

Mar.

二月初三
星期日

3月 02

多层加权日内信噪比

高频因子-收益分布类

$$\text{新 } SNR = Weight_{Vol} \times SNR_{layer2} + (1 - Weight_{Vol}) \times SNR_{layer3}$$

$$Weight_{Vol} = 0.5 + \frac{Vol_{std} - 0.5}{\delta}$$

$$Vol_{std} = \frac{Vol - \min(Vol)}{\max(Vol) - \min(Vol)}$$

其中

- ① SNR_{layern} 为分解至 n 层后计算得到的信噪比，计算逻辑见相关日历页；
- ② Vol 为日内价格波动率，先在横截面上对其做极值处理和归一化处理，然后将标准化后的波动率缩至合适的权重区间。

说明

新的日内信噪比因子由日内价格波动为权重对第二层和第三层信噪比因子加权得到；日内价格波动越大，剥离更多噪声，赋予第三层更大的权重；加权后的复合信噪比的表现有所提升。

[参考文献]

严佳伟, 朱定毅. 信息提纯, 寻找高质量反转因子. 华安证券.



Mar.

二月初四

星期一

3月 03

每日最大异常成交量

行为金融因子-投资者注意力

$$ABNVOLD = \max_t \left(\frac{VOL_{i,t}}{\overline{VOL}_{i,t}} \right)$$

其中

- ① $VOL_{i,t}$ 为股票 i 在 t 交易日的成交量；
- ② $\overline{VOL}_{i,t}$ 为股票 i 在 t 交易日滚动 1 年的平均成交量；
- ③ m 为交易日窗口，一般取为月度。

说明

每日最大异常成交量为日频成交量除以过去一年平均成交量的最大值，属于极端交易量类因子；异常成交量也是投资者注意力的常用代理指标，因子值越高，投资者对该股票的关注度越高；而有限注意力导致投资者更倾向于交易引起他们关注的股票，投资者关注度越高的股票通常意味着更多的投资购买；但 A 股市场的做多与做空不对称，使得投资者非理性买入高于非理性卖出，进而导致关注度高的股票由于投资者净买入而存在溢价，未来也更可能出现反转。

[参考文献]

- ① Cosemans, M., & Frehen, R, 2021, Salience theory and stock prices: Empirical evidence. *Journal of Financial Economics*, 140(2), 460-483.
- ② 陈升锐, 2024, 投资者有限关注及注意力捕捉与溢出, 中信迎投.



Mar.

二月初五

星期二

3月

04

价格拐点

期货-高频因子

$$inflection_point = inflection_t - \frac{\sum_{t-5}^{t-1} inflection_t}{5}$$

其中

$inflection_t$ 为 t 交易日的拐点数，拐点的定义为：当最新价格不是持续向上或者向下的，也就是当前最新价变动方向与上一次变动方向不一致时即为一次拐点。

说明

拐点因子衡量了短期内期货品种相较过去的交易分歧度；当市场分歧小（因子值小，做多），当日的拐点数可能相对较少；当市场分歧大（因子值大，做空），博弈频繁时，拐点数就可能相对较多。

[参考文献]

赵妍, 高智威, 郭子峰, 2024, 基于日内高频博弈信息的商品 cta 策略, 国金证券.



星期三

二月初六

星期三

3月

05

回归版账面市值比

基本面因子改进

下述横截面回归的残差即为回归版的账面市值比：

$$\ln(Book) = \beta * \ln(MV) + \alpha + \varepsilon$$

其中

- ① $Book$ 为最新报告期归属于母公司股东权益合计；
- ② MV 为总市值。

说明

通过截面回归计算的基本面因子不再受除法分母接近 0 导致的除数效应的影响，抗噪能力更强；能够剥离截面市场行情对因子的影响，时序可比性会更纯粹；更强调截面的相对可比性，因此更适合于截面选股问题；财务指标 A 和 B 的数量级不统一甚至不可比时也可以计算相对的财务取值，以此能够挖掘更多的相对基本面因子。

[参考文献]

杨怡玲, 2024, 基本面因子的重构, 东方证券.



Mar.

二月初七

星期四



3月

06

业务残余波动率

另类因子-因子改进

计算公司各项产品的收入占比，剔除掉“其他”和“内部抵消”产品占比 30 %以上的公司，并剔除“其他”和“内部抵消”两项业务收入：

$$product_ratio_{j,i} = \frac{product_income_{j,i}}{\sum_k product_income_{k,i}}$$

筛选业务较为纯粹的公司（单项业务收入占比 50%以上），用市值加权计算每日平均收益作为绝对业务动量：

$$product_ret_{j,t} = \frac{\sum_i mkt_{i,t} \times ret_{i,t}}{\sum_i mkt_{i,t}}$$

计算每日业务绝对动量，即公司产品占比加权的平均产品收益：

$$product_abs_mom_{i,t} = \sum_j product_ratio_{j,i} \times product_ret_{j,t}$$

对过去 20 个交易日的股票收益与股票业务绝对动量进行时序回归取残差，残差的标准差即为业务残余波动率：

$$\begin{aligned} ret_{i,t} &= a + b * product_abs_mom_{i,t} + \varepsilon \\ product_resvol_{i,t} &= std(\varepsilon) \end{aligned}$$

其中

- ① $product_income_{j,i}$ 为公司 i 在产品 j 上的业务收入；
- ② $ret_{i,t}$ 、 $mkt_{i,t}$ 分别为股票 i 在 t 日的收益率和自由流通市值。

说明

将传统特质波动率计算过程中使用的市场收益率替换成业务收益，以业务收益作为系统风险，计算得到特质波动率即为业务残余波动率。

[参考文献]

王琦, 张栋梁, 2023, 基于主营业务拆分收益差的选股因子, 东北证券.



阿尔法

二月初八

星期五

3月

07

行业动量溢出效应

另类因子-动量溢出

$$RET_{it}^{IND} = Industrial_weight_{ij,t} \times ReT_{jt}$$

$$Industrial_weight_{ij,t} = \frac{IMV_{j,t}}{\sum_{j=1 \text{ and } j \neq i}^N IMV_{j,t}}$$

其中

- ① ReT_{jt} 为股票 j 在 t 时刻的月度收益率；
- ② $IMV_{j,t}$ 为 t 时期同行业内公司 j 的市值。

说明

行业动量溢出效应因子刻画了与单只股票同属于相同行业下其余股票涨跌对该股票涨跌的联动效应（领先滞后效应）。

[参考文献]

Usman Ali, David Hirshleifer, 2020, Shared analyst coverage: Unflying momentum spillover effects, Journal of Financial Economics, 6(136), 649-675.



Mar.

二月初九
星期六

3月 08

早盘主动净流入率稳定性

量价因子改进

剔除当月早盘振幅最高的 1/5 标的，对剩余标的计算如下因子：

$$\text{早盘主动净流入率稳定性因子} = \frac{\text{mean}(\text{当月每日早盘主动净流入率}_{i,t})}{\text{std}(\text{当月每日早盘主动净流入率} + i, t)}$$

$$\text{早盘主动净流入率}_{i,t} = \frac{\text{开盘10:00前净主动买入额}_{i,t}}{\text{10:00前总成交额}_{i,t}}$$

说明

该因子衡量了早盘主动资金流入的稳定性，可用于筛选主动资金近期具有稳定偏好的标的；使用早盘振幅是为了剔除非稳定性资金干扰较高的标的。

【参考文献】

王琦, 2023, JASON's alpha: 基本面+量价复合策略, 东北证券.



Mar.

二月初十
星期日

3月 09

成交量“潮汐”的价格变动速率

高频因子-量价相关性类

- ① 对股票 i ，计算 T 日内每分钟的领域成交量，即第 n 分钟及其前后 4 分钟（共计 9 分钟）的成交量总和；
- ② 将领域成交量最高点所在分钟 t 为“顶峰时刻”；将该时刻之前（第 $5 \sim t-1$ 分钟）领域成交量最低点所在分钟 m 为“涨潮时刻”，该分钟的领域成交量为 V_m ，收盘价为 C_m ；将该时刻之后（第 $t+1 \sim 233$ 分钟）领域成交量最低点所在分钟 n 为“退潮时刻”，该分钟的领域成交量为 V_n ，收盘价为 C_n ；
- ③ 上述“涨潮时刻”~“退潮时刻”的全过程记为一次“潮汐”，计算全潮汐过程的价格变动速率 $(C_n - C_m)/C_m/(n - m)$ ；对其计算最近 20 个交易日的均值即得到“全潮汐”因子。

其中

- ① 计算因子需剔除开盘和收盘数据，仅考虑日内分钟频数据；
- ② V_m 和 V_n 大小将“潮汐”过程进行强弱拆分，进一步构建“强势半潮汐”因子和“弱势半潮汐”因子。

说明

“全潮汐”因子衡量了投资者出售或购买股票意愿的强烈程度，与未来收益呈反比关系；在潮汐过程中，若股价快速下跌，表明部分投资者急于抛售，易导致过度反应，未来易发生补涨；若股价快速上涨，表明部分投资者急于建仓买入，同样易导致过度反应，未来也易发生补跌。

[参考文献]

蔚春晓, 2022, 个股成交量的潮汐变化及“潮汐”因子构建, 万正证券.



Mar.

二月十一

星期一

3月

10

技术密集度

另类因子-招聘信息

$$TechnicalIntensity_{xm_{i,t}} = \frac{\sum_{j=t-x+1}^t Number_of_technical_job_posting_{i,j}}{\sum_{j=t-x+1}^t Number_of_job_posting_{i,j}}$$

其中

- ① $Number_of_technical_job_posting_{i,j}$ 为公司 i 在 t 月的关于技术类职位的招聘广告总数；
- ② $Number_of_job_posting_{i,j}$ 为公司 i 在 t 月的招聘广告总数；
- ③ x 为回溯窗口长度，可以是 $x = 1, 3, 6, 12$ 个月；
- ④ 也可进一步计算技术密集度的短期变化（如近 1 个月相比过去 3 个月）或长期变化（如近 1 个月相比过去 1 年）：

$$TechnicalIntensity_{xm_{i,t}} - TechnicalIntensity_{ym_{i,t}}$$

说明

技术密集度为过去一段时间内技术类职位发布数量占总职位数量的比例；有研究表明拥有更高常规任务劳动份额的公司通常表现出较低的预期回报，而那些需要更多技术性劳动力的公司可能更能长期内保持竞争优势；但因子测试 IC 为负，说明较大的技术性人力成本投入可能也会导致成本过高，降低经济效益。

[参考文献]

Yin, Jenny, and the QES Team, 2024, Alpha Insights From Global Job Positions Data, Wolfe Research.



Mar.

二月十二

星期二

3月

11

基于换手率的注意力捕捉

行为金融因子-投资者注意力

对每个行业内分别统计每日行业成分股的平均换手率作为市场关注代理指标：

$$indus_turn_d = std(\text{行业内成分股的日度换手率序列} turnover_{i,d})$$

将市场关注代理指标与股票日收益进行月度时序回归，代理指标前的系数绝对值 $|\beta_{i,t}|$ 即为股票对市场关注的敏感程度：

$$\begin{aligned} r_{i,t} = & u_{i,t} + \beta_{i,t} indus_turn_d + \rho_{1,i,t} Mkt_d + \rho_{2,i,t} SMB_d + \rho_{3,i,t} HML_d \\ & + \rho_{4,i,t} UMD_d + \varepsilon_{i,t} \end{aligned}$$

其中

- ① $n_{upper,d}$ 、 $n_{lower,d}$ 、 N_d 分别为 d 日涨停股票数量、跌停股票数量、当日交易股票总数；
- ② $r_{i,d}$ 为个股日度收益；
- ③ Mkt_d 、 SMB_d 、 HML_d 为经典的 Fama-French 三因子；
- ④ UMD_d 为动量因子收益，即全市场过去 11 个月累积收益前后 30% 的组合在 d 日的收益差。

说明

不同股票对同一市场关注度事件的反映程度往往是不同的，可借助个股收益对市场关注度事件的敏感程度来衡量股票对该事件的注意力捕捉效应；捕捉能力越强（敏感程度越高），表明个股更能吸引市场注意，从而导致买盘压力增大，股价高估；上述因予以换手率作为市场关注度事件，换手率衡量投资者交易热度，反映关注水平。

[参考文献]

- ① Fengjiao LIN, Zhigang QIU, Zheng W, 2023, Cranes among chickens: The general-attention - grabbing effect of daily price limits in China's stock market, Journal of Banking & Finance, 150, 106010.
- ② 陈升锐, 2024, 投资者有限关注及注意力捕捉与溢出, 中信理财.



清华大学经济管理学院



Mar.

二月十三

星期三

3月

12

成交量支撑区域下限与收盘价差异

高频因子-成交分布类

- ① 计算同价成交量：将日内相同分钟收盘价的成交量累加至一起，得到在当前价格上的成交量总和；
- ② 将成交量累计最大的价格定义为该股当日的成交量支撑点 Volume Support Price，包含该支撑点的附近区域定义为成交量支撑区域 Volume Support Area；
- ③ 逐步计算 VSP 周围的成交量累计值（价格由近及远，成交量由大及小），该累计值与全天成交量总和的比值超 50% 的最小区域，即为成交量支撑区域 VSA，该区域上限价格为 VSA_High，下限价格为 VSA_Low；
- ④ 计算下限价格 VSA_Low 与当日收盘价之间的差异即得到 vsa_ratio。

说明

成交量支撑区域是判断个股日内公允价格区域的重要依据，过高/过低于该区域的价格将难以维持；若下限价格与当日收盘价之间差异较大，说明该股收盘后处于异常推动状态，当日大幅跌穿公允价格区域，其未来预计会回到公允价格区域，出现回弹上涨。

[参考文献]

郑兆麟, 2022, 成交量分布中的 Alpha, 兴业证券.



尚道九

二月十四

星期四

3月

13

回归版市盈率倒数

基本面因子改进

下述横截面回归的残差即为回归版的市盈率倒数：

$$\ln(Quart_E) = \beta * \ln(MV) + \alpha + \varepsilon$$

其中

- ① $Quart_E$ 为单季度净利润；
- ② MV 为总市值。

说明

通过截面回归计算的基本面因子不再受除法分母接近 0 导致的除数效应的影响，抗噪能力更强；能够剥离截面市场行情对因子的影响，时序可比性会更纯粹；更强调截面的相对可比性，因此更适合于截面选股问题；财务指标 A 和 B 的数量级不统一甚至不可比时也可以计算相对的财务取值，以此能够挖掘更多的相对基本面因子。

[参考文献]

杨怡玲, 2024, 基本面因子的重构, 东方证券.





改进的客户动量因子 (基于锚定偏差)

另类因子-动量溢出

利用供应链的相关数据，计算如下客户动量因子：

$$cmom_i^{1M} = \sum_{j=1}^{N_i} w_{ij}^{sales} mom_j^{1M}, i = 1, 2, \dots, N$$

计算最近 52 周最高价因子作为衡量“锚定偏差”的代理变量：

$$nearness\text{to}52\text{-weekhigh} = \frac{\text{月末收盘价}}{\text{过去52周内最高价}}$$

先基于客户动量因子（正向）进行分组，在各组内基于 nearness to 52-week high 因子（正向）进行分组，以此构建投资组合。

其中

① mom_j^{1M} 为公司 i 的客户 j 过去一个月收益率；

② w_{ij}^{sales} 为销售占比。

说明

客户动量因子可以从“投资者有限注意力或信息反应不足”这个维度来解释，投资者对客户信息反应越不足，客户动量效应可能越强；“锚定偏差”会延缓投资者对新信息（包括客户等各种关联公司信息）的反应，将其与客户动量结合使用会提升客户动量因子的表现。

[参考文献]

- ① Cohen, L., and A. Frazzini, 2008, Economic Links and Predictable Returns, *Journal of Finance*, 63 (4), 1977-2011.
- ② Huang, Shiyang, Tse-Chun Lin, and Hong Xiang, 2021, Psychological barrier and cross-firm return predictability, *Journal of Financial Economics* 142(1), 338-356.



Mar.

二月十六

星期六

3月

15

套保者对冲压力因子

期货-持仓类因子

$$hp_t = \frac{\# \text{of short hedge positions} - \# \text{of long hedge positions}}{\text{total} \# \text{of hedge positions}}$$

说明

套保者对冲压力因子为套保者净空头持仓数量与总持仓数量的比值；因子取值为正（净空头），表明当前市场套期保值者空头单占比高，市场很可能呈期货贴水状态（backwardated market），现货溢价，投机者因承担了套保者转移的现货价格而获得风险补偿。

[参考文献]

- ① Szymanowska, M., de Roon, F. A., Nijman, T. E., & van den Goorbergh, R. W. J., 2014, An anatomy of commodity futures risk premia, *The Journal of Finance*, 69(1), 453-482.
- ② Basu D, Miffre J., 2013, Capturing the risk premium of commodity futures: The role of hedging pressure, *Journal of Banking & Finance*, 37(7), 2652-2664.



Mar.

二月十七

星期日

3月

16

跳跃强度

高频因子-波动跳跃类

计算检验是否存在跳跃的检验统计量，原假设为不存在跳跃：

$$T_{SwV} = \frac{BV_t}{N^{-1}\sqrt{\hat{\Omega}_{SwV}}}(1 - \frac{RV_t}{SwV_t}) \xrightarrow{d} N(0, 1)$$

$$\hat{\Omega}_{SwV} = \frac{\mu_6}{9} \frac{N^3}{N-5} \int_{t-1}^t \sigma_s^6 ds = \frac{\mu_6}{9} \frac{N^3}{N-3} \mu_{3/2}^{-4} \sum_{i=0}^{N-3} \left(\prod_{k=1}^4 |r_{t_i+k}|^{3/2} \right)$$

$$SwV_t = 2 \sum_{i=1}^N (R_{t_i} - r_{t_i})$$

构建用于表示日内价格是否存在跳跃的示性函数： $I_{jump_SwV} = I(T_{SwV} > \Phi_{1-\alpha}^{-1})$

计算观察期内跳跃强度，如月度的 $D=20$ ： $JArr = \frac{1}{D} \sum_{t=1}^D I_{jump_SwV,t}$

其中

- ① R_{t_i} 和 r_{t_i} 分别为交易日 t 内第 i 个时间区间内的简单收益率和对数收益率；
- ② Φ 为标准正态分布累计分布函数， α 为显著性水平；
- ③ 可根据日内收益的正负方向，分开统计正向跳跃强度和负向跳跃强度。

说明

跳跃强度为一段时间内出现价格跳跃的天数占比，衡量了股价异常波动的持续性，可以用于刻画信息冲击的强度。

[参考文献]

- ① Tauchen, G., & Zhou, H., 2011, Realized jumps on financial markets and predicting credit spreads, Journal of Econometrics, 160(1), 102-118.
- ② 周飞鹏, 罗军, 安宁宁, 2022, 再谈股价跳跃因子, 广发证券.



Mar.

二月十八
星期一

3月 17

时间网络相对中心度

另类因子-文本分析

计算股票 i 在 t 时刻截面上相对于样本股票池的平均收益的偏离程度：

$$z_{i,t} = \frac{r_{i,t} - \bar{r}_{m,t}}{\sigma_{m,t}}$$

计算过去 20 个交易日股票 i 对其他股票收益的偏离程度的平均值：

$$\bar{z}_{i,t} = \sqrt{\frac{1}{\Delta t} \sum_{\tau=1}^{\Delta t} z_{i,t-\tau}^2}$$

计算时间维度的网络相对中心度：

$$TCC_{i,t} = \frac{1}{\bar{z}_{i,t}^2}$$

说明

TCC 因子衡量了股票的网络位置随着时间推移的稳定性，因子值越高，说明过去时间段内股票 i 对网络中其它股票的收益偏离程度越小，股票在网络中的位置越稳定。

【参考文献】

- ① 张自力, 闫红苗, 张楠, 2020, 股票网络、系统性风险与股票定价, 经济学, 1, 329-350.
- ② 岳海晓, 杨国平, 2021, 股票网络与网络中心度因子研究, 华西证券.



Mar.

二月十九
星期二

3月

18

能量指标

高频技术因子

$$HLCC_{i,t} = \frac{High_{i,t} + Low_{i,t} + Close_{i,t}}{3}$$

$$CR_{1,i,t} = \frac{\sum_{t=238}^t \max\{0, High_{i,t} - HLCC_{i,t-1}\}}{239}$$

$$CR_{2,i,t} = \frac{\sum_{t=238}^t \max\{0, HLCC_{i,t-1} - Low_{i,t}\}}{239}$$

$$CR_{i,t} = \frac{CR_{1,i,t}}{CR_{2,i,t}} * 100$$

其中

将过去 20 个交易日的日度高频因子进行时间衰减加权即可降频为月度因子。

说明

CR 指标认为中间价是股市中最有代表性的价格，比中间价高的价位其能量为“强”，比中间价低的价位其能量为“弱”，强弱之比反映了股票的能量。

[参考文献]

姚紫微, 2024, 技术指标因子高频率, 中信建投.



量化投资与机器学习



Mar.

二月二十
星期三

3月

19

仓单因子

期货-基本面因子

$$S_{Warehouse}(L, K) = \frac{\sum_{t=T-L+1}^T Warehouse_t}{\sum_{t=T-L-K+1}^{T-L} Warehouse_t} - 1$$

其中

- ① $Warehouse_t$ 为注册仓单数量，通常滚动 $L=20$ 天做移动平均值；
- ② 剔除注册仓单数量为 0 或数据缺失的品种。

说明

仓单因子即为注册仓单数量的环比变化率，仓单变化反映了库存变化（对应期货库存），进而反映了商品实物供给情况，仓单数量下降（做多），库存减少，供给减少，商品价格通常会上涨。

[参考文献]

王冬黎, 梁海晴, 2022, 商品多因子模型框架再探究, 东证期货.



Mar.

二月廿一

星期四

3月 20

新闻热度异常因子

另类因子-文本分析

- ① 对股票 i 在 t 交易日，对过去一周内 15 分钟频数的新闻热度序列计算均值和标准差，将那些新闻热度值超过均值 ± 5 倍标准差的时刻定义为热度的显著升高或降低时刻，再从这些异常时刻中筛选出热度变化与股价涨跌方向不一致的时刻 n ，共计符合条件的有 $n=1,2,3,\dots,N$ 个 15 分钟时刻；
- ② 将 N 个 15 分钟时刻对应的 15 分钟涨跌幅进行等权合成，即得到新闻热度异常因子：

$$abnormal_news_pop_{i,t} = \frac{1}{N} \sum_{n=1}^N ret_{i,n}$$

观察

新闻热度异常因子通过追踪股票涨跌方向与热度变化方向不一致时股价涨跌情况，从而捕捉投资机会；因子整体呈动量趋势，因子值越低，股价越可能在新闻热度上升，往往预示着极端消息已被充分消化，未来股价会进一步下跌。

“新闻热度异常因子”是《量化投资》中的有限注意力和“凸性效应”因子构建，方正证券。





“国家队”入场择时因子

分析师、机构持仓

计算“国家队”持仓市值相较于中证800的自由流通市值比，用以衡量“国家队”相对中证800的走势：

$$\text{自由流通市值比} = \frac{\text{国家队持仓市值}}{\text{中证800自由流通市值}}$$

滚动三年计算自由流动市值比的中位数 Zscore，用以衡量“国家队”资金强度变化：

$$Zscore = \frac{x - median}{\sigma}$$

滚动一个季度计算“国家队”市值比的 Zscore 的斜率：换仓日取近一个季度国家队资金 Zscore 对时间 t 做 OLS 回归，回归系数 β 即为“国家队”入场择时因子：

$$Zscore = \beta * t + \alpha, t = 1, 2, \dots, 63$$

数据

“国家队”持仓数据从上市公司的年报、半年报及季报披露的“十大股东”及“十大流通股东”信息中提取得到，属于“国家队”的机构类型包括：中央汇金、社保基金及养老金、国家集成电路、外管局、国新控股及国新资管、证监会等。

逻辑

“国家队”入场择时因子衡量了“国家队”资金入场强度；当因子值 >0 ，认为该区间为“国家队”资金入场时期；当因子值 ≤ 0 ，认为该区间为“国家队”资金出场时期；可基于该择时逻辑对国家队相关因子进行择时构建复合因子。

[参考文献]

刘策, 2024. 如何追寻“国家队”的进场和出场组合?(或图表 2.0). 中银证券.



Mar.

二月廿三

星期六

3月 22

逐档订单失衡率

高频因子-流动性类

$$SOIR_t = \frac{\sum_{i=1}^5 w_i SOIR_{i,t}}{\sum_{i=1}^5 w_i} \quad SOIR_{i,t} = \frac{V_{i,t}^B - V_{i,t}^A}{V_{i,t}^B + V_{i,t}^A}$$

$$w = 1 - (i - 1)/5, i = 1, 2, 3, 4, 5$$

其中

$V_{i,t}^B$ 和 $V_{i,t}^A$ 分别为 t 时刻第 i 档的委买量和委卖量。

说明

SOIR 是对各档订单失衡率的加权，可以避免某一档订单量过大对总体比率的影响；其综合衡量了盘口各档买卖委托量的不均衡程度，SOIR 为正说明市场买压大于卖压，未来价格趋向上涨，且 SOIR 的值越大，上涨的概率越高，反之亦然。

【参考文献】

丁彦明, 陈升锐, 2021, 高频订单失衡及价差因子, 中信建投证券.



Mar.

二月廿四
星期日

3月 23

业务动量

另类因子-因子改进

计算公司各项产品的收入占比，剔除掉“其他”和“内部抵消”产品占比 30 %以上的公司，并剔除“其他”和“内部抵消”两项业务收入：

$$product_ratio_{j,i} = \frac{product_income_{j,i}}{\sum_k product_income_{k,i}}$$

筛选业务较为纯粹的公司（单项业务收入占比 50%以上），用市值加权计算每日平均收益作为绝对业务动量：

$$product_ret_{j,t} = \frac{\sum_i mkt_{i,t} \times ret_{i,t}}{\sum_i mkt_{i,t}}$$

计算每日业务相对动量，即公司产品占比加权的平均产品收益与公司股票当日收益的差值：

$$product_mom_{i,t} = \sum_j product_ratio_{j,i} \times (product_ret_{j,t} - ret_{i,t})$$

计算相应频率的累计相对收益，如月频业务相对动量因子为过去 20 个交易日每日业务动量加一累乘值：

$$product_mom_month_{i,t} = \prod_{k=0}^{19} (product_mom_{i,t-k} + 1)$$

其中

- ① $product_income_{j,i}$ 为公司 i 在产品 j 上的业务收入；
- ② $ret_{i,t}$ 、 $mkt_{i,t}$ 分别为股票 i 在 t 日的收益率和自由流通市值。

说明

由公司业务动量减去公司自身动量得到的相对业务动量即为业务动量因子（正向）；即使业务结构相同的公司，由于公司个体差异（如投资者关注度、分析师覆盖度或机构持股比例等不同），信息反映速度也会存在差异，相对业务动量相比绝对业务动量可能更具有预测意义。

[参考文献]

王琦, 张栋梁, 2023, 基于主营业务拆分收益差的选股因子, 东北证券



Mar.

二月廿五
星期一

3月 24

CSAD 模型

行为金融因子-羊群效应

利用个股收益 r_i 和市场收益 R_m 计算股票收益的横截面绝对偏差：

$$CSAD = \frac{\sum_{i=1}^N |r_i - R_m|}{N}$$

以 CSAD 为 y，以市场收益绝对值和市场收益平方为 x，进行如下时序回归：

$$CSAD = \alpha + \beta_1|R_{mt}| + \beta_2|R_{mt}|^2 + \varepsilon$$

说明

CSAD 模型对羊群效应的检验主要基于时序回归上市场收益平方项前回归系数正负及显著性；当 β_1 显著小于 0 时，说明市场极端上涨或下跌时，CSAD 取值明显下降，投资者的投资行为越来越一致，市场上羊群效应越明显；当 β_1 和 β_2 都显著小于 0 时，说明市场羊群效应非常显著。

[参考文献]

Chang, E., J. Cheng, and A. Khorana, 2000, An Examination of Herd Behavior InEquity Markets: An International Perspective, *Journal of Banking&Finance*, 24(10), 1651-1679.



二月廿六
星期二

3月 25

账面价值的内涵价值/价格

基本面因子改进

基于剩余收益估值模型计算股票内涵价值：

$$V_0 = B_0 + \frac{RE_1}{(1+r_E)} + \frac{RE_2}{(1+r_E)^2} + \cdots + \frac{RE_N}{(1+r_E)^N} + \frac{CV_N}{(1+r_E)^N}$$

$$RE_n = Earn_n - r_E * B_{n-1}$$

其中

- ① B 为账面价值；
- ② RE_n 为第 n 期的剩余价值，即当期利润中超过股东必要报酬率之外的溢价；
- ③ CV_N 为最终的持续价值；
- ④ r_E 为必要报酬率；
- ⑤ 上述指标需根据公司基本面状况进行估算。

说明

该因子用来表现公司为股东创造的超额收益折现价值相对于市场价格的水平，与静态 BP 近似，但包含了公司的盈利预期信息，更为公允的展示了公司内涵价值相对于价格的高低。

[参考文献]

王琦, 2023, JASON's alpha: 基本面+量价复合策略, 东北证券.





时间序列动量

期货-动量趋势因子

$$TSMOM_{t,t+1}^s = \text{sign}(r_{t-h,t}^s) \frac{40\%}{\sigma_t^s} r_{t,t+1}^s$$

其中

① $r_{t,t+1}^s$ 为品种 s 在 $t+1$ 月相对时间序列上的超额收益，计算方法为月收益率减月收益率的指数移动平均值：

$$\hat{r}_{t,s} = r_{t,s} - \bar{r}_{t,s}$$

$$\bar{r}_{t,s} = \sum_{j=0}^{\infty} (1-\delta) * \delta^j * r_{t-j,s}$$

$$\hat{\sigma}_{t,s}^2 = \sum_{j=0}^{\infty} (1-\delta) * \delta^j * (r_{t-1-j,s} - \bar{r}_{t-1-j,s})^2$$

② $r_{t-h,t}^s$ 为品种 s 在过去 h 个月的收益率，根据其符号确定交易方向，正则看多，负则看空，滞后月份 $h=1、3、12$ 分别对应 1 个月、3 个月和 1 年等；

③ $\frac{40\%}{\sigma_t^s}$ 为品种 s 在趋势策略组合中的权重，即波动率倒数加权，用于控制权重和杠杆。

注释

净动量是对单个资产历史趋势的跟踪，长期动量效应指的是前期上涨的商品未来更可能继续上涨，前期下跌的商品未来更可能继续下跌。

{参考文献}

Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H., 2012. Time series momentum. Journal of Financial Economics 104(2), 229-250



Mar.

二月廿八

星期四

3月

27

改进的客户动量因子 (基于信息离散度)

另类因子-动量溢出

利用供应链的相关数据，计算如下客户动量因子：

$$cmom_i^{1M} = \sum_{j=1}^{N_i} w_{ij}^{sales} mom_j^{1M}, i = 1, 2, \dots, N$$

对每个客户 c 计算信息离散度 ID：

$$ID_{c,t} = sign(CR_{c,t}) \times [\%neg_{c,t} - \%pos_{c,t}]$$

先基于信息离散度 ID (负向) 进行分组，在各组内基于客户动量因子 (正向) 进行分组，以此构建投资组合。

其中

- ① mom_j^{1M} 为公司 i 的客户 j 过去一个月收益率；
- ② w_{ij}^{sales} 为销售占比；
- ③ $CR_{c,t}$ 为客户 c 在过去 3 个月的累计收益率；
- ④ $\%neg_{c,t}$ 、 $\%pos_{c,t}$ 分别为客户 c 在过去 3 个月内负收益天数占比和正收益天数占比。

说明

客户动量因子可以从“投资者有限注意力或信息反应不足”这个维度来解释，投资者对客户信息反应越不足，客户动量效应可能越强，信息离散度统计了收益形成期间日度股价涨跌天数占比的差异，间接反映了投资者对信息的反应程度，当信息连续时（即收益由小幅但大量的上涨天数来推动），投资者可能由于有限注意力而未能及时反应，股票动量效应更强；当信息离散时（即收益由大幅但少量的上涨天数来推动），投资者能够更快反应，股票动量效应更弱。同样的，投资者对经济相关公司（如供应链、竞争对手等）提供的连续信息反应不足，而对离散信息则反应迅速，所以可以用信息离散度来改进客户动量等关联性因子。

[参考文献]

- ① Huang, Shiyang, Charles M.C. Lee, Yang Song, and Hong Xiang, 2022, A Frog in Every Pan: Information Discreteness and the Lead-lag Returns Puzzle, *Journal of Financial Economics*, 145(2), 83-102.
- ② Zhi Da, Umit G.Gurun, Mitch Warachka, 2014, Frog in the Pan: Continuous Information and Momentum, *Review of Financial Studies*, Society for Financial Studies, 27(7), 528-554.



Mar.

二月廿九
星期五

3月 28

买卖意愿

高频技术因子

$$BR_{i,t} = \frac{\sum_{t=239}^t \max\{0, High_{i,t} - Close_{i,t-1}\}}{\sum_{t=239}^t \max\{0, Close_{i,t-1} - Low_{i,t-1}\}} * 100$$

其中

将过去 20 个交易日的日度高频因子进行时间衰减加权即可降频为月度因子：

$$\widehat{\text{Factor}}_j = \frac{1}{\sum_{j=1}^{20} \frac{j}{20}} \times \sum_{i=1}^{20} \widehat{\text{Factor}}_{j,k} \times \frac{j}{20}$$

说明

BR 指标反映的是市场交易意愿，以“反市场心理”为基础，买在无人问津处，卖在人声鼎沸时。

【参考文献】

姚紫璇, 陈升锐, 2024, 技术指标因子高阶化, 中信建投.



Mar.

三月初一
星期六

3月 29

近远月价差因子

期货-期限结构因子

$$-\frac{1}{K} \sum_{t=T-K+1}^T \frac{P_{next,t} - P_{front,t}}{P_{front,t}} * \frac{365}{M_{next,t} - M_{front,t}}$$

其中

① $P_{front,t}$ 、 $P_{next,t}$ 分别为近月合约和次近月合约的 t 日结算价；

② $M_{front,t}$ 、 $M_{next,t}$ 分别为近月合约和次近月合约距到期日天数。

说明

近远月价差因子为次近月合约较近月合约年化升贴水率在过去 K 个交易日的均值；做多贴水高的品种，做空贴水低的品种，通过基差收敛获得展期收益。

[参考文献]

王冬黎, 常海晴, 2022, 商品价因子模型框架再探究, 东证期货.



职位地区重要性得分

另类因子-招聘信息

① 计算 TF 得分用于衡量职位地区 j 在公司职位发布中的频率：

$$TF_{i,j,t} = \log \left(\frac{\text{Number of Postings with Location } j \text{ in Firm } i's \text{ Documents at month } t}{\text{Number of Postings in Firm } i's \text{ Documents at month } t} \right)$$

② 计算 IDF 得分用于衡量职位地区 j 在所有公司职位发布中的普遍程度（独特性）：

$$IDF_{j,t} = \log \left(\frac{\text{Total number of Postings}_t}{\text{Number of Postings with Location } j, t} \right)$$

③ 结合 TF 和 IDF 计算职位地区 j 在公司 i 职位发布中的 TF-IDF 评分，并计算每个地区 j 的最终重要性得分（N 为月份 t 涉及职位地区 j 的公司总数）：

$$TF_IDF_{i,j,t} = IDF_{j,t} \times TF_{i,j,t}$$

$$\text{方法1: } TF_IDF_Score_{j,t} = \frac{1}{N} \sum_{i=1}^N (TF_IDF_{i,j,t})^2$$

$$\text{方法2: } TF_IDF_Score_{j,t} = \frac{1}{N} \sum_{i=1}^N TF_IDF_{i,j,t}$$

④ 对每个职位 j 涉及的所有地区（共计 S 个）的 TF-IDF 评分求均值，并将每家公司 i 的所有职位的 TF-IDF 评分均值加总得到公司 i 在 t 月的技能重要性因子值：

$$Location_Importance_Scores_{i,t} = \sum_j \frac{\sum_{j=1}^S TF_IDF_Score_{i,j,t}}{S}$$

说明

职位地区重要性因子衡量了公司职位地区对公司整体招聘策略的独特贡献和重要性，因子值越高，说明公司职位地理分布合理，职位地点与公司业务发展需求相一致，招聘质量较高，与未来收益正相关。

[参考文献]

Yan Jeremy and the QES Team. 2024. Alpha Insights From Global Job Postings Data - March Edition.



Mar.

三月初三

星期一

31

3月

31

买单集中度

高频因子-资金流类

$$\frac{1}{T} \sum_{n=t}^{t-T+1} \frac{\sum_{j=1}^N \text{买单成交额}_{i,j,n}^2}{\text{总成交额}_{i,n}^2}$$

其中

- ① 基于逐笔成交数据中的叫买与叫卖单号将逐笔成交数据合成为买卖单数据；
- ② i、j、n 分别表示第 i 只股票在第 n 个交易日内的第 j 分钟的成交额；
- ③ 月度选股下，T=20 个交易日；
- ④ 将买单替换成卖单即可得卖单集中度。

说明

成交集中度类因子刻画了日内买卖单成交金额分布的均匀程度，集中度越高，未来超额收益表现越好；买单集中度因子与卖单集中度因子的 IC 方向相同，并未呈现出“买单集中度看多，卖单集中度看空”的现象。

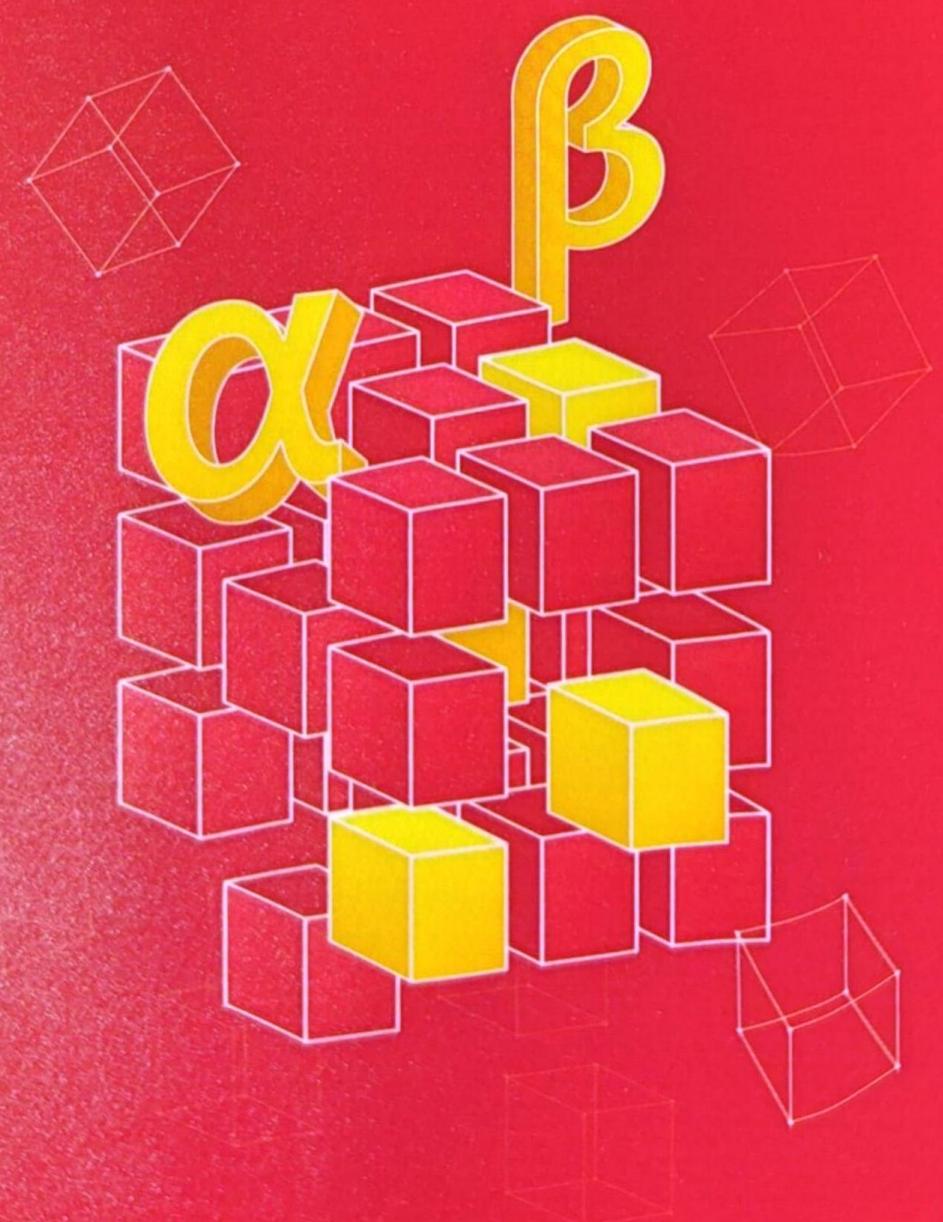
[参考文献]

冯佳蔚,袁林蔚,2019,买卖单数据中的Alpha,海通证券.



4月

April



Apr.

三月初四

星期二

4月

01

展期收益

期货-期限结构因子

$$\text{除法形式: } \frac{P_{N,t} - P_{F,t}}{P_{F,t}} \times \frac{365}{T_{F,t} - T_{N,t}}$$

$$\text{对数形式: } [\ln(P_{N,t}) - \ln(P_{F,t})] \times \frac{365}{T_{F,t} - T_{N,t}}$$

其中

- ① $P_{N,t}$ 、 $P_{F,t}$ 分别为近月合约和次近月合约的 t 日结算价；
- ② $T_{N,t}$ 、 $T_{F,t}$ 分别为近月合约和次近月合约的到期期限；
- ③ 也可用远月主力合约来代替展期后欲持有的远月合约。

说明

期货的超额收益可以由现货收益和展期收益组成，展期收益反映了现货溢价程度；根据现货溢价理论或对冲压力假说，做多现货溢价程度更高（展期大于0、期货贴水高）的期货合约可以获得更高的预期收益；现货溢价市场的期货价格相对预期现货价格有上行趋势，期货多头承担了现货多头套期保值者面临的价格下行风险，可获得正的风险补偿。

[参考文献]

Erb, C., and Harvey, C., 2006, The strategic and tactical value of commodity futures, *Financial Analysts Journal*, 62, 69-97.

