

对审稿意见的修改说明

注: 由于文中公式的输入使用插件, 可能会造成 Word 版本出现乱码的现象,

我们将论文以及此修改说明输出 PDF 格式, 以便阅读。

1. 论文应该是应用创新, 并没有机器学习方法的创新, 用的还是比较经典的自编码器方法, 因此不要在论文中强调自己提出了新的自编码器或者改进了。

回复:

非常感谢您的宝贵建议! 根据您的建议, 我们对本文中的用词进行了规范, 删去了“改进的”等字眼。原文的本意是因为模型不是单纯的使用自编码模型, 后面加入了 logistic 回归, 所以使用了“改进的”一词, 表达不当, 引起了不必要的误解。

2. 论文引言部分的研究动机写得不够明确, 尤其是为何需要用自编码器来提炼股票市场的因子呢?

回复:

非常感谢您的宝贵建议! 根据您的建议, 我们本文引言部分进行了重新的修改, 调整了行文逻辑, 着重突出研究动机, 同时文献综述部分放在了引言的最后, 见第 2 页起标蓝处, 同时摘录如下:

“股票市场的实证资产定价模型在近几十年来已取得了长足的发展, 国内外学者到目前为止已发现了四百多个能预测股票收益的市场异常指标 (Hou 等^[1]), 即能够提供超额收益的独特因子。Cochrane^[2]在美国金融学年会的主旨演讲中, 将这些海量定价指标称为因子动物园 (Factor Zoo)。这些定价指标中可能饱含着同一类定价信息, 并且单个指标的预测能力随时间逐渐减弱。McLean 和 Pontiff^[3]发现, 揭示因子的以往论文在发表后超额收益相对之前下降了三分之二。因此, 只依靠发现单个定价指标来进行资产价格预测和解释难以满足目前研究和实际需要。

另一方面, 依赖传统方法进行多指标金融大数据的资产定价研究也遇到了不小的困难。第一, Bali 等^[4]指出, 由于变量过多, 若用传统的回归方法对数百个股票因子同时进行回归, 将会带来“维数灾难”。同时, 传统的组合分析 (Portfolio Analysis) 和 Fama-MacBeth 回归并未综合考虑因子间的交互作用; 第二, 无论是传统资本资产定价模型 (Capital Asset Pricing Model, CAPM) 还是以 Fama-French 模型为代表的多因子模型, 其主要目的在于找到不同风险溢价下对应的资产收益, 而受制于市场时变性和信息获取难度等问题, 资产风险往往难以用唯一的模型有效度量。Cochrane^[2]认为在存在大量噪声和高度相关的预测因子的金融市场, 研究人员应该使用新颖的计量经济学方法, 使用综合方法进行信息提取。

为了解决上述问题所面临的挑战, 本文尝试采用基于自编码机器学习方法从公司特征中提取共同信息, 构建定价和预测模型来解释截面收益。对于提取因子中的共同信息, Kozak 等^[5]和 Kelly 等^[6]分别运用主成分分析 (Principal Components Analysis, PCA) 和增量主成分分析 (Instrumented Principal Components Analysis, IPCA) 方法来进行提取, 发现基于只依赖于少数主成分的模型就可以较好地解释和预测截面收益。然而, PCA 只能通过对指标进行线性变化来获得静态线性因子, 但越来越多的研究发现资产收益和指标之间还存在非线性关系 (Campbell 和 Cochrane^[7], Bansal 和 Yaron^[8]以及 He 和 Krishnamurthy^[9]等)。Pohl^[10]等表明, 当存在非线性关系时, 线性近似会导致预测估计的巨大误差。因此, 在提取多指标的公共信息时, 我们需要构造非线性动态模型来进行非线性降维变换, 从而更好地提取金融数据中的非线性信息。

.....

本文的研究有一定的现实意义和理论贡献。近年来，金融学开始探讨机器学习对现有研究范式的价值和意义，美国市场的大量研究尝试使用包括神经网络机器学习方法在内的大数据方法构建预测和定价模型，来解决因子大量涌现而给传统研究方法带来的挑战。例如 Light 等[13]采用“偏最小二乘”(Partial Least Squares, PLS)来检验公司特征对预期收益的预测能力；Gu 等[11]对比了包括主成分分析、偏最小二乘法、LASSO、随机森林和神经网络在内的各类机器学习模型在美国股票市场的应用，结果发现非线性的机器学习模型的预测表现可以有效地超越传统线性模型，同时浅层网络模型要优于深度学习模型；Feng 等[14]在进行股价预测时引入无套利约束并结合了深度学习网络模型。

不少研究也着眼于中国市场。洪永淼和汪寿阳[15]认为大数据和机器学习对经济学的研究范式和研究方法带来了机遇和挑战；苏治等[16]对近年来在金融实证里应用的深度学习模型进行了总结归纳，并对未来中国金融市场中机器学习的应用给出了评述；陈卫华等[17]基于深度学习和股票论坛数据对股市波动率进行了预测，发现深度学习模型显著提升了预测精度；李斌等[18]基于96个异象因子，运用12种机器学习算法，构建股票收益预测模型及投资组合，发现非线性方法能获得较高的投资绩效。本文紧跟前沿，实证上从动态非线性的视角丰富了中国市场股票截面收益影响因素的研究，丰富了资产定价实证研究的机器学习工具，通过动态非线性机器学习算法探索横截面股票收益的决定因素，还尝试去探究自编码算法在收益可预测性方面的经济解释以及与宏观经济之间的关联。”

3. 论文对标准自编码器的描述有些是不正确的，不规范的，而且可读性较差，符号的使用，解释方面很多不规范。

回复：

再次感谢您的宝贵建议，根据您的建议，我们认真地对文中对标准自编码器的描述进行了梳理改正，对其解释进行规范，增强可读性。并结合下面的“建议4”做了部分删减，以达到缩减字数的目的。

修改见第4页起的标蓝处，同时摘录如下：

“标准自编码器可以用数学表达式表示。 N 是输入和输出层的神经元的个数， x_k 是输入层 N 个神经元的输入，以及输入向量为 $x = (x_1, x_2, \dots, x_N)^T$ 。 \tilde{x}_k 是输出层 N 个神经元的输出，以及输出向量为 $\tilde{x} = (\tilde{x}_1, \tilde{x}_2, \dots, \tilde{x}_N)^T$ 。

在压缩层的向量为 $z = (z_1, z_2, \dots, z_K)^T$ ， K 是在压缩层的神经元的个数，前一层的输入向量经过一个非线性的激活函数 $g(\cdot)$ 传递到输出层。初始化网络，输入层使用公司特征的横截面数据为原始输入 $x = (x_1, x_1, \dots, x_N)^T$ 。

自编码器的编码过程为：

$$z = g(b + Wx), \quad (1)$$

其中， W 是 $K \times N$ 维的权重矩阵， b 是一个被称作是偏差参数的 $K \times 1$ 维向量。本文使用线性整流函数（Rectified linear unit, ReLU）作为输入层的非线性激活函数。

自编码器的解码过程为：

$$\tilde{x} = \tilde{g}(W^T z + b^T), \quad (2)$$

其中，解码层的激活函数 $\tilde{g}(\cdot)$ 可以和编码层的不同，这里使用Sigmoid函数。Sigmoid函数是一种常见的S型函数，可以把变量映射到 $[0, 1]$ 之间。这里激活函数的选取与后续使用logistic回归模型有关。

对于自编码器，我们真正关心的是压缩层的编码，或者说是输入层到压缩层的映射。一般而言，压缩层神经元的个数 K 小于输入变量神经元的个数 N ，表示网络学习到输入样本的特征，试图以更小的维度去描述原始数据而尽量不损失数据信息，学习到关于输入的低维表达，以进行压缩表示。实际上，自主学习原始数据的特征表达是神经网络和深度学习的核心之一。本文就是将压缩层的信息作为后续机器学习算法的输入，这里采用的是 Logistic 回归模型。”

4. 论文对堆积自编码器的描述也不对，不规范，比如式子(8)通过惩罚项约束使得参数稀疏，惩罚项里的参数前面根本没定义，根据前文的描述，惩罚项里的参数应该是 w ，很多下标使用也不规范或者不对。

回复：

非常感谢您提出的宝贵建议，根据您的建议，我们对编码器的描述，不仅是公式，包括文字，都进行了认真地检查，并对不规范的地方进行了修改。例如，将输出变量矩阵由 $X^{(l)} = (X_1^{(l)}, X_1^{(l)}, \dots, X_{K^{(l)}}^{(l)})^T$ 修改为 $x^{(l)} = (x_1^{(l)}, x_1^{(l)}, \dots, x_{K^{(l)}}^{(l)})^T$ ，对(3)式增加了代表层数的上标，对向量 $b^{(l-1)}$ 的结构进行了描述，将(8)式惩罚项的参数修改为 W ，等等。更多修改见第5页起的标蓝处，同时摘录如下：

“ L 个隐藏层的自编码器可以用递推公式表示。设 $K^{(l)}$ 是第 l 层神经元的数量， $l=1,2,\dots,L$ ， $x_k^{(l)}$ 是 l 层 k 个神经元的输出，则这一层所有的输出变量组成的矩阵为 $x^{(l)} = (x_1^{(l)}, x_1^{(l)}, \dots, x_{K^{(l)}}^{(l)})^T$ 。在每一个隐藏层中，前一层的输入在传入下一层前经过非线性激活函数 $g(\cdot)$ 。递归输出公式为

$$x^{(l)} = g(b^{(l-1)} + W^{(l-1)}x^{(l-1)}), \quad (4)$$

在每一隐藏层 l 中参数的数量为 $K^{(l)}(1 + K^{(l-1)})$ ， $b^{(l-1)}$ 是一个 $K^{(l)} \times 1$ 维的列向量。自编码器的最终的输出为

$$G(x, b, W) = b^{(L)} + W^{(L)}x^{(L)}, \quad (5)$$

利用与输入相同的维数来逼近 x 。

非线性 x 函数的递推公式为：

$$x_{i,t-1}^{(0)} = x_{i,t-1}, \quad (6)$$

$$x_{i,t-1}^{(l)} = g(b^{(l-1)} + W^{(l-1)}x_{i,t-1}^{(l-1)}), l=1,2,\dots,L, \quad (7)$$

$$z_{i,t-1} = b^{(L)} + W^{(L)}x_{i,t-1}^{(L)}. \quad (8)$$

(5)式表达的是初始化网络模型的为第 i 家公司在 $t-1$ 期的公司特征数据 $x_{i,t-1}$ ，(6)式的方程描述了特征数据通过隐藏层神经元传播时的非线性(和交互的)转换，(7)式描述了一组 K 维因子如何从终端输出层产生。

为了防止过拟合，最常见的机器学习方法是在目标函数上附加一个惩罚项，以得到更简洁的设定。这种“正则化”方法减少模型的样本内性能，以期提高模型的样本外稳定性。惩罚项的设定减少模型的拟合噪声，防止过拟合。

设 $L(W; \cdot)$ 为优化自编码器的损失函数，用 W 来概括因子网络模型(5)式到(7)式到中的权重参数。将评估目标定义为：

$$L(W; \cdot) = \frac{1}{NT} \sum_{t=1}^T \sum_{i=1}^N \|x_{i,t} - \tilde{x}_{i,t}\|^2 + \phi(W; \cdot). \quad (9)$$

其中， $\phi(W; \cdot)$ 是对模型进行正则化的惩罚函数，称为正则化项。有很多惩罚函数 $\phi(\cdot)$ 可供选择，它们虽然使用了不同的正则化项，但最终都实现了约束参数从而防止过拟合的效果。这里用 LASSO 或 “ l_1 ” 惩罚，它采取的形式为

$$\phi(W; \lambda) = \lambda \sum_j |W_j|. \quad (10)$$

这里的参数 λ 就是正则化系数，可以取大于 0 的任意值。 λ 的值越大，对模型中参数的惩罚力度越大，就会有更多的参数被训练为 0。”

5. 式子 (12) 里应该用前一期的 $x_{i,t-1}$ 预测下一期的 $y_{i,t}$ ，而在式子 (13) 怎么就没有了时间 t 呢？很容易让人误解或者混淆，而且 p_i 应该就是式子 (11) 定义的 $P(y_{i,t}=1|Z_{i,t-1})$ 。

回复：

非常感谢您的宝贵意见！根据您的意见，我们对公式进行重新的梳理编写，但同时考虑到表达简洁、突出重点，我们将 (11) 式定义为 $P_{i,t}$ ，并对 (13) 式以及 (13) 式的解释说明进行了修改，见第 6 页标蓝处，同时摘录如下：

“对本文而言，Logistic 回归建立股票上涨概率模型

$$P_{i,t} \equiv P(y_{i,t} = 1 | z_{i,t-1}) = \frac{1}{1 + e^{-z^T(x_{i,t-1})\beta_t}}, \quad (11)$$

$$z^T(x_{i,t-1}) = x_{i,t-1}^T \Gamma. \quad (12)$$

其中 $x_{i,t-1}$ 是原始横截面公司特征， $z_{i,t-1}$ 是 $x_{i,t-1}$ 作为输入变量经过自编码器学习的压缩层的变量。

对于 t 时刻的模型，对数损失函数的计算公式如下：

$$L(y_{i,t}, P_{i,t}) = -\frac{1}{N} \sum_{i=1}^N (y_{i,t} \log p_{i,t} + (1 - y_{i,t}) \log(1 - p_{i,t})). \quad (13)$$

其中， $y_{i,t}$ 是所属的每个类别的概率值， $z_{i,t}$ 是逻辑回归模型的输入变量， N 为输入样本量， $p_{i,t}$ 为预测输入实例 $z_{i,t-1}$ 属于类别 1 的概率。对所有样本的对数损失表示对每个样本的对数损失的平均值，对于完美的分类器，对数损失为 0。”

6. 论文分析结果的展示也令人眼花缭乱，没有抓重点。比如表 5 那个基本上看不清楚了。

回复：

再次感谢您的建议，根据您的建议，我们将表 3 和表 4 进行了重新布局，抓住主要信息，删减和结论无关的结果，并用图和表相结合的方式展示，使结果尽量清晰。针对表 5，我们将逐月的权重绝对数 6 个月进行加总平均，得到新的表 5。除此之外，参考下面的“建议 4”，仅用颜色代表类别，将文字去掉，并将此表附于附件 1。

修改见 11 页起的表格，同时摘录如下：

表 1 不同编码器结构的横截面股票收益（滚动窗口为 12 个月）

自编码器结构	AE(32)			AE(64)			AE(128)		
投资组合	Low	High	High-Low	Low	High	High-Low	Low	High	High-Low
月平均收益	-0.27	1.41	1.67	-0.26	1.42	1.68	-0.31	1.34	1.65
t 值	[-0.32]	[1.81]	[3.36]	[-0.31]	[1.84]	[3.36]	[-0.37]	[1.71]	[3.32]
夏普比率	-0.2	0.41	0.99	-0.2	0.42	0.99	-0.22	0.39	0.98
FF5- α	-0.81	0.23	1.04	-0.81	0.20	1.01	-0.84	0.10	0.94
t 值	[-2.62]	[0.78]	[2.16]	[-2.71]	[0.67]	[2.13]	[-2.89]	[0.33]	[2.00]
自编码器结构	AE(32,16)			AE(64,32)			AE(128,64)		
投资组合	Low	High	High-Low	Low	High	High-Low	Low	High	High-Low
月平均收益	-0.21	1.33	1.54	-0.17	1.29	1.46	-0.28	1.54	1.82
t 值	[-0.24]	[1.74]	[3.06]	[-0.20]	[1.68]	[2.87]	[-0.34]	[1.97]	[3.54]
夏普比率	-0.18	0.39	0.90	-0.17	0.37	0.85	-0.21	0.46	1.05
FF5- α	-0.75	0.16	0.91	-0.7	0.06	0.77	-0.79	0.32	1.12
t 值	[-2.54]	[0.55]	[1.90]	[-2.35]	[0.23]	[1.60]	[-2.63]	[1.09]	[2.29]
自编码器结构	AE(32,16,8)			AE(64,32,16)			AE(128,64,32)		
投资组合	Low	High	High-Low	Low	High	High-Low	Low	High	High-Low
月平均收益	-0.09	1.22	1.32	-0.05	1.41	1.47	-0.22	1.36	1.58
t 值	[-0.11]	[1.61]	[2.69]	[-0.06]	[1.8]	[2.96]	[-0.26]	[1.75]	[3.09]
夏普比率	-0.14	0.35	0.79	-0.13	0.41	0.87	-0.19	0.40	0.91
FF5- α	-0.66	0.13	0.79	-0.60	0.14	0.74	-0.72	0.14	0.86
t 值	[-2.34]	[0.46]	[1.78]	[-2.09]	[0.48]	[1.58]	[-2.43]	[0.47]	[1.80]

表 2 不同滚动窗口的自编码器的多空组合在横截面的股票收益

自编码器序号	1	2	3	4	5	6	7	8	9
6 个月滚动窗口									
月平均收益	1.13	1.08	1.08	0.93	1.11	1.12	0.6	1.03	1.18
t 值	[2.23]	[2.10]	[2.08]	[1.82]	[2.11]	[2.09]	[1.1]	[1.95]	[2.20]
夏普比率	0.66	0.62	0.61	0.54	0.62	0.62	0.33	0.57	0.65
24 个月滚动窗口									
月平均收益	1.25	1.30	1.43	1.25	1.35	1.33	1.47	1.29	1.30
t 值	[2.74]	[2.86]	[3.11]	[2.81]	[2.93]	[2.86]	[3.49]	[2.89]	[2.80]
夏普比率	0.81	0.84	0.92	0.83	0.86	0.84	1.03	0.85	0.83
36 个月滚动窗口									
月平均收益	0.96	1.06	1.11	1.15	1.08	1.02	0.91	0.95	0.94
t 值	[2.09]	[2.34]	[2.45]	[2.63]	[2.28]	[2.20]	[2.14]	[2.10]	[2.09]
夏普比率	0.62	0.69	0.72	0.78	0.67	0.65	0.63	0.62	0.62
48 个月滚动窗口									
月平均收益	0.95	1.07	1.12	0.86	1.00	1.16	1.13	0.66	1.26
t 值	[2.08]	[2.41]	[2.48]	[1.89]	[2.09]	[2.52]	[2.47]	[1.37]	[2.75]

夏普比率	0.61	0.71	0.73	0.56	0.62	0.74	0.73	0.41	0.81
60 个月滚动窗口									
月平均收益	0.98	1.14	1.17	0.86	1.01	1.13	0.67	0.80	1.01
t值	[2.04]	[2.49]	[2.53]	[1.77]	[2.12]	[2.41]	[1.45]	[1.64]	[2.16]
夏普比率	0.6	0.74	0.74	0.52	0.63	0.71	0.43	0.48	0.64

注：为简明起见，自编码结构用序号代替，自编码器序号所对应的自编码结构见表 1；12 个月滚动窗口的横截面收益如表 3 所示。

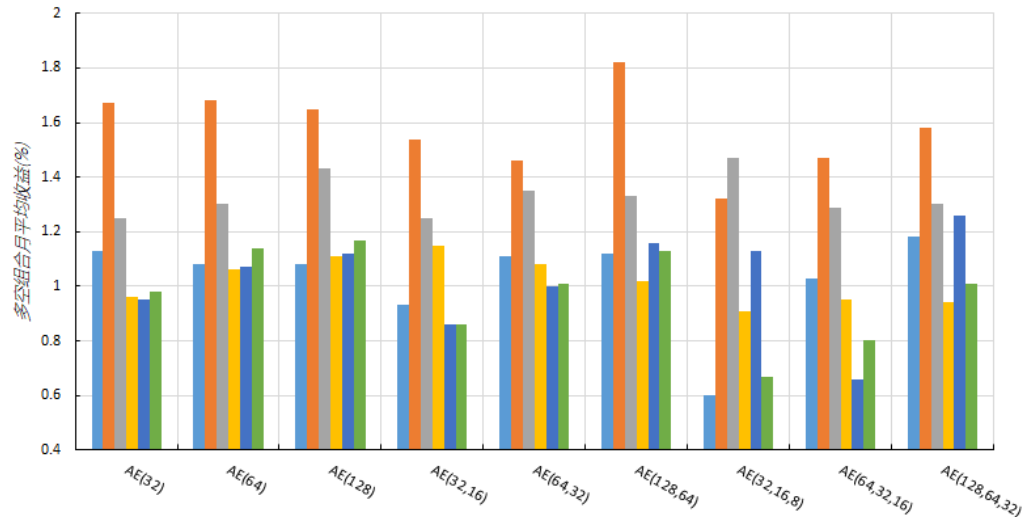


图 1 不同滚动窗口的自编码器的多空组合在横截面的股票收益

表 6 自编码器输出层因子重要度排序

year	1	2	3	4	5	6	7	8	9	10	
2001	MOM1M	PRCDEL	TAN	GM	CHMOM	PRC	ROA	FOE	ROE	MAXRET	<div>估值与成长类</div> <div>投资类</div> <div>盈利类</div> <div>惯性类</div> <div>交易摩擦类</div> <div>无形资产类</div>
	MOM1M	TAN	PRCDEL	ROE	FOE	EPS	MOM6M	CHMOM	PRC	MAXRET	
2002	ROE	EPS	PRCDEL	FOE	MOM6M	TAN	MOM1M	CHMOM	AGE	CHTX	
	INDMOM	FOE	EPS	MOM6M	CHMOM	AGE	PRCDEL	TAN	ROE	MAXRET	
2003	INDMOM	EPS	TAN	FOE	MAXRET	CHMOM	MOM6M	IVOL	AGE	PRCDEL	
	INDMOM	EPS	TAN	PRCDEL	CHMOM	FOE	MAXRET	MOM6M	AGE	MOM1M	
2004	INDMOM	EPS	PRCDEL	CHMOM	MOM6M	AGE	FOE	TAN	MOM1M	ROE	
	INDMOM	CHMOM	MOM6M	ROE	PRCDEL	AGE	MOM1M	EPS	FOE	FM	
2005	INDMOM	ROE	CHMOM	AGE	MOM1M	PRCDEL	FOE	dPIA	MOM6M	GM	
	GM	PRCDEL	DOLVOL	FOE	CHMOM	AGE	INDMOM	TBI	CFOA	MOM1M	
2006	PRCDEL	DOLVOL	GM	FOE	CFOA	AGE	TAN	CHMOM	IVC	BETASQ	
	PRCDEL	TAN	GM	AGE	DOLVOL	CFOA	CHMOM	FOE	CFP	EP	
2007	PRCDEL	GM	AGE	TAN	CHMOM	EPS	CFOA	CFD	INDMOM	EY	
	EPS	CHMOM	TAN	GM	EY	PRCDEL	ROE	TBI	IP	AGE	
2008	EPS	ROE	EY	PRCDEL	EP	GM	TAN	IVOL	CHMOM	RNA	
	PRCDEL	ROE	EPS	IVOL	EY	TAN	FOE	DOLVOL	CFOA	ROA	
2009	ROA	TAN	PRCDEL	EY	EPS	IVOL	CFOA	DOLVOL	ROE	CFD	
	ROA	TAN	EY	CFOA	PRCDEL	EPS	GM	DOLVOL	ROE	IVOL	
2010	ROA	PRCDEL	GM	EPS	EY	CFOA	TAN	ROE	dPIA	DOLVOL	
	PRCDEL	GM	EPS	dPIA	CFOA	ROA	TAN	EY	FM	FOE	
2011	CFOA	GM	TAN	PRCDEL	dPIA	EPS	FM	ROA	EY	CHMOM	
	CFOA	TAN	GM	FOE	dPIA	EY	PRCDEL	FM	ROA	DOLVOL	
2012	CFOA	GM	TAN	PRCDEL	FM	FOE	EY	IVC	DOLVOL	dPIA	
	FM	CFOA	PRCDEL	TAN	GM	dPIA	IVC	DOLVOL	OCFP	CAC	
2013	FM	PRCDEL	TAN	GP	GM	dPIA	DOLVOL	IVC	OCFP	CAC	
	FM	TAN	GP	PRCDEL	IVC	GM	DOLVOL	PRC	ROIC	CFP	
2014	FM	TG	TAN	IVC	GM	DOLVOL	CFP	EY	EP	EPS	
	TAN	GM	DOLVOL	CFP	CP	PRCDEL	EY	TG	INDMOM	FM	
2015	PRCDEL	TAN	CFP	DOLVOL	GM	CP	TBI	OCFP	CHMOM	INDMOM	
	PRCDEL	DOLVOL	TAN	CFP	CINVEST	TBI	IP	OCFP	IVC	CHMOM	
2016	PRCDEL	DOLVOL	TAN	BETASQ	GP	CFP	IP	EY	OCFP	PRC	
	TAN	PRCDEL	TG	DOLVOL	GP	CFP	OCFP	OCFP	EP	BETASQ	
2017	TAN	EY	DOLVOL	PRCDEL	EPS	OCFP	CFP	PRC	CHMOM	EP	
	TAN	EY	PRCDEL	SG	IVC	GM	DOLVOL	EPS	CHTX	CFP	
2018	EY	PRCDEL	TAN	TG	IVC	CFD	GM	EP	CHTX	EPS	
	EY	CFD	PRCDEL	TAN	SIZE	GM	TG	EP	IVC	CHTX	

注：表中的英文缩写所代表的公司特征指标详见论文附录 1 公司特征指标名称、定义以及文献出处。

7. 对于数据分析的结论分析和解释不够深入，对背后的机理理解得不够透彻。

回复：

感谢您的宝贵建议，我们对结论部分的解释和分析进行了修改和重新撰写。修改见第 10 页起标蓝处，同时摘录如下：

“分析表 2 的结果可知，各投资组合的月平均收益随着上市公司的预期上涨概率的增加而增加。具体而言，月度平均收益从预期收益最低组的-0.28% ($t=-0.34$) 增加至预期收益最高组的 1.54% ($t=1.97$)，预期收益最高与最低组的月度平均收益之差为 1.82%， t 值为 3.54，在 1% 水平上显著。这表明通过投资于按预期收益大小所构建的多空对冲组合，平均每年将获 21.84% 的收益。同时，夏普比率也存在着从低到高的单调递增规律，多空对冲组合的夏普比率更是高达 1.05 以上。此外，用 CAPM 模型、Fama-French 三因子模型、Fama-French 五因子模型衡量的超额收益都随着各组预期收益的增加而增加。多空对冲组合的月平均 CAPM- α 、FF3- α 、FF5- α 分别为 1.84 ($t=3.64$)、1.26 ($t=3.63$)、1.12% ($t=2.29$)，均在 5% 水平以上显著。这说明自编码方法能够显著预测中国股票市场横截面收益。

.....

接下来探讨基于 12 个月的滚动窗口，不同自编码器结构的横截面股票收益情况。根据表 3 的结果显示，包含一个隐藏层和两个隐藏层的自编码器能够获得较好的横截面收益预测结果，其多空对冲组合月收益最高达 1.82% ($t=3.54$)，夏普比率为 1.05，多种模型衡量的超额收益也在 5% 水平显著，特别的，标准（一层）自编码器随着压缩层节点的变化横截面收益结果显著且保持相对稳定，分别为 1.67% ($t=3.36$)、1.68 ($t=3.36$) 和 1.65 ($t=3.32$)。但是，随着神经网络法中隐藏层不断增加，模型复杂度逐渐上升，不仅没有提高模型的预测能力，反而多空对冲组合的月平均收益率有所下降。其中，三层自编码器月平均收益率分别为 1.32 ($t=2.69$)、1.47 ($t=2.96$) 和 1.58 ($t=3.09$)。这说明在使用神经网络法进行金融大数据预测时，并不是模型越复杂越好。与深度学习在如计算机视觉领域的蓬勃发展不同，这些领域往往拥有更海量的数据集。但金融市场中包含着各类公司信息和交易信号，这些信息中难免有许多与未来资产价格无关的噪音。当隐含层层数越多时，输入变量所包含的这些噪音一定程度上会影响最终输出结果的预测精度。未来的研究若能在神经网络中加入经济约束（依据不同的定价模型），更精确地衡量输入层变量之间的相互关系，或许能改善深层学习的预测结果。

.....

结合表 2 的结果可以发现，以 12 个月为训练样本得到的多空组合的月平均收益最高达到 1.81%，除此之外的其他结构也在 1.36% 到 1.68% 之间的范围，明显高于其他的滚动窗口。此结果从图 5 展示的更加清晰，12 个月为训练样本的结果（图中为橙色）的多空组合月平均收益最好，24 个月（灰色）次之，而更短或更长时间的训练样本都没有达到更好的效果，这与李斌等^[19]在运用机器学习算法时最优的滚动窗口长度一致。前文提到，金融市场中包含着各类公司信息和交易信号，这些信息中难免有许多与未来资产价格无关的噪音。训练集的长度设置的过大，反而会引入更多的噪音，从而影响预测结果。

”

建议：

1. 2.2 节，中国市场有其特色，采用美国市场的因子数据可能并不合适，请针对该问题做些讨论。

回复：

非常感谢您的宝贵建议！我们虽然主要是使用美国市场中发现在主流因子文献，整理后应用于中国市场，使用公司财务报表等公开数据（主要来自于 CSMAR 数据库）计算得来，但不仅限于国外文献，还参考了很多针对中国市场研究的主流文献，计算在中国市场中发现的因子指标。在此附上中国市场因子指标构建说明，里面包含具体的计算方法，详见附件 2。

2. 建议探讨下模型编码-解码后的经济含义，公司特征编码-解码后的东西是什么？

回复：

再次感谢您的宝贵建议！根据您的意见，我们输出每月每只股票的解码层信息，并和原始编码信息计算相关系数，结果发现相关系数接近于 1，全部在 0.99 左右，说明原始信息经过自编码网络先压缩后解码，输出的信息和原始信息几乎一致。

另外，由于解码层的信息过多，每一个公司都有 89 个数据（输出层有 89 个神经元），不方便处理，所以我们去探究了经过自编码网络和 logistic 回归后的预测值的经济含义。除了文中第 4 部分研究根据预测值构造的投资组合与宏观经济变量之间的联系外，我们还研究了预测值和各个因子指标的相关系数。

摘录如下：

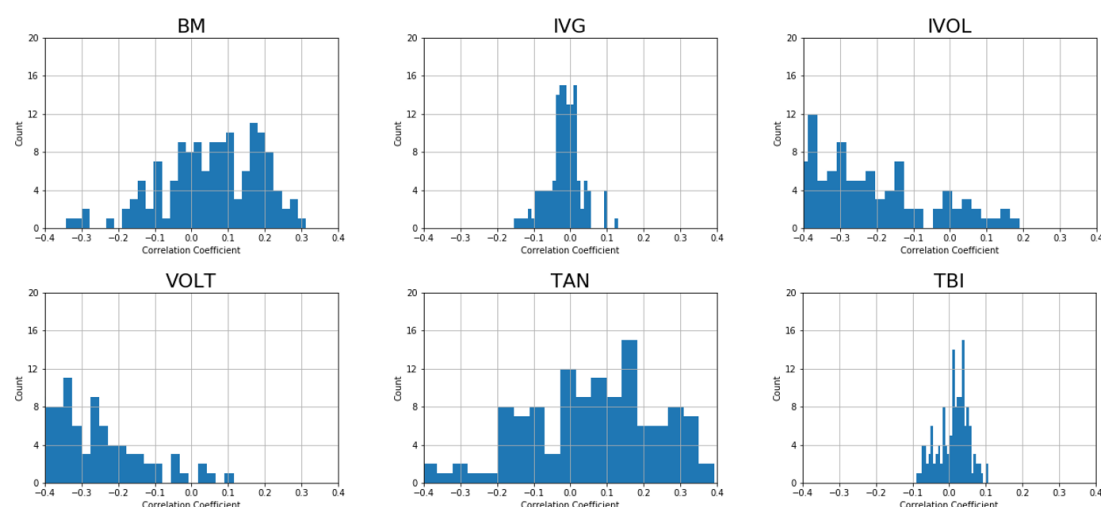


图 2 预测值与各因子指标的相关系数分布图

每个月，我们将指标经过自编码网络和 logistic 回归模型后得到的预测值与各因子指标计算相关系数，目的是探究预测值和个因子指标的相关关系及其经济含义。将所有月度相关系数的分布展示于图 1，图 1 的横坐标是相关系数的数值，纵坐标是出现的频数，标题是指标的缩写，图中选取了有代表性的指标，全部指标的相关系数详见附件 3。结果显示，预测值与异质性收益波动率 (IVOL) 和交易量趋势 (VOLT) 的相关系数分布集中在图像的左侧，数值为负，平均值分别为 -0.31 和 -0.35，可以认为预测值与异质性收益波动率 (IVOL) 和交易量趋势 (VOLT) 为负相关关系；与账面市值比 (BM) 和资产有形度 (TAN) 在大多数月的相关系数都为正数，数值集中在图像的右边，经计算平均值分别为 0.06 和 0.09，可以认为预测值与账面市值比 (BM) 和资产有形度 (TAN) 为正相关；而存货增长率 (IVG) 与应

税所得与账面资产比（TBI）的相关系数集中在 0 附近，可以认为预测值与这两个变量的相关性几乎为 0，存货增长率（IVG）与应税所得与账面资产比（TBI）这两个指标在预测中的作用不大。

3. 针对中国特色做一些实证研究，比如股改前后等。

非常感谢您的宝贵建议！非常抱歉，关于中国股票市场股改前后的研究，在我们的论文可能难以实现，原因如下：我们数据的原始样本是从 2000 年 1 月至 2018 年 12 月，但作为机器学习算法，需要一定量的训练样本和交叉验证样本，导致股改（2005 年）之前的数据非常少，训练的次数少会增加模型的不稳定性，得不到令人信服的结果。

但根据您的建议，我们仍针对中国特色做了一些实证研究，分为两部分。首先，我们将股票市场中的公司分行业分企业性质进行分类，进行 Single Sorts 构造多空投资组合，探究不同行业的投资组合月度收益和不同企业性质的投资组合月度收益差距。其次，我们选取了有代表性的因子指标，如公司规模（SIZE）、账面市值比（BM）等，基于自编码因子和这指标进行双变量排序（Double Sorts），研究投资组合的表现。再次感谢您的建设性建议！

摘录如下：

本文依据中国证监会《上市公司行业分类指引》中的二级行业分类代码，将所有公司按行业分类，并剔除可交易公司小于 50 家的行业，分行业进行投资组合的构建，探究各行业之间投资组合收益水平的不同。表 2 展示了在各个不同的行业根据股票上涨概率的预测值构建投资组合的度平均收益，分析可知，交通运输、仓储和邮政业（G）和房地产业（K）表现较好，多空组合的月平均收益分别 2.69%（ $t=4.37$ ）和 2.17%（ $t=3.32$ ），高于全要本平均值 1.82%。而信息传输、软件和信息技术服务业（I）表现最差，其次是建筑业（E），空组合的月平均收益分别 0.21%（ $t=0.25$ ）和 0.91%（ $t=1.02$ ）。另外，本文根据企业性质，对国营或国有控股、私营企业和中外合资企业进行分别研究，结果如表 3 所示，三类企业的多空对冲组合均能获得显著的月平均收益，其中中外合资企业收益最好，每月平均能达到 1.71%（ $t=2.04$ ）。

表 3 不同各行业横截面股票收益

投资组合	B	C	D	E	F	G	I	J	K
Low	-0.17	-0.15	-0.77	-0.06	-0.58	-0.64	0.68	-0.95	-0.90
	[-0.16]	[-0.18]	[-0.82]	[-0.05]	[-0.59]	[-0.75]	[0.57]	[-1.02]	[-0.94]
High	1.60	1.42	0.69	0.85	0.63	2.05	0.89	0.59	1.27
	[1.59]	[1.64]	[0.84]	[0.93]	[0.77]	[2.35]	[0.94]	[0.68]	[1.49]
High-Low	1.77	1.57	1.46	0.91	1.20	2.69	0.21	1.54	2.17
	[2.18]	[3.51]	[2.40]	[1.02]	[2.00]	[4.37]	[0.25]	[2.07]	[3.32]

注：行业分类依据中国证监会《上市公司行业分类指引》中的二级行业分类代码。B 采矿业，C 制造业，D 电力、热力、燃气及水生产和供应业，E 建筑业，F 批发和零售业，G 交通运输、仓储和邮政业，I 信息传输、软件和信息技术服务业，J 金融业，K 房地产业。

表 4 不同企业性质横截面股票收益

投资组合	国营或国有控股	私营企业	中外合资
Low	-0.38	-0.08	-0.07
	[-0.46]	[-0.09]	[-0.07]
5	0.70	0.60	0.61
	[0.89]	[0.70]	[0.62]

	1.15	1.37	1.64
High	[1.59]	[1.56]	[1.69]
	1.53	1.45	1.71
High-Low	[3.17]	[3.09]	[2.04]

有关自编码因子与其他因子指标的双变量排序结果，已增加至论文正文中，见论文 13 页起，同时摘录如下：

“本小节，我们进一步考察基于自编码预测值和其他指标的双变量排序的投资组合的表现，因子指标选取了在股票市场上具有代表性的因子指标，分别为公司规模（SIZE）、账面市值比（BM）、净运营资本（NOA）、资产收益率（ROA）、非流动性（ILL）和换手率（TURN）。

根据Hou 等^[1]的研究，每月对自编码预测值和其他因子值进行独立的双变量排序，并在月度频率上进行再平衡。具体而言，每月初按照自编码预测值从小到大对样本中全体公司进行排序，等分成 10 组投资组合；按照因子指标的数值从小到大进行排序，按 30 和 70 百分位点将全体公司分为三组投资组合，我们持有这些组合一个月并计算市值加权的投资组合月度收益。

对于账面市值比（BM），低于 30% 分位数点公司的称为“成长型（Growth）”公司，30% 到 70%之间的公司称为“中性”（Neutral），高于 70%的公司称为“价值型（Value）”公司。根据公司规模（size）的三个分位点可分为“微型（Micro）”公司、“小型（Small）”公司和“大型（Large）”公司。如此之外，我们还选取了衡量投资水平的指标净运营资本（NOA）、盈利类指标资产收益率（ROA）以及非流动性（ILL）和换手率（TURN）四个指标，使研究更加全面。

表 5 基于自编码预测值和其他因子指标双变量排序（Double Sorts）的投资组合表现

投资 组合	公司规模（SIZE）			账面市值比(BM)			净运营资本（NOA）		
	微型	小型	大型	成长	中	价值	低	中	高
Low	1.43	0.42	-0.36	-0.04	-0.38	-0.47	-0.08	-0.24	-0.44
	[1.35]	[0.43]	[-0.42]	[-0.05]	[-0.43]	[-0.49]	[-0.09]	[-0.28]	[-0.53]
High	2.35	1.24	0.91	1.19	1.37	1.79	1.17	1.39	1.73
	[2.48]	[1.45]	[1.26]	[1.40]	[1.64]	[2.16]	[1.51]	[1.79]	[1.97]
H-L	0.91	0.82	1.26	1.23	1.75	2.26	1.25	1.63	2.17
	[1.96]	[1.91]	[2.35]	[2.37]	[3.41]	[3.21]	[2.32]	[2.91]	[4.01]
投资 组合	资产收益率(ROA)			非流动性（ILL）			换手率（TURN）		
	低	中	高	低	中	高	低	中	高
Low	-0.53	-0.46	0.02	-0.51	0.83	1.00	0.37	-0.21	-0.58
	[-0.58]	[-0.53]	[0.02]	[-0.59]	[0.94]	[1.08]	[0.42]	[-0.25]	[-0.67]
High	1.90	1.41	1.13	1.03	1.22	2.07	1.59	1.55	1.28
	[2.33]	[1.71]	[1.41]	[1.37]	[1.52]	[2.34]	[2.10]	[1.75]	[1.45]
H-L	2.43	1.87	1.11	1.54	0.39	1.17	1.22	1.76	1.84
	[3.67]	[3.37]	[2.26]	[2.59]	[0.87]	[2.47]	[1.84]	[3.44]	[3.12]

表 5 展示了自编码预测值和其他因子指标的双变量排序（Double Sorts）的投资组合表现。结果显示，自编码预测值的溢价在不同水平上有不同的表现，具体来看，在大型公司中多空投资组合的月平均收益为 1.26%（t=2.35），高于微型公司和小型公司，分别为 0.91%和 0.82%；在价值型公司中多空投资组合的月平均收益为 2.26%（t=3.21），高于微型公司和小型公司的 1.23%和 1.75%，总的来说，基于自编码的多空组合收益主要集中在大型公司和价值型公司。另外，投资组合随投资水平的提高单调增加，范围从低净运营资本的 1.25%到高

净运营资本 2.17% ($t=4.01$)；而自编码投资在所有盈利能力的公司中均有超额收益，但在盈利能力低的公司中对收益率的区分更好，从 Low 组的-0.53%到 High 组 1.90%，多空组合获得 2.43%月平均收益。除此之外，对于非流动性和换手率两个因子的双变量排序可以看出，自编码多空投资组合在高流动性和高换手率的公司种能获得更高的回报。”

4. P16 的图字体太小，数据太多，无法看清，建议去掉，仅用颜色表示类别，可能更好些。

非常感谢您的建设性建议，根据您的建议，我们将指标名称去掉，仅用颜色表示列表，见附录 1。除此之外，为缩减表格长度，突出重点，我们对权重每 6 个月进行加总平均，重新绘制表 6。

表 6 自编码器输出层因子重要度排序

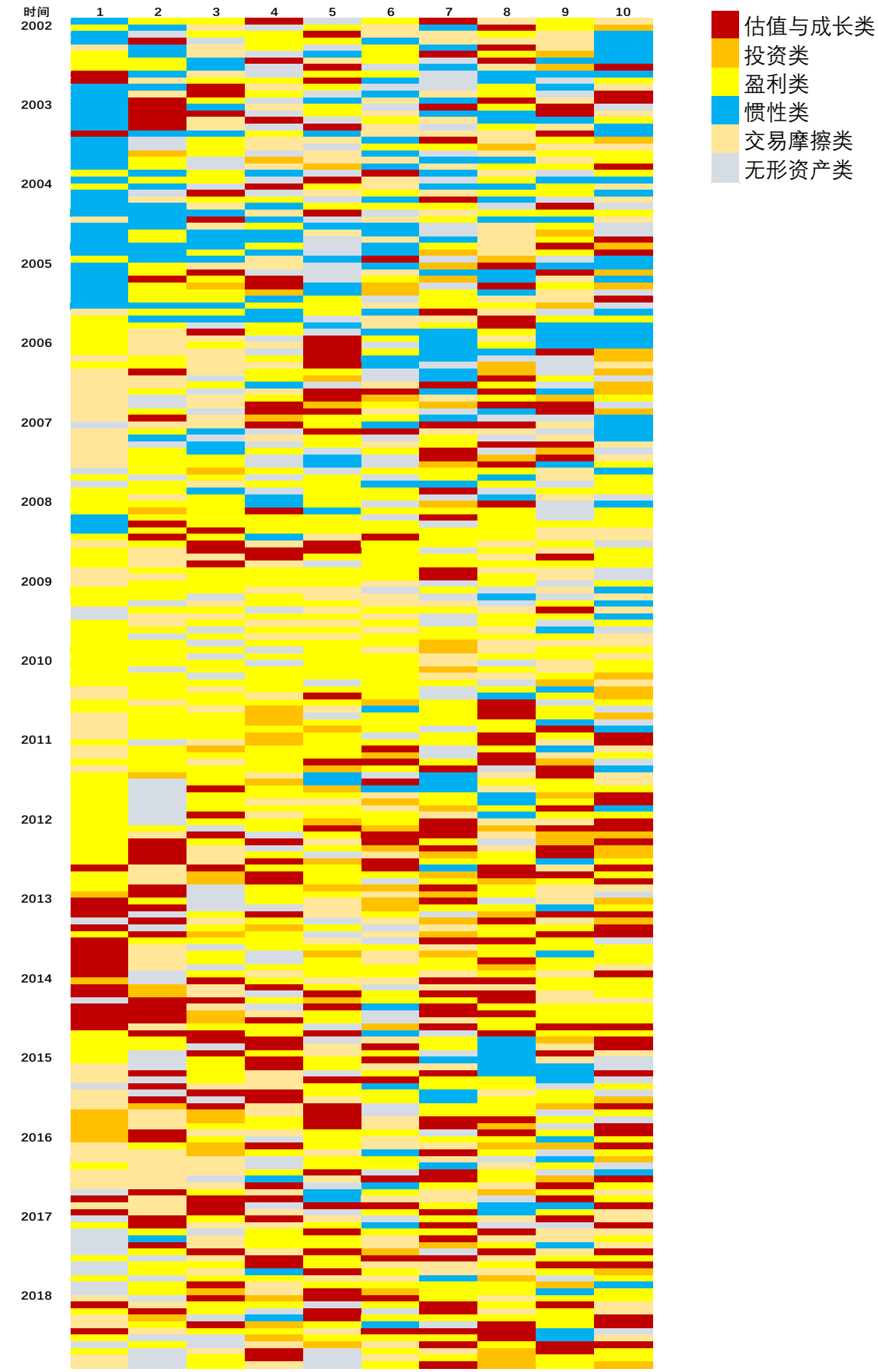
year	1	2	3	4	5	6	7	8	9	10	
2001	MOM1M	PRCDEL	TAN	GM	CHMOM	PRC	ROA	FOE	ROE	MAXRET	估值与成长类
	MOM1M	TAN	PRCDEL	ROE	FOE	EPS	MOM6M	CHMOM	PRC	MAXRET	
2002	ROE	EPS	PRCDEL	FOE	MOM6M	TAN	MOM1M	CHMOM	AGE	CHTX	投资类
	INDMOM	FOE	EPS	MOM6M	CHMOM	AGE	PRCDEL	TAN	ROE	MAXRET	
2003	INDMOM	EPS	TAN	FOE	MAXRET	CHMOM	MOM6M	IVOL	AGE	PRCDEL	盈利类
	INDMOM	EPS	TAN	PRCDEL	CHMOM	FOE	MAXRET	MOM6M	AGE	MOM1M	
2004	INDMOM	EPS	PRCDEL	CHMOM	MOM6M	AGE	FOE	TAN	MOM1M	ROE	惯性类
	INDMOM	CHMOM	MOM6M	ROE	PRCDEL	AGE	MOM1M	EPS	FOE	FM	
2005	INDMOM	ROE	CHMOM	AGE	MOM1M	PRCDEL	FOE	dPIA	MOM6M	GM	交易摩擦类
	GM	PRCDEL	DOLVOL	FOE	CHMOM	AGE	INDMOM	TBI	CFOA	MOM1M	
2006	PRCDEL	DOLVOL	GM	FOE	CFOA	AGE	TAN	CHMOM	IVC	BETASQ	无形资产类
	PRCDEL	TAN	GM	AGE	DOLVOL	CFOA	CHMOM	FOE	CFP	EP	
2007	PRCDEL	AGE	TAN	CHMOM	EPS	CFOA	CFD	INDMOM	EY		
	EPS	CHMOM	TAN	GM	EY	PRCDEL	ROE	TBI	IP	AGE	
2008	EPS	ROE	EY	PRCDEL	EP	GM	TAN	IVOL	CHMOM	RNA	
	PRCDEL	ROE	EPS	IVOL	EY	TAN	FOE	DOLVOL	CFOA	ROA	
2009	ROA	TAN	PRCDEL	EY	EPS	IVOL	CFOA	DOLVOL	ROE	CFD	
	ROA	TAN	EY	CFOA	PRCDEL	EPS	GM	DOLVOL	ROE	IVOL	
2010	ROA	PRCDEL	GM	EPS	EY	CFOA	TAN	ROE	dPIA	DOLVOL	
	PRCDEL	GM	EPS	dPIA	CFOA	ROA	TAN	EY	FM	FOE	
2011	CFOA	GM	TAN	PRCDEL	dPIA	EPS	FM	ROA	EY	CHMOM	
	CFOA	TAN	GM	FOE	dPIA	EY	PRCDEL	FM	ROA	DOLVOL	
2012	CFOA	GM	TAN	PRCDEL	FM	FOE	EY	IVC	DOLVOL	dPIA	
	FM	CFOA	PRCDEL	TAN	GM	dPIA	IVC	DOLVOL	OCFP	CAC	
2013	FM	PRCDEL	TAN	GP	GM	dPIA	DOLVOL	IVC	OCFP	CAC	
	FM	TAN	GP	PRCDEL	IVC	GM	DOLVOL	PRC	ROIC	CFP	
2014	FM	TG	TAN	IVC	GM	DOLVOL	CFP	EY	EP	EPS	
	TAN	GM	DOLVOL	CFP	CP	PRCDEL	EY	TG	INDMOM	FM	
2015	PRCDEL	TAN	CFP	DOLVOL	GM	CP	TBI	OCFP	CHMOM	INDMOM	
	PRCDEL	DOLVOL	TAN	CFP	CINVEST	TBI	IP	OCFP	IVC	CHMOM	
2016	PRCDEL	DOLVOL	TAN	BETASQ	GP	CFP	IP	EY	OCFP	PRC	
	TAN	PRCDEL	TG	DOLVOL	GP	CFP	EPS	OCFP	EP	BETASQ	
2017	TAN	EY	DOLVOL	PRCDEL	EPS	OCFP	CFP	PRC	CHMOM	EP	
	TAN	EY	PRCDEL	SG	IVC	GM	DOLVOL	EPS	CHTX	CFP	
2018	EY	PRCDEL	TAN	TG	IVC	CFD	GM	EP	CHTX	EPS	
	EY	CFD	PRCDEL	TAN	SIZE	GM	TG	EP	IVC	CHTX	

注：表中的英文缩写所代表的公司特征指标详见论文附录 1 公司特征指标名称、定义以及文献出处。

5. 文章太长，建议缩减。

非常感谢您的建议，我们对文章进行了大幅的缩减，从之前的 18000 字缩减至 15000 字左右，特别是对第 4 节“自编码因子的经济学理论”的删减，使用了更加精炼的语言书写。除此之外，我们对行文进行了多处修改，增强其可读性。再次感谢您的宝贵建议！

附录 1 自编码器输出层因子重要度排序



附录 2 中国市场因子指标构建说明

一、 估值与成长类指标（20 个）

1. 资产市值比（Assets-to-market, AM）

根据 Fama and French(1992)计算资产市值比，资产市值比等于总资产合计除以 A 股流通市值。

$AM = \text{总资产} / A \text{ 股流通市值}$

2. 账面市值比（Book-to-market-equity, BM）

根据 Rosenberg, Reid, and Lanstein(1985)计算账面市值比，账面市值比等于股东权益除以 A 股流通市值。

$BM = \text{账面市值} / \text{公司市值}$

3. 现金流股价比（Cashflow-to-price, CFP）

根据 Lakonishok, Shleifer, and Vishny(1994)计算现金流股价比，现金流股价比等于经营活动产生的现金流量净额除以 A 股流通市值。经营活动产生的现金流量净额等于经营活动产生的现金流入与经营活动产生的现金流出之差，A 股流通市值为个股的流通股数与月收盘价的乘积。

$CFP = \text{经营活动产生的现金流量净额} / A \text{ 股流通市值}$

4. 债务股本比（Debt-to-equity-ratio, DER）

根据 Bhandari(1988)，债务股本比等于总负债除以总市值。总负债等于总资产与所有者权益合计值的差额，总市值等于个股的发行总股数与月收盘价的乘积。

$\text{总负债} = \text{总资产} - \text{所有者权益合计值}$

$DER = \text{总负债} / \text{总市值}$

5. 红利价格比（Dividend-to-price-ratio, DP）

根据 Litenberger and Ramaswamy(1982)，红利价格比（DP）等于 t 期应付股利除以 t-1 期 A 股流通市值。应付股利为公司经董事会或股东大会决议确定分配的、尚未支付的现金股利。A 股流通市值为个股的流通股数与月收盘价的乘积。

$DP = \text{应付股利} / A \text{ 股流通市值}$

6. 市盈率（Earnings-to-price, EP）

根据 Basu(1983)，市盈率等于净利润除以 A 股流通市值。净利润是指公司实现的净利润。A 股流通市值为个股的流通股数与月收盘价的乘积。

$EP = \text{净利润} / A \text{ 股流通市值}$

7. 债务增长（Liability-growth, LG）

根据 Richardson, Sloan, Soliman, and Tuna(2005)，债务增长等于总负债的变动额除以上一期总负债。总负债的变动额等于当期总负债与上一期总负债的差值。

$LG = \text{总负债的变动额} / \text{上一期总负债}$

$\text{总负债的变动额} = \text{当期总负债} - \text{上一期总负债}$

8. 营运现金流价格比（Operating-cashflow-to-price, OCFP）

根据 Desai, Rajgopal, and Venkatachalam(2004)，营运现金流价格比等于经营活动产生的现金流量净额除以 A 股流通市值。经营活动产生的现金流量净额为经营活动产生的现金流入与经营活动产生的现金流出之差，A 股流通市值为个股的流通股数与月收盘价的乘积。

$OCFP = \text{经营活动产生的现金流量净额} / A \text{ 股流通市值}$

9. 股息股价比（Payout-yield, PY）

根据 Boudoukh, Michaely, Richardson, and Roberts(2007)，股息股价比等于净利润与所有者权益合计值变动额的差值除以 A 股流通市值。所有者权益合计值变动额为当期所有者权益合计值与上一期所有者权益合计值的差值，A 股流通市值为个股的流通股数与月收盘价的乘积。

$PY = (\text{净利润} - \text{所有者权益合计值变动额}) / A \text{ 股流通市值}$

股东权益变动额=当期所有者权益合计值-上一期所有者权益合计值

10. 反转 (Reversal, Revl)

根据 DeBondt and Thaler(1985)计算反转因子, 主要根据 t-60~t-13 的月度数据, 将期间的累计收益率当做反转因子。累计收益率等于以 e 为底数、1 加收益率的自然对数之和为指数, 最后减 1。

$$\text{累计收益} = \exp\left(\sum \log(1 + \text{ret})\right)$$

11. 可持续增长率 (Sustainable-growth, SG)

根据 Lockwood and Prombutr(2010), 可持续增长率等于当期所有者权益合计值与上一期所有者权益合计值的比值减 1。

$$\text{可持续增长率 (SG)} = (\text{当期所有者权益合计值} / \text{上一期所有者权益合计值}) - 1$$

12. 销量增长与存货增长差 (Sales-growth-minus-inventory-growth, SMI)

根据 Abarbanell and Bushee(1998), 销量增长与存货增长差等于销量增长与存货增长的差额。销量增长等于营业收入变动额除以上一期营业收入, 营业收入变动额为当期营业收入与上一期营业收入的差额。存货增长为存货净额的变动额除以上一期存货净额, 存货净额的变动额为当期存货净额与上一期存货净额的差额。

$$\text{销量增长与存货增长差 (SMI)} = \text{销量增长} - \text{存货增长}$$

$$\text{销量增长} = \text{营业收入变动额} / \text{上一期营业收入}$$

$$\text{营业收入变动额} = \text{当期营业收入} - \text{上一期营业收入}$$

$$\text{存货增长} = \text{存货净额的变动额} / \text{上一期存货净额}$$

$$\text{存货净额的变动额} = \text{当期存货净额} - \text{上一期存货净额}$$

13. 销量价格比 (Sales-to-price, SP)

根据 Barbee, Mukherji, and Raines(1996), 销量价格比等于营业收入除以 A 股流通市值。A 股流通市值为个股的流通股数与月收盘价的乘积。

$$\text{销量价格比 (SP)} = \text{营业收入} / \text{A 股流通市值}$$

14. 纳税增长 (Tax-growth, TG)

根据 Thomas and Zhang(2011), 纳税增长等于应交税费的变动额除以上一期应交税费。应交税费的变动额等于当期应交税费与上一期应交税费的差额。

$$\text{纳税增长 (TG)} = \text{应交税费的变动额} / \text{上一期应交税费}$$

$$\text{应交税费的变动额} = \text{当期应交税费} - \text{上一期应交税费}$$

15. 财务杠杆 (Long-Term-Debt-to-Sum-of-Long-Term-Debt-and-Equity, DA)

根据苏冬蔚(2005), 财务杠杆等于长期借款除以长期借款与所有者权益合计值的和。长期借款为公司向银行或其他金融机构借入的期限在一年期以上(不含一年)的各项借款。

$$\text{财务杠杆 (DA)} = \text{长期借款} / (\text{长期借款} + \text{所有者权益合计值})$$

16. 股利支付率 (Dividend-payout-ratio, DPR)

根据蒋志强, 田婧雯, 周炜星(2019), 股利支付率等于应付股利除以利润总额。应付股利为公司经董事会或股东大会决议确定分配的、尚未支付的现金股利, 利润总额为公司实现的利润总额。

$$\text{应付股利 (DPR)} = \text{应付股利} / \text{利润总额}$$

17. 现金流量价格比 (Net-cash-flow-to-price, NCF)

根据赵春光(2004), 现金流量价格比等于现金及现金等价物净增加额的变动额除以 A 股流通市值。现金及现金等价物净增加额的变动额等于当期现金及现金等价物净增加额与上一期现金及现金等价物净增加额的差额。A 股流通市值为个股的流通股数与月收盘价的乘积。

$$\text{现金流量价格比 (NCF)} = \text{现金及现金等价物净增加额的变动额} / \text{A 股流通市值}$$

$$\text{现金及现金等价物净增加额的变动额} = \text{当期现金及现金等价物净增加额} - \text{上一期现金及现金等价物净增加额}$$

18. 基于现金流的估值因子 (Flow-to-market, FM)

根据干伟明, 张涤新(2018), 基于现金流的估值因子等于投资活动产生的现金流量净额与经营活动产生的现金流量净额之和除以 A 股流通市值。投资活动产生的现金流量净额为投资活动产生的现金流入与投资活动产生的现金流出之差, 经营活动产生的现金流量净额为经营活动产生的现金流入与经营活动产生的现金流出之差, A 股流通市值为个股的流通股数与月收盘价的乘积。

基于现金流的估值因子 (FM) = (投资活动产生的现金流量净额 + 经营活动产生的现金流量净额) / A 股流通市值

19. 基于现金流构建的权益估值因子 (Flow on Equity)

根据干伟明, 张涤新(2018), 基于现金流构建的权益估值因子等于现金流量净额与经营活动产生的现金流量净额之和除以公司的所有者权益。投资活动产生的现金流量净额为投资活动产生的现金流入与投资活动产生的现金流出之差, 经营活动产生的现金流量净额为经营活动产生的现金流入与经营活动产生的现金流出之差, 所有者权益则是股东各项权益项目的总和。

基于现金流的权益估值因子 (FOE) = (投资活动产生的现金流量净额 + 经营活动产生的现金流量净额) / 所有者权益

20. 每股净资产 (Net-Assets-Per-Share, NAPS)

根据陈信元, 陈冬华, 朱红军 (2002) 计算每股净资产, 每股净资产是股东权益与总股数的比率, 每股净资产等于所有者权益合计除以股本总额。

每股净资产 = 所有者权益合计 / 股本总额

二、 投资类指标 (12 个)

1. 应计收入 (Accruals, ACC)

根据 Sloan(1996) 计算应计收入, 即用流动资产总额变动减去流动负债总额变动的差, 除以资产总计平均值。前者为当期流动性资产与上一期流动性资产的差值, 减去当期流动负债与上一期流动负债的差值, 再减去资产折旧。后者为当期资产总计与上一期资产总计的平均值。

应计收入 = 流动资产总额变动 - 流动负债总额变动 / 资产总计平均值

流动资产总额变动 - 流动负债总额变动 = 当期流动资产 - 上一期流动资产 - 流动负债 + 上一期流动负债 - 固定资产折旧、油气资产折耗、生产性生物资产折旧

资产总计平均值 = $\frac{1}{2}$ (资产总计 + 上一期资产总计)

2. 百分比应计收入 (Percent-accruals, PACC)

根据 Hafzalla, Lundholm, and VanWinkle(2011) 计算百分比应计收入, 百分比应计收入等于利润总额减去营业现金流, 再除以净利润。

百分比应计收入 = (利润总额 - 经营活动产生的现金流量净额) / 净利润

3. 资本开销增长率 (Capital-expenditure-growth, CAPXG)

根据 Mcconnell and Muscarella(1985) 计算资本开销增长率, 资本开销增长率是资本开销年变动额与上一年的资本开销的比率。资本开销增长率等于固定资产净额变动加上无形资产净额变动, 除以上期固定资产净额与上期无形资产净额之和。前者为当期固定资产净额减去上期固定资产净额之差加上当期无形资产净额减去上期无形资产净额之差。固定资产净额为固定资产原价除去累计折旧和固定资产减值准备之后的净额。无形资产净额为无形资产的原价扣除摊销和减值准备后的净额。

资本开销增长率 = (固定资产净额变动 + 无形资产净额变动) / (上期固定资产净额 + 无形资产净额)

固定资产净额变动 + 无形资产净额变动 = 当期固定资产净额 - 上期固定资产净额 + 当期无形资产净额 - 上期无形资产净额

4. 股东权益变化 (Change-in-shareholders' equity, dBe)

根据 Richardson, Sloan, Soliman, and Tuna(2005) 计算股东权益变化, 股东权益变化等于所有者权益净额变动除以上一期总资产

股东权益变化=所有者权益净额变动/上一期总资产

所有者权益净额变动=当期所有者权益净额-上期所有者权益净额

5. 固定资产与存货变化率 (Changes-in-PPE-and-inventory-to-assets, dPIA)

根据 Lyandres, Sun, and Zhang(2008)计算固定资产与存货变化率, 固定资产与存货变化率为固定资产以及存货的年变动额与上期总资产的比率。固定资产与存货变化率等于固定资产净额变动加上存货净额变动之和, 除以上期资产总额。前者为本期固定资产净额减去上期固定资产净额之差加上本期存货净额减去上期存货净额。固定资产净额为固定资产原价除去累计折旧和固定资产减值准备之后的净额。存货净额为存货与存货跌价准备之差。

固定资产与存货变化率=(固定资产净额变动+存货净额变动)/上期资产总额

固定资产净额变动+存货净额变动=当期固定资产净额-上期固定资产净额+当期存货净额-上期存货净额

6. 投资资产比 (Investment-to-assets, IA)

根据吴世农,冉孟顺,肖珉,李雅莉(1999)计算投资资产比(总资产增长率), 投资资产比是公司总资产年变动百分比。投资资产比等于当期总资产减去上期总资产之差, 除以上期总资产, 即本期总资产除以上期总资产之比减一。

投资资产比=(当期总资产-上期总资产)/上期总资产=当期总资产/上期总资产-1

7. 存货变化率 (Inventory-change, IVC)

根据 Thomas and Zhang(2002)计算存货变化率, 存货变化率为存货变动额与总资产平均值的比率。存货变化率等于存货净额变动除以资产总计平均值。前者为当期存货净额减去上一期存货净额之差, 后者为当期资产总计与上一期资产总计的平均值。

存货变化率=存货净额变动/资产总计平均值

存货变动额=当期存货净额-上期存货净额

资产总计平均值= $\frac{1}{2}$ (资产总计+上一期资产总计)

8. 存货增长率 (Inventory-growth, IVG)

根据 Belo and Lin(2011)计算存货增长率, 存货增长率等于存货净额变动除以上期存货净额。前者为当期存货净额减去上期存货净额之差。

存货增长率=存货净额变动/上期存货净额

存货净额变动=当期存货净额-上期存货净额

9. 净运营资本 (Net-operating-assets, NOA)

根据 Hirshleifer, Hou, Teoh, and Zhang(2004)计算净运营资本, 净运营资本是净运营资本与总资产的比值。净运营资产等于运营资产减去运营负债之差, 除以总资产。前者为流动资产减去货币资金的差值, 减去流动负债减去交易性金融负债减去一年内到期非流动负债的差值。

净运营资本=净运营资本/资产合计

净运营资本=运营资产-运营负债

运营资产=流动资产-货币资金

运营负债=流动负债-交易性金融负债-一年内到期非流动负债

10. 房地产投资量 (Realestate-holdings, REALESTATE)

根据 Tuzel(2010)计算房地产投资量, 房地产投资量等于投资性房地产净额除以固定资产净额。

房地产投资量=投资性房地产净额/固定资产净额

11. 公司投资 (Corporate-investment, CINVEST)

根据 Titman, Wei, and Xie(2004)计算公司投资, 公司投资等于购建固定资产、无形资产和其他长期资产支付的现金加上支付其他与经营活动有关的现金减去处置固定资产、无形资产和其他长期资产收回的现金, 除以总资产。

公司投资=(购建固定资产、无形资产和其他长期资产支付的现金+支付其他与经营活动有关的现金-处置固定资产、无形资产和其他长期资产收回的现金)/总资产

12. 投资资本比 (Investment-to-Net-Capital, IC)

根据王茵田,朱英姿(2011)计算投资资本比,投资资本比是投资支出对固定资产净值的比率。投资资本比等于投资支付的现金除以固定资产净额。

投资资本比=投资支付的现金/固定资产净额

13. 可持续增长率 (Sustainable-growth, SG)

根据 Lockwood and Prombutr(2010),可持续增长率等于当期所有者权益合计值与上一期所有者权益合计值的比值减1。

可持续增长率 (SG) = (当期所有者权益合计值/上一期所有者权益合计值) - 1

三、 盈利类指标 (23 个)

1. 资产换手率 (Asset-turnover, ATO)

根据 Soliman(2008)计算资产换手率,ATO 衡量资产的利用率和效率。资产换手率等于营业总收入除以资产总计平均值。后者为当期资产总计与上一期资产总计的平均值。

资产换手率=营业总收入/资产总计平均值

营业总收入=企业经营过程中所有收入之和

资产总计平均值 = $\frac{1}{2}$ (资产总计+上一期资产总计)

2. 现金流资产比 (Cash-flow-over-assets, CFOA)

根据 Asness, Frazzini, and Pedersen(2017)计算现金流资产比,即现金流量除以资产总计。前者为税后经营净利润加上折旧与摊销减去经营营运资本增加再减去资本支出。

现金流资产比=(净利润+固定资产折旧+无形资产摊销-(流动资产合计-上一期流动资产合计-流动负债合计+上一期流动负债合计-固定资产净值+上一期固定资产净值-无形资产净值+上一期无形资产净值))/资产总计

3. 现金生产率 (Cash-productivity, CP)

根据 Chandrashekai and Rao(2009)计算现金生产率,现金生产率为滞后一期流通市值加上当期长期负债再减去当期总资产合计,最后除以当期货币资金。

现金生产率=(流通市值+长期借款-资产总计)/货币现金

4. 现金资产比 (Cash-to-assets, CTA)

根据 Palazzo(2012)计算现金资产比,现金资产比是指货币资金除以资产总计平均值,后者为当期资产总计与上一期资产总计的平均值。

现金资产比=货币现金/资产总计平均值

资产总计平均值 = $\frac{1}{2}$ (资产总计+上一期资产总计)

5. 资本换手率 (Capital-turnover, CTO)

根据 Haugen and Baker(1996)计算资本换手率,资本换手率等于当期营业收入除以滞后一期资产总计。

资本换手率=营业收入(t)/资产总计(t-1)

6. 息税前收益 (Earnings-before-interests-and-taxes, EBIT)

根据 Greenblatt(2006)计算息税前收益,息税前收益等于净利润加上所得税费用再加上财务费用。

息税前收益=净利润+所得税费用+财务费用

7. 企业收益率 (Earnings-yield, EY)

根据 Greenblatt(2006)计算企业收益率,企业收益率等于息税前利润除以企业价值。前者为净利润加上所得税费用再加上财务费用。后者为资产总计减去所有者权益加上流通市值。

企业收益率=息税前利润/企业价值

息税前利润=净利润+所得税费用+财务费用

企业价值=资产总计-所有者权益+流通市值

8. 边际毛利 (Gross-margins, GM)

根据 Novy-Marx(2013)计算边际毛利, 边际毛利是指经营利润与营业收入的比值。即边际毛利等于营业收入与营业成本的差值除以营业收入。

边际毛利= (营业收入-营业成本) / 营业收入

9. 毛利率 (Gross-profitability, GP)

根据 Novy-Marx(2013);Jiang, Qi, and Tang(2018)计算毛利率, 毛利率等于营业收入与营业成本的差值除以资产总计。

毛利率= (营业收入-营业成本) / 资产总计

10. 净利润 (Net-payout-over-profits, NPOP)

根据 Asness, Frazzini, and Pedersen(2017)计算净利润, 即净利润与所有者权益变动之差除以利润总额。其中, 所有者权益变动等于当期所有者权益减去上一期所有者权益。

净利润= (净利润-当期所有者权益+上一期所有者权益) / 利润总额

11. 净运营资产收益 (Return-on-net-operating-assets, RNOA)

根据 Soliman(2008)计算净运营资产收益, 净运营资产收益率等于净利润除以流动资产与流动负债的差值。

净经营收益率=净利润/ (流动资产合计-流动负债合计)

12. 资产收益率 (Return-on-assets, ROA)

根据 Balakrishnan, Bartov and Faurel (2010) ;Jiang,Qi,andTang(2018)计算资产收益率, 资产收益率等于当期净利润除以滞后一期总资产合计。

资产收益率=净利润/资产总计

13. 股权收益率 (Return-on-equity, ROE)

根据 Hou, Xue and Zhang(2015); Jiang, Qi, and Tang(2018)计算股权收益率, 股权收益率等于当期净利润除以滞后一期所有者权益合计。

股权收益率=净利润/所有者权益

14. 投资型资产收益率 (Return-on-invested-capital, ROIC)

根据 Greenblatt(2006)计算投资型资产收益率, 资产收益率是指息税前利润与投资资本的比值。前者为净利润加上所得税费用再加上财务费用。后者为流动资产减去流动负债加上固定资产净额。

投资型资产收益率=息税前利润/投资资本

息税前利润=净利润+所得税费用+财务费用

投资资本=流动资产总计-流动负债总计+固定资产净额

15. 应税所得与账面资产比 (Tax-able-income-to-book-income, TBI)

根据 Green, Hand, and Zhang(2013)计算应税所得与账面资产比, 应税所得与账面资产比是指应税所得除以账面收益。前者为税前利润。后者为净利润。

应税所得与账面资产比= (净利润+所得税费用) / 净利润

16. Z 评分 (Z-score, Z)

根据 Asness, Frazzini, and Pedersen(2017)计算 Z 评分, Z 评分是总资产中的营运资金, 未分配利润, 未计利息和税项的收入, 市场份额和销售的加权平均值。等于营运资本的 1.2 倍加上未分配利润的 1.4 倍加上息税前利润的 3.3 倍加上上个月总市值 0.6 倍加上营业收入的 0.99 倍所得之和除以资产总计。

Z 评分= (1.2*营运资本+1.4*留存收益+3.3*息税前利润+0.6*月个股总市值+0.99*营业收入) / 资产总计

17. 销售收入比 (Sales-to-receivables, SALEREV)

根据 Ou and Penman(1989)计算销售收入比, 即营业收入与应收账款的比值。销售收入比=营业收入/应收账款

18. 纳税变化额 (Change-in-tax-expense, CHTX)

根据 Thomas and Zhang(2011)计算纳税变化额,即用本季度的纳税费用减去年同一季度纳税费用得到的值,除以去年同一季度的资产总计。

纳税变化额=(本季度纳税额-去年同一季度纳税额)/去年同一季度资产

19. 每股盈余 (Earnings-per-share, EPS)

根据赵宇龙(1998)计算每股盈余,等于公司净利润除以流通在外的普通股股数,即净利润与实收资本(或股本)的比值。

每股盈余=净利润/实收资本(或股本)

20. 永久经济盈余 (Consensus-Earnings-Per-Share, CEPS)

根据赵宇龙,易琮(1999)计算永久经济盈余,永久经济盈余等于营业收入、营业成本与营业税金及附加三者的差值与实收资本(或股本)的比值。

永久经济盈余=(营业收入-营业成本-营业税金及附加)/实收资本(或股本)

21. 有形资产与营业收入比(Tangible-Assets-to-Operating-Revenue, TA)

根据罗婷,朱青,李丹(2009),有形资产与营业收入比等于有形资产除以营业收入。有形资产等于总资产减去无形资产净额再减去长期待摊费用。无形资产净额是指公司各项无形资产的原价扣除摊销和减值准备后的净额。

有形资产与营业收入比=有形资产/营业收入

有形资产=总资产-无形资产净额-长期待摊费用

22. 投资收益占比(Investment-gains-to-net-profit, IP)

根据翟进步,罗玫(2014)计算投资收益占比,这个因子反映了投资收益对净利润的贡献程度。投资收益占比较小的公司可能还有其它大量能够持续经营的业务,盈利持续性较强。投资收益占比等于本期投资收益除以净利润。

投资收益占比=投资收益/净利润

23. 主营业务利润率增长速度(TMT)

根据干伟明,张涤新(2018)计算主营业务利润增长速率,也就是主营业务利润率环比增速。该因子等于主营业务利润率的变动除以上一年主营业务利润率。主营业务利润率为营业收入减去营业总成本与营业税金。主营业务利润率的变动则等于本期主营业务利润率减去上一期主营业务利润率。

主营业务利润率=营业收入-营业总成本-营业税金

主营业务利润率增长速率=(主营业务利润率-上一期主营业务利润率)/上一期主营业务利润率

四、趋势类指标(7个)

1. 6个月惯性变化(Change-in-6-month-momentum, CHMOM6)

根据Gittleman and Marks(2006)计算6个月惯性变化,该因子反映的是动量因子的变动。t月份的6个月惯性变化因子等于t-6月到t-1月份的累计月收益减t-12月到t-7月份的累计月收益。

6个月惯性变化=t-6月到t-1月份的总收益-t-12月到t-7月份的总收益

2. 交易量趋势(Volume-trend, VOLT)

根据Haugen and Baker(1996)计算交易量趋势因子,这是一个流动性因子。t月的交易量趋势因子等于t月的交易量除以过去五年的累计交易量。

过去五年的累计交易量=从股票上市到t月累计交易量-t-60月累计交易量

交易量趋势因子=第t月交易量/过去五年的累计交易量

3. 行业惯性(Industry-momentum, INDMOM)

根据Moskowitz and Grinblatt(1999)计算行业惯性,该因子可以理解为在每个行业内构造的等权投资组合的惯性。行业惯性等于行业内所有公司12个月惯性的均值。t月的12个月惯性等于t-12月到t-2月的总收益。

4. 1个月惯性(1-month-momentum, MOM1M)

根据 Jegadeesh and Titman(1993)计算 1 个月惯性, 即股票上一个月的总收益。t 月的 1 个月惯性等于 t-1 月的收益。

5. 6 个月惯性 (6-month-momentum, MOM6M)

根据 Jegadeesh and Titman(1993)计算 6 个月惯性, 即股票过去 6 个月的总收益。t 月的 5 个月惯性等于 t-6 月到 t-2 月的总收益。

6. 12 个月惯性 (12-month-momentum, MOM12M)

根据 Jegadeesh and Titman(1993)计算 12 个月惯性, 即股票过去一年的总收益。t 月的 12 个月惯性等于 t-12 月到 t-2 月的总收益。

7. 36 个月惯性 (36-month-momentum, MOM36M)

根据 Jegadeesh and Titman(1993)计算 36 个月惯性, 即过去三年中前两年的总收益。t 月的 36 个月惯性等于 t-36 月到 t-13 月的总收益。

五、 交易摩擦类指标 (15 个)

1. 系统性风险系数 (Market-beta, BETA)

根据 Fama and MacBeth(1973)计算系统性风险系数, 即用 t 月市值加权的市场日度超额收益率对股票日度超额收益率回归, 估计得出的系数作为 t 月 BETA 的估计。

2. 系统性风险系数方差 (Beta-squared, BETASQ)

根据 Fama and MacBeth(1973)计算系统性风险系数方差, 即系统性风险系数的二次方项。

3. 异质性收益波动率 (Idiosyncratic-return-volatility, IVOL)

根据 Ali, Hwang, and Trombley(2003)计算异质性收益波动率因子, 反映公司间波动率的差异, 等于资产定价模型中不能被市场或行业解释的部分。即用三因子定价模型对市场超额收益率进行回归得到残差项, 残差项的标准差就是异质性收益波动率因子。

4. 非流动性 (Illiquidity, ILLIQ)

根据 Amihud (2002) 计算非流动性, ILLIQ 因子反映了在单位成交额下, 证券价格波动的大小。即用个股日收益率绝对值, 除以每日交易额之比, 最后取平均。前者是收益率绝对值, 后者是每月收盘价绝对值乘以流通股数。

股价变动幅度=收益率的绝对值

5. 最大日收益 (Maximum-daily-returns, MAXRET)

根据 Bali, Cakici, and Whitelaw(2011)计算最大日收益因子, 该因子是个股在一个月内的最大日收益。

6. 股价 (Price, PRC)

根据 Blume and Husic(1973)计算股价, 指股票的交易价格, 即用每日收盘价的对数表示个股股价。

7. 延迟股价 (Price-delay, PRCDEL)

根据 Hou and Moskowitz(2005)计算延迟股价, 即 36 个月内周度股票收益的波动率除以过去四周的市场收益增量。

8. 人民币交易量 (RMB-trading-volume, RVOL)

根据 Chordia, Subrahmanyam, and Anshuman(2001)计算人民币交易量因子, 即用滞后两期的月度交易额, 再取对数。其中滞后两期的月度交易额用滞后两期的月个股交易股数乘以滞后两期的月收盘价计算。

人民币交易量=滞后两期的月度交易额取自然对数

滞后两期的月度交易额=滞后两期的月个股交易股数*滞后两期的月收盘价

9. 公司规模 (Firm-size, SIZE)

根据 Banz(1981)计算公司规模, 即用每个月月末的(最后一个交易日的)股票收盘价格(不复权价格)乘以每个月月末的 A 股流通股本, 也称其为 A 股流通市值。

10. 人民币交易量波动率 (Volatility-of-RMB-trading-volume, STD_RVOL)

根据 Chordia, Subrahmanyam, and Anshuman(2001)计算人民币交易量波动率因子, 即用日度交易额的绝对值取对数, 然后求出该数的标准差。日度交易额由日个股交易股数乘日收盘价得出。

人民币交易量波动率=日度交易额绝对值的自然对数求标准差

日度交易额=每日股市收盘价*日个股交易股数

11. 换手率波动率 (Volatility-of-turnover, STD_TURN)

根据 Chordia, Subrahmanyam, and Anshuman(2001)计算换手率波动率因子, 换手率波动率因子是指每日换手率的标准差。每日换手率用当日成交量除以该股发行总股数计算得出。

换手率波动性因子=每日换手率的标准差

每日换手率=当日成交量/该股发行总量

12. 收益率波动率 (Return-volatility, RETVOL)

根据 Ang, Hodrick, Xing, and Zhang(2006)计算收益率波动率, 即日收益率的标准差。

13. 股票换手率 (Share-turnover, TURN)

根据 Datar, Naik, and Radcliffe(1998)计算股票换手率, 股票换手率指的是一定时间内市场中股票转手买卖的频率, 是反映股票流通性强弱的指标之一。即先计算出每日换手率的值, 然后再来求和。每日换手率等于当日成交量除以该股发行总股数。

股票换手率=每日换手率的和

每日换手率=当日成交量/该股发行总量

14. 零交易天数 (Zero-trading-days, ZEROTRADE)

根据 Liu (2006) 计算零交易天数因子, 该因子是考虑股票零交易天数在内的一个流动性指标。即用有关零交易天数的部分加上有交易量天数的部分, 最后标准化计算。零交易天数部分是每月没有流通股数交易日的总和, 有交易量天数部分用股票换手率的倒数除以一个缩减指数 480000, 使得所有股票的有交易天数部分范围在 0~1 中间。最后由于每月交易日从 15 到 23 天不等, 用 21 除以每月交易总天数, 对前面相加的值进行标准化处理, 使得每月交易日固定为 21 天。

零交易天数因子=(零交易天数+(1/股票换手率)/480000)*21/交易总天数

15. 销售收入 (Total-operating-revenue, TOR)

根据陈信元,陈冬华,朱红军(2002), 销售收入为企业经营过程中所有收入之和

六、 无形资产类指标 (12 个)

1. 公司年龄 (Firm-age, AGE)

根据 Jiang, Lee, and Zhang(2005)计算公司年龄, 认为从公司首次招股日期到现在总共的年份数。即用当前年份减去首次招股日期的年份可得到公司年龄。

公司年龄=当期年份-上市年份

2. 现金流负债比 (Cashflow-to-debt, CFD)

根据 Ou and Penman(1989)计算现金流负债比, 即用当期净利润除以负债总计的平均值。负债总计的平均值为当前负债总计与上一期负债总计的平均值。

现金流负债比=当期净利润/负债总计平均值

负债总计平均值=
$$\frac{1}{2}$$
 (当期负债总计+上一期负债总计)

3. 流动比率 (Current-ratio, CR)

根据 Ou and Penman(1989)计算流动比率, 即为流动资产与流动负债的比率, 用于衡量企业短期的偿债能力。

流动比率=流动资产/流动负债

4. 流动比率增长 (Current-ratio-growth, CRG)

根据 Ou and Penman(1989)计算流动比率增长, 反映的是流动比率的增长情况。即用当期流动比率与上一期流动比率的差除以上一期流动比率。

流动比率增长=(当期流动比率-上一期流动比率)/上一期流动比率

5. 速动比率 (Quick-ratio, QR)

根据 Ou and Penman(1989)计算速动比率, 即为企业速动资产与流动负债的比率, 是衡量短期偿债能力的指标之一。其中速动资产为流动资产与存货净额之间的差值。

速动比率=速动资产/流动负债

速动资产=流动资产合计-存货净额

6. 速动比率增长 (Quick-ratio-growth, QRG)

根据 Ou and Penman(1989)计算速动比率增长, 反映的是速动比率的变化情况。即用当期速动比率与上一期速动比率的差值除以上一期速动比率。

速动比率=速动资产/流动负债

速动资产=流动资产合计-存货净额

7. 销量现金比 (Sales-to-cash, SC)

根据 Ou and Penman(1989)计算销售现金比, 是用来反映企业销售质量高低的指标。即用营业收入除以货币资金。

销量现金比=营业收入/货币资金

8. 销量存货比 (Sales-to-inventory, SI)

根据 Ou and Penman(1989)计算销量存货比, 用于检测库存量是否合理。即用营业收入除以存货净额。

销量存货比=营业收入/存货净额

9. 产品销售成本与销售、行政及管理费用之和与总资产之比(OL)

根据 Novy-Marx (2011) 计算年度产品销售成本与销售、行政及管理费用之和与总资产之比, 是公司的年度运营成本除以账面资产 (乘以任意的规模常数)。

年度产品销售成本与销售、行政及管理费用之和与总资产之比=年度产品销售成本+销售、行政以及管理费用/总资产

10. 资产有形度 (TAN)

根据 Hahn and Lee(2009)计算 TAN 年度, TAN 是指公司的预期资产清算价值, 当公司销售诸如应收账款, 存货, 固定资产等主要资产时, 一单位美元的账面价值平均产生 72 美分的应收账款, 55 美分的存货和 54 美分的财产及设备。TAN 年度等于现金及等价物加上应收账款的 0.715 倍加上存货的 0.547 加上财产及设备总额的 0.535 倍之和除以总资产。

TAN 年度=(现金及等价物+0.715×应收账款+0.547×存货+0.535×财产及设备总额)/总资产

11. 每股无形资产(IBV)

根据薛云奎、王志台(2001)计算每股无形资产, 每股无形资产为期末无形资产净额除以期末股本总额。

每股无形资产=无形资产/股本

12. 现金及其等价物对固定资产净值比(Cash-asset-to-Capital, CAC)

根据王茵田、朱英姿(2011), 现金及其等价物对固定资产净值比等于货币资金除以所有者权益合计值。

现金及其等价物对固定资产净值比 (CAC) =货币资金/所有者权益合计值

附录3 预测值与各因子指标相关系数的分布图

