收盘价的乘积作为日度交易额,将过去 12 个月的日度交易额累加作为以人民币计价的交易额;第四,股票价格(PRC).当股票价格低时,买卖价差相对较高,套利成本大.当月收盘价的自然对数即为股票价格.

简言之,异质性波动越强、非流动性越高、交易额越低、股票收盘价越低,则有限套利程度越大,股票越可能遭受错误定价.

1.5 可读性代理变量

AI TEXTS 的年报文本数据库中汇总了自 2001 年开始的 A 股上市公司年报信息.本文从年报长度视角度量可读性,选择年报总页数($PAGE$)、总字数($WORD$)、总词数($TERM$)等指标测度年报可读性.认为页数越多、字数越多、词数越多,则可读性越差^⑤.

1.6 金融市场与经济增长代理变量

在检验金融周期与公司质量预测能力的关系

时,从股票市场、利率市场和汇率市场分别选择沪深 300 指数($CSI300$)、一年定期存款利率(IR)和美元兑人民币汇率(ER)三个变量.通常认为指数上涨,股市利好;利率越低,货币政策越宽松;汇率高时,人民币处于升值周期.在经济周期驱动机制方面,选择国民生产总值(GDP)增长率、消费者物价指数(CPI)和生产价格指数(PPI)三个变量.其中, CPI 采用环比指数, PPI 为采掘工业当月同比指数^⑥.认为 GDP 增速放缓、通货紧缩、 PPI 下跌是经济下行的表现.

2 综合质量指数构建

本文采用主成分分析法(PCA)、珙玛麦克白回归法(FM)、预测组合法(FC)、缩放主成分分析法(SPCA)、标准化法(NA)和偏最小二乘法(PLS),分别构建综合质量指数.

2.1 主成分分析法

PCA 是最常见的降维方法.应用此方法提取公司基本面指标中的第一个主成分,作为综合质量测度 Q^{PCA} .具体步骤如下:第一,将公司基本面指标进行标准化,得到 t 期基本面指标集合 $X_t = (x_{1,t}, x_{2,t}, \dots, x_{N,t})'$,其中 $N = 94$;第二,计算第一个主成分与每个公司特征变量的相关系数 $\beta_{i,t}^{PCA}$;第三,将 $(t-11)$ 期至 t 期的相关系数取均值 $\beta_{i,(t-11,t)}^{PCA}$,作为 t 期第 i 个公司特征的最终权重,计算综合质量指标如下

$$\beta_{i,(t-11,t)}^{PCA} = \frac{1}{12} \sum_{j=t-11}^t \beta_{i,j}^{PCA} \quad (3)$$

$$Q_t^{PCA} = \sum_{i=1}^N (\beta_{i,(t-11,t)}^{PCA} \times x_{i,t}) \quad (4)$$

2.2 珙玛麦克白回归法

FM 是资产定价领域常见的回归方法,本文应用此方法度量公司特征对股票未来收益的贡献度.首先在横截面上,应用式(5)对标准化后的公司基本面指标进行 FM 回归,得到 t 期第 i 个公司

⑤ 中国证监会对年度报告的内容与格式有明确规定,例如《公开发行证券的公司信息披露内容与格式准则第 2 号——年度报告的内容与格式(2001 年修订稿)》等,因此能用简短的语言描述规定的内容表明年报的可读性越高.

⑥ 替换代理变量,后文检验仍可得到类似结论.

特征 $x_{i,t}$ 对股票预期收益的敏感度 $\delta_{i,t}$; 然后同样将包括当期在内的过去 12 期回归系数取均值得到 $\delta_{i,(t-11,t)}^{\text{FM}}$, 并以此为权重计算综合质量指标(式 7)

$$R_{t+1} = C_t + \sum_{i=1}^N (\delta_{i,t} \times x_{i,t}) + \varepsilon_{t+1} \quad (5)$$

$$\delta_{i,(t-11,t)}^{\text{FM}} = \frac{1}{12} \sum_{j=t-11}^t \delta_{i,j} \quad (6)$$

$$Q_t^{\text{FM}} = \sum_{i=1}^N (\delta_{i,(t-11,t)}^{\text{FM}} \times x_{i,t}) \quad (7)$$

2.3 预测组合法

预测组合法的核心思想是基于每一个公司基本面指标的预测能力, 构建该基本面指标的份额, 然后将每个基本面指标的份额一视同仁, 取算术平均构建综合质量指数. 具体而言, 首先在横截面上, 应用式(8)分别对标准化后的公司基本面指标进行 FM 回归, 分别得到每个公司特征的回归系数 $\delta'_{i,t}$; 然后同样取过去 12 期的移动平均值得到 $\delta_{i,(t-11,t)}^{\text{FC}}$, 以此为权重计算每个公司基本面指标的“拟合值” $\hat{x}_{i,t}$; 第三, 将所有 $\hat{x}_{i,t}$ 取算术平均, 构建综合质量指标(式 10)

$$R_{t+1} = C_{i,t} + \delta'_{i,t} \times x_{i,t} + \varepsilon_{t+1} \quad (8)$$

$$\delta_{i,(t-11,t)}^{\text{FC}} = \frac{1}{12} \sum_{j=t-11}^t \delta'_{i,j} \quad (9)$$

$$Q_t^{\text{FC}} = \frac{\sum_{i=1}^N \hat{x}_{i,t}}{N}, \hat{x}_{i,t} = \delta_{i,(t-11,t)}^{\text{FC}} \times x_{i,t} \quad (10)$$

2.4 缩放主成分分析法

缩放主成分分析法是 FC 方法与 PCA 方法的结合. 首先, 应用预测组合法, 计算每个公司基本面指标的“拟合值” $\hat{x}_{i,t}$ (式 10), 得到“拟合值”数据集 $\hat{X}_t = (\hat{x}_{1,t}, \hat{x}_{2,t}, \dots, \hat{x}_{N,t})'$; 然后对 \hat{X}_t 采用主成分分析法, 提取第一个主成分与每个公司特征“拟合值”的相关系数 $\beta_{i,j}^{\text{SPCA}}$, 取 12 个月的移动平均值, 并以此为权重构建综合质量指标(式 12)

$$\beta_{i,(t-11,t)}^{\text{SPCA}} = \frac{1}{12} \sum_{j=t-11}^t \beta_{i,j}^{\text{SPCA}} \quad (11)$$

$$Q_t^{\text{SPCA}} = \sum_{i=1}^N (\beta_{i,(t-11,t)}^{\text{SPCA}} \times \hat{x}_{i,t}) \quad (12)$$

2.5 标准化法

Asness 等^[1]构建质量指数的方法核心思想是多次利用标准化与分类加和, 具体分为两步. 首

先, 依次将第 j 类 t 期第 i 个基本面指标的排序 $\text{Rank}_{i,t}^j$ 进行标准化得到 $z(\text{Rank}_{i,t}^j)$, 再进行同一类别中 k_j 个标准化后指标的加和运算; 然后, 将第一步的结果继续进行标准化得到第 j 类的综合指标 z_t^j (式 13), 再将 5 个类别的综合指标 z_t^j 进行加和, 并对此再次应用标准化, 得到最终的质量测度(式 14). 其中当基本面指标 $x_{i,t}^j$ 是股票收益的正向预测因子时, 则采用升序排列得到 $\text{Rank}_{i,t}^j$; 反之, 采用降序排列

$$z_t^j = z \left[\sum_{i=1}^{k_j} z(\text{Rank}_{i,t}^j) \right] \quad (13)$$

$$Q_t^{\text{NA}} = z \left(\sum_{j=1}^5 z_t^j \right) \quad (14)$$

2.6 偏最小二乘法

PLS 可以通过两个最小二乘法 (ordinary least squares, OLS) 来实现^[19]. 首先, 在横截面上, 将 t 期股票收益 R_t 对每个滞后一期基本面指标 $x_{i,t-1}$ 进行回归.

$$R_t = \pi_{i,0} + \pi_i x_{i,t-1} + u_{i,t}, t = 1, \dots, T \quad (15)$$

其中 π_i 表示公司特征 $x_{i,t-1}$ 对股票预期回报 R_t 的敏感度. 假设股票回报 R_t 可预测的部分是由公司质量 Q_{t-1} 所致, 基本面指标与股票收益可预测部分相关, 与无法预测部分不相关, 因此 π_i 描述了每个公司特征对综合质量指数的贡献度.

第二步, 在时间序列上, 运行每个基本面指标 $x_{i,t}$ 关于 π_i 估计值的线性回归

$$x_{i,t} = \omega_t + Q_t^{\text{PLS}} \hat{\pi}_i + v_{i,t}, i = 1, \dots, N \quad (16)$$

由式(16)回归得到的 Q_t^{PLS} 就是 PLS 方法构建的综合质量指数. 第一步的横截面回归, 先获得 $\hat{\pi}_i$, 即每个公司特征对最终指数的贡献度; 第二步的时间序列回归, 则是将每个公司特征对这种贡献度进行回归, 得到最终的校准指数. 换言之, PLS 方法使用预期股票回报对基本面进行降维, 筛除个体噪音, 从而提取可以预测回报的公司质量信息.

理论上, PCA 和 SPCA 中都应用了主成分分析法, 该方法虽然可以解释绝大多数单一公司特征的信息, 但无法区分公共信息和公共误差. 实际上, 公共误差的偏差越大, 其在 PCA 中的作用就越强, 所以导致构建的综合质量指数中含有大量

与股票收益预测无关的信息,失去对股票收益的预测能力. FM 方法未考虑不同基本面指标的相关性问题. FC 方法中认为单一指标对股票收益预测的敏感度具有无差异性. NA 方法中简单加和的方式并未考虑单一指标贡献度上的不同. 而 PLS 是一个基于 OLS 的两步法,首先提取单一基本面指标对股票收益预测的敏感性,滤掉无关信息,然后通过再次回归从基本面指标中提取上述敏感性的公共信息来构建指标,达到高效、准确提取与股票收益预测最相关信息的效果. 因此,PLS 方法的终极目标是从一系列预测指标中提取与股票预期收益协方差最大的公共因子,该因子也许并不是预测因子中最重要的公共元素,但具有最强的预测能力. 总的来说,PLS 方法不仅可以从经济学上理解每一步得到的回归系数的含义,也能避免将不同指标杂糅回归、主观赋权等问题,因而理论上应是实证表现最好的方法.

3 综合质量指数与股价表现

3.1 不同方法构建的质量指数表现分析

对于单一基本面指标而言,本文发现每类公司特征在横截面股票收益上都具有一定的预测能力,但预测能力的差异较大^⑦,再次说明降维构建综合质量测度的重要性. 因此,应用单变量投资组合分析法,比较偏最小二乘等 6 种方法构建的综合质量指数 Q^{PLS} 、 Q^{NA} 、 Q^{PCA} 、 Q^{FM} 、 Q^{FC} 和 Q^{SPCA} 的预测能力. 在每月月初,由最新公布的季报计算综合质量指数,据此将股票分为 10 组,以等权重方式构建投资组合,同时通过买入高质量、卖出低质量的股票构建多空对冲投资组合,计算每个投资组合的平均月度原始收益率. 然后基于 CAPM、三因子^[17]和五因子模型^[16],分别计算对冲组合的超额收益率 $CAPM \alpha$ 、 $FF3 \alpha$ 和 $FF5 \alpha$. 表 2 为单变量投资组合分析法的实证结果.

表 2 不同方法构建综合质量指数预测能力比较

Table 2 Performance of firm quality predictors aggregated by different approaches

收益	投资组合	(1) Q^{PLS}	(2) Q^{NA}	(3) Q^{PCA}	(4) Q^{FM}	(5) Q^{FC}	(6) Q^{SPCA}
Ret	L	1.01 [1.23]	0.93 [1.13]	1.87 [2.02]	1.54 [1.83]	1.37 [1.63]	1.71 [2.07]
	H	4.17 [4.60]	3.37 [3.88]	2.13 [2.93]	2.54 [3.05]	2.86 [3.41]	2.46 [2.88]
	H-L	3.16 [6.65]	2.44 [6.92]	0.26 [0.59]	1.00 [3.08]	1.48 [3.38]	0.76 [1.61]
$CAPM \alpha$	H-L	3.10 [6.50]	2.51 [7.11]	0.37 [0.85]	1.02 [3.11]	1.51 [3.41]	0.69 [1.46]
$FF3 \alpha$	H-L	2.96 [5.25]	1.78 [5.23]	-0.35 [-1.21]	0.96 [2.48]	1.47 [2.81]	0.89 [1.60]
$FF5 \alpha$	H-L	2.77 [5.46]	1.85 [6.95]	0.88 [3.72]	0.78 [2.29]	1.22 [2.60]	0.76 [1.56]

注: 该表汇报由综合质量指数构建的投资组合的月度平均原始收益率(Ret , %), 对冲投资组合的超额收益率($CAPM \alpha$ 、 $FF3 \alpha$ 和 $FF5 \alpha$, %)及其 t 统计量([·]). “L”(“H”)指具有最低(最高)综合质量的组合, “H-L”指多空对冲投资组合.

结果表明,除(缩放)主成分分析法外,其他四种方法构建的综合质量指数均有一定程度的预测能力. 例如,由 Q^{PLS} 、 Q^{NA} 、 Q^{FM} 、 Q^{FC} 构建的对冲组合均可以获得稳健的正收益,但 Q^{PLS} 表现更胜一筹. 当控制市场因子后, Q^{PLS} 的对冲组合收益率仍然在 3% 以上. 在中国三因子模型下, Q^{NA} 的投资

策略可以获得 1.78% 的超额收益,较原始收益率下降了近 0.6%,而 Q^{PLS} 仅减少约 0.2%. Q^{NA} 投资策略收益的绝对数值虽然仅次于 Q^{PLS} ,但风险调整收益下降幅度更大,说明通用的定价因子可以在更大的程度上解释该策略的收益. 当进一步控制投资和盈利因子后,各组合的超额收益均有一

⑦ 单一基本面指标的投资组合分析结果有兴趣者可向作者索要.

定幅度的减少,但由 Q^{PLS} 得到的 $FF5 \alpha$ 在数值上仍然大于 Q^{NA} 的原始收益率. 若投资 Q^{PLS} 的对冲组合,平均每年可以获得高于 30% 的超额收益,显著高于其他投资策略. 同时,相较于同样应用标准化法构建变量,从而构建对冲组合的原始收益率 0.42%^[1]、0.99%^[7]、1.17%^[6],本文在更大基本面指标集基础上构建的质量指数 Q^{NA} 的收益率 (2.44%) 显著更高.

综上,后文将以 Q^{PLS} 作为公司基本面质量的测度指标,分析公司综合质量在股票横截面收益预测中的作用.

3.2 价值控制下的质量投资策略分析

每个月月初,以最新公布的季报计算公司账面市值比,将公司分为 3 组,从低至高账面市值比依次为成长型企业、中间型企业和价值型企业;同时以综合质量指数的五分位数,将公司

分为 5 组. 因此在每个月得到 15 (3 × 5) 个以等市值方式构建的投资组合. 同时,在成长型企业、中间型企业和价值型企业中,通过买入高质量、卖出低质量的股票,分别构建基于 Q^{PLS} 的多空对冲投资组合. 表 3 是公司账面市值比和 Q^{PLS} 独立双变量投资组合分析的结果. 结果表明,当控制账面市值比后,关于 Q^{PLS} 的对冲组合有显著的正收益. 而且随账面市值比增加,对冲组合的平均收益率依次升高,表明基于公司基本面质量指数的投资策略趋向于“价值型”策略. 同时也发现,在不同类型的公司中,多头和空头在投资策略中的贡献度有显著差异. 例如,在成长型公司中对冲组合的 $FF3 \alpha$ 大多来自空头,而在价值型公司中大部分来自多头,表明在价值型公司中采用该投资策略极可能不受市场卖空限制影响.

表 3 公司账面市值比与综合质量指数的投资组合分析结果

Table 3 Returns of the portfolios double sorted on book-to-market and aggregate firm quality

投资组合	成长型			中间型			价值型		
	<i>Ret</i>	<i>FF3 α</i>	<i>FF5 α</i>	<i>Ret</i>	<i>FF3 α</i>	<i>FF5 α</i>	<i>Ret</i>	<i>FF3 α</i>	<i>FF5 α</i>
L	1.09 [1.31]	-0.84 [-3.03]	-0.75 [-2.59]	1.40 [1.70]	-0.66 [-2.23]	-0.55 [-1.78]	1.59 [1.90]	-0.54 [-1.87]	-0.35 [-1.20]
H	2.65 [3.02]	0.40 [1.32]	0.50 [1.59]	3.13 [3.47]	0.68 [2.81]	0.80 [3.23]	3.60 [3.99]	1.07 [3.67]	1.14 [3.79]
H-L	1.57 [3.68]	1.24 [2.88]	1.25 [2.77]	1.73 [4.46]	1.33 [3.49]	1.35 [3.35]	2.01 [4.85]	1.60 [3.92]	1.49 [3.48]

注: 该表汇报的是公司账面市值比与综合质量指数双变量投资组合的月度平均初始收益 (*Ret*, %)、超额收益 (*FF3 α* 和 *FF5 α*, %) 及其 *t* 统计量 ([·]). 账面市值比 ($\ln(B/M)$) 为账面价值与市场价值比值的自然对数. “L” (“H”) 指具有最低 (最高) 综合质量的组合, “H-L” 指某一账面市值比水平上的多空对冲投资组合.

3.3 规模控制下的质量投资策略分析

根据公司流通市值,将样本分为微型企业、小型企业和大型企业;根据综合质量指数将样本分为 5 组. 在不同市值股票中,分别构建基于 Q^{PLS} 的对冲组合并计算各投资组合的初始收益率和超额收益率. 如表 4 所示,基于质量的对冲投资策略属于非大型企业投资策略. 在大型企业中,对冲组合的平均收益率均不显著,甚至在三因子模型下只能获得负的超额收益率. 同样的投资策略在小、微企业中不仅收益率高,显著

性也明显提高. 而且在流通市值越小的股票中,投资策略收益越大,多头对超额收益的贡献度越高,策略受卖空限制的可能性越低. 因此,在中国股票市场中,基于公司综合质量的投资策略极可能受小公司驱动,与我国上市公司中小公司占比较高的实际情况一致;而小公司由于价格低、交易额小、流动性差等因素,极可能妨碍套利者交易,难以修正市场上存在的错误定价,从而导致公司基本面综合质量具有较强的预测能力.

表 4 公司规模与综合质量指数的投资组合分析结果

Table 4 Returns of the portfolios double sorted on firm size and aggregate firm quality

投资组合	微型			小型			大型		
	Ret	$FF3\ \alpha$	$FF5\ \alpha$	Ret	$FF3\ \alpha$	$FF5\ \alpha$	Ret	$FF3\ \alpha$	$FF5\ \alpha$
L	2.49 [2.58]	-0.21 [-0.69]	-0.16 [-0.51]	1.40 [1.64]	-0.85 [-3.21]	-0.76 [-2.75]	1.27 [1.64]	-0.39 [-1.35]	-0.26 [-0.85]
H	4.21 [4.26]	1.38 [4.73]	1.55 [5.30]	2.23 [2.58]	-0.13 [-0.53]	0.05 [0.21]	1.45 [1.73]	-0.45 [-1.33]	-0.17 [-0.49]
H-L	1.73 [4.58]	1.59 [4.12]	1.71 [4.30]	0.84 [2.54]	0.72 [2.13]	0.81 [2.28]	0.18 [0.44]	-0.06 [-0.14]	0.09 [0.20]

注：该表汇报的是公司规模与综合质量指数双变量投资组合的月度平均初始收益(Ret , %)、超额收益($FF3\ \alpha$ 和 $FF5\ \alpha$, %)及其 t 统计量([·]). 公司规模($\ln(ME)$)为月底总流通市值的自然对数。“L”(“H”)指具有最低(最高)综合质量的组合,“H-L”指某一规模公司的多空对冲投资组合。

4 经济机制检验

本节从行为金融和宏观周期两个视角,检验驱动公司基本面质量对股票横截面收益预测能力的经济机制。

4.1 有限套利

当股票被错误定价时,盈利机会将吸引理性投资者通过套利交易,修正市场中的错误定价。理想环境中,套利是无风险的、显而易见的,并且无成本,那么股价应该准确、无误的反应所有可得的信息,即使存在错误定价,也会被立刻修正。然而,实际市场中不仅套利有风险、成本高,而且能够执行的套利交易也极其有限。虽然套利者会针对错误定价不断的进行交易,但当有限套利程度高时,也会花费更长的时间来消除错误定价。由于套利者通常缺乏多样性,受股票的非系统性波动率、流动性等因素影响,优化的交易策略极可能遭受损失,因而阻碍套利者寻求交易机会^[8, 20]。综上,当套利风险更大、成本更高时,由公司基本面综合质量带来的套利机会,会降低对套利者的吸引力。如果股价仅能逐步反应公司质量信息,未及时反应至股价中的信息造成了错误定价,那么当套利摩擦更大时,套利者通过交易修正错误定价的动力更低,因此公司质量信息的预测能力应该更强。

应用独立双变量投资组合分析法,对有限套利机制进行检验。在每月月初,根据上月有限套利

代理变量将样本分为 2 组,同时根据公司综合质量指标分为 5 组,以等市值方式构建 10 个初始投资组合,且在不同套利受限水平下构建公司质量的对冲组合。表 5 为有限套利与质量指数独立双变量投资组合分析的结果。

结果表明,对于不同非系统性风险的股票,基于质量指数的投资策略收益均显著,并且差异较小,在分别控制市场因子、规模和价值因子后,呈现类似的规律。而非流动性、交易额、收盘价的实证结果,与有限套利理论预期一致。股票非流动性高(交易额小、收盘价低),则套利受限水平高,投资者可以通过构建对冲策略,获得平均每月 1.38%(1.89%、1.89%)的初始收益率, t 值为 4.20(4.90、5.36),在 1% 的水平上显著;而非流动性低(交易额大、收盘价高)的股票,则套利限制低,同样的投资策略平均每月仅收入 0.46%(0.72%、1.21%),并且统计上不显著(在 10% 水平上显著、在 1% 水平上显著),表明套利限制低时,交易者一旦发现公司基本面质量的价值,能及时通过套利交易减少甚至消除市场上的错误定价,降低综合质量的预测能力。从初始收益中剔除风险定价因子后,超额收益率也呈现相同的规律,公司基本面质量指数的投资策略均在套利摩擦高的公司中可以获得显著的风险调整收益。因此,对于非流动性高、交易额小、股价低的股票,套利交易受限程度越大,错误定价水平越高,导致基本面综合质量的预测能力越强。

表 5 有限套利与综合质量指数的投资组合分析结果

Table 5 Returns of the portfolios double sorted on limits-to-arbitrage and aggregate firm quality

套利摩擦变量	投资组合	低套利摩擦				高套利摩擦			
		<i>Ret</i>	<i>CAPM</i> α	<i>FF3</i> α	<i>FF5</i> α	<i>Ret</i>	<i>CAPM</i> α	<i>FF3</i> α	<i>FF5</i> α
<i>IVOL</i>	L	1.34 [1.72]	0.10 [0.24]	-0.69 [-2.52]	-0.56 [-2.01]	1.34 [1.56]	0.01 [0.02]	-0.70 [-2.44]	-0.58 [-1.93]
	H	3.16 [3.66]	1.84 [3.84]	0.77 [3.05]	0.88 [3.45]	3.05 [3.24]	1.64 [3.06]	0.55 [1.96]	0.70 [2.46]
	H-L	1.81 [5.01]	1.74 [4.80]	1.46 [4.09]	1.44 [3.84]	1.72 [3.89]	1.63 [3.69]	1.25 [2.87]	1.28 [2.81]
<i>ILLIQ</i>	L	1.02 [1.27]	-0.28 [-0.75]	-0.86 [-3.06]	-0.73 [-2.51]	1.84 [2.18]	0.60 [1.15]	-0.44 [-1.67]	-0.33 [-1.21]
	H	1.49 [1.74]	0.13 [0.31]	-0.65 [-2.15]	-0.50 [-1.58]	3.22 [3.58]	1.87 [3.61]	0.69 [3.50]	0.91 [4.71]
	H-L	0.46 [1.17]	0.41 [1.03]	0.21 [0.51]	0.23 [0.55]	1.38 [4.20]	1.27 [3.92]	1.13 [3.42]	1.24 [3.58]
<i>RVOL</i>	L	1.12 [1.36]	-0.21 [-0.53]	-0.82 [-2.84]	-0.66 [-2.20]	1.79 [2.16]	0.58 [1.11]	-0.43 [-1.47]	-0.43 [-1.45]
	H	1.84 [2.12]	0.48 [1.07]	-0.36 [-1.20]	-0.20 [-0.63]	3.68 [3.97]	2.32 [4.18]	1.08 [4.15]	1.27 [4.91]
	H-L	0.72 [1.75]	0.69 [1.66]	0.45 [1.08]	0.46 [1.04]	1.89 [4.90]	1.75 [4.62]	1.50 [3.96]	1.70 [4.31]
<i>PRC</i>	L	1.11 [1.38]	-0.16 [-0.38]	-0.78 [-2.79]	-0.77 [-2.62]	1.67 [1.88]	0.30 [0.63]	-0.68 [-2.09]	-0.36 [-1.17]
	H	2.31 [2.66]	1.00 [2.01]	0.14 [0.48]	0.17 [0.56]	3.56 [3.86]	2.20 [4.02]	0.93 [3.35]	1.22 [4.69]
	H-L	1.21 [3.21]	1.16 [3.07]	0.92 [2.43]	0.94 [2.37]	1.89 [5.36]	1.89 [5.31]	1.61 [4.60]	1.59 [4.31]

注：该表汇报的是有限套利与综合质量指数的双变量投资组合的月度平均初始收益(*Ret*, %)、超额收益(*CAPM* α 、*FF3* α 和 *FF5* α , %) 及其 *t* 统计量([·])。 “L”(“H”)指具有最低(最高)综合质量的组合, “H-L”指某一错误定价水平下的多空对冲投资组合。

4.2 公司决策

以下从财务管理视角,分析公司质量与年报可读性、公司投资决策的关系. 表 6 汇报了全样本(2003 年—2017 年)及 4 个子样本(2003 年—2006 年, 2007 年—2010 年, 2011 年—2014 年, 2015 年—2017 年)中,不同质量公司在年报页数、字数、词数及资产增长率(*AG*)上的差异. 如表 6 所示,高质量的公司总能以较短的年报披露必要信息,年报可读性较高,从而降低公司与投资者间的信息不对称. 例如在全样本中,基本面综合质量高的公司平均年报长度(字数、词数)为 122.12 页(11.08 万字、5.67 万词),而质量低的公司为

134.30 页(12.39 万字、6.32 万词),表明高质量的公司可以减少约 12 页(1.31 万字、6 000 余词)的赘述,报告总体更精炼,尽可能少的分散投资者注意力,增强年报的可读性,利于增强质量溢价. 随时间推移,无论公司质量高低,年报在页数、字数和词数上均有所上升;2003 年—2006 年间,不同质量公司的年报可读性差异并不显著,但随后数值上逐渐显现差异,2007 年—2014 年间这种差异在统计意义上显著. 表 6 的最后一组显示了公司质量与投资的关系. 结果显示,高质量公司资产增长率为 19%,低质量为 63%,差异在统计意义上显著($t = -2.10$),表明质量低的公司有盲目扩

张的嫌疑. 样本初期公司非但没有扩大规模反而缩减, 表明资产增长率极可能是金融危机到来的先行指标; 2007 年起, 高低质量水平的公司都开始不同程度的投资, 其中低质量水平的公司投资

增长明显, 3 个子样本期分别为 34%、143% 和 94%, 而高质量的公司仅为 25%、25% 和 47%, 说明质量低的公司徒有其表, 不仅基本面质量差, 还存在过度投资现象, 导致股票预期收益较低.

表 6 公司质量与年报可读性、公司投资

Table 6 Readability of annual reports, firm investment and aggregate firm quality

变量	公司质量	2003 年—2017 年	2003 年—2006 年	2007 年—2010 年	2011 年—2014 年	2015 年—2017 年
PAGE	L	134.30	77.81	133.42	156.81	180.77
	H	122.12	76.41	114.14	139.88	170.03
	H-L	-12.18 [-3.88]	-1.39 [-0.80]	-19.28 [-6.49]	-16.93 [-3.86]	-10.75 [-0.86]
WORD	L	123 878.99	87 324.92	122 787.13	137 096.69	156 449.95
	H	110 794.69	85 183.64	102 778.93	118 281.15	145 648.48
	H-L	-13 084.30 [-3.28]	-2 141.28 [-1.19]	-20 008.20 [-5.98]	-18 815.54 [-3.46]	-10 801.47 [-0.59]
TERM	L	63 234.35	36 169.12	65 392.77	72 765.30	83 735.51
	H	56 709.03	35 418.14	55 224.00	63 546.09	77 960.84
	H-L	-6 525.33 [-3.52]	-750.99 [-1.11]	-10 168.77 [-5.93]	-9 219.21 [-3.71]	-5 774.67 [-0.72]
AG	L	0.63	-0.10	0.34	1.43	0.94
	H	0.19	-0.16	0.25	0.25	0.47
	H-L	-0.45 [-2.10]	-0.06 [-0.51]	-0.09 [-0.60]	-1.18 [1.74]	-0.47 [-1.86]

注: 该表汇报的是全样本及分时段子样本中, 不同质量的公司的年报可读性及投资情况. “L” (“H”) 指具有最低 (最高) 综合质量的公司; “H-L” 指质量高、低的公司中年报可读性或公司投资的差异, [·] 为相应的 t 统计量.

4.3 市场情绪

投资者情绪理论认为噪音交易者比较容易受情绪影响, 从而导致股票价格脱离基本面^[12]. 理论上, 理性投资者会寻求市场上由被高估和低估造成的盈利机会, 通过不断交易, 逐渐减少甚至消除错误定价. 因此, 在有效市场中, 较难发现被低估的股票, 反而由于卖空等限制往往较多被高估的现象. 但当理性投资者无法找到或利用全部机会, 极可能为情绪在市场中发挥作用带来契机. 股票价格往往受持乐观态度的投资者影响, 这是由于即使悲观投资者认为股价过高, 也无法进行卖空交易, 从而导致情绪作用的非对称性. 因此根据 Stambaugh 等^[13] 和 Yan 等^[21] 的分析, 推断公司基本面综合质量指标具有预测能力是质优股被低估、劣质股被高估的体现, 那么由综合质量产生的溢价应该在情绪高涨时更强.

具体地, 将市场情绪高于样本期中位数的月

份定义为高情绪的月份, 低于中位数的月份定义为低情绪的月份^[12-14]. 在每月月初, 根据上月市场情绪将样本划分为情绪高涨期和情绪低迷期; 在两个子样本中, 分别由基本面综合质量构建对冲投资组合. 表 7 为投资组合分析的结果.

结果表明, 无论市场情绪高低, 基于 Q^{PLS} 的对冲投资策略都可以获得显著的正收益. 当市场乐观时, 对冲组合的平均月度原始收益率为 3.72% ($t = 4.83$); 而市场悲观时, 组合的收益率为 2.59% ($t = 4.70$). 可见, 控制市场情绪后, 基本面综合质量的预测能力仍然稳健; 同时, 上述收益率也表明高情绪时期, 由综合质量产生的溢价显著高于低情绪期, 年化收益率相差大于 13%, 与 Stambaugh 等^[13] 和 Cheema 等^[14] 结果一致. 当从原始收益率中滤掉通用风险因子的影响后, 高情绪期的超额收益比原始收益下降幅度更大, 例如经过五因子调整后的收益为 2.16% ($t = 2.18$), 此时风险因子

可以解释超过 4 成的策略收益;而低情绪期的 $FF5$ 15%. 上述结果表明,市场情绪乐观时,风险因子定价模型对更高的质量溢价有一定的解释能力。 α 为 2.21% ($t = 3.54$), 相较原始收益仅下降不足

表 7 市场情绪与综合质量指数的预测能力^⑧

Table 7 Performance of aggregate firm quality conditional on market sentiment

投资组合	低情绪时期				高情绪时期			
	Ret	$CAPM\ \alpha$	$FF3\ \alpha$	$FF5\ \alpha$	Ret	$CAPM\ \alpha$	$FF3\ \alpha$	$FF5\ \alpha$
L	-2.39 [-2.02]	0.14 [0.27]	-1.04 [-3.01]	-1.04 [-2.87]	4.42 [4.36]	0.05 [0.06]	-0.62 [-1.05]	-0.24 [-0.37]
H	0.20 [0.17]	2.57 [4.42]	1.41 [3.34]	1.17 [2.73]	8.14 [6.48]	2.97 [2.78]	2.01 [3.15]	1.92 [2.67]
H-L	2.59 [4.70]	2.43 [4.24]	2.45 [4.05]	2.21 [3.54]	3.72 [4.83]	2.92 [3.24]	2.62 [2.98]	2.16 [2.18]

注：该表描述在不同市场情绪时期，由综合质量指数构建的投资组合的月度平均原始收益率（ Ret ，%），超额收益率（ $CAPM\ \alpha$ 、 $FF3\ \alpha$ 和 $FF5\ \alpha$ ，%）及其 t 统计量（ $[\cdot]$ ）。“L”（“H”）指具有最低（最高）综合质量的组合，“H-L”指某一投资者情绪时期多空对冲投资组合。

对于各投资组合而言,市场整体悲观时,大部分投资组合的原始收益率都为负值,即使综合质量最高的组合也未能获得显著正收益,此时空头对投资策略的贡献度达到 90% 以上;市场乐观时,各投资组合均能获得显著正收益,而且随基本面质量提高呈单调递增趋势;同时,无论市场情绪处于上行或下行周期,多头经风险调整后的收益均显著为正,而且多头在策略收益中的贡献度也有所上升. 由于我国股票市场机构投资者少、投资者金融素养匮乏等特征,投资者无法精准瞄定套利机会,使得市场中同时存在被高估和低估的股票. 实证结果说明在我国股票市场中,情绪高涨使市场整体被低估,在短期内会迎来股市上涨. 表 7 的结果也发现乐观期多头、空头收益都显著高于悲观期,表明我国股市投资中,质量好的公司在市场乐观时被严重低估、质量差的公司悲观期被严重高估,与 Cheema 等^[14] 结论一致. 而 Stambaugh 等^[13] 认为相对空头,多头被低估,市场上的低估可以由套利交易者通过买入行为消除,并且发现美国股票市场中,无论情绪高低多头收益无显著差异,这与我国市场不同.

4.4 金融周期与经济周期

为检验金融市场周期、经济周期与综合质

量溢价的关系,采用如下回归进行实证分析,模型(17)为基础模型,(模型 18~模型 20)分别控制不同的风险因子进一步检验经济周期机制的稳健性. 其中 Ret_t^Q 为基于质量指数构建的对冲组合在 t 月的收益率; $MKTRF$ 、 SMB 、 HML 、 RMW 、 CMA 分别为市场、规模、价值、盈利和投资因子,回归系数为相应的风险载荷;本文选择沪深 300 指数($CSI300$)、利率(IR)和汇率(ER), GDP 增长率、 CPI 和 PPI 分别作为宏观金融、经济基本面, $Dummy_t^H$ 为根据这些变量构造的虚拟变量,当 t 月的变量数值高于样本中位数时取值为 1,表示股市利好、货币政策紧缩、人民币升值或经济上行,否则为 0. 实证结果如表 8

$$Ret_t^Q = \alpha + \varphi Dummy_t^H + \varepsilon_t \tag{17}$$

$$Ret_t^Q = \alpha + \beta_{MKT}MKTRF_t + \varphi Dummy_t^H + \varepsilon_t \tag{18}$$

$$Ret_t^Q = \alpha + \beta_{MKT}MKTRF_t + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \varphi Dummy_t^H + \varepsilon_t \tag{19}$$

$$Ret_t^Q = \alpha + \beta_{MKT}MKTRF_t + \beta_{SMB}SMB_t + \beta_{HML}HML_t + \beta_{RMW}RMW_t + \beta_{CMA}CMA_t + \varphi Dummy_t^H + \varepsilon_t \tag{20}$$

⑧ 限于篇幅,正文未汇报中等质量投资组合的收益率,结果的完整版可向作者邮件索要.

表 8 金融周期、经济周期与综合质量溢价

Table 8 Quality premium concerning on the financial and business cycle

模型	φ_{CSI300}	φ_{IR}	φ_{ER}	φ_{GDP}	φ_{CPI}	φ_{PPI}
(17)	0.47 [0.48]	-1.29 [-1.35]	0.00 [0.00]	-1.58 [-1.65]	-3.17 [-3.44]	-2.07 [-2.20]
(18)	0.44 [0.45]	-1.26 [-1.33]	-0.05 [-0.05]	-1.74 [-1.81]	-3.18 [-3.45]	-1.94 [-1.99]
(19)	0.44 [0.44]	-1.22 [-1.27]	-0.08 [-0.08]	-1.71 [-1.76]	-3.23 [-3.48]	-1.87 [-1.91]
(20)	0.32 [0.33]	-1.05 [-1.10]	-0.09 [-0.09]	-1.42 [-1.47]	-3.45 [-3.70]	-1.68 [-1.73]

注：该表汇报模型(17)~模型(20)中虚拟变量 $Dummy_t^H$ 的回归系数 $\varphi(\%)$ 及其 t 统计量 $([\cdot])$ 。

实证结果发现,质量溢价不受金融周期影响。

如表 8 所示,在上述 4 个回归模型中,由股市指数、利率和汇率得到的虚拟变量 $Dummy_t^H$ 的回归系数均不显著,表明股市上涨与否、利率调整、汇率浮动都与质量溢价无显著相关关系。在经济周期方面, $Dummy_t^H$ 的回归系数 φ 均显著为负,表明当 GDP 增速放缓、CPI 和 PPI 下降时,经济处于下行期,导致不同基本面表现的公司产生的错误定价被放大,增强公司质量的预测能力。经济下行期更显著的质量溢价既可以理解成对宏观风险的补偿,也可以理解成是投资者相较于经济上行反应不足的结果,与 Yan 等^[21] 和 Chen 等^[22] 研究发现一致。从表 8 中可以看出,由 CPI 得到的经济周期敏感度较高,下行期的综合质量溢价超出经济利好时约 3.1%,而对于 GDP 增长率和 PPI 而言,仅超出不足 2.1%。根据李拉亚^[23] 的分析, CPI 下降即通货紧缩时,厂商获得的流动性会下降,因此会选择缩小生产规模、减少支出,使得经济下行, PPI 增长动力不足,导致经济进一步下滑。因此,可以认为 CPI 是 PPI 和 GDP 的先行指标,所以由 CPI 判定的经济周期对综合质量溢价的影响更大。综合上述分析,公司基本面综合质量的预测能力免疫于宏观金融市场的系统性风险,但在不同的宏观经济周期中预测能力差异显著。

5 进一步分析

5.1 其他公司综合质量指标

首先将对比其他公司质量测度与 Q^{PLS} 测度的预测能力。具体包括: Asness 等^[1] 构建的质量测度 Q^{AFP} , 胡熠等^[6] 构建的综合指标 Q_B^{HG} 和质量指标 Q_Q^{HG} , 尹力博等^[7] 构建的质量测度 Q^{YL} , 这些指标均使用标准化法, 分别应用 16 个、8 个、3 个、14 个基本面指标进行构建。单变量投资组合分析结果如表 9 所示, 由 Q^{AFP} 、 Q_Q^{HG} 、 Q^{YL} 构建的对冲组合均能获得显著的正收益, 表明我国股票市场上存在质量溢价, 但都低于表 2 中 Q^{NA} 的结果, 说明基本面指标集的扩大能为横截面股票收益预测带来更多的增量信息。另一方面, Q^{PLS} 对冲组合收益率显著高于上述四个替换指标, 再次说明偏最小二乘法构建的综合质量测度表现更好。本文同时考虑将基本面指标库的容量从 94 个扩大至 139 个^⑩, 对比偏最小二乘法与标准化法构建综合质量测度 Q_{EX}^{PLS} 和 Q_{EX}^{NA} 的预测能力。结果表明, 通过买入 Q_{EX}^{PLS} 的对冲组合可以获得更高的收益。因此, 偏最小二乘法在公司综合质量指标构建上更具优势。

⑨ 胡熠等^[6] 从 3 个维度构建基本面综合指标 B-score, 为保持符号一致性, 文中将 B-score 表示为 Q_B^{HG} , 将质量维度的综合测度表示为 Q_Q^{HG} 。

⑩ 扩展后的指标库可向作者邮件索要。

表 9 其他公司质量测度的稳健性检验结果
Table 9 Robustness tests with alternative quality predictors

公司质量指标	Ret			CAPM α	FF3 α	FF5 α
	L	H	H - L	H - L	H - L	H - L
Q^{AFP}	1.56 [1.75]	2.66 [3.34]	1.10 [2.91]	1.26 [3.47]	0.70 [1.66]	1.11 [3.36]
Q_B^{HG}	1.87 [1.99]	2.10 [2.89]	0.23 [0.61]	0.46 [1.28]	-0.30 [-0.91]	0.59 [2.18]
Q_Q^{HG}	1.72 [1.97]	2.31 [2.97]	0.58 [2.49]	0.70 [3.14]	0.35 [1.44]	0.83 [3.80]
Q^{YL}	1.57 [1.70]	2.42 [3.32]	0.84 [2.51]	1.06 [3.46]	0.33 [1.23]	1.13 [4.60]
Q_{EX}^{PLS}	1.10 [1.34]	4.11 [4.46]	3.01 [6.38]	2.91 [6.18]	2.85 [5.12]	2.65 [5.26]
Q_{EX}^{NA}	0.74 [0.87]	3.34 [4.09]	2.61 [7.66]	2.71 [8.08]	1.77 [4.83]	2.27 [8.28]

注：该表汇报由其他综合质量指数构建的投资组合的月度平均原始收益率(Ret , %),对冲投资组合的超额收益率($CAPM \alpha$ 、 $FF3 \alpha$ 和 $FF5 \alpha$, %)及其 t 统计量([·]).“L”(“H”)指具有最低(最高)质量的组合,“H-L”指多空对冲投资组合。

5.2 质量投资策略与指数表现

本文还将对比 Q^{PLS} 对冲策略与常见指数的表现,具体包括:上证 A 股指数,上证 50 指数,上证全指价值指数,上证基本面 200 指数、300 指数、500 指数,沪深 300 指数,深证 700 价值指数。如图 1 所示,上述 8 个指数的累积收益率趋

势基本一致,其中沪深 300 指数的累积月度收益率最高,最高可达近 5 倍收益。因此,本文将进一步对比 Q^{PLS} 对冲策略与沪深 300 指数的表现。图 2 表明,基于公司质量的投资策略远远赶超大盘,说明本文的研究结果亦可用于实际长期投资。

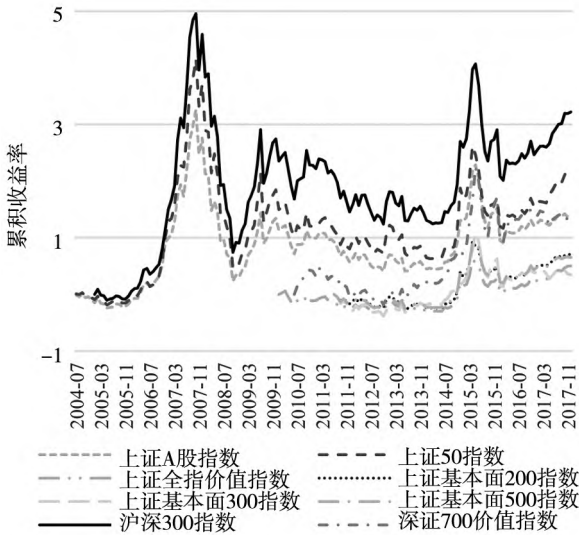


图 1 股票指数累积收益率①
Fig. 1 Accumulative returns of Chinese indexes

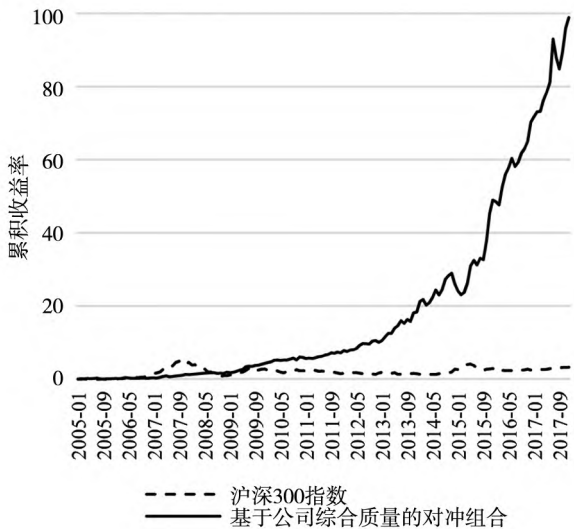


图 2 公司综合质量策略与沪深 300 指数
Fig. 2 Performance of aggregate firm quality and CSI 300 Index

① 由于数据可得性,各指数样本起始时间如下:上证 A 股指数、上证 50 指数为 2 004.07,上证全指价值指数为 2 010.01,上证基本面 200、300、500 指数为 2 011.07,沪深 300 指数为 2 005.01,深证 700 价值指数为 2 010.06。

5.3 极端市场下质量投资策略表现

如图 1 所示,2007 年—2008 年、2015 年—2016 年我国股市有两次较大的动荡,分别是全球金融危机和国内股市暴涨暴跌时期.根据胡熠等^[6]、尹力博等^[7]的研究,分别采用 2007 年 1

月—2008 年 12 月、2015 年 1 月—2016 年 12 月为样本,分析 Q^{PLS} 的预测能力.如表 10 所示,在两次重大市场动荡时期,基于公司综合质量的投资策略仍然可以获得显著的收益,再次证明了质量投资策略收益的稳健性.

表 10 极端市场下质量投资策略的表现

Table 10 Performance of firm aggregate quality under market turbulence

事件	短样本区间	<i>Ret</i>			<i>CAPM</i> α	<i>FF3</i> α	<i>FF5</i> α
		L	H	H - L	H - L	H - L	H - L
全球金融危机	2007 - 01—2008 - 12	1.86 [0.56]	5.47 [1.40]	3.60 [2.40]	3.58 [2.42]	3.26 [1.82]	3.44 [2.19]
国内股灾	2015 - 01—2016 - 12	2.17 [0.88]	6.50 [2.02]	4.33 [3.12]	4.19 [3.24]	2.09 [1.20]	0.58 [0.47]

注:该表汇报 2008 年全球金融危机和 2015 年国内股灾期间,由综合质量指数构建的投资组合的月度平均原始收益率(*Ret*, %),对冲投资组合的超额收益率(*CAPM* α 、*FF3* α 和 *FF5* α , %)及其 *t* 统计量([·]).

6 结束语

本文从盈利能力、成长能力、营运策略、潜在价值和安全性五个方面的 94 个公司基本面指标出发,分别应用主成分分析法、珐玛麦克白回归法、预测组合法、缩放主成分分析法、标准化法和偏最小二乘法构建基本面综合质量指数.结果发现,基于偏最小二乘的综合质量指标与股价表现关联性最强.相较于绝大多数单一基本面指标,投资者通过该综合质量指标构建

投资策略,获得的收益更大、更稳定.该投资策略在控制公司账面市值比后仍然显著,但属于小盘股投资策略.据此通过检验经济机制发现,有限套利、公司决策、市场情绪和经济周期,有助于理解质量溢价,但基本面质量的预测能力不受金融市场的周期性影响.在进一步分析中发现,由偏最小二乘构建的综合质量测度显著优于以往文献中的综合质量指标,质量投资策略的盈利性高于市场指数,同时在极端市场条件下也可获得显著的收益.综上,我国股票市场已进入优胜劣汰的高质量发展阶段.

参 考 文 献:

- [1] Asness C S, Frazzini A, Pedersen L H. Quality minus junk[J]. *Review of Accounting Studies*, 2019, 24(1): 34 - 112.
- [2] 蒋志强, 田婧雯, 周炜星. 中国股票市场收益率的可预测性研究[J]. *管理科学学报*, 2019, 22(4): 92 - 109.
Jiang Zhiqiang, Tian Jingwen, Zhou Weixing. Return predictability in the Chinese stock markets[J]. *Journal of Management Sciences in China*, 2019, 22(4): 92 - 109. (in Chinese)
- [3] Hou K, Mo H, Xue C, et al. An augmented Q-factor model with expected growth[J]. *Review of Finance*, 2021, 25(1): 1 - 41.
- [4] Jiang F, Tang G, Zhou G. Firm characteristics and Chinese stocks[J]. *Journal of Management Science and Engineering*, 2018, 3(4): 259 - 283.
- [5] Huang D, Jiang F, Li K, et al. Scaled PCA: A new approach to dimension reduction[J]. *Management Science*, 2022, 68(3): 1678 - 1695.
- [6] 胡 熠, 顾 明. 巴菲特的阿尔法: 来自中国股票市场的实证研究[J]. *管理世界*, 2018, 34(8): 41 - 54, 191.
Hu Yi, Gu Ming. Buffett's Alpha: Evidence from China stock market[J]. *Management World*, 2018, 34(8): 41 - 54,

191. (in Chinese)
- [7] 尹力博, 廖辉毅. 中国 A 股市场存在品质溢价吗? [J]. 金融研究, 2019, (10): 170–187.
Yin Libo, Liao Huiyi. Is there a quality premium in the Chinese A-share market? [J]. Journal of Financial Research, 2019, (10): 170–187. (in Chinese)
- [8] Jiang F, Qi X, Tang G. Q-theory, mispricing, and profitability premium: Evidence from China [J]. Journal of Banking & Finance, 2018, (87): 135–149.
- [9] 尹力博, 聂婧. “现金牛”比“铁公鸡”更有价值吗? [J]. 管理科学学报, 2021, 24(10): 58–81.
Yin Libo, Nie Jing. Are “cash cow” more valuable than “penny pinchers”? [J]. Journal of Management Sciences in China, 2021, 24(10): 58–81. (in Chinese)
- [10] 赵胜民, 刘笑天. 特质风险、投资者偏好与股票收益——基于前景理论视角的分析[J]. 管理科学学报, 2020, 23(3): 100–115.
Zhao Shengmin, Liu Xiaotian. Idiosyncratic volatility, investor preference and stock returns: An analysis based on prospect theory[J]. Journal of Management Sciences in China, 2020, 23(3): 100–115. (in Chinese)
- [11] 姜富伟, 马甜, 张宏伟. 高风险低收益? 基于机器学习的动态 CAPM 模型解释[J]. 管理科学学报, 2021, 24(1): 109–126.
Jiang Fuwei, Ma Tian, Zhang Hongwei. High risk low return? Explanation from machine learning based conditional CAPM model[J]. Journal of Management Sciences in China, 2021, 24(1): 109–126. (in Chinese)
- [12] Baker M, Wurgler J. Investor sentiment and the cross-section of stock returns [J]. Journal of Finance, 2006, 61(4): 1645–1680.
- [13] Stambaugh R F, Yu J, Yuan Y. The short of it: Investor sentiment and anomalies [J]. Journal of Financial Economics, 2012, 104(2): 288–302.
- [14] Cheema M A, Nartea G V, Man Y. Maxing out in China: Optimism or attention? [J]. International Review of Finance, 2020, 20(4): 961–971.
- [15] 汪昌云, 武佳薇. 媒体语气、投资者情绪与 IPO 定价[J]. 金融研究, 2015, (9): 174–189.
Wang Changyun, Wu Jiawei. Media tone, investor sentiment and IPO pricing [J]. Journal of Financial Research, 2015, (9): 174–189. (in Chinese)
- [16] Fama E F, French K R. A five-factor asset pricing model [J]. Journal of Financial Economics, 2015, 116(1): 1–22.
- [17] Liu J, Stambaugh R F, Yuan Y. Size and value in China [J]. Journal of Financial Economics, 2019, 134(1): 48–69.
- [18] Amihud Y. Illiquidity and stock returns: Cross-section and time-series effects [J]. Journal of Financial Markets, 2002, 5(1): 31–56.
- [19] Light N, Maslov D, Rytchkov O. Aggregation of information about the cross section of stock returns: A latent variable approach [J]. Review of Financial Studies, 2017, 30(4): 1339–1381.
- [20] 马甜, 姜富伟, 唐国豪. 深度学习与中国股票市场因子投资——基于生成式对抗网络方法 [J]. 经济学(季刊), 2022, 22(3): 819–842.
Ma Tian, Jiang Fuwei, Tang Guohao. Deep learning and factor investing in Chinese stock market: Based on generative adversarial networks [J]. China Economic (Quarterly), 2022, 22(3): 819–842. (in Chinese)
- [21] Yan X S, Zheng L. Fundamental analysis and the cross-section of stock returns: A data-mining approach [J]. Review of Financial Studies, 2017, 30(4): 1382–1423.
- [22] Chen Y, Da Z, Huang D. Short selling efficiency [J]. Journal of Financial Economics, 2022, 145(2): 387–408.
- [23] 李拉亚. 通货紧缩测度指标及数据变化研究 [J]. 经济学动态, 2016, (5): 74–86.
Li Laya. Research on deflation index and change of data [J]. Economic Perspectives, 2016, (5): 74–86. (in Chinese)

Survival of the fittest or fattest: The relationship between quality of listed firms and performance of stock price

*JIN Fu-jing*¹, *JIANG Fu-wei*², *TANG Guo-hao*^{3, 4*}

1. School of Economics and Management, Beijing Jiaotong University, Beijing 100044, China;
2. Center for Macroeconomic Research/School of Economics/Wang Yanan Institute for Studies in Economics, Xiamen University, Xiamen 361000, China;
3. College of Finance and Statistics, Hunan University, Changsha 410006, China;
4. Key Laboratory of High-Performance Distributed Ledger Technology and Digital Finance (Hunan University), Ministry of Education, Changsha 410082, China

Abstract: Building a high-quality capital market with survival of the fittest is the key for finance to serve the real economy and the supply-side reform. This study extracts fundamental information from a high-dimensional big dataset of over 90 accounting indicators by means of several approaches in machine learning and econometrics, evaluates the quality of listed firms with an aggregate fundamental index, and investigates the correlation between the index and the performance of stock price. Results show that the quality index can significantly and positively predict stock returns. Among them, the partial-least-squares-based measurement has the strongest predicting power, with an annualized return of approximately 38%, which cannot be explained away by CAPM, three-factor or five factor models. Furthermore, this paper explores the impact mechanism from the perspective of behavioral finance and macroeconomic cycle, and finds that market sentiment, limits to arbitrage, firm policy, and macro business cycle can help dissect the quality premium among listed firms. Our study suggests that the pricing efficiency in Chinese stock market has steadily improved, and it has stepped into the stage of survival of the fittest and value investment.

Key words: fundamental quality; stock returns; textual sentiment; limits to arbitrage; big data and machine learning