

相关研究

《选股因子系列研究（二十八）——一致预期质量分析》2017.10.27

《分析师荐股报告是否具有超额收益》2017.10.19

选股因子系列研究（二十九）——因子降维

1: 底层因子降维方法对比

随着投资者对于选股因子挖掘的深入，构建选股模型时可供选择的因子越来越多，随之而来的问题也逐渐增多。其中有两个问题值得关注：1) 因子在数量较多的情况下，无法高效地进行正交化处理；2) 对于某一类高相关因子，正交的手段无法较好地提炼它们的信息。

对于上述问题，投资者可通过因子降维进行处理。也即，对于高相关因子分配因子降维权重，并加权求和计算降维后的因子值。由于因子降维方法多种多样，故而本报告为投资者提供了几种基本降维方法的效果回测对比。

- **因子降维的思路多种多样。**本报告从“选取类别内选股效果最强的因子”、“基于类别内因子选股效果加权”以及“最大化类别内因子选股效果波动”三条思路出发构建了相应的降维方法。
- **“选取类别内选股效果最强的因子”降维效果较好。**模型复合因子选股能力以及模型月度多空收益在降维后都出现了改善。使用正交因子 IC 作为因子选股效果的衡量指标，降维后的模型选股能力出现了进一步的提升。
- **“基于类别内因子选股效果加权”降维效果同样较好。**模型复合因子选股能力以及模型月度多空收益在降维后都出现了改善。使用正交因子 IC 作为因子选股效果的衡量指标，降维后的模型选股能力出现了进一步的提升。
- **“最大化解释类别内因子选股效果波动”降维效果较差。**模型复合因子选股能力在降维后改善较少，使用正交因子后模型表现并无改善。
- **相比而言，“选取类别内选股效果最强的因子”降维后，模型具有最强的选股效果。**对比来看，“取类别内最强”能够对于模型产生最大幅度的提升，同时各类因子在降维后也能保持较强的选股能力。若投资者想尽可能多地使用到类别内各指标的信息，则可使用“类别内因子选股效果加权”。从结果上看，并不推荐使用“最大化解释类别内因子波动”进行降维。
- **风险提示。**市场系统性风险、资产流动性风险以及政策变动风险会对策略表现产生较大影响。

分析师:冯佳睿

Tel:(021)23219732

Email:fengjr@htsec.com

证书:S0850512080006

分析师:袁林青

Tel:(021)23212230

Email:ylq9619@htsec.com

证书:S0850516050003

目 录

1. 选股因子的降维	4
2. 降维前后模型表现对比	5
2.1 选取类别内选股效果最强	6
2.2 选取类别内选股效果最强-改进	7
2.3 类别内选股效果加权	9
2.4 类别内选股效果加权-改进	11
2.5 类别内 PCA 降维	14
2.6 类别内 PCA 降维-改进	16
3. 降维方法对比	17
4. 总结与展望	19
5. 风险提示	19

表目录

表 1	因子截面相关性均值（2009 年 1 月至 2017 年 7 月）	4
表 2	因子计算方法说明	5
表 3	“取类别内选股效果最强”降维前后因子模型表现	6
表 4	“取类别内选股效果最强”降维权重分配情况	7
表 5	“取类别内选股效果最强”（改进）降维前后因子模型表现	8
表 6	“取类别内选股效果最强”（改进）降维权重分配情况	9
表 7	“类别内选股效果加权”降维前后因子模型表现	10
表 8	“类别内选股效果加权”降维权重分配情况	11
表 9	“类别内选股效果加权”（改进）降维前后因子模型表现	12
表 10	“类别内选股效果加权”（改进）降维权重分配情况	13
表 11	“类别内 PCA 第一主成分”降维前后因子模型表现	14
表 12	“类别内 PCA 第一主成分”降维权重分配情况	15
表 13	“类别内 PCA 第一主成分”（改进）降维前后因子模型表现	16
表 14	“类别内 PCA 第一主成分”（改进）降维权重分配情况	17
表 15	不同降维方法下因子模型表现	18
表 16	不同降维方法下的各类因子降维前后的 IC 与 ICIR	19

随着投资者对于选股因子挖掘的深入，构建选股模型时可供选择的因子越来越多，随之而来的问题也逐渐增多。其中有两个值得关注：1) 因子在数量较多的情况下，无法高效地进行正交化处理；2) 对于某一类高相关因子，正交的手段无法较好地提炼它们的信息。

对于上述问题，投资者可通过因子降维进行应对。也即，可对于高相关因子分配因子降维权重，并加权求和计算降维后的因子值，然后使用降维后的因子进行收益模型的构建。由于因子降维方法多种多样，故而本报告为投资者提供了几种基本降维方法的效果回测对比以及相关改进方向。

报告第一章简要介绍了因子降维的整体思路，第二章对比了降维前后多因子模型的表现，第三章对比了不同降维思路下多因子模型的表现。

1. 选股因子的降维

随着因子库中因子数量的增多，因子之间的相关性不可避免，而逻辑类似的因子更有可能具有较高的相关性。下表给出了一个较为简单的因子库中，各因子在 2009 年 1 月至 2017 年 7 月底截面线性相关性的均值。

表 1 因子截面相关性均值 (2009 年 1 月至 2017 年 7 月)

		市值	非线性	换手			反转			特异度			估值			成长			盈利				偿付能力				
		市值	市值平方	1M	3M	6M	1M	3M	6M	1M	3M	6M	PE	PB	PS	营收	营业利润	净利润	ROE	ROA	ROIC	毛利率	净利率	D2E	A2D	流动	速动
市值	市值	1.00	0.50	-0.34	-0.36	-0.38	0.04	0.08	0.10	-0.03	-0.03	-0.02	0.02	-0.33	-0.04	0.15	0.13	0.14	0.41	0.30	0.35	0.11	0.33	0.16	0.15	-0.11	-0.09
	非线性	0.50	1.00	-0.23	-0.24	-0.25	-0.01	-0.01	-0.01	0.05	0.07	0.08	-0.03	-0.19	-0.02	0.04	0.01	0.02	0.20	0.08	0.12	0.04	0.16	0.14	0.14	-0.07	-0.06
	1M	-0.34	-0.23	1.00	0.92	0.84	0.17	0.22	0.21	-0.20	-0.21	-0.19	0.20	0.25	0.20	0.01	0.00	0.00	-0.13	-0.05	-0.09	-0.01	-0.08	-0.16	-0.14	0.15	0.14
	3M	-0.36	-0.24	0.92	1.00	0.94	0.04	0.16	0.20	-0.08	-0.19	-0.20	0.20	0.26	0.21	0.01	0.00	-0.01	-0.14	-0.05	-0.09	-0.01	-0.09	-0.17	-0.15	0.17	0.16
换手	6M	-0.38	-0.25	0.84	0.94	1.00	0.01	0.06	0.15	-0.04	-0.09	-0.17	0.21	0.26	0.22	0.01	0.00	-0.01	-0.13	-0.04	-0.08	0.00	-0.08	-0.19	-0.16	0.19	0.18
	1M	0.04	-0.01	0.17	0.04	0.01	1.00	0.51	0.35	-0.31	-0.12	-0.07	0.11	0.07	0.06	0.01	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	
	3M	0.08	-0.01	0.22	0.16	0.06	0.51	1.00	0.65	-0.30	-0.34	-0.19	0.20	0.11	0.10	0.03	0.04	0.04	0.02	0.02	0.02	0.01	0.00	0.00	0.01	0.00	
	6M	0.10	-0.01	0.21	0.20	0.15	0.35	0.65	1.00	-0.28	-0.35	-0.36	0.29	0.14	0.15	0.07	0.09	0.09	0.05	0.05	0.05	0.03	0.02	-0.01	0.00	0.02	0.01
特异度	1M	-0.03	0.05	-0.20	-0.08	-0.04	-0.31	-0.30	-0.28	1.00	0.62	0.48	-0.26	-0.14	-0.12	-0.02	0.00	0.00	0.02	0.01	0.01	-0.02	0.05	-0.01	-0.01	0.00	0.01
	3M	-0.03	0.07	-0.21	-0.19	-0.09	-0.12	-0.34	-0.35	0.62	1.00	0.78	-0.35	-0.19	-0.17	-0.03	-0.01	0.00	0.03	0.01	0.01	-0.02	0.07	-0.01	-0.02	0.01	0.02
	6M	-0.02	0.08	-0.19	-0.20	-0.17	-0.07	-0.19	-0.36	0.48	0.78	1.00	-0.40	-0.21	-0.20	-0.04	-0.02	-0.02	0.03	0.00	0.00	-0.04	0.08	-0.01	-0.01	0.00	0.02
	PE	0.02	-0.03	0.20	0.20	0.21	0.11	0.20	0.29	-0.26	-0.35	-0.40	1.00	0.34	0.50	0.12	0.09	0.09	0.15	0.21	0.22	0.22	0.05	-0.05	-0.03	0.10	0.10
估值	PB	-0.33	-0.19	0.25	0.26	0.26	0.07	0.11	0.14	-0.14	-0.19	-0.21	0.34	1.00	0.31	-0.13	-0.16	-0.17	-0.56	-0.46	-0.50	-0.10	-0.38	-0.12	-0.13	0.06	0.06
	PS	-0.04	-0.02	0.20	0.21	0.22	0.06	0.10	0.15	-0.12	-0.17	-0.20	0.50	0.31	1.00	-0.03	0.02	0.01	0.00	0.16	0.12	0.58	0.33	-0.50	-0.46	0.41	0.39
成长	营收	0.15	0.04	0.01	0.01	0.01	0.01	0.03	0.07	-0.02	-0.03	-0.04	0.12	-0.13	-0.03	1.00	0.41	0.39	0.34	0.29	0.29	0.08	0.23	0.03	0.03	0.03	0.03
	营业利润	0.13	0.01	0.00	0.00	0.00	0.01	0.04	0.09	0.00	-0.01	-0.02	0.09	-0.16	0.02	0.41	1.00	0.78	0.39	0.38	0.39	0.15	0.32	-0.01	-0.02	0.03	0.02
	净利润	0.14	0.02	0.00	-0.01	-0.01	0.01	0.04	0.09	0.00	0.00	-0.02	0.09	-0.17	0.01	0.39	0.78	1.00	0.46	0.36	0.37	0.15	0.37	-0.01	-0.01	0.02	0.02
	ROE	0.41	0.20	-0.13	-0.14	-0.13	0.00	0.02	0.05	0.02	0.03	0.03	0.15	-0.56	0.00	0.34	0.39	0.46	1.00	0.83	0.86	0.35	0.70	-0.08	-0.09	0.14	0.15
盈利	ROA	0.30	0.08	-0.05	-0.05	-0.04	0.00	0.02	0.05	0.01	0.01	0.00	0.21	-0.46	0.16	0.29	0.38	0.36	0.83	1.00	0.96	0.44	0.70	-0.38	-0.37	0.32	0.36
	ROIC	0.35	0.12	-0.09	-0.09	-0.08	0.00	0.02	0.05	0.01	0.01	0.00	0.22	-0.50	0.12	0.29	0.39	0.37	0.86	0.96	1.00	0.40	0.69	-0.31	-0.31	0.28	0.31
	毛利率	0.11	0.04	-0.01	-0.01	0.00	0.00	0.01	0.03	-0.02	-0.02	-0.04	0.22	-0.10	0.58	0.08	0.15	0.15	0.35	0.44	0.40	1.00	0.63	-0.38	-0.32	0.33	0.30
	净利率	0.33	0.16	-0.08	-0.09	-0.08	0.00	0.00	0.02	0.05	0.07	0.08	0.05	-0.38	0.33	0.23	0.32	0.37	0.70	0.70	0.69	0.63	1.00	-0.31	-0.27	0.32	0.34
偿付能力	D2E	0.16	0.14	-0.16	-0.17	-0.19	0.00	0.00	-0.01	-0.01	-0.01	-0.01	-0.05	-0.12	-0.50	0.03	-0.01	-0.01	-0.08	-0.38	-0.31	-0.38	-0.31	1.00	0.92	-0.74	-0.77
	A2D	0.15	0.14	-0.14	-0.15	-0.16	0.00	0.00	0.00	-0.01	-0.02	-0.01	-0.03	-0.13	-0.46	0.03	-0.02	-0.01	-0.09	-0.37	-0.31	-0.32	-0.27	0.92	1.00	-0.58	-0.63
	流动	-0.11	-0.07	0.15	0.17	0.19	0.01	0.01	0.02	0.00	0.01	0.00	0.10	0.06	0.41	0.03	0.03	0.02	0.14	0.32	0.28	0.33	0.32	-0.74	-0.58	1.00	0.89
	速动	-0.09	-0.06	0.14	0.16	0.18	0.00	0.00	0.01	0.01	0.02	0.02	0.10	0.06	0.39	0.03	0.02	0.02	0.15	0.36	0.31	0.30	0.34	-0.77	-0.63	0.89	1.00

资料来源：Wind，海通证券研究所

从上表不难发现，对于同一类因子，各因子之间的截面相关性极高。例如，对于盈利类的因子，各指标之间的截面相关性普遍高于 0.50，甚至有的指标之间截面线性相关性均值高于 0.80。若想在构建多因子模型时将这一类因子的信息全部纳入到模型中，投资者需要通过因子降维来进行信息的精炼。简单来说，因子降维是将高相关因子加权计算为复合因子的过程。

对于某一类因子 i ，因子降维的实现方法可具体表述为下式：

$$\text{因子}_i = \sum_{k=1}^K w_k^i * \text{子因子}_k^i$$

其中，因子 i 为降维后的因子，子因子 k 为降维前类别内的各因子， w_k^i 为降维权重。降维权重是因子降维的关键，不同的降维方法会给出不同的降维权重，从而影响降维后因子的表现以及模型最终的表现。本报告主要从以下三种思路出发，构建了相关降维方法：

1) 选取类别内选股效果最强的因子。由于 IC 是一个较好的衡量因子选股能力的指

标，因此可在同一类选股因子中选取 IC 均值绝对值最高的因子。

2) 按照类别内各因子选股能力加权。由于 IC 是一个较好的衡量因子选股能力的指标，对于同一类因子，可基于各因子的 IC 进行加权处理。

3) 最大化解释类别内各因子选股能力波动。可基于各因子的 IC 序列进行主成分分析，并选取第一主成分的权重作为该类因子的降维权重。

本报告将在后文中对于上述三种降维方法下的模型表现进行回测分析，并给出相关改进方法。

2. 降维前后模型表现对比

本章主要对于不同降维方法下，降维前后的多因子模型的表现进行对比。在进行回测前，首先需要构建因子库。下表给出了本报告所使用的因子库中，各因子的计算方法。（在实际操作中，投资者可进一步扩充因子库。）

表 2 因子计算方法说明

因子	子因子	因子说明
市值	市值	股票总市值
非线性	市值平方	对数总市值截面标准化后的平方
换手	1M	股票前 1 个月日均换手率
	3M	股票前 3 个月日均换手率
	6M	股票前 6 个月日均换手率
反转	1M	股票前 1 个月涨跌幅
	3M	股票前 3 个月涨跌幅
	6M	股票前 6 个月涨跌幅
特异度	1M	股票前 1 个月日收益 Fama-French 三因素回归 R 方
	3M	股票前 3 个月日收益 Fama-French 三因素回归 R 方
	6M	股票前 6 个月日收益 Fama-French 三因素回归 R 方
估值	PE	股票总市值 / (归属母公司股东的净利润 TTM)
	PB	股票总市值 / 最新公告股东权益(不含少数股东权益)
	PS	股票总市值 / 营业总收入(TTM)
成长	营收	公司营业收入同比增速
	营业利润	公司营业利润同比增速
	净利润	公司净利润同比增速
盈利	ROE	归属母公司股东净利润*2 / (期初归属母公司股东权益+期末归属母公司股东权益)
	ROA	净利润(含少数股东损益)*2 / (期初总资产+期末总资产)
	ROIC	归属于母公司股东的净利润*2 / (期初全部投入资本+期末全部投入资本)
	毛利率	(营业收入-营业成本) / 营业收入
	净利率	净利润 / 营业收入
偿付	D2E	负债合计 / 股东权益总额
	A2D	资产合计/负债合计
	流动	流动资产 / 流动负债
	速动	(流动资产 - 存货净额) / 流动负债

资料来源：海通证券研究所

本报告使用 2008 年 12 月 31 日至 2017 年 7 月 31 日之间的数据进行回测。在设定选股范围时剔除 ST、停牌以及上市不满 6 个月的股票。

在构建多因子模型时，使用最大化复合因子单期 IC 的方法。该方法的具体实现细节可参考系列前期专题报告《从最大化复合因子单期 IC 角度看因子权重》。

在进行模型对比时，未降维模型使用因子库中的所有因子进行最大化单期复合因子加权构建复合因子。而降维模型则先进行降维，并在降维后进行因子正交以及最大化单期复合因子加权。因子正交方法的相关细节可参考系列前期专题报告《选股因子系列研究(十七)——选股因子的正交》。

需要注意的是，降维前的模型未进行因子间的正交处理，而降维后的模型进行了因子正交处理。之所以对于未降维模型不进行因子正交，是因为在因子数量过多且因子间存在较高相关性的情况下，逐步正交的方法实现起来存在诸多困难。面对数量较多的选股因子时，投资者面临的选择往往是：“不降维+不正交 VS 降维+正交”。因子降维不可避免地会带来信息的损耗，但也会使得因子间的正交成为可能。

2.1 选取类别内选股效果最强

基于“选取类别内最强”这一思路，可使用 IC 衡量因子选股效果的强弱。故而可使用以下方法确定因子降维权重：

- 1) 对于第 i 类因子，计算其中每个因子历史 N 期 IC 的均值；
- 2) 取均值绝对值最大的因子的降维权重为 1，类别内其他因子的降维权重为 0；
- 3) 基于各因子的因子值与降维权重计算降维后的因子 i 。

本报告在回测时， N 取值为 24，也即比较因子 IC 时使用的是过去 24 个月 IC 的均值。下表给出了降维前后，复合因子的表现以及月度 Top Bottom 10% 组合的多空收益。

表 3 “取类别内选股效果最强”降维前后因子模型表现

	类型	IC	ICIR	胜率	显著比例	多空	多头	空头
降维前	Pearson	0.087	3.195	86.4%	68.9%	3.05%	1.55%	-1.50%
	Spearman	0.107	3.492	86.4%	76.7%			
降维后	Pearson	0.098	3.346	83.5%	72.8%	3.62%	2.03%	-1.59%
	Spearman	0.119	3.619	87.4%	78.6%			

资料来源：Wind，海通证券研究所

观察上表可知，降维后的复合因子相比于未降维的复合因子有更好的选股能力。复合因子 IC、Rank IC、ICIR 以及 Rank ICIR 都出现了上升，且因子 TOP BOTTOM 10% 多空收益、多头收益以及空头收益的绝对值也在降维后显著上升。

下表给出了各类因子内部降维权重的分配情况。观察历史权重可以发现：

- 1) 流动性类因子中，1 个月日均换手率以及 6 个月日均换手率被选出的频率最高。其中，6 个月日均换手率在 2010 年下半年以及 2014 年至 2015 年上半年中被选出；
- 2) 反转类因子中，1 个月反转以及 3 个月反转出现频率较高。其中，3 个月反转在 2015 年下半年以来持续被选出；
- 3) 波动率类因子中，1 个月回归 R 方在整个回测区间持续被选出；
- 4) 估值类因子中，PB 以及 PS_TTM 被选出的频率较高。其中，PS_TTM 在 2010 年初至 2012 年上半年以及 2013 年下半年至 2015 年上半年持续被选出；
- 5) 成长能力类因子中，营业收入同比增速、营业利润同比增速以及净利润同比增速出现频率较为接近；

关性，直接计算出的 IC 可比性会受到影响。考虑到这一问题，可将类别内因子对于前序降维后的大类因子进行正交处理，从而使得因子 IC 更加具有可比性。

在实际操作中可按照以下流程进行降维权重的确定：

- 1) 对于第 i 类因子，将类别中的每个因子与前序大类降维后的因子进行正交处理，计算得到正交后的因子值；
- 2) 计算各正交后的因子在过去 N 期的 IC 均值；
- 3) 取均值绝对值最大的因子的降维权重为 1，其他类别内因子降维权重为 0；
- 4) 基于正交后的因子值与降维权重计算降维后的因子 i 。

本报告在回测时， N 取值为 24，也即比较因子 IC 时使用的是过去 24 个月 IC 的均值。下表给出了降维前后，复合因子的表现以及月度 Top Bottom 10% 组合的多空收益。

表 5 “取类别内选股效果最强”（改进）降维前后因子模型表现

	类型	IC	ICIR	胜率	显著比例	多空	多头	空头
降维前	Pearson	0.087	3.195	86.4%	68.9%	3.05%	1.55%	-1.50%
	Spearman	0.107	3.492	86.4%	76.7%			
降维后	Pearson	0.098	3.346	83.5%	72.8%	3.62%	2.03%	-1.59%
	Spearman	0.119	3.619	87.4%	78.6%			
降维后 (改进)	Pearson	0.102	3.519	87.4%	73.8%	3.78%	2.05%	-1.73%
	Spearman	0.123	3.827	88.3%	77.7%			

资料来源：Wind，海通证券研究所

观察上表可知，在使用了正交因子 IC 后，降维后复合因子的选股能力得到了进一步的增强。复合因子 IC、Rank IC、ICIR 以及 Rank ICIR 都得到了进一步增强，且因子 TOP BOTTOM 10% 多空收益、多头收益以及空头收益的绝对值也在降维后有所提高。

下表给出了各类因子内部降维权重的分配情况。观察历史权重可以发现：

- 1) 流动性类指标中，1 个月日均换手率持续被选出；
- 2) 反转类指标中，1 个月反转与 3 个月反转被选出的频率较高。其中，3 个月反转因子在 2010 年下半年、2011 年上半年以及 2015 年初至 2016 年年底间持续被选出；
- 3) 波动率类指标中，1 个月回归 R 方持续被选出；
- 4) 估值类指标中，PE_TTM 与 PB 被选出的频率较高。其中，PB 在 2009 年至 2010 年初以及 2015 年初至 2015 年末间持续被选出；
- 5) 成长能力类指标中，营业收入同比增速持续被选出；
- 6) 盈利能力类指标中，ROE、毛利率与净利率被选出的频率都较高；
- 7) 清偿能力类指标中，D2E 与速动比率被选出的频率较高。

2) 按照均值的大小进行权重分配, 降维权重 w 的具体计算方法如下式:

$$w_k = \frac{IC_i}{\sum_{k=1}^K |IC_i|}$$

3) 基于各因子的因子值与降维权重计算降维后的因子 i 。

本报告在回测时, N 取值为 24, 也即比较因子 IC 时使用的是过去 24 个月 IC 的均值。下表给出了降维前后, 复合因子的表现以及月度 Top Bottom 10% 组合的多空收益。

表 7 “类别内选股效果加权”降维前后因子模型表现

	类型	IC	ICIR	胜率	显著比例	多空	多头	空头
降维前	Pearson	0.087	3.195	86.4%	68.9%	3.05%	1.55%	-1.50%
	Spearman	0.107	3.492	86.4%	76.7%			
降维后	Pearson	0.095	3.174	81.6%	70.9%	3.43%	1.99%	-1.44%
	Spearman	0.117	3.515	88.3%	74.8%			

资料来源: Wind, 海通证券研究所

观察上表可知, 模型在进行了因子降维后, 选股能力得到了提升。复合因子 IC 、Rank IC 、Rank $ICIR$ 都出现了上升, 且因子 TOP BOTTOM 10% 多空收益、多头收益在降维后有所上升。值得注意的是, 模型月度多空收益的改善更多来源于多头收益的提升, 而非空头收益的改善。

下表给出了各类因子内部降维权重的分配情况。观察历史权重可以发现:

- 1) 流动性类因子中, 1 个月日均换手率在大部分时间具有较高的降维权重, 3 个月日均换手率次之。6 个月日均换手率因子在 2014 年年初至 2015 年年中具有较高的降维权重。
- 2) 反转类因子中, 1 个月反转以及 3 个月反转的降维权重相对较高。2014 年年初前, 1 个月反转的降维权重较高, 而 2014 年年初以来, 3 个月反转的降维权重普遍高于 40%。
- 3) 波动率类因子中, 1 个月回归 R 方具有较高的降维权重, 降维权重接近 50%;
- 4) 估值类因子中, PB 以及 PS_TTM 具有较高的降维权重。其中, PS_TTM 在 2013 年下半年至 2015 年上半年中持续被选出;
- 5) 成长能力类因子中, 营业收入同比增速、营业利润同比增速以及净利润同比增速在大部分时间段上的降维权重较为接近;
- 6) 盈利能力类因子中, ROE 与毛利率在部分时间段上具有较高的降维权重。其中, ROE 在 2014 年年中至 2016 年年末的降维权重较高, 毛利率在 2012 年年中至 2014 年年末的降维权重较高;
- 7) 清偿能力类因子中, 各个因子降维权重分配较为均衡。

表 8 “类别内选股效果加权” 降维权重分配情况

代码	市值	非线性	换手			反转			动量			PB	估值			成长			盈利			ROE			ROA			ROIC			毛利			净利			D/E			清偿能力		
			市值	市值平方	换手1M	换手3M	换手6M	反转1M	反转3M	反转6M	动量1M		动量3M	动量6M	PE	TTM	PS	TTM	营业收入	净利润	净利润	ROE	ROA	ROIC	毛利率	净利率	ROE	ROA	ROIC	毛利率	净利率	D/E	A2D	流动比率	速动比率							
20081231	100%	100%	53%	33%	14%	34%	34%	32%	43%	30%	27%	65%	-3%	-9%	13%	26%	-61%	27%	25%	28%	-1%	20%	0%	12%	49%	33%																
20090123	100%	100%	65%	31%	-4%	33%	34%	33%	43%	30%	27%	61%	-9%	30%	18%	-21%	-61%	26%	25%	28%	-2%	20%	3%	12%	49%	33%																
20090227	100%	100%	62%	30%	-4%	33%	32%	32%	44%	30%	26%	58%	-11%	31%	20%	-25%	-55%	25%	26%	27%	-2%	20%	3%	12%	49%	33%																
20090331	100%	100%	68%	27%	-4%	34%	32%	31%	43%	29%	28%	56%	-14%	30%	13%	-20%	-67%	24%	25%	29%	7%	15%	3%	12%	49%	33%																
20090430	100%	100%	61%	30%	-9%	33%	32%	30%	42%	30%	28%	51%	-20%	30%	15%	3%	-82%	13%	19%	22%	4%	15%	-1%	1%	49%	33%																
20090527	100%	100%	73%	27%	0%	33%	32%	31%	41%	31%	28%	56%	-14%	31%	23%	28%	-48%	25%	26%	27%	1%	20%	3%	12%	49%	33%																
20090630	100%	100%	69%	11%	20%	34%	32%	30%	38%	32%	30%	63%	-7%	30%	29%	35%	-37%	25%	28%	28%	-1%	17%	6%	2%	49%	33%																
20090731	100%	100%	73%	25%	2%	34%	32%	31%	37%	32%	31%	60%	-10%	30%	22%	40%	-38%	25%	29%	29%	-3%	15%	5%	2%	49%	33%																
20090831	100%	100%	73%	26%	1%	33%	32%	32%	41%	31%	29%	58%	-11%	31%	22%	40%	-35%	25%	26%	27%	1%	21%	-2%	23%	49%	33%																
20090930	100%	100%	68%	27%	-5%	33%	31%	31%	43%	31%	27%	58%	-11%	30%	19%	43%	-38%	24%	24%	24%	8%	20%	-2%	17%	49%	33%																
20091030	100%	100%	68%	16%	31%	33%	32%	32%	43%	30%	27%	61%	-4%	35%	22%	41%	-37%	27%	24%	24%	2%	24%	-22%	21%	23%	28%	33%															
20091130	100%	100%	66%	23%	11%	33%	32%	31%	44%	30%	27%	58%	-4%	34%	15%	43%	-42%	24%	21%	24%	7%	24%	-22%	23%	31%	24%	33%															
20091231	100%	100%	68%	24%	8%	33%	31%	32%	45%	28%	27%	57%	-9%	34%	13%	42%	-44%	26%	22%	25%	1%	24%	-20%	19%	33%	43%	33%															
20100129	100%	100%	66%	15%	38%	33%	32%	30%	47%	29%	25%	58%	-8%	34%	7%	44%	-48%	26%	19%	23%	8%	24%	-24%	22%	27%	27%	33%															
20100226	100%	100%	72%	21%	7%	33%	32%	30%	49%	28%	23%	61%	-9%	30%	2%	44%	-54%	25%	17%	20%	1%	20%	-22%	20%	37%	27%	33%															
20100331	100%	100%	69%	12%	12%	33%	32%	31%	47%	30%	22%	58%	-12%	31%	8%	39%	-53%	26%	16%	20%	1%	27%	-22%	19%	31%	29%	33%															
20100430	100%	100%	70%	16%	14%	33%	32%	31%	47%	30%	23%	60%	-3%	37%	20%	33%	-47%	29%	17%	19%	5%	29%	-24%	21%	27%	28%	33%															
20100531	100%	100%	66%	12%	22%	33%	32%	30%	51%	29%	20%	57%	-8%	35%	13%	36%	-51%	28%	-3%	-9%	27%	32%	-22%	28%	28%	33%																
20100630	100%	100%	63%	-7%	31%	33%	32%	31%	49%	30%	21%	55%	-13%	32%	30%	28%	-42%	27%	-5%	10%	29%	33%	-23%	20%	33%	27%	33%															
20100730	100%	100%	67%	14%	14%	33%	32%	33%	47%	30%	22%	51%	-17%	32%	3%	20%	-42%	25%	11%	13%	17%	33%	-22%	19%	33%	43%	33%															
20100831	100%	100%	65%	25%	25%	33%	31%	32%	48%	28%	16%	52%	-4%	44%	41%	26%	-33%	27%	13%	3%	40%	33%	-22%	21%	29%	27%	33%															
20100930	100%	100%	61%	31%	21%	33%	32%	31%	49%	28%	27%	41%	-13%	34%	15%	42%	-44%	26%	22%	25%	1%	24%	-20%	19%	33%	43%	33%															
20101029	100%	100%	69%	12%	12%	33%	32%	31%	47%	30%	22%	58%	-12%	31%	8%	39%	-53%	26%	16%	20%	1%	27%	-22%	19%	31%	29%	33%															
20101130	100%	100%	14%	34%	34%	33%	32%	31%	62%	27%	12%	34%	-6%	60%	37%	1%	2%	11%	4%	8%	3%	64%	-23%	24%	26%	29%	33%															
20101231	100%	100%	12%	38%	38%	33%	32%	31%	59%	28%	17%	37%	-12%	51%	34%	6%	11%	7%	5%	0%	16%	72%	-20%	24%	29%	33%																
20110131	100%	100%	66%	7%	27%	33%	32%	31%	49%	28%	16%	37%	-13%	50%	33%	1%	-10%	12%	13%	12%	15%	49%	-19%	25%	29%	33%																
20110228	100%	100%	38%	22%	40%	34%	32%	31%	57%	29%	14%	31%	-2%	67%	66%	14%	20%	11%	-7%	-9%	18%	54%	-21%	29%	24%	26%	33%															
20110331	100%	100%	78%	-9%	13%	33%	32%	31%	59%	30%	11%	30%	-6%	65%	78%	11%	11%	13%	16%	15%	16%	41%	-14%	26%	8%	22%	33%															
20110429	100%	100%	73%	21%	-5%	33%	32%	31%	60%	30%	10%	30%	-6%	64%	62%	18%	20%	12%	17%	17%	35%	25%	12%	33%	30%	33%																
20110531	100%	100%	56%	27%	17%	33%	32%	31%	60%	30%	6%	22%	-14%	62%	59%	27%	24%	1%	15%	15%	42%	23%	11%	35%	29%	33%																
20110630	100%	100%	80%	20%	1%	33%	32%	31%	72%	27%	-1%	14%	-11%	75%	44%	32%	24%	-6%	3%	12%	22%	47%	-2%	48%	30%	33%																
20110729	100%	100%	79%	15%	6%	33%	32%	31%	72%	13%	-15%	10%	-2%	88%	36%	36%	28%	-7%	6%	-5%	11%	70%	-12%	50%	33%	33%																
20110831	100%	100%	75%	-9%	1%	33%	32%	31%	72%	12%	-6%	12%	5%	83%	35%	37%	28%	15%	2%	5%	0%	59%	-19%	52%	6%	2%	33%															
20110930	100%	100%	77%	20%	3%	33%	32%	31%	64%	27%	9%	17%	7%	77%	33%	40%	27%	-9%	31%	10%	14%	55%	-5%	43%	40%	-11%	33%															
20111031	100%	100%	76%	21%	-4%	33%	32%	30%	60%	28%	11%	25%	-2%	73%	34%	39%	27%	13%	14%	14%	-4%	54%	-20%	46%	21%	13%	33%															
20111130	100%	100%	73%	20%	-7%	34%	32%	30%	75%	24%	0%	22%	37%	51%	36%	36%	28%	4%	15%	3%	25%	43%	-20%	36%	10%	28%	33%															
20111230	100%	100%	57%	26%	3%	33%	32%	31%	68%	30%	9%	4%	25%	61%	34%	35%	31%	14%	22%	22%	3%	0%	24%	33%	14%	29%	33%															
20120131	100%	100%	51%	28%	21%	33%	32%	31%	68%	30%	14%	14%	32%	48%	32%	38%	34%	18%	3%	0%	18%	3%	23%	31%	23%	33%																
20120229	100%	100%	53%	28%	19%	33%	32%	30%	54%	31%	15%	18%	35%	47%	31%	36%	32%	19%	1%	19%	7%	32%	24%	22%	22%	22%	33%															
20120330	100%	100%	45%	31%	24%	33%	32%	31%	59%	30%	15%	16%	39%	45%	32%	36%	33%	33%	25%	1%	-4%	30%	19%	25%	26%	29%	33%															
20120427	100%	100%	45%	30%	25%	33%	32%	31%	51%	31%	18%	28%	34%	39%	28%	37%	35%	43%	13%	9%	14%	21%	28%	22%	21%	29%	33%															
20120531	100%	100%	46%	30%	24%	33%	32%	31%	47%	32%	22%	34%	32%	34%	24%	39%	36%	36%	32%	-7%	2%	23%	29%	20%	29%	33%																
20120629	100%	100%	46%	30%	24%	33%	32%	31%	47%	32%	21%	34%	33%	32%	24%	40%	36%	41%	1%	1%	40%	1%	22%	20%	29%	33%																
20120731	100%	100%	44%	31%	26%	33%	32%	31%	50%	32%	18%	28%	44%	28%	28%	40%	36%	42%	1%	1%	40%	1%	22%	20%	29%	33%																
20120831	100%	100%	45%	31%	25%	33%	32%	31%	49%	33%	21%	39%	42%	19%	31%	35%	34%	42%	1%	1%	40%	1%	22%	20%	29%	33%																
20120930	100%	100%	42%	32%	22%	33%	32%	31%	45%	33%	22%	43%	42%	15%	30%	35%	34%	42%	1%	1%	40%	1%	22%	20%	29%	33%																
20121031	100%	100%	46%	31%	24%	33%	32%	31%	49%	33%	21%	39%	42%	17%	31%	35%	34%	42%	1%	1%	40%	1%	22%	20%	29%	33%																
20121130	100%	100%	40%	32%	27%	33%	32%	31%	44%	33%	23%	42%	36%	22%	29%	36%	35%	45%	1%	1%	40%	1%	22%	20%	19%	27%	33%															
20121231	100%	100%	42%	32%	26%	33%	32%	31%	45%	33%	21%	42%	36%	19%	24%	38%	38%	44%	1%	1%	40%	1%	22%	20%	16%	27%	33%															
20130131	100%	100%	44%	32%	24%	33%	32%	31%	47%	33%	18%	46%	44%	0%	30%	38%	32%	42%	1%	1%	40%	1%	22%	20%	12%	28%	33%															
20130228	100%	100%	43%	32%	24%	33%	32%	31%	47%	33%	18%	48%	52%	0%	33%	39%	31%	45%	1%</																							

正交后的因子值;

2) 计算各正交后的因子在过去 N 期的 IC 均值;

3) 按照均值的大小进行权重分配, 降维权重 w 具体计算方法如下式:

$$w_k = \frac{IC_i}{\sum_{k=1}^K |IC_i|}$$

4) 基于正交后的因子值与降维权重计算降维后的因子 i 。

本报告在回测时, N 取值为 24, 也即比较因子 IC 时使用的是过去 24 个月 IC 的均值。下表给出了降维前后, 复合因子的表现以及月度 Top Bottom 10% 组合的多空收益。

表 9 “类别内选股效果加权”(改进)降维前后因子模型表现

	类型	IC	ICIR	胜率	显著比例	多空	多头	空头
降维前	Pearson	0.087	3.195	86.4%	68.9%	3.05%	1.55%	-1.50%
	Spearman	0.107	3.492	86.4%	76.7%			
降维后	Pearson	0.095	3.174	81.6%	70.9%	3.43%	1.99%	-1.44%
	Spearman	0.117	3.515	88.3%	74.8%			
降维后 (改进)	Pearson	0.100	3.369	84.5%	70.9%	3.59%	2.05%	-1.54%
	Spearman	0.122	3.707	88.3%	78.6%			

资料来源: Wind, 海通证券研究所

通过使用正交因子, 在类别内分配因子权重时能够剔除其他类别因子的影响, 从而得到更好的分配结果。从复合因子的角度看, 降维后复合因子的 IC、ICIR、胜率以及显著比例都优于改进前的降维方法以及未降维的模型。

下表给出了各类因子内部降维权重的分配情况。观察历史权重可以发现:

- 1) 流动性类因子中, 1 个月日均换手率的降维权重相对更高, 3 个月日均换手率次之;
- 2) 反转类因子中, 1 个月反转以及 3 个月反转的降维权重相对较高;
- 3) 波动率类因子中, 1 个月回归 R 方在整个回测区间都具有较高的降维权重, 权重在 50%~70% 之间波动;
- 4) 估值类因子中, PB 在 2009 年年初至 2010 年年中以及 2015 年具有较高的降维权重, PE_TTM 在 2010 年年中至 2012 年年末具有较高的降维权重, PS_TTM 在 2014 年具有较高的降维权重;
- 5) 成长能力类因子中, 营业收入同比增速的降维权重相对较高, 营业利润同比增速与净利润同比增速次之;
- 6) 盈利能力类因子中, ROE 与毛利率具有较高的降维权重, ROA 与 ROIC 在 2015 年年中至 2016 年年中具有较高的降维权重;
- 7) 清偿能力类因子中, 各因子降维权重较为接近。

表 10 “类别内选股效果加权”（改进）降维权重分配情况

	市值	非线性	换手			反转			特异度			估值			成长			盈利			ROE			ROA			ROIC			毛利率			净利润率			D2E			A2D			流动比率			速动比率																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																										
			市值平方	换手1M	换手3M	换手6M	反转1M	反转3M	反转6M	特异度1M	特异度3M	特异度6M	PB	PE	TTM	PS	TTM	营业收入	营业收入	净利润	净利润	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA	ROIC	ROE	ROA

2.5 类别内 PCA 降维

基于“最大化解释类别内因子选股效果波动”这一思路，可使用主成分分析进行因子降维。在实际操作中，可对于因子的历史 IC 序列进行 PCA 分析，并取第一主成分权重作为降维权重。具体步骤如下所示：

- 1) 对于第 i 类因子，计算其中每个因子历史 N 期 IC 的均值；
- 2) 对于类别内因子 IC 序列进行主成分分析，并取第一主成分权重 w' ，该类别各因子降维权重计算方法如下：

$$w_k = \frac{w'_k}{\sum_{k=1}^K |w'_k|}$$

- 3) 基于各因子的因子值与降维权重计算降维后的因子 i。

本报告在回测时，N 取值为 24，也即比较因子 IC 时使用是过去 24 个月 IC 的均值。下表给出了降维前后，复合因子的表现以及月度 Top Bottom 10% 组合的多空收益。

表 11 “类别内 PCA 第一主成分”降维前后因子模型表现

	类型	IC	ICIR	胜率	显著胜率	多空	多头	空头
降维前	Pearson	0.087	3.195	86.4%	68.9%	3.05%	1.55%	-1.50%
	Spearman	0.107	3.492	86.4%	76.7%			
降维后	Pearson	0.094	3.148	80.6%	68.9%	3.43%	1.99%	-1.44%
	Spearman	0.115	3.467	87.4%	75.7%			

资料来源：Wind，海通证券研究所

在 PCA 方法下，降维后的复合因子提升幅度有限，仅在因子的 IC 均值、多空收益率上有所提升，而因子 ICIR 相比于降维前出现了回落。

下表给出了各类因子内部降维权重的分配情况。在 PCA 降维法下，各类因子的降维权重分配较为均衡，并无明显区别。

表 12 “类别内 PCA 第一主成分” 降维权重分配情况

	市值	非线性	换手			反转			特异度			估值			成长			盈利			清偿能力						
			市值平方	换手1M	换手3M	换手6M	反转1M	反转3M	反转6M	特异度1M	特异度3M	特异度6M	PB	PE TTM	PS TTM	营业收入	营业利润	净利润	ROE	ROA	ROIC	毛利率	净利率	D2E	A2D	流动性	速动比
20081231	100%	100%	35%	34%	31%	27%	38%	34%	27%	35%	38%	46%	-28%	26%	31%	33%	36%	25%	23%	23%	10%	19%	26%	24%	21%	28%	28%
20090123	100%	100%	35%	34%	31%	27%	38%	35%	27%	35%	38%	48%	-26%	26%	30%	33%	37%	25%	23%	23%	10%	19%	26%	24%	21%	28%	28%
20090227	100%	100%	35%	34%	31%	29%	38%	34%	27%	35%	38%	47%	-27%	26%	30%	34%	36%	26%	23%	23%	9%	19%	26%	24%	21%	29%	29%
20090331	100%	100%	34%	35%	32%	28%	39%	33%	27%	35%	38%	47%	-25%	28%	30%	34%	36%	26%	23%	23%	10%	19%	26%	24%	21%	29%	29%
20090430	100%	100%	33%	34%	32%	28%	38%	34%	27%	35%	38%	48%	-23%	29%	30%	34%	36%	26%	23%	23%	9%	19%	26%	24%	21%	29%	29%
20090527	100%	100%	33%	35%	32%	27%	38%	34%	27%	35%	38%	47%	-24%	29%	30%	35%	36%	26%	23%	23%	10%	19%	26%	24%	21%	29%	29%
20090630	100%	100%	32%	35%	33%	28%	38%	34%	29%	34%	37%	51%	-15%	34%	29%	35%	36%	25%	23%	22%	10%	20%	26%	24%	21%	29%	29%
20090731	100%	100%	33%	35%	33%	26%	38%	36%	27%	34%	39%	51%	-13%	35%	28%	35%	36%	25%	23%	22%	10%	20%	26%	24%	21%	29%	29%
20090831	100%	100%	33%	35%	33%	25%	38%	36%	26%	34%	39%	52%	-14%	35%	28%	35%	37%	25%	23%	22%	10%	20%	26%	24%	21%	29%	29%
20090930	100%	100%	33%	35%	33%	25%	39%	37%	26%	34%	40%	53%	-15%	34%	29%	35%	38%	25%	23%	22%	10%	20%	26%	24%	21%	29%	29%
20091030	100%	100%	32%	35%	33%	26%	38%	36%	25%	35%	40%	57%	-7%	36%	29%	35%	36%	25%	23%	23%	10%	19%	25%	23%	21%	30%	30%
20091130	100%	100%	33%	35%	32%	26%	41%	33%	22%	36%	42%	53%	10%	37%	28%	34%	38%	24%	23%	23%	10%	19%	25%	23%	22%	30%	30%
20091231	100%	100%	33%	35%	32%	26%	41%	33%	22%	36%	42%	50%	12%	38%	28%	34%	38%	24%	23%	23%	10%	20%	24%	22%	23%	31%	31%
20100129	100%	100%	33%	35%	33%	26%	40%	34%	24%	35%	41%	50%	12%	38%	28%	34%	37%	24%	23%	23%	11%	19%	26%	24%	21%	30%	30%
20100226	100%	100%	32%	35%	33%	27%	38%	35%	28%	35%	38%	42%	24%	34%	28%	34%	38%	25%	23%	23%	10%	19%	26%	24%	20%	30%	30%
20100331	100%	100%	32%	35%	33%	26%	38%	35%	28%	35%	38%	44%	22%	34%	29%	34%	37%	24%	23%	23%	11%	20%	26%	24%	21%	30%	30%
20100430	100%	100%	32%	35%	33%	25%	38%	37%	27%	35%	38%	46%	22%	32%	28%	36%	36%	22%	23%	24%	12%	19%	26%	24%	20%	31%	31%
20100531	100%	100%	32%	35%	32%	26%	38%	36%	27%	35%	38%	50%	18%	32%	29%	35%	36%	22%	23%	24%	13%	18%	26%	24%	20%	30%	30%
20100630	100%	100%	32%	35%	32%	26%	38%	36%	27%	35%	38%	50%	20%	31%	31%	34%	35%	22%	23%	24%	13%	18%	26%	24%	20%	30%	30%
20100730	100%	100%	32%	35%	32%	22%	39%	39%	27%	35%	38%	48%	21%	31%	30%	34%	36%	22%	24%	24%	14%	17%	26%	24%	20%	30%	30%
20100831	100%	100%	33%	35%	32%	21%	38%	41%	27%	35%	38%	48%	24%	27%	30%	35%	35%	22%	24%	24%	13%	17%	26%	24%	19%	30%	30%
20100930	100%	100%	33%	35%	32%	22%	37%	40%	28%	35%	37%	49%	24%	27%	32%	34%	34%	22%	24%	24%	14%	17%	26%	24%	21%	30%	30%
20101029	100%	100%	32%	35%	32%	23%	37%	40%	27%	35%	37%	49%	25%	28%	32%	33%	34%	22%	24%	24%	13%	17%	26%	24%	21%	30%	30%
20101130	100%	100%	32%	36%	33%	25%	35%	40%	27%	35%	38%	46%	26%	28%	34%	33%	33%	21%	24%	24%	14%	16%	26%	24%	21%	29%	29%
20101231	100%	100%	32%	35%	33%	25%	35%	40%	27%	35%	38%	45%	28%	27%	34%	33%	33%	22%	24%	24%	13%	17%	27%	24%	20%	28%	28%
20110131	100%	100%	31%	35%	34%	22%	38%	40%	26%	36%	38%	44%	28%	28%	33%	32%	33%	21%	24%	23%	14%	17%	27%	24%	21%	28%	28%
20110228	100%	100%	30%	35%	34%	21%	39%	40%	26%	35%	38%	42%	31%	27%	33%	32%	32%	21%	23%	23%	15%	17%	26%	24%	22%	28%	28%
20110331	100%	100%	30%	35%	35%	22%	37%	41%	27%	36%	38%	42%	30%	28%	33%	33%	33%	20%	24%	24%	15%	17%	26%	24%	22%	28%	28%
20110429	100%	100%	29%	35%	36%	23%	38%	38%	27%	36%	37%	42%	29%	29%	34%	33%	33%	20%	24%	24%	15%	17%	26%	24%	23%	28%	28%
20110531	100%	100%	29%	35%	36%	25%	37%	38%	26%	36%	38%	42%	29%	29%	37%	31%	33%	20%	24%	24%	15%	17%	26%	24%	23%	28%	28%
20110630	100%	100%	29%	35%	36%	25%	37%	37%	27%	36%	37%	42%	29%	29%	36%	31%	33%	20%	24%	24%	15%	17%	26%	24%	23%	28%	28%
20110729	100%	100%	29%	35%	36%	26%	36%	38%	28%	35%	37%	41%	28%	31%	38%	30%	32%	19%	24%	23%	16%	17%	26%	24%	23%	27%	27%
20110831	100%	100%	29%	35%	36%	27%	36%	37%	29%	35%	36%	41%	28%	31%	40%	29%	32%	19%	24%	23%	17%	17%	26%	24%	23%	27%	27%
20110930	100%	100%	30%	35%	36%	27%	36%	37%	29%	35%	37%	41%	27%	31%	37%	30%	32%	18%	24%	23%	18%	17%	26%	24%	23%	27%	27%
20111031	100%	100%	30%	35%	36%	27%	36%	37%	29%	35%	37%	41%	27%	31%	37%	30%	32%	17%	24%	23%	18%	17%	26%	24%	23%	27%	27%
20111130	100%	100%	30%	35%	35%	27%	37%	37%	28%	35%	37%	40%	27%	32%	38%	30%	32%	16%	24%	22%	20%	17%	26%	24%	23%	27%	27%
20111230	100%	100%	30%	35%	35%	26%	37%	37%	26%	36%	38%	40%	31%	30%	37%	30%	33%	18%	23%	22%	19%	18%	26%	24%	23%	27%	27%
20120131	100%	100%	30%	35%	35%	23%	37%	40%	26%	36%	38%	38%	30%	32%	37%	30%	33%	17%	23%	22%	20%	19%	26%	24%	23%	25%	27%
20120229	100%	100%	30%	34%	36%	22%	38%	40%	24%	36%	39%	38%	30%	32%	36%	31%	34%	17%	23%	22%	20%	19%	26%	24%	23%	25%	27%
20120330	100%	100%	31%	34%	35%	21%	38%	41%	25%	36%	39%	38%	30%	32%	34%	32%	34%	17%	22%	22%	20%	19%	26%	24%	23%	25%	27%
20120427	100%	100%	31%	34%	35%	22%	38%	40%	25%	36%	40%	38%	30%	32%	36%	32%	33%	17%	23%	22%	20%	19%	26%	24%	23%	25%	27%
20120531	100%	100%	30%	34%	36%	18%	39%	43%	24%	36%	40%	36%	31%	33%	34%	33%	33%	17%	22%	22%	19%	20%	25%	23%	25%	27%	27%
20120629	100%	100%	30%	34%	36%	19%	39%	42%	24%	35%	40%	36%	30%	34%	35%	32%	33%	16%	22%	21%	21%	20%	25%	23%	25%	27%	27%
20120731	100%	100%	30%	34%	36%	22%	36%	42%	24%	35%	40%	36%	30%	34%	35%	32%	33%	16%	22%	21%	21%	20%	25%	23%	25%	27%	27%
20120831	100%	100%	30%	34%	36%	22%	38%	40%	25%	35%	40%	33%	29%	37%	33%	33%	34%	16%	22%	21%	21%	20%	25%	23%	25%	27%	27%
20120930	100%	100%	30%	34%	36%	21%	41%	38%	25%	35%	40%	33%	29%	37%	33%	33%	34%	16%	22%	21%	21%	20%	25%	23%	25%	27%	27%
20121031	100%	100%	30%	34%	36%	21%	40%	39%	25%	35%	40%	33%	29%	38%	31%	34%	36%	16%	22%	21%	21%	20%	25%	23%	25%	27%	27%
20121130	100%	100%	31%	34%	35%	22%	40%	38%	26%	35%	40%	34%	30%	37%	29%	35%	36%	16%	22%	21%	21%	20%	25%	23%	25%	27%	27%
20121231	100%	100%	31%	34%	35%	23%	40%	37%	26%	35%	39%	34%	29%	37%	26%	36%	38%	16%	22%	21%	20%	20%	25%	23%	25%	27%	27%
20130131	100%	100%	31%	34%	35%</																						

2.6 类别内 PCA 降维-改进

考虑到因子 IC 的可比性受到因子相关性的影响，可使用正交因子 IC 序列替代原始因子的 IC 序列作为因子选股效果的衡量指标。改进后的因子降维步骤如下所示：

- 1) 对于第 i 类因子，将类别中的每个因子与前序大类因子进行正交处理，计算得到正交后的因子值；
- 2) 计算各正交后的因子在过去 N 期的 IC 均值；
- 3) 对于类别内因子 IC 序列进行主成分分析并取第一主成分权重 w' ，该类别各因子降维权重计算方法如下：

$$w_k = \frac{w'_k}{\sum_{k=1}^K |w'_k|}$$

- 4) 基于各因子的因子值与降维权重计算降维后的因子 i。

本报告在回测时，N 取值为 24，也即比较因子 IC 时使用的是过去 24 个月 IC 的均值。下表给出了降维前后，复合因子的表现以及月度 Top Bottom 10% 组合的多空收益。

表 13 “类别内 PCA 第一主成分”（改进）降维前后因子模型表现

	类型	IC	ICIR	胜率	显著胜率	多空	多头	空头
降维前	Pearson	0.087	3.195	86.4%	68.9%	3.05%	1.55%	-1.50%
	Spearman	0.107	3.492	86.4%	76.7%			
降维后	Pearson	0.094	3.148	80.6%	68.9%	3.43%	1.99%	-1.44%
	Spearman	0.115	3.467	87.4%	75.7%			
降维后 (改进)	Pearson	0.094	3.129	80.6%	69.9%	3.37%	1.98%	-1.39%
	Spearman	0.114	3.442	87.4%	74.8%			

资料来源：Wind，海通证券研究所

使用正交因子后，降维后的复合因子选股效果反而出现降低，复合因子在 IC、ICIR、月度胜率以及多空收益上都弱于类别内 PCA 降维下的复合因子。

下表给出了各类因子内部降维权重的分配情况。在 PCA 降维法下，各类因子的降维权重分配较为均衡，并无明显区别。

表 14 “类别内 PCA 第一主成分” (改进) 降维权重分配情况

日期	市值	非线性	换手			反转			特异度			PB	估值			成长			盈利			清偿能力					
			市值平方	换手1M	换手3M	换手6M	反转1M	反转3M	反转6M	特异度1M	特异度3M		特异度6M	PE	TTM	PS	TTM	营业收入	营业利润	净利润	ROE	ROA	ROIC	毛利率	净利率	D/E	A2D
20081231	100%	100%	107%	37%	34%	29%	30%	31%	32%	25%	34%	40%	37%	48%	-16%	24%	38%	38%	21%	26%	26%	13%	14%	27%	26%	21%	26%
20090123	100%	100%	107%	37%	34%	29%	31%	36%	33%	25%	34%	41%	36%	49%	-15%	22%	39%	39%	21%	26%	26%	13%	14%	28%	26%	21%	25%
20090227	100%	100%	107%	36%	34%	30%	31%	36%	32%	25%	34%	41%	36%	47%	-17%	23%	38%	39%	20%	27%	26%	13%	15%	28%	26%	21%	25%
20090331	100%	100%	107%	35%	35%	30%	30%	37%	33%	25%	33%	42%	36%	46%	-18%	23%	38%	39%	20%	26%	25%	14%	14%	28%	26%	21%	25%
20090430	100%	100%	107%	35%	34%	31%	32%	36%	33%	23%	34%	43%	35%	45%	-20%	22%	38%	39%	20%	27%	26%	12%	14%	27%	25%	22%	26%
20090527	100%	100%	107%	35%	34%	31%	31%	36%	33%	24%	34%	42%	33%	50%	-16%	22%	39%	39%	19%	27%	26%	14%	14%	27%	24%	24%	25%
20090630	100%	100%	104%	34%	34%	31%	32%	36%	32%	24%	33%	43%	38%	40%	-21%	22%	38%	39%	22%	27%	26%	11%	14%	27%	24%	24%	25%
20090731	100%	100%	107%	35%	34%	31%	30%	36%	34%	22%	34%	44%	39%	40%	-22%	22%	40%	38%	22%	28%	27%	10%	13%	27%	25%	24%	24%
20090831	100%	100%	107%	34%	35%	31%	31%	36%	34%	25%	33%	42%	40%	40%	-20%	21%	41%	38%	22%	28%	27%	10%	13%	28%	26%	24%	25%
20090930	100%	100%	104%	34%	35%	31%	30%	36%	34%	25%	33%	42%	39%	43%	-18%	22%	40%	38%	24%	29%	29%	8%	10%	27%	25%	23%	25%
20091030	100%	100%	107%	34%	35%	31%	31%	36%	33%	25%	33%	42%	41%	41%	-18%	25%	37%	38%	25%	29%	31%	7%	8%	27%	25%	23%	24%
20091130	100%	100%	107%	34%	35%	31%	31%	36%	32%	25%	33%	42%	43%	39%	-18%	25%	36%	39%	25%	29%	31%	7%	9%	27%	25%	24%	24%
20091231	100%	100%	107%	34%	35%	31%	32%	36%	32%	27%	32%	40%	35%	54%	-10%	26%	36%	38%	21%	29%	30%	10%	9%	27%	25%	23%	26%
20100226	100%	100%	104%	34%	35%	31%	32%	35%	32%	29%	32%	39%	36%	57%	-7%	26%	36%	38%	23%	28%	30%	10%	10%	27%	27%	21%	25%
20100331	100%	100%	107%	33%	35%	32%	32%	37%	32%	29%	32%	39%	39%	47%	-15%	28%	36%	36%	24%	28%	29%	9%	10%	26%	26%	21%	26%
20100430	100%	100%	103%	33%	35%	32%	31%	37%	32%	28%	32%	39%	45%	37%	-18%	28%	37%	35%	24%	29%	30%	8%	10%	26%	27%	20%	27%
20100531	100%	100%	107%	33%	35%	32%	29%	37%	33%	26%	32%	41%	52%	23%	25%	30%	36%	35%	25%	28%	29%	7%	11%	28%	29%	19%	25%
20100630	100%	100%	103%	33%	35%	32%	30%	37%	33%	26%	32%	42%	49%	32%	19%	29%	37%	34%	25%	27%	29%	8%	11%	27%	27%	20%	26%
20100730	100%	100%	103%	33%	35%	32%	29%	38%	34%	26%	32%	42%	49%	32%	19%	29%	37%	35%	24%	28%	29%	9%	10%	28%	27%	20%	26%
20100831	100%	100%	104%	34%	35%	31%	27%	37%	36%	26%	33%	41%	54%	20%	25%	29%	38%	23%	26%	28%	8%	11%	27%	27%	20%	25%	
20100930	100%	100%	107%	33%	35%	32%	28%	37%	35%	27%	33%	41%	56%	17%	27%	29%	38%	23%	25%	28%	8%	12%	27%	23%	23%	27%	
20101029	100%	100%	107%	33%	35%	32%	29%	36%	35%	26%	33%	41%	55%	-17%	29%	28%	37%	35%	24%	27%	27%	9%	12%	26%	24%	23%	27%
20101130	100%	100%	103%	33%	35%	32%	29%	36%	35%	26%	33%	41%	58%	-5%	38%	30%	36%	34%	23%	27%	26%	10%	14%	26%	24%	23%	27%
20101231	100%	100%	107%	33%	35%	32%	29%	35%	35%	27%	34%	39%	47%	18%	35%	26%	37%	37%	24%	25%	25%	11%	15%	28%	26%	21%	25%
20110131	100%	100%	103%	33%	34%	33%	26%	37%	36%	27%	34%	39%	48%	15%	37%	26%	36%	37%	23%	26%	25%	11%	15%	28%	26%	22%	25%
20110228	100%	100%	103%	33%	34%	33%	27%	37%	36%	27%	34%	39%	44%	24%	32%	26%	37%	37%	22%	24%	24%	14%	16%	25%	24%	20%	27%
20110331	100%	100%	103%	33%	34%	33%	26%	38%	35%	29%	35%	37%	46%	18%	36%	26%	37%	37%	22%	24%	24%	14%	16%	24%	24%	19%	28%
20110429	100%	100%	102%	32%	34%	34%	27%	40%	34%	30%	35%	39%	48%	15%	37%	24%	39%	37%	22%	25%	24%	14%	16%	24%	24%	18%	29%
20110531	100%	100%	102%	32%	34%	34%	27%	39%	34%	29%	35%	38%	47%	15%	37%	24%	39%	38%	21%	25%	24%	14%	16%	24%	24%	18%	29%
20110630	100%	100%	101%	31%	34%	35%	28%	39%	33%	28%	35%	38%	47%	15%	37%	24%	39%	38%	21%	25%	24%	14%	16%	24%	24%	18%	29%
20110729	100%	100%	101%	31%	34%	35%	28%	37%	34%	31%	35%	39%	44%	15%	41%	26%	36%	38%	21%	25%	24%	14%	16%	24%	24%	19%	28%
20110831	100%	100%	101%	31%	34%	35%	28%	37%	34%	27%	36%	37%	44%	13%	43%	29%	34%	37%	21%	24%	24%	14%	17%	24%	24%	18%	29%
20110930	100%	100%	107%	32%	34%	34%	28%	37%	35%	28%	35%	36%	46%	11%	43%	26%	36%	38%	21%	24%	23%	14%	18%	25%	24%	19%	28%
20111031	100%	100%	102%	32%	34%	33%	29%	37%	35%	27%	36%	37%	46%	12%	43%	26%	36%	38%	21%	24%	23%	14%	18%	25%	24%	19%	28%
20111130	100%	100%	103%	32%	34%	33%	28%	37%	35%	24%	37%	39%	44%	11%	46%	24%	37%	39%	21%	24%	23%	14%	17%	25%	24%	19%	28%
20111230	100%	100%	102%	32%	34%	33%	27%	37%	36%	21%	39%	40%	43%	16%	40%	23%	37%	40%	20%	23%	22%	16%	19%	25%	24%	18%	29%
20120131	100%	100%	107%	32%	34%	33%	26%	36%	38%	21%	38%	41%	41%	17%	42%	18%	41%	41%	20%	23%	22%	16%	19%	25%	24%	17%	29%
20120229	100%	100%	103%	33%	34%	34%	24%	36%	39%	18%	38%	44%	40%	17%	42%	18%	43%	39%	19%	23%	22%	17%	19%	25%	24%	18%	29%
20120330	100%	100%	103%	33%	34%	33%	24%	36%	41%	19%	38%	45%	40%	17%	43%	20%	43%	37%	19%	23%	22%	17%	19%	25%	24%	18%	29%
20120427	100%	100%	103%	33%	34%	33%	24%	36%	41%	19%	37%	44%	40%	17%	43%	21%	43%	39%	19%	23%	22%	17%	19%	25%	24%	17%	29%
20120531	100%	100%	103%	33%	34%	34%	24%	34%	41%	21%	37%	42%	39%	16%	44%	19%	44%	37%	19%	23%	22%	17%	19%	25%	24%	17%	29%
20120629	100%	100%	107%	33%	34%	33%	25%	34%	41%	21%	37%	42%	39%	15%	46%	23%	41%	35%	18%	23%	21%	18%	20%	25%	24%	18%	29%
20120731	100%	100%	103%	33%	34%	34%	26%	33%	42%	21%	37%	42%	39%	15%	45%	24%	41%	35%	18%	22%	21%	18%	20%	25%	24%	18%	29%
20120831	100%	100%	103%	33%	34%	34%	25%	34%	41%	21%	37%	42%	38%	14%	48%	20%	41%	38%	18%	22%	21%	18%	20%	25%	24%	18%	29%
20120928	100%	100%	103%	33%	34%	33%	24%	35%	41%	21%	37%	43%	38%	14%	48%	19%	41%	39%	19%	22%	21%	18%	20%	25%	24%	18%	29%
20121031	100%	100%	102%	32%	34%	34%	23%	35%	41%	21%	37%	43%	38%	13%	49%	20%	37%	36%	19%	22%	21%	18%	20%	25%	24%	18%	29%
20121130	100%	100%	102%	32%	34%	34%	24%	35%	40%	21%	37%	43%	39%	16%	45%	26%	37%	37%	18%	22%	21%	18%	20%	25%	24%	17%	30%
20121231	100%	100%	102%	32%	34%	35%	25%	36%	39%	21%	37%	42%	40%	15%	45%	26%	37%	37%	19%	22%	21%	18%	20%	25%	24%	17%	30%
20130131	100%</																										

显而易见，使用正交因子 IC 选取类别内效果最强的因子进行降维，能够得到具有更强选股能力的复合因子。考虑到“类别内正交因子 IC 最高”仅仅是“类别内正交因子 IC 加权”的一种特殊形式，投资者在实际构建模型时也可使用正交因子 IC 加权法进行因子降维。从最终模型复合因子表现上看，“类别内正交因子 IC 加权法”下的复合因子同样具有较强的选股效果，其表现仅次于“类别内正交因子 IC 最高法”。

表 15 不同降维方法下因子模型表现

	类型	IC	ICIR	胜率	显著胜率	多空	多头	空头
降维前	Pearson	0.087	3.195	86.4%	68.9%			
	Spearman	0.107	3.492	86.4%	76.7%	3.05%	1.55%	-1.50%
IC 最高	Pearson	0.098	3.346	83.5%	72.8%	3.62%	2.03%	-1.59%
	Spearman	0.119	3.619	87.4%	78.6%			
正交 IC 最高	Pearson	0.102	3.519	87.4%	73.8%	3.78%	2.05%	-1.73%
	Spearman	0.123	3.827	88.3%	77.7%			
IC 加权	Pearson	0.095	3.174	81.6%	70.9%	3.43%	1.99%	-1.44%
	Spearman	0.117	3.515	88.3%	74.8%			
正交 IC 加权	Pearson	0.100	3.369	84.5%	70.9%	3.59%	2.05%	-1.54%
	Spearman	0.122	3.707	88.3%	78.6%			
IC 序列 PCA 降维	Pearson	0.094	3.148	80.6%	68.9%	3.43%	1.99%	-1.44%
	Spearman	0.115	3.467	87.4%	75.7%			
正交 IC 序列 PCA 降维	Pearson	0.094	3.129	80.6%	69.9%	3.37%	1.98%	-1.39%
	Spearman	0.114	3.442	87.4%	74.8%			

资料来源：Wind，海通证券研究所

为了能够进一步对比各降维方法的不同，可对比不同类别因子降维前后选股效果的不同。下表对比了“类别内正交因子 IC 最高”、“类别内正交因子 IC 加权”以及“类别内正交因子 IC 序列 PCA 降维”下，各类因子降维前后的 IC 与 ICIR。

使用“类别内正交因子 IC 最高”降维后，换手率类因子、特异度类因子、盈利能力类因子以及偿付能力类因子的 IC 与 ICIR 皆相对于降维前各子类因子出现了提升。降维后的反转类因子选股效果弱于 1 个月反转，降维后的估值类因子选股效果弱于 PB，降维后的成长能力类因子选股效果弱于营业收入同比增速。

使用“类别内正交因子 IC 加权”降维后，仅有成长能力类因子以及偿付能力类因子的 IC 与 ICIR 相对于各子类因子出现全面提升。其余各类因子在降维后通常会弱于类别内最强的因子。

使用“类别内正交因子 IC 序列 PCA 降维法”降维后，各类因子虽相对于部分子类因子有所提升，但表现通常会弱于子类因子中最强的因子。

对比来看，“类别内正交因子 IC 最高降维法”能够对于多因子选股模型产生最大幅度的提升，同时各类因子在降维后也能保持较强的选股能力。若投资者想尽可能多地使用到类别内各指标的信息，则可使用“类别内正交因子 IC 加权降维法”。从结果上看，并不推荐使用“类别内正交因子 IC 序列 PCA 降维法”

表 16 不同降维方法下的各类因子降维前后的 IC 与 ICIR

		类别内取正交 IC 最高				类别内正交 IC 加权				类别内正交 IC 序列 PCA 降维			
		降维前		降维后		降维前		降维后		降维前		降维后	
		IC	ICIR	IC	ICIR	IC	ICIR	IC	ICIR	IC	ICIR	IC	ICIR
市值	市值	-0.066	-1.893	-0.066	-1.893	-0.066	-1.893	-0.066	-1.893	-0.066	-1.893	-0.066	-1.893
非线性	市值平方	0.037	2.049	0.037	2.049	0.037	2.049	0.037	2.049	0.037	2.049	0.037	2.049
换手	1M	-0.062	-2.037			-0.062	-2.037			-0.062	-2.037		
	3M	-0.049	-1.625	-0.062	-2.037	-0.049	-1.625	0.054	1.700	-0.049	-1.625	-0.052	-1.632
	6M	-0.043	-1.400			-0.043	-1.400			-0.043	-1.400		
反转	1M	-0.039	-1.435			-0.045	-1.686			-0.045	-1.716		
	3M	-0.034	-1.203	-0.036	-1.317	-0.039	-1.398	0.044	1.476	-0.040	-1.429	-0.043	-1.424
	6M	-0.021	-0.771			-0.022	-0.833			-0.023	-0.848		
特异度	1M	0.031	1.990			0.032	2.231			0.033	2.352		
	3M	0.020	1.187	0.032	2.087	0.017	1.110	0.029	1.842	0.017	1.126	0.020	1.207
	6M	0.012	0.655			0.007	0.404			0.006	0.345		
估值	PE	-0.009	-0.381			-0.001	-0.048			-0.002	-0.110		
	PB	-0.023	-1.338	-0.018	-0.907	-0.019	-1.157	0.006	0.308	-0.020	-1.238	0.004	0.170
	PS	-0.005	-0.215			-0.001	-0.036			-0.001	-0.062		
盈利增长	营收	0.023	1.985			0.024	2.212			0.026	2.387		
	营业利润	0.012	1.090	0.021	1.898	0.014	1.200	0.024	2.239	0.017	1.571	0.024	2.170
	净利润	0.015	1.500			0.017	1.621			0.020	2.013		
盈利	ROE	0.008	0.451			0.009	0.595			0.016	0.914		
	ROA	0.001	0.048			0.004	0.214			0.012	0.684		
	ROIC	0.001	0.063	0.014	0.893	0.004	0.235	0.012	0.616	0.012	0.695	0.016	0.830
	毛利率	0.009	0.610			0.011	0.865			0.018	1.403		
	净利率	0.005	0.319			0.006	0.448			0.014	0.936		
清偿能力	D2E	0.004	0.265			0.003	0.188			0.002	0.154		
	A2D	0.004	0.334	0.015	0.888	0.003	0.247	0.011	0.765	0.003	0.282	0.005	0.368
	流动	0.002	0.125			0.004	0.283			0.005	0.373		
	速动	0.002	0.122			0.004	0.291			0.005	0.374		

资料来源：Wind，海通证券研究所

4. 总结与展望

本文对于因子降维中的底层因子降维方法进行了讨论与回测。在回测中，报告从“类别内取最强”、“类别内选股效果加权”、“最大化解释类别内因子选股效果波动”三种思路出发，分别构建了具体降维方法以及对应的改进方向。

通过对比可知，基于“类别内取最强”、“类别内选股效果加权”构建的降维方法能够对于未降维模型产生提升，但是基于“最大化解释类别内因子选股效果波动”构建的降维方法并未能够产生提升效果。

随着模型中因子数量的增多，投资者不仅仅需要注意底层因子的降维方法，还需要考虑因子的分类方法。本报告基于因子逻辑对于因子进行了分类，但是投资者可使用其他方法进行因子分类。相关回测结果请关注海通量化团队后续发布的专题报告。

5. 风险提示

市场系统性风险、资产流动性风险以及政策变动风险会对策略表现产生较大影响。

信息披露

分析师声明

冯佳睿 金融工程研究团队
袁林青 金融工程研究团队

本人具有中国证券业协会授予的证券投资咨询执业资格，以勤勉的职业态度，独立、客观地出具本报告。本报告所采用的数据和信息均来自市场公开信息，本人不保证该等信息的准确性或完整性。分析逻辑基于作者的职业理解，清晰准确地反映了作者的研究观点，结论不受任何第三方的授意或影响，特此声明。

法律声明

本报告仅供海通证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因接收人收到本报告而视其为客户。在任何情况下，本报告中的信息或所表述的意见并不构成对任何人的投资建议。在任何情况下，本公司不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告所载的资料、意见及推测仅反映本公司于发布本报告当日的判断，本报告所指的证券或投资标的的价格、价值及投资收入可能会波动。在不同时期，本公司可发出与本报告所载资料、意见及推测不一致的报告。

市场有风险，投资需谨慎。本报告所载的信息、材料及结论只提供特定客户作参考，不构成投资建议，也没有考虑到个别客户特殊的投资目标、财务状况或需要。客户应考虑本报告中的任何意见或建议是否符合其特定状况。在法律许可的情况下，海通证券及其所属关联机构可能会持有报告中提到的公司所发行的证券并进行交易，还可能为这些公司提供投资银行服务或其他服务。

本报告仅向特定客户传送，未经海通证券研究所书面授权，本研究报告的任何部分均不得以任何方式制作任何形式的拷贝、复印件或复制品，或再次分发给任何其他人，或以任何侵犯本公司版权的其他方式使用。所有本报告中使用的商标、服务标记及标记均为本公司的商标、服务标记及标记。如欲引用或转载本文内容，务必联络海通证券研究所并获得许可，并需注明出处为海通证券研究所，且不得对本文进行有悖原意的引用和删改。

根据中国证监会核发的经营证券业务许可，海通证券股份有限公司的经营经营范围包括证券投资咨询业务。

海通证券股份有限公司研究所

路 颖 所长
(021)23219403 luying@htsec.com

高道德 副所长
(021)63411586 gaodd@htsec.com

姜 超 副所长
(021)23212042 jc9001@htsec.com

江孔亮 副所长
(021)23219422 kljiang@htsec.com

邓 勇 所长助理
(021)23219404 dengyong@htsec.com

荀玉根 所长助理
(021)23219658 xyg6052@htsec.com

钟 奇 所长助理
(021)23219962 zq8487@htsec.com

宏观经济研究团队

姜 超(021)23212042 jc9001@htsec.com
顾潇啸(021)23219394 gxx8737@htsec.com
于 博(021)23219820 yb9744@htsec.com
梁中华(021)23154142 lzh10403@htsec.com
联系人
李金柳(021)23219885 lij11087@htsec.com
宋 潇(021)23154483 sx11788@htsec.com
陈 兴(021)23214504 cx12025@htsec.com

金融工程研究团队

高道德(021)63411586 gaodd@htsec.com
冯佳睿(021)23219732 fengjr@htsec.com
郑雅斌(021)23219395 zhengyb@htsec.com
罗 蕾(021)23219984 ll9773@htsec.com
沈泽承(021)23212067 szc9633@htsec.com
余浩淼(021)23219883 yhm9591@htsec.com
袁林青(021)23212230 ylq9619@htsec.com
联系人
周一洋(021)23219774 zyy10866@htsec.com
姚 石(021)23219443 ys10481@htsec.com
吕丽颖(021)23219745 lly10892@htsec.com
张振岗(021)23154386 zzg11641@htsec.com
颜 伟(021)23219914 yw10384@htsec.com
史霄安 sx11398@htsec.com
梁 镇(021)23219449 lz11936@htsec.com

金融产品研究团队

高道德(021)63411586 gaodd@htsec.com
倪韵婷(021)23219419 niyt@htsec.com
陈 瑶(021)23219645 chenyaoyao@htsec.com
唐洋运(021)23219004 tangyy@htsec.com
宋家骥(021)23212231 sjj9710@htsec.com
薛 涵 xh11528@htsec.com
皮 灵(021)23154168 pl10382@htsec.com
联系人
谈 鑫(021)23219686 tx10771@htsec.com
王 毅(021)23219819 wy10876@htsec.com
蔡思圆(021)23219433 csy11033@htsec.com
徐燕红(021)23219326 xyh10763@htsec.com
庄梓恺 zzk11560@htsec.com

固定收益研究团队

姜 超(021)23212042 jc9001@htsec.com
周 霞(021)23219807 zx6701@htsec.com
朱征星(021)23219981 zzx9770@htsec.com
姜珣珊(021)23154121 jps10296@htsec.com
联系人
杜 佳(021)23154149 dj11195@htsec.com
李 波 lb11789@htsec.com

策略研究团队

荀玉根(021)23219658 xyg6052@htsec.com
钟 青(010)56760096 zq10540@htsec.com
高 上(021)23154132 gs10373@htsec.com
李 影(021)23154147 ly11082@htsec.com
联系人
姚 佩(021)23154184 yp11059@htsec.com
唐一杰 021-23219406 tyj11545@htsec.com

中小市值团队

张 宇(021)23219583 zy9957@htsec.com
钮宇鸣(021)23219420 ymniu@htsec.com
刘 宇(021)23219608 liuy4986@htsec.com
孔维娜(021)23219223 kongwn@htsec.com
潘莹练(021)23154122 pyl10297@htsec.com
联系人
王鸣阳(021)23219356 wmy10773@htsec.com
程碧升(021)23154171 cbs10969@htsec.com
相 姜(021)23219945 xj11211@htsec.com

政策研究团队

李明亮(021)23219434 lml@htsec.com
陈久红(021)23219393 chenjiuhong@htsec.com
吴一萍(021)23219387 wuyiping@htsec.com
朱 蕾(021)23219946 zl8316@htsec.com
周洪荣(021)23219953 zhr8381@htsec.com
王 旭(021)23219396 wx5937@htsec.com

石油化工行业

邓 勇(021)23219404 dengyong@htsec.com
朱军军(021)23154143 zjj10419@htsec.com
毛建平(021)23154134 mjp10376@htsec.com
联系人
胡 歆(021)23154505 hx11853

医药行业

余文心(0755)82780398 ywx9461@htsec.com
郑 琴(021)23219808 zq6670@htsec.com
孙 建(021)23154170 sj10968@htsec.com
师成平(010)50949927 scp10207@htsec.com
联系人
贺文斌(010)68067998 hwb10850@htsec.com
刘 浩(010)56760098 lh11328@htsec.com
吴佳松 01056760092 wjs11852@htsec.com

汽车行业

邓 学(0755)23963569 dx9618@htsec.com
杜 威 0755-82900463 dw11213@htsec.com
谢亚彤(021)23154145 xyt10421@htsec.com
王 猛(021)23154017 wm10860@htsec.com

公用事业

张一弛(021)23219402 zyc9637@htsec.com
张 磊(021)23212001 zl10996@htsec.com
赵树理(021)23219748 zsl10869@htsec.com
联系人
陈佳彬(021)23154509 cjb11782@htsec.com
傅逸帆(021)23154398 fyf11758@htsec.com

批发和零售贸易行业

汪立亭(021)23219399 wanglt@htsec.com
王 晴(021)23154116 wq10458@htsec.com
李宏科(021)23154125 lkh11523@htsec.com
联系人
史 岳(021)23154135 sy11542@htsec.com

互联网及传媒

钟 奇(021)23219962 zq8487@htsec.com
郝艳辉(010)58067906 hyh11052@htsec.com
许樱之 xyz11630@htsec.com
孙小雯(021)23154120 sxw10268@htsec.com
联系人
强超廷(021)23154129 qct10912@htsec.com
毛云聪(010)58067907 myc11153@htsec.com
刘 欣(010)58067933 lx11011@htsec.com
唐 宇(021)23219389 ty11049@htsec.com
陈星光 cxg11774@htsec.com

有色金属行业

施 毅(021)23219480 sy8486@htsec.com
联系人
杨 娜(021)23154135 yn10377@htsec.com
李妹醒(021)23219401 lsx11330@htsec.com
陈晓航(021)23154392 cxh11840@htsec.com
李 骥(021)23154513 lj11875@htsec.com

房地产行业

涂力磊(021)23219747 tll5535@htsec.com
谢 盐(021)23219436 xiey@htsec.com
联系人
杨 凡(021)23219812 yf11127@htsec.com
金 晶(021)23154128 jj10777@htsec.com

电子行业

陈 平(021)23219646 cp9808@htsec.com
 联系人
 谢 磊(021)23212214 xl10881@htsec.com
 张天闻 ztw11086@htsec.com
 尹 苓(021)23154119 yl11569@htsec.com
 石 坚 010-58067942 sj11855@htsec.com

煤炭行业

吴 杰(021)23154113 wj10521@htsec.com
 李 淼(010)58067998 lm10779@htsec.com
 戴元灿(021)23154146 dyc10422@htsec.com

电力设备及新能源行业

房 青(021)23219692 fangq@htsec.com
 徐柏乔(021)32319171 x bq6583@htsec.com
 张向伟(021)23154141 z xw10402@htsec.com
 曾 彪(021)23154148 zb10242@htsec.com

基础化工行业

刘 威(0755)82764281 lw10053@htsec.com
 刘 强(021)23219733 lq10643@htsec.com
 刘海荣(021)23154130 lhr10342@htsec.com
 联系人
 张翠翠 zcc11726@htsec.com

计算机行业

郑宏达(021)23219392 zhd10834@htsec.com
 谢春生(021)23154123 xcs10317@htsec.com
 鲁 立 ll11383@htsec.com
 黄竞晶(021)23154131 hjj10361@htsec.com
 杨 林(021)23154174 yl11036@htsec.com
 联系人
 洪 琳(021)23154137 hl11570@htsec.com

通信行业

朱劲松(010)50949926 zjs10213@htsec.com
 余伟民(010)50949926 ywm11574@htsec.com
 联系人
 庄 宇(010)50949926 zy11202@htsec.com
 张峥青 zzq11650@htsec.com

非银行金融行业

孙 婷(010)50949926 st9998@htsec.com
 何 婷(021)23219634 ht10515@htsec.com
 联系人
 夏昌盛(010)56760090 xcs10800@htsec.com
 李芳洲(021)23154127 lfz11585@htsec.com

交通运输行业

虞 楠(021)23219382 yun@htsec.com
 张 杨(021)23219442 zy9937@htsec.com
 联系人
 童 宇(021)23154181 ty10949@htsec.com
 李 丹 021-23154401 ld11766@htsec.com

纺织服装行业

梁 希(021)23219407 lx11040@htsec.com
 于旭辉(021)23219411 yxh10802@htsec.com
 联系人
 马 榕(021)23219431 mr11128@htsec.com

建筑建材行业

邱友锋(021)23219415 qyf9878@htsec.com
 冯晨阳(021)23212081 fcy10886@htsec.com
 钱佳佳(021)23212081 qjj10044@htsec.com

机械行业

余炜超(021)23219816 swc11480@htsec.com
 耿 耘(021)23219814 gy10234@htsec.com
 杨 震(021)23154124 yz10334@htsec.com
 沈伟杰(021)23219963 swj11496@htsec.com

钢铁行业

刘彦奇(021)23219391 liuyq@htsec.com
 联系人
 刘 璇(021)23219197 lx11212@htsec.com
 周慧琳(021)23154399 zhl11756@htsec.com

建筑工程行业

杜市伟 dsw11227@htsec.com
 毕春晖(021)23154114 bch10483@htsec.com

农林牧渔行业

丁 频(021)23219405 dingpin@htsec.com
 陈雪丽(021)23219164 cxl9730@htsec.com
 陈 阳(010)50949923 cy10867@htsec.com
 联系人
 关 慧(021)23219448 gh10375@htsec.com
 夏 越(021)23212041 xy11043@htsec.com

食品饮料行业

闻宏伟(010)58067941 whw9587@htsec.com
 成 珊(021)23212207 cs9703@htsec.com

军工行业

徐志国(010)50949921 xzg9608@htsec.com
 刘 磊(010)50949922 ll11322@htsec.com
 蒋 俊(021)23154170 jj11200@htsec.com
 联系人
 张宇轩 zyx11631@htsec.com
 张恒晖 zhx10170@htsec.com

银行行业

林媛媛(0755)23962186 lyy9184@htsec.com
 联系人
 谭敏沂 tmy10908@htsec.com

社会服务行业

李铁生(010)58067934 lts10224@htsec.com
 联系人
 陈扬扬(021)23219671 cyy10636@htsec.com
 顾熹闽 021-23154388 gxm11214@htsec.com

家电行业

陈子仪(021)23219244 chenzy@htsec.com
 联系人
 李 阳 ly11194@htsec.com
 朱默辰 zmc11316@htsec.com
 刘 璐 ll11838@htsec.com

造纸轻工行业

曾 知(021)23219810 zz9612@htsec.com
 赵 洋(021)23154126 zy10340@htsec.com

研究所销售团队

深广地区销售团队

蔡铁清(0755)82775962 ctq5979@htsec.com
伏财勇(0755)23607963 fcy7498@htsec.com
辜丽娟(0755)83253022 gulj@htsec.com
刘晶晶(0755)83255933 liujj4900@htsec.com
王雅清(0755)83254133 wyq10541@htsec.com
饶伟(0755)82775282 rw10588@htsec.com
欧阳梦楚(0755)23617160
oymc11039@htsec.com
巩柏含 gbh11537@htsec.com
宗亮 zl11886@htsec.com

上海地区销售团队

胡雪梅(021)23219385 huxm@htsec.com
朱健(021)23219592 zhuj@htsec.com
季唯佳(021)23219384 jiwj@htsec.com
黄毓(021)23219410 huangyu@htsec.com
漆冠男(021)23219281 qgn10768@htsec.com
胡宇欣(021)23154192 hyx10493@htsec.com
黄诚(021)23219397 hc10482@htsec.com
蒋炯 jj10873@htsec.com
毛文英(021)23219373 mwy10474@htsec.com
马晓男 mxn11376@htsec.com
杨祎昕(021)23212268 yyx10310@htsec.com
方烨晨(021)23154220 fyc10312@htsec.com
慈晓聪 021-23219989 cxc11643@htsec.com
王朝领 wcl11854@htsec.com
张思宇 zsy11797@htsec.com

北京地区销售团队

殷怡琦(010)58067988 yyq9989@htsec.com
吴尹 wy11291@htsec.com
陆铂锡 lbx11184@htsec.com
张丽莹(010)58067931 zlx11191@htsec.com
陈铮茹 czr11538@htsec.com
杨羽莎(010)58067977 yys10962@htsec.com
杜飞 df12021@htsec.com

海通证券股份有限公司研究所

地址：上海市黄浦区广东路 689 号海通证券大厦 9 楼
电话：(021) 23219000
传真：(021) 23219392
网址：www.htsec.com