

CTA 产品风格与收益归因体系初探

一、引言

近年来，随着国内 CTA 产品的迅猛发展，加之其与传统资产的低相关性特征，使得该类产品成为基金投资中重要的配置工具。但考虑到产品数目众多及收益风险结构多样，如何对其表现进行归因与评价成为筛选流程中的关键一环。

由于 CTA 产品主要以私募形式存在，披露的信息极度匮乏（净值数据不全面、持仓信息较难获得），从而给该类型产品的深入解析与筛选带来了较大困难。因此本篇报告尝试基于净值数据出发，构建了 CTA 产品的风格与收益归因体系。

（注：即便投顾有意提供每日持仓信息，但由于该类型产品的高换手特性（如对于日内套利及投机型策略，不留隔夜敞口，因此不存在每日的持仓信息），也使得基于持仓的归因难以实行（除非可以拿到交易流水）。）

图 1：CTA 产品数量

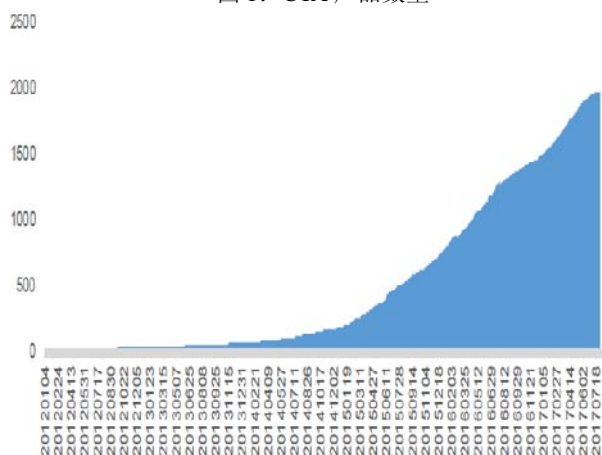
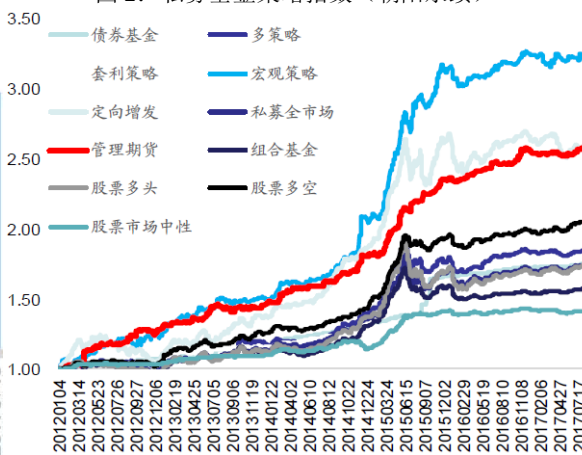


图 2：私募基金策略指数（朝阳永续）



二、CTA 产品净值数据预处理

由于 CTA 产品多以私募形式存在，业绩信息披露无标准要求，可获取的净值数据无论在频率还是质量上均不统一。具体来看：

(1)产品披露频率方面，进行日周、周度和月周披露的产品占比分别为 9.3%、37%和 4.2%，而其余将近 50%的产品并未声明披露频率，净值披露时点并无规律可寻。

(2)产品披露时点方面，在声明披露频率的产品中，披露时点也存在较大不确定性。选择在周五进行披露的产品数量最多，但占比也仅为 44%。

因此，需要在具体分析每一只 CTA 产品前先进行数据的预处理。为了尽可能地降低数据不规律与人为因素对归因分析造成的影响，我们不对 CTA 产品设定固定的数据采样频率（如日频或者周频等），而是根据各个产品的真实净值披露时点，获取相对应的因子数据值进行分析。

（注 1：由于从 Wind 数据库接口中导出的产品净值为日频，存在大量地无效重复数据，因此需要先进行数据的去重操作，从而获得预处理后的产品净值与对应时点因子数据值。）

（注 2：后续实证部分也验证了数据是否去重在模型解释度上的显著差异。）

图 3：CTA 产品净值披露频率

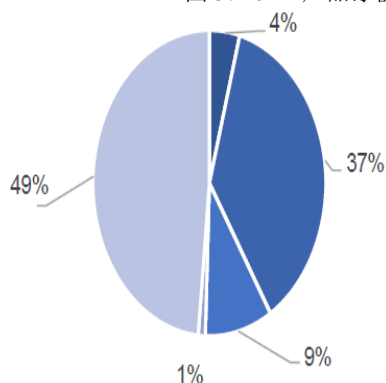
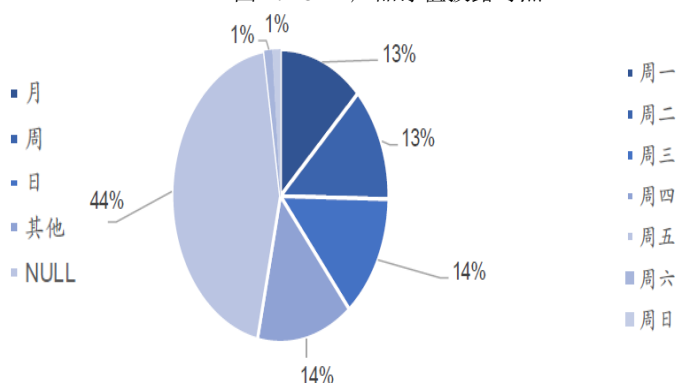


图 4：CTA 产品净值披露时点



三、归因体系构建初探

如何选取公认且合理的因子是构建归因体系的关键环节。因子构建一般遵循三个原则：一是因子定义的逻辑清晰、二是因子之间的相关性较低、三是对于常规 CTA 策略具有较好的解释力度。

不过由于 CTA 产品收益风险结构的多样性（如单边投机型策略、套利型策略及复合策略等）以及狭窄的横截面宽度，因此相较于股票端耳熟能详的归因模型（如横截面归因-Barra 模型、时间序列归因-FF5 因子模型等），国内外文献对此的研究相对较少且不存在较为公认的因子体系。

基于此，参考 CTA 策略开发过程中常用的信号指标，尝试构建了多维因子体系并结合具体产品进行归因分析。

1、因子定义

目前国内期货市场上共有三大类品种：股指期货、国债期货和商品期货。考虑到不同类型品种的基本面属性差异较大，因此构建因子的期货标的暂时仅包括商品的 48 个品种，使用各品种的指数连续数据（包括收盘价、成交量和持仓量）。

以 5 天为调仓周期，在每个调仓时点剔除上市不足 3 个月且近 20 个交易日的平均持仓量低于 5000 张的品种，并以当日收盘价为调仓价格。未考虑交易费用与冲击成本。

（注 1：指数连续品种的收盘价构建方法：在每一交易日，基于各合约成交量/持仓量的加权平均。）

（注 2：指数连续品种的成交量构建方法：在每一交易日，基于各合约成交量的加总。）

（注 3：指数连续品种的持仓量构建方法：在每一交易日，基于各合约持仓量的加总。）

(1) 市场因子

为衡量 CTA 产品对整体商品资产系统性风险的暴露程度，选取南华商品指数为市场因子的代理变量。

(2) 风格因子

➤ 动量因子

由于期货具有 T+0、多空双向交易等特征，使得趋势类策略成为 CTA 管理人策略池中的标准配置，正常占比大约为 60%-70%。因此首先构建了动量因子来解释趋势类策略的收益，以期更好地描述该类型策略在各个期货品种上的“追涨杀跌”特征。

具体计算方法如下：

- ✓ 指标计算：过去 20 个交易日的涨跌幅；
- ✓ 指标降序排列；
- ✓ 选取排名靠前的 30% 做为多头组合、靠后的 30% 做为空头组合；
- ✓ 多空组合等权重加权后相减即为该因子收益率。

➤ 活跃度因子

该因子描述了将更多风险敞口配置于流动性更好的品种所带来的影响，同时又反映了交易情绪对各个品种的趋势冲击。

具体计算方法如下：

- ✓ 指标计算：过去 20 个交易日平均成交量/过去 60 个交易日平均成交量 - 1；
- ✓ 指标降序排列；
- ✓ 选取排名靠前的 30% 做为多头组合、靠后的 30% 做为空头组合；
- ✓ 多空组合等权重加权后相减即为该因子收益率。

➤ 波动率因子

标的波动率高低是影响仓位及盈亏的核心因素之一，因此该因子部分描述了仓位控制与止盈止损能力的高低，同时也可以反映交易情绪对各个品种的影响。

具体计算方法如下：

- ✓ 指标计算：过去 60 个交易日收益率的标准差；
- ✓ 指标降序排列；
- ✓ 选取排名靠前的 30% 做为多头组合、靠后的 30% 做为空头组合；
- ✓ 多空组合等权重加权后相减即为该因子收益率。

➤ 相关度因子

相关性衡量多品种之间的联动效应。CTA 策略开发的核心三原则为多策略、多品种、多周期。相同策略在相关性较低品种上的应用可以有效降低整体组合的波动。

该因子描述了将品种之间的相关性考虑到配置当中所带来的影响。

具体计算方法如下：

- ✓ 指标计算：过去 60 个交易日各品种与其它品种的平均相关性；
- ✓ 指标降序排列；
- ✓ 选取排名靠前的 30% 做为多头组合、靠后的 30% 做为空头组合；
- ✓ 多空组合等权重加权后相减即为该因子收益率。

➤ 偏度因子

偏度是统计数据的三阶矩，衡量了变量分布的偏斜与非对称程度。因此该因子描述了品种走势的极端可能情况以及盈亏的不确定性。

具体计算方法如下：

- ✓ 指标计算：过去 60 个交易日收益率的偏度；

- ✓ 指标降序排列；
- ✓ 选取排名靠前的 30% 做为多头组合、靠后的 30% 做为空头组合；
- ✓ 多空组合等权重加权后相减即为该因子收益率。

➤ 极端因子

VaR（Value at Risk）衡量随机变量的尾部风险，因此该因子描述了品种可能的极端振荡幅度。具体计算方法如下：

- ✓ 指标计算：过去 60 个交易日收益率 95% 分位值 - 5% 分位值；
- ✓ 指标降序排列；
- ✓ 选取排名靠前的 30% 做为多头组合、靠后的 30% 做为空头组合；
- ✓ 多空组合等权重加权后相减即为该因子收益率。

➤ 投机因子

持仓量的高低反映了市场上套保与投资力量的强弱，而成交量的高低则反映了套利与投机力量的强弱。该因子描述了市场上不同类型交易者相互博弈所带来的影响。

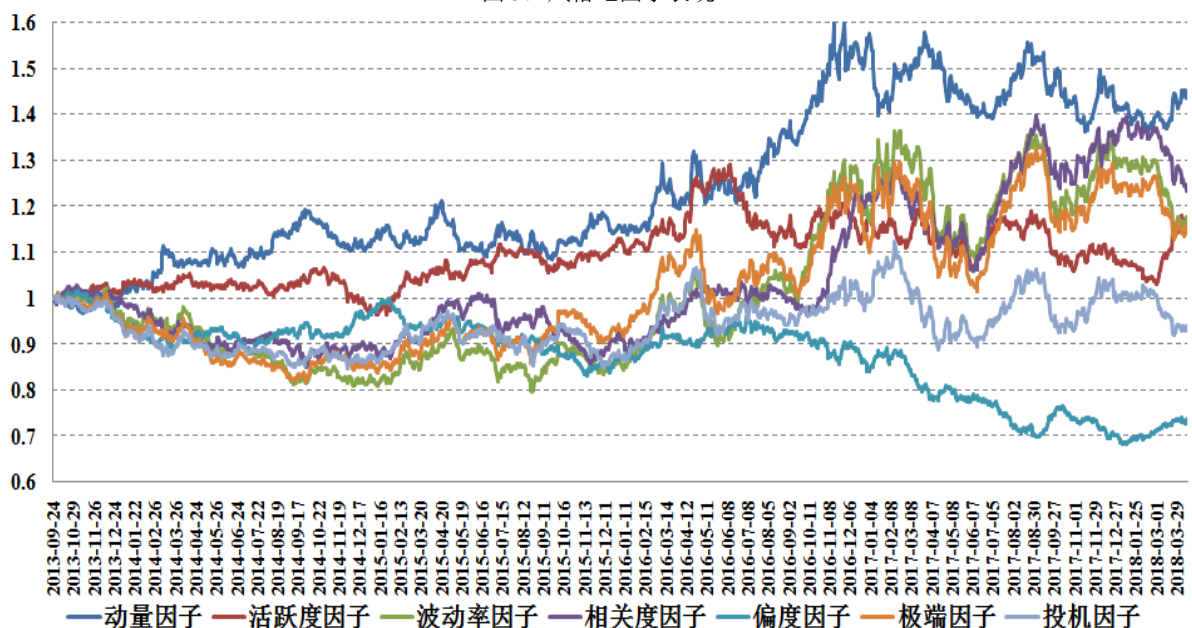
具体计算方法如下：

- ✓ 指标计算：过去 20 个交易日平均成交量/当前交易日持仓量；
- ✓ 指标降序排列；
- ✓ 选取排名靠前的 30% 做为多头组合、靠后的 30% 做为空头组合；
- ✓ 多空组合等权重加权后相减即为该因子收益率。

（注 1：后续可以尝试基于其它维度开发更多样化的风格因子（如期限结构因子等），以增强对不同 CTA 产品的解释力度）

（注 2：上述风格因子完全根据经验自行开发，逻辑上可能存在不准确之处，需要不断验证与改进。）

图 5：风格七因子表现



2、因子相关性分析与筛选

根据上述七个风格因子的相关性可以将其分为两大类：低相关度因子（动量因子、活跃度因子、偏度因子）和高相关度因子（波动率因子、相关度因子、极端因子、投机因子）。

为了降低后续回归中的共线性问题，在高相关度因子中剔除极端因子与投机因子，保留波动率因子与相关度因子。因此筛选后的风格因子依次为：动量因子、活跃度因子、波动率因子、相关度因子和偏度因子。

（注 1：保留依据-最大分散化原则：(1)首先相关度因子与其他因子相关性均值最低，因此优先保留相关度因子；(2)剩余的三个因子均具有较高的相关性，只能保留其中一个，而波动率因子与已经保留的相关度因子的相关性最低，进而保留波动率因子。）

（注 2：此处仅简单地根据因子间的相关性进行了取舍，后续还可以考虑其他更为复杂的方法进行处理，如因子逐步提纯、正交化等。但此类方法的弊端是修正后因子的逻辑不再清晰。）

表 1：风格七因子相关矩阵

	动量因子	活跃度因子	波动率因子	相关度因子	偏度因子	极端因子	投机因子	与其他因子相关性均值
动量因子		0.187	0.153	-0.041	0.099	0.156	0.104	0.110
活跃度因子	0.187		0.230	0.214	0.022	0.240	0.335	0.205
波动率因子	0.153	0.230		0.643	-0.051	0.978	0.838	0.465
相关度因子	-0.041	0.214	0.643		-0.088	0.646	0.674	0.341
偏度因子	0.099	0.022	-0.051	-0.088		-0.031	-0.046	-0.016
极端因子	0.156	0.240	0.978	0.646	-0.031		0.835	0.471
投机因子	0.104	0.335	0.838	0.674	-0.046	0.835		0.457

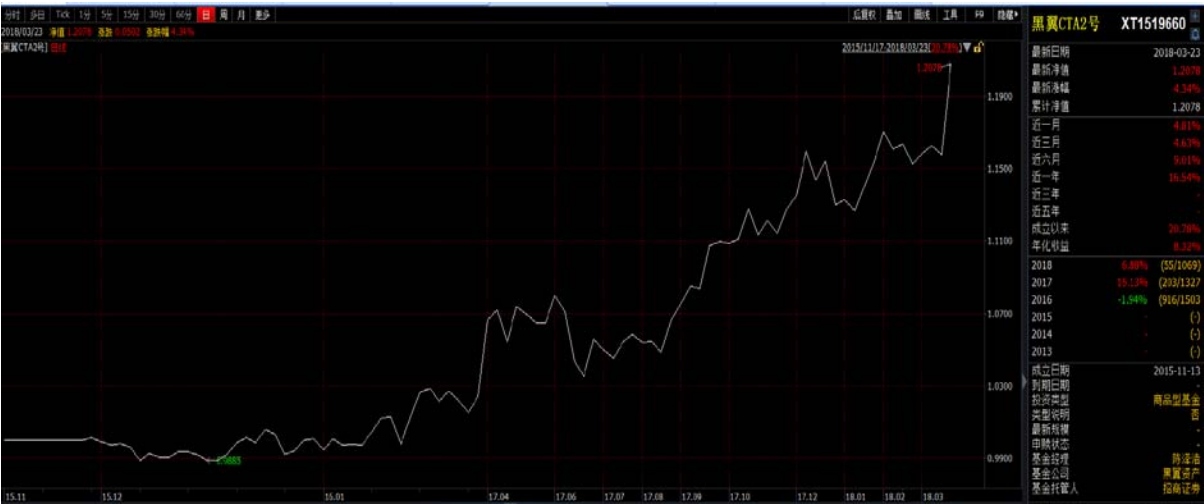
表 2：高相关度因子的相关性均值

	与高相关度因子相关性均值	与低相关度因子相关性均值
波动率因子	0.820	0.110
相关度因子	0.654	0.028
极端因子	0.820	0.122
投机因子	0.782	0.131

四、归因体系实证分析

基于上述构造的多维因子体系（1 个市场因子+5 个风格因子），本节针对国内某只商品 CTA 产品进行回归分析。该产品历史净值表现如下图所示。

图 6：产品历史净值表现



1、净值数据预处理方法差异显著

首先验证前述不同数据预处理方法在模型解释度方面的差异。基于全样本净值数据，以产品收益率为因变量，6 个因子为自变量进行回归分析。

从回归结果可以看出，净值数据进行去重操作后，模型的解释力度明显上升（R2 从 0.015 上升到 0.158、p 值从 0.1844 下降到 0.0001），从而最大程度地降低了数据不规律与人为因素对归因分析造成的影响。

后续基于去重操作后的净值进行归因分析。

表 3：净值去重操作前后的回归结果对比

对象	R2	调整后R2	F统计量	p值	残差估计误
去重操作前净值	0.0150	0.0048	1.4749	0.1844	0.0000
去重操作后净值	0.1586	0.1266	4.9627	0.0001	0.0001

2、回归时间窗口对模型解释力度与统计显著性存在较大影响

期货 T+0 的交易规则导致了 CTA 产品的高换手特征，持仓周期相对较短，不同时点发出信号的品种也可能大相径庭，从而导致归因时的时间窗口长度对模型的解释力度可能产生较大影响。

经过对时间窗口期遍历发现，时间窗口期越短则模型的解释力度越强，但噪音数据可能更多，存在局部时段模型难以解释的情况，即 p 值相对较大、统计上并不显著。反则反之。

因此，在进行归因分析时选择合适的时间窗口期是至关重要的，需要综合考虑，不能过长，也不能过短。基于下图各统计量的对比，选取 60 个交易日的窗口期分析因子暴露的稳定性。

图 7：不同时间窗口期 R2 与调整后 R2

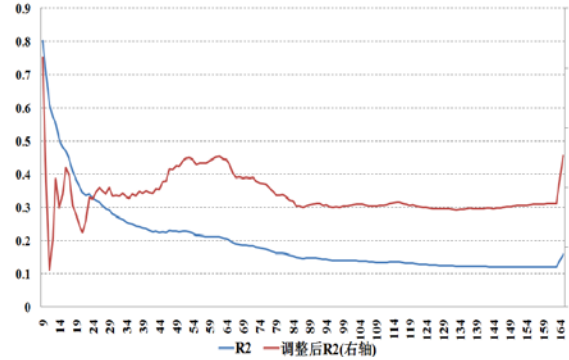
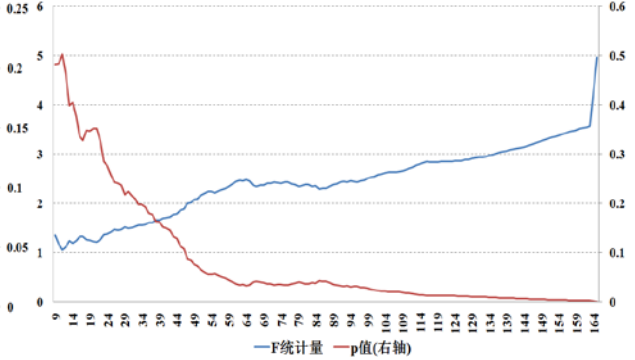


图 8：不同时间窗口期 F 统计量与 p 值



3、具体案例分析

该产品整体对市场因子及动量因子有显著正向暴露，对波动率因子、相关度因子和偏度因子有显著的负向暴露，对活跃度因子有轻微负向暴露，但并不显著。

在具体策略及品种应用上的推断如下：

- (1)以趋势型策略为主，收益来源大部分由整个市场趋势上涨的多头部分贡献；
- (2)可能存在部分反转型策略，偏左侧交易，在趋势判断出错后止损，因此偏度因子的收益贡献为负，但占比较低；
- (3)进行多品种分散配置，可能将波动性相对较小的品种包含其中，导致部分品种在区间内多以震荡行情为主，趋势性机会较少，从而反复出信号止损，使得收益贡献为负。

图 9：产品平均风格暴露

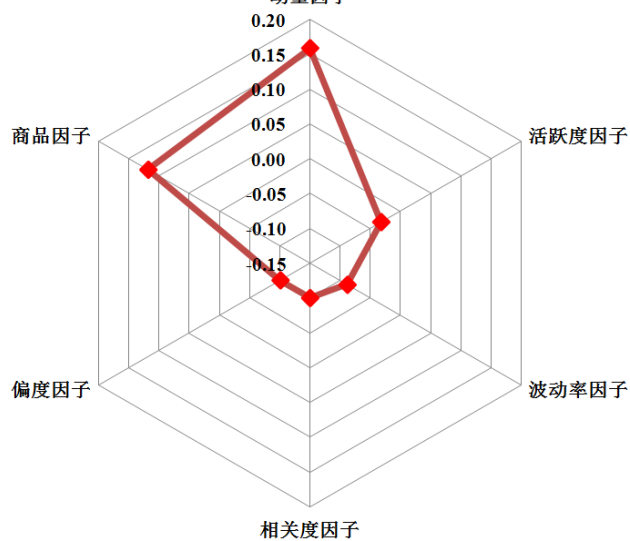


图 10：产品收益分解（单位：%）

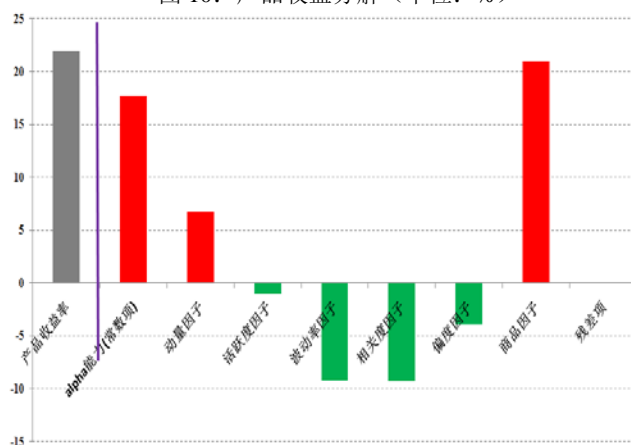


图 11：收益贡献占比（单位：%）

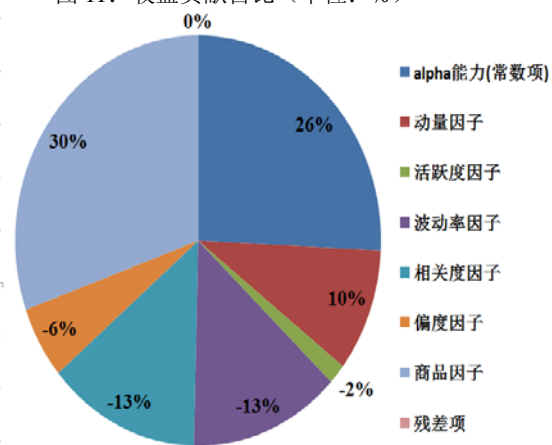
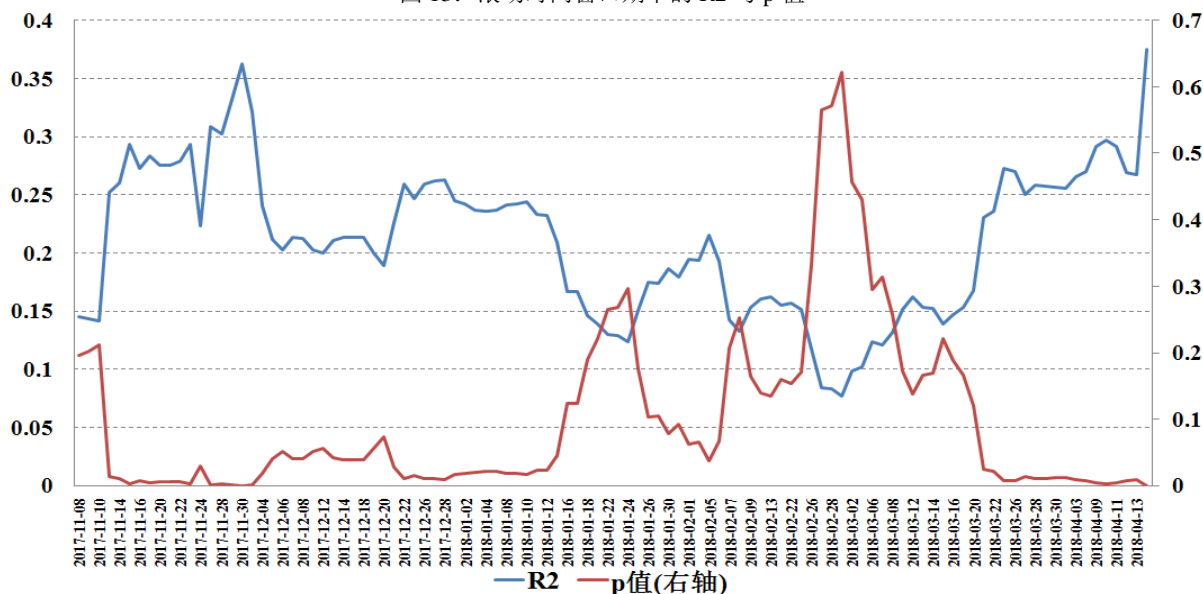


图 12：南华商品指数阶段走势



虽然模型解释力度与显著性整体稳定，但存在难以被模型解释的局部阶段（如 2018 年第一季度，可据此有针对性地和产品管理人沟通期间的操作细节）。

图 13：滚动时间窗口期下的 R2 与 p 值



五、说明

由于期货具有 T+0、多空双向交易等高频特征（属于流动性交易，主要捕捉盘口变化），使得 CTA 产品收益风险结构更具有多样性（多策略•多周期•多品种、日内与日间、投机与套利叠加等），收益来源也更复杂多变（以绝对收益目标为主），从而使得目前所构建的归因体系对于 CTA 产品的解释力度十分有限（R2 在 0.1-0.4 之间），收益端仍有大量难以解释的部分。

1、后续改进方向

(1)基于其它维度开发更多样化的风格因子（如期限结构因子等），以增强对 CTA 产品的解释力度；

(2)基于净值归因的核心问题在于因子间可能存在多重共线性现象，因此需要研究如何在保留原有因子逻辑的前提下尽可能地降低因子间的多重共线性；

(3)因子开发新思路：策略型因子。

上述方法论主要基于 FF3 因子模型框架，但期货端的策略模式与股票端的相差甚远，简单套用股票端的归因模型可能并不适用。因此可以尝试基于策略开发的角度构建因子。

究其原因，我们做收益归因的目的其实是想根据结果反推决策，了解产品管理人的投资逻辑与意图，从而让投资人可以更清晰地把握产品的运行规律。

比如，通过逻辑信号+止损止赢规则构建各种策略（如应该于多品种+多周期的区间突破型策略、均线型策略、振荡型策略、波动率突破型策略、反转型策略、跨品种套利策略、跨期套利策略等），策略的净值即为因子表现。再将各类策略净值对某产品进行回归，根据暴露程度判断产品的因子（策略）属性。对于暴露程度较高的因子（策略），由于策略逻辑完全清晰，因此可以从侧面了解产品管理人的逻辑出发点。

（注：策略间可能存在较高的相关性，可考虑根据逐步回归法依次确定因子（策略）的重要程度。）

2、基于持仓的 CTA 产品归因思路

由于 CTA 产品的持仓较难获取，加之持仓数据具有更强的时间序列属性，因此本文并未对基于持仓的 CTA 产品构建归因体系。查阅国内外学术文献，也未找到相关的参考资料。

假设可以拿到产品的交易流水数据，仅大胆推测相关归因思路。对于每笔完整的交易（建仓与平仓），对应区间品种的波动率、交易前后价格的延续性及建/平仓时点的买卖价差及盘口深度等都可以做为分析的因子维度。

图 14：CTA 策略交易流水示例

交易明细													
#	公式应用	类型	商品	建仓时间	建仓价格	平仓时间	平仓价格	数量	交易成本	净利	累计净利	收益率	累计收益率
1	tradayHLRange_ch	多头	i9000	2014/12/31 09:00	508.0	2014/12/31 09:15	506.0	4	40.56	(840.56)	(840.56)	(16.55%)	(16.55%)
2	tradayHLRange_ch	多头	rb000	2014/12/31 09:00	2605.0	2014/12/31 09:15	2599.0	5	26.02	(326.02)	(1166.58)	(12.52%)	(29.08%)
3	tradayHLRange_ch	多头	hc000	2014/12/31 09:00	2969.0	2014/12/31 09:15	2951.0	5	29.60	(929.60)	(2096.18)	(31.31%)	(60.37%)
4	tradayHLRange_ch	多头	bu000	2014/12/31 09:00	3522.0	2014/12/31 09:30	3514.0	7	49.25	(609.25)	(2705.43)	(17.30%)	(77.67%)
5	tradayHLRange_ch	多头	j9000	2014/12/31 09:00	1059.5	2014/12/31 10:45	1055.0	1	21.15	(471.14)	(3176.58)	(4.45%)	(82.12%)
6	tradayHLRange_ch	空头	j9000	2014/12/31 10:45	1055.0	2014/12/31 14:15	1059.0	1	21.14	(421.14)	(3597.72)	(3.99%)	(86.11%)
7	tradayHLRange_ch	空头	i9000	2014/12/31 09:15	506.0	2014/12/31 14:30	508.0	4	40.56	(840.56)	(4438.28)	(16.61%)	(102.72%)
8	tradayHLRange_ch	多头	j9000	2014/12/31 14:15	1059.0	2014/12/31 14:30	1054.5	1	21.13	(471.13)	(4909.41)	(4.45%)	(107.17%)
9	tradayHLRange_ch	多头	i9000	2014/12/31 14:30	508.0	2014/12/31 14:45	505.0	4	40.52	(1240.52)	(6149.93)	(24.42%)	(131.59%)
10	tradayHLRange_ch	空头	j9000	2014/12/31 14:30	1054.5	2014/12/31 14:45	1059.0	1	21.13	(471.13)	(6621.07)	(4.47%)	(136.06%)
11	tradayHLRange_ch	多头	j9000	2014/12/31 14:45	1059.0	2015/01/05 09:00	1044.0	1	21.03	(1521.03)	(8142.10)	(14.36%)	(150.42%)
12	tradayHLRange_ch	空头	j9000	2015/01/05 09:00	1044.0	2015/01/05 09:15	1060.0	1	21.04	(1621.04)	(9763.14)	(15.53%)	(165.95%)
13	tradayHLRange_ch	多头	pp000	2014/12/31 13:45	7966.0	2015/01/05 10:30	7717.0	4	31.37	(5011.37)	(14774.50)	(125.82%)	(291.77%)
14	tradayHLRange_ch	空头	i9000	2014/12/31 14:45	505.0	2015/01/09 14:15	511.0	4	40.64	(2440.64)	(17215.14)	(48.33%)	(340.10%)
15	tradayHLRange_ch	多头	i9000	2015/01/09 14:15	511.0	2015/01/12 09:00	501.0	4	40.48	(4040.48)	(21255.62)	(79.07%)	(419.17%)
16	tradayHLRange_ch	多头	j9000	2015/01/05 09:15	1060.0	2015/01/12 09:00	1064.5	1	21.25	428.76	(20826.87)	4.04%	(415.12%)
17	tradayHLRange_ch	空头	i9000	2015/01/12 09:00	501.0	2015/01/15 10:00	498.0	4	39.96	1160.04	(19666.83)	23.15%	(391.97%)
18	tradayHLRange_ch	多头	i9000	2015/01/15 10:00	498.0	2015/01/19 09:15	494.0	4	39.68	(1639.68)	(21306.51)	(32.93%)	(424.89%)
19	tradayHLRange_ch	空头	i9000	2015/01/19 09:15	494.0	2015/01/19 10:00	500.5	4	39.78	(2639.78)	(23946.29)	(53.44%)	(478.33%)
20	tradayHLRange_ch	空头	m9000	2014/12/31 09:00	2892.0	2015/01/19 10:00	2732.0	6	33.74	9566.26	(14380.03)	330.78%	(147.55%)
21	tradayHLRange_ch	多头	m9000	2015/01/19 10:00	2732.0	2015/01/19 14:15	2677.0	6	32.45	(3332.45)	(17712.49)	(121.96%)	(269.52%)
22	tradayHLRange_ch	空头	rb000	2014/12/31 09:15	2599.0	2015/01/19 14:15	2527.0	5	25.63	3574.37	(14138.12)	137.53%	(132.00%)
23	tradayHLRange_ch	多头	i9000	2015/01/19 10:00	500.5	2015/01/21 14:00	490.5	4	39.64	(4039.64)	(18177.76)	(80.71%)	(212.71%)
24	tradayHLRange_ch	多头	rb000	2015/01/19 14:15	2527.0	2015/01/26 10:45	2461.0	5	24.94	(3324.94)	(21502.70)	(131.58%)	(344.28%)
25	tradayHLRange_ch	空头	i9000	2015/01/21 14:00	490.5	2015/01/27 09:00	471.0	4	38.46	7761.54	(13741.16)	158.24%	(186.05%)
26	tradayHLRange_ch	空头	j9000	2015/01/12 09:00	1064.5	2015/01/27 09:00	1031.0	1	20.96	3329.05	(10412.11)	31.27%	(154.77%)
27	tradayHLRange_ch	多头	i9000	2015/01/27 09:00	471.0	2015/01/27 09:30	465.0	4	37.44	(2437.44)	(12849.55)	(51.75%)	(206.52%)
28	tradayHLRange_ch	空头	i9000	2015/01/27 09:30	465.0	2015/01/28 14:45	469.0	4	37.36	(1637.36)	(14486.91)	(35.21%)	(241.74%)
29	tradayHLRange_ch	空头	v9000	2014/12/31 09:30	5445.0	2015/02/02 09:00	5070.0	6	31.55	11218.46	(3268.46)	412.06%	170.33%