

《中国资本市场中的创新质量溢价：风险补偿还是错误定价》

附录

附录 1 创新质量各维度指标及定义

附表 1 创新质量各维度指标及定义

变量名称	变量含义	变量定义
<i>Citations</i>	前向引用数	引用别的专利个数，衡量借鉴知识范围
<i>Inventors</i>	申请人经验	该专利申请年份-该专利申请人第一次申请专利的年份，衡量知识经验
<i>Claims</i>	权利要求项数	专利申请中对于发明的特征和技术范围的具体描述的数量
<i>Words-Claims</i>	权利要求项字数	专利权利要求项总字数，代表界定的法律范围越详细
<i>Abstract</i>	摘要字数	对专利申请中的发明内容、技术特征和实施方式进行简明扼要的描述，代表技术含量的高低
<i>Classification-Number</i>	IPC 分类号数	标识专利文件所涉及的技术领域，分类号越多代表可被应用的技术范围越广
<i>Cite-Mean</i>	前向引用平均数	平均每个专利引用数，由审查员要求添加，作为审查质量的衡量指标
<i>Generality</i>	通用性	衡量一项专利被后续专利引用的技术范围的平均值，衡量应用的广度
<i>Citation After Grant</i>	后向引用数	专利公开后截止到收益率时间的被引次数，衡量后续的知识传播和影响力

注：本文的核心解释变量“创新质量”通过专利文本与引证信息两类方法，围绕四个关键维度构建而成，具体包括：技术质量、申请文件撰写质量、审查质量与应用质量。在此基础上，本文综合设计九项指标，系统衡量企业创新的深度与价值。上表给出了各指标的具体定义。

附录 2 非监督的线性降维算法详解

本文采用非监督的线性降维方法提取创新质量指标，主要包括主成分分析（PCA）和稀疏主成分分析（SPCA）。该方法基于一个核心假设：高质量专利在特定特征组合上呈现出规律性聚集，如后向引用数、权利要求数量等指标往往协同上升，而低质量专利则偏离这种组合趋势。例如，PCA 通过正交变换提取方差最大的主成分，将高维特征映射到低维空间，使质量相近的专利在降维后更容易聚类，从而体现高质量专利特征的稳定性。

线性降维的优势在于整合全局信息与增强可解释性。PCA 通过协方差矩阵分解，提取能代表技术复杂性与技术影响力的主成分，既压缩冗余信息，也提炼出具有经济含义的核心维度。例如，第一主成分可能综合权利要求数和后向引用量，第二主成分则与普遍性指数和前向引用相关。SPCA 在此基础上引入正则化约束，进一步筛选对质量影响显著的特征，剔除次要变量，突出关键维度，从而提升质量评估的精度与解释力。

（一）主成分分析（PCA）

主成分分析（PCA）是一种线性降维方法，通过将数据投影到低维空间，保留数据的主要方差信息，从而实现数据的降维和特征提取。

对于给定的数据集 $\{x_1, x_2, \dots, x_N\}$ ，首先计算数据的协方差矩阵 C ，其元素 C_{ij} 表示样本点 x_i 和 x_j 之间的协方差。协方差矩阵的计算方式如下：

$$C = \frac{1}{N} \sum_{i=1}^N (x_i - \mu)(x_i - \mu)^T \quad (A1)$$

其中， μ 是数据的均值向量。对协方差矩阵 C 进行特征值分解，得到特征值 $\{\lambda_1, \lambda_2, \dots, \lambda_N\}$ 和对应的特征向量 $\{v_1, v_2, \dots, v_N\}$ 。根据特征值的大小选择前 k 个最大的特征值及其对应的特征向量作为主成分。

将原始数据通过选取的主成分进行投影，从高维空间映射到低维空间。假设我们选择了前 k 个特征向量 $\{v_1, v_2, \dots, v_k\}$ ，则降维后的数据可以表示为：

$$Y = [v_1, v_2, \dots, v_k]^T X \quad (A2)$$

其中， X 是原始数据矩阵。主成分分析能够在保持数据线性结构的基础上实现对高维数据的降维表示。

(二) 核主成分分析 (Kernel PCA)

稀疏主成分分析 (SPCA) 是主成分分析 (PCA) 的一种扩展，通过引入稀疏性约束，使得主成分的载荷向量具有稀疏性，从而更容易解释和提取关键特征。

对于给定的数据集 $\{x_1, x_2, \dots, x_N\}$ ，首先计算数据的协方差矩阵 C ，其元素 C_{ij} 表示样本点 x_i 和 x_j 之间的协方差。稀疏主成分分析的目标是求解以下优化问题：

$$\min_v \|X - Xvv^T\|_F^2 + \lambda \|v\|_1 \quad (A3)$$

其中， v 是主成分的载荷向量， λ 是稀疏性参数， $\|v\|_1$ 表示 v 的 $L1$ 范数，用于引入稀疏性。通过求解该优化问题，得到稀疏的主成分载荷向量 $\{v_1, v_2, \dots, v_k\}$ 。将原始数据通过选取的稀疏主成分进行投影，从高维空间映射到低维空间。假设我们选择了前 k 个稀疏特征向量 $\{v_1, v_2, \dots, v_k\}$ ，则降维后的数据为：

$$Y = [v_1, v_2, \dots, v_k]^T X \quad (A4)$$

稀疏主成分分析能够在保持数据线性结构的基础上，实现对高维数据的降维表示，并提取具有稀疏性的关键特征。

附录 3 稳健性检验

附表 2 按创新质量进行十分组的平均月度收益和风险调整后收益

Panel A: QualityPCA 构造下的单变量分组												
	等权	Low (L)	2	3	4	5	6	7	8	9	High (H)	H-L
原始	Alpha	0.009	0.008	0.01	0.008	0.01	0.011	0.012	0.012	0.013	0.013	0.004
	T 值	1.182	1.188	1.451	1.161	1.386	1.477	1.694	1.634	1.809	1.662	1.855
FF-3	Alpha	0.008	0.008	0.01	0.008	0.01	0.011	0.012	0.012	0.013	0.013	0.005
	T 值	1.269	1.264	1.527	1.274	1.526	1.574	1.785	1.669	1.873	1.782	2.348
FF-5	Alpha	0.013	0.014	0.016	0.014	0.015	0.016	0.017	0.017	0.019	0.019	0.006
	T 值	1.77	1.971	2.207	1.942	2.306	2.179	2.484	2.412	2.544	2.459	3.131
	加权	Low (L)	2	3	4	5	6	7	8	9	High (H)	H-L
原始	Alpha	0.006	0.004	0.008	0.008	0.005	0.003	0.012	0.008	0.009	0.013	0.007
	T 值	0.891	0.743	1.27	1.053	0.848	0.824	2.006	1.446	1.589	2.06	2.342
FF-3	Alpha	0.006	0.005	0.008	0.008	0.006	0.004	0.012	0.009	0.009	0.015	0.009
	T 值	1.097	0.951	1.589	1.241	1.124	1.052	2.57	1.737	1.84	2.478	2.807
FF-5	Alpha	0.007	0.009	0.012	0.011	0.009	0.006	0.015	0.012	0.012	0.019	0.011
	T 值	1.288	1.481	2.143	1.764	1.686	1.457	2.806	2.353	2.044	2.946	3.389
Panel B: QualitySPCA 构造下的单变量分组												
	等权	Low (L)	2	3	4	5	6	7	8	9	High (H)	H-L
原始	Alpha	0.008	0.009	0.011	0.008	0.01	0.011	0.012	0.012	0.013	0.012	0.004
	T 值	1.147	1.258	1.454	1.125	1.359	1.529	1.696	1.637	1.821	1.624	1.774
FF-3	Alpha	0.008	0.009	0.011	0.008	0.009	0.011	0.012	0.012	0.014	0.013	0.005
	T 值	1.217	1.371	1.525	1.221	1.485	1.629	1.783	1.685	1.883	1.747	2.282
FF-5	Alpha	0.013	0.014	0.016	0.013	0.015	0.016	0.016	0.017	0.019	0.018	0.005
	T 值	1.739	2.038	2.195	1.851	2.337	2.248	2.445	2.464	2.554	2.413	2.906

	加权	Low (L)	2	3	4	5	6	7	8	9	High (H)	H-L
原始	Alpha	0.005	0.005	0.009	0.007	0.006	0.006	0.009	0.008	0.009	0.013	0.008
	T 值	0.804	0.82	1.37	0.981	1.025	1.181	1.713	1.331	1.705	1.986	2.486
FF-3	Alpha	0.005	0.006	0.009	0.007	0.007	0.007	0.01	0.008	0.01	0.014	0.009
	T 值	0.982	1.075	1.638	1.213	1.333	1.395	2.313	1.575	2.006	2.41	2.999
FF-5	Alpha	0.007	0.009	0.013	0.01	0.011	0.008	0.013	0.011	0.013	0.018	0.011
	T 值	1.233	1.619	2.134	1.694	1.995	1.625	2.584	2.086	2.236	2.901	3.474

附表 3 基于过去一年企业层面专利信息构造的创新质量溢价

	QualityPCA	QualitySPCA	QualityPLS-3	QualityPLS-5
Quality	0.0078	0.0076	0.0198	0.019
	2.282	2.227	4.699	5.287
控制变量	Yes	Yes	Yes	Yes
Intercept	0	0	0	0
	-1.019	-1.02	-1.017	-1.01
R	0.103	0.103	0.105	0.104
mean N	152496	152496	152496	152496

附表 4 加入创新效率和创新投入后的创新质量溢价

	QualityPCA	QualitySPCA		
Quality	0.0156	0.0158	0.0153	0.0155
	2.801	2.999	2.76	2.987
Quantity	0.0107		0.0107	
	2.669		2.677	
InnovationEff		0.0148		0.0148
		2.798		2.799
InnovationInv		0.0029		0.003
		0.334		0.338
控制变量	Yes	Yes	Yes	Yes
Intercept	0	0	0	0
	-1.009	-1.002	-1.009	-1.003
R	0.112	0.116	0.119	0.119
mean N	128880	126144	128880	126144

附录 4 错误定价检验的指标定义

附表 5 错误定价检验的指标定义

驱动因素	变量	变量名称	构建说明
有限注意力	<i>Grant-Num</i>	专利授权数	特定时间段内（如月/季/年）企业获得的专利授权总数
	<i>Sustainability</i>	创新质量的持续	通过计算企业过去一段时间其专利后向引用次数的标准差衡量
套利限制	<i>Ivol</i>	流动性	采用 Amihud（2002）指标
	<i>Turnm</i>	季度换手率	采用 Datar et al.,（1998）指标
信息不对称与透明度不足	<i>D1</i>	价格滞后指标 <i>D1</i>	参考（Hou & Moskowitz, 2005; Boehmer & Wu, 2013），通过比较包含与不包含滞后市场收益的回归模型决定系数差异计算得出
	<i>D2</i>	价格滞后指标 <i>D2</i>	通过计算个股收益回归模型中滞后市场收益系数绝对值之和占所有（同期+滞后）市场收益系数绝对值之和的比例得出
市场情绪	<i>Em</i>	操控性应计项目	采用张庆和朱迪辛（2014）的行为代理变量
	<i>At</i>	月度异常换手率	计算当期换手率与过去 12 个月移动平均值的标准化残差

参考文献

- [1] 张庆和朱迪星, 2021, 《投资者情绪、管理层持股与企业实际投资——来自中国上市公司的经验证据》, 《管理世界》第 3 期, 第 130~145 页。
- [2] Amihud, Y., 2002. "Illiquidity and Stock Returns: Cross-Section and Time-Series Effects," *Journal of Financial Markets*, 5(1), pp.31~56.
- [3] Datar, V. T., N. Y. Naik and R. Radcliffe, 1998, "Liquidity and stock returns: An alternative

test,” *Journal of Financial Markets*, 1(2), pp.203~219.

[4] Hou, K. and T. J. Moskowitz, 2005, “Market Frictions, Price Delay, and the Cross-section of Expected Returns,” *Review of Financial Studies*, 18(3), pp.981~1020.

[5] Boehmer, E. and J. Wu, 2013, “Short Selling and the Price Discovery Process,” *Review of Financial Studies*, 26(2), pp.287~322.