跳跃因子系列研报复现-上

一、跳跃因子逻辑概述

股票市场中，由于利好或利空消息的不断到达，叠加投资者情绪的影响，股票价格经常会表现出突然的大幅波动。**股价跳跃**常常被用来描述和衡量这种价格的突然变动，这一变动，因其大多数情况下由投资者的过度反应导致，而往往被认为是股价走势可能发生转折的征兆之一。

二、文献综述

学术研究中越来越多的成果表明，资产价格的变化不能用简单的连续时间模型来刻画，而应该将资产价格中的不连续变化（即跳跃）考虑在内。随着Barndorff和Shephard（2004）提出股价的跳跃—扩散模型，自此从理论上证明了可以将资产价格波动率中的跳跃部分分离出来，从而使得研究股价波动率中的跳跃部分对资产收益率的预测成为可能。在此基础上，随后的许多学术研究尝试从不同方面对股价波动率的跳跃部分进行分解，以寻找其中更显著的资产定价信息。

股价跳跃往往与短期的信息冲击有关。Fama（1991）认为短期内股票价格的大幅变化主要是由意外信息引起的。基于这一理论，近年来，学者们开始使用股价跳跃作为信息冲击的代理变量，研究投资者对信息冲击的反应。例如，Jiang and Zhu (2017) 研究了美股投资者对信息冲击的反应情况，发现投资者在1-3个月的短期内对股价跳跃存在反应不足的现象。

在A股市场中，徐龙炳和吴文彬（2023）研究了股价跳跃与动量/反转之间的关系。虽然动量效应在不同国家和不同资产上普遍存在，但在A股市场却并非如此。学者们认为，我国股票市场中散户投资者的占比较大。受限于散户投资者对信息的判断和反应能力，其非理性程度较高，对于信息过度敏感。对于信息冲击，A股市场的投资者往往过度反应，这导致A股市场反转效应比动量效应更明显。徐龙炳和吴文彬（2023）发现A股市场的反转效应主要是由于股价跳跃所带来的，扣除股价跳跃后的非跳跃收益则呈现动量效应。

此外，Meng et al. (2024)发现股价跳跃与投资者关注息息相关。当股价发生跳跃时，可能会吸引博彩偏好型投资者的关注，引起短期关注过热的现象。同时，市场关注度较高的个股，股价跳跃与未来收益的相关性越强。博彩偏好，即投资者偏好具有博彩性质的股票，高估那些有较低概率获得高收益的股票的价值。在交易过程中，那些股价突然发生跳跃上涨的个股，会更容易吸引博彩偏好型投资者前来购买，犹如耀眼的火光更容易吸引飞蛾，然而这些突然跳跃上涨的个股，往往由于被过分关注而被超买，从而导致未来股价下跌，博彩偏好型投资者的结局也常常如扑火的飞蛾一般。相反，那些真正基本面向好、股价能够持续上涨的股票，其上涨的过程往往是均匀平缓非跳跃的，因此我们希望通过对股价变化过程的跳跃程度加以衡量，从而对股票未来收益走势在一定程度上进行预测。

三、股价跳跃分解与识别-研究框架

1.股价跳跃分解

首先给出股价运动所满足的一般化模型，假设股票价格遵循几何布朗运动，且在考虑股价变化存在跳跃的情况下，对数股票价格满足如下跳跃——扩散模型：

上式中右边第一、二、三项分别代表漂移、扩散及跳跃带来的股价变化，其中表示瞬时价格漂移速率，表示没有跳跃时的瞬时标准差，表示标准布朗运动，表示股票价格的跳跃，表示具有瞬时强度的计数过程。

基于上述模型，Andersen等（2001）推导出利用高频数据计算的已实现方差𝑅𝑉𝑡（也称作已实现波动率）收敛于𝑝𝑡的二次变分：

𝑅𝑉𝑡 =

此处 表示第t天第i个高频数据时间段内（实际计算中使用日内minK线替代）股价的对数收益率，𝑁表示一天中高频收益率的个数。而之前的研究已证明度量了股价的波动率。的表达式中，右侧第一项表示积分波动率𝐼𝑉𝑡=，刻画了股价连续变化部分的波动水平，而第二项则表示跳跃波动，刻画的是股价跳跃部分带来的波动水平。

随后，Barndorff和Shephard（2004）通过理论推导证明上述积分波动率𝐼𝑉𝑡 可以用多幂次变差𝐼𝑉𝑡来进行有效地估计：

𝜇𝑚 ≡ 2𝑚/2Γ((𝑚 + 1)/2)/Γ(1/2)

𝛤(𝛼) = ，𝑘𝑚 = 2

据此我们可以构建出已实现波动率的跳跃分量。

2.股价跳跃的识别

（1）统计量检验

基于这一非参数估计的方法，Barndorff和Shephard（2004，2006）给出了一种用于检验日内股价是否出现跳跃的方法。具体地，他们通过理论推导证明在日内股票价格不存在跳跃的情况下，检验统计量收敛于标准正态分布:

其中𝐵𝑉𝑡表示修正后的双幂次变差，收敛于积分波动率𝐼𝑉𝑡：

表示修正后的三幂次变差，收敛于：

根据此检验统计量可以构建用于指示日内价格是否存在跳跃的示性变量：

其中𝛷表示标准正态分布累计分布函数，𝛼表示显著性水平。

（2）𝑇𝐽𝑂,𝑡统计量检验

除此之外，Jiang和Oomen（2008）基于方差互换提出了另一种跳跃检验的方法。从跳跃—扩散模型出发，利用伊藤引理可以推导出股价𝑆𝑡满足：

将此式与跳跃—扩散模型相减并积分可以得到：

可以看到此式右端第二项只与股价的跳跃成分有关，当日内股价不存在跳跃时，此项为0。Jiang和Oomen（2008）据此构建出了统计检验量𝑇𝐽𝑂,𝑡：

同样依分布收敛于标准正态分布，式中表示第t天第i个高频数据时间段内股价的简单收益率。类似地，利用这一检验统计量可以构建用于指示日内价格是否存在跳跃的示性变量：

四、跳跃因子构建

1.已实现跳跃波动类因子

将个股日内已实现波动率的跳跃分量按照大小、方向（上行、下行）和幅度（大程、小程）等维度分解，检验不同的分量对于股票收益率预测的效果。计算过程中，利用5分钟数据频率构建因子，合计26个。

|  |  |  |
| --- | --- | --- |
| 因子名 | 因子描述 | 因子定义 |
| RJV（RRJV，括号中代表归一化因子，即除以RV，下同） | 已实现跳跃波动 |  |
| SJ（RSJ） | 已实现跳跃波动不对称 | ,其中 |
| RJVP（RRJVP） | 已实现上行跳跃波动 |  |
| RJVN（RRJVN） | 已实现下行跳跃波动 |  |
| SRJV（RSRJV） | 上下行跳跃波动不对称 |  |
| RLJV（RRLJV） | 已实现大程跳跃波动 | 其中, |
| RSJV（RRSJV） | 已实现小程跳跃波动 |  |
| RLJVP（RRLJVP） | 已实现上行大程跳跃波动 |  |
| RLJVN（RRLJVN） | 已实现下行大程跳跃波动 |  |
| SRLJV（RSRLJV） | 上下行大程跳跃波动不对称 |  |
| RSJVP（RRSJVP） | 已实现上行小程跳跃波动 |  |
| RSJVN（RRSJVN） | 已实现下行小程跳跃波动 |  |
| SRSJV（RSRSJV） | 上下行小程跳跃波动不对称 |  |

2.跳跃到达率类

以过去一个月中被判断为出现显著跳跃的交易日频率构造因子，直观上以股价异常波动的持续时间来刻画信息冲击的强度。共构建8个因子。

|  |  |  |
| --- | --- | --- |
| 因子名 | 因子描述 | 因子定义 |
| Jarr | 观察期内跳跃到达率 |  |
| Jarr+ | 观察期内正向跳跃到达率 |  |
| Jarr- | 观察期内负向跳跃到达率 |  |
| JT | 观察期内跳跃检验统计量的均值 |  |

3.跳跃识别与跳跃波动结合类

通过加入观察期内跳跃显著程度的信息，提高跳跃波动因子对股票收益的预测能力，合计构建4个因子。

|  |  |  |
| --- | --- | --- |
| 因子名 | 因子描述 | 因子定义 |
| TSRJV | 观察期内以跳跃显著性程度为权重加权的上下行跳跃波动不对称 |  |
| TSRJVP | 观察期内根据跳跃显著与否进行方向筛选的跳跃波动 |  |

五、单因子测试结果汇总

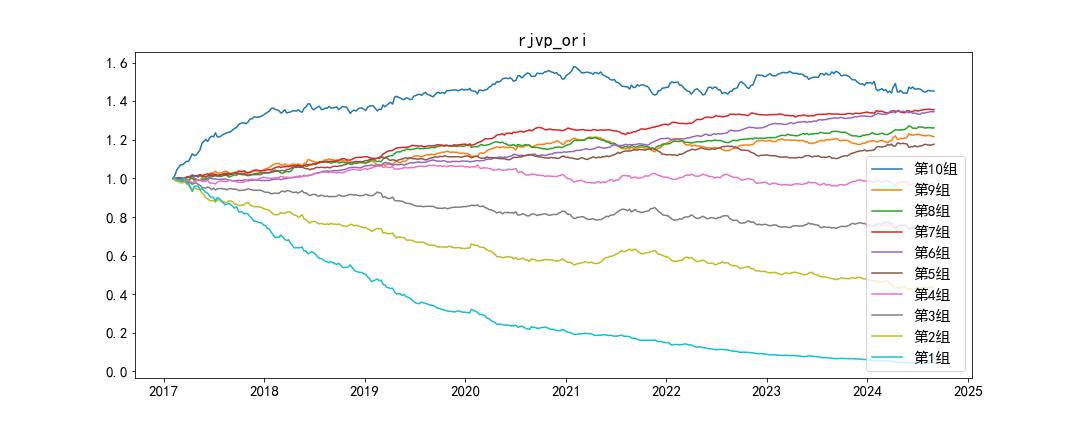
本文计算因子的日频数据，并进行周频回测，股票池为全部A股，回测起止日期为2017.2-2024.8。

受限于篇幅，本文仅列出前文涉及因子的全样本区间回测结果，各类别因子仅选取表现较好的因子进行展示，仅供参考。因子构建完整代码见知识星球，欲获取全部因子的回测统计图表，请联系本文作者。

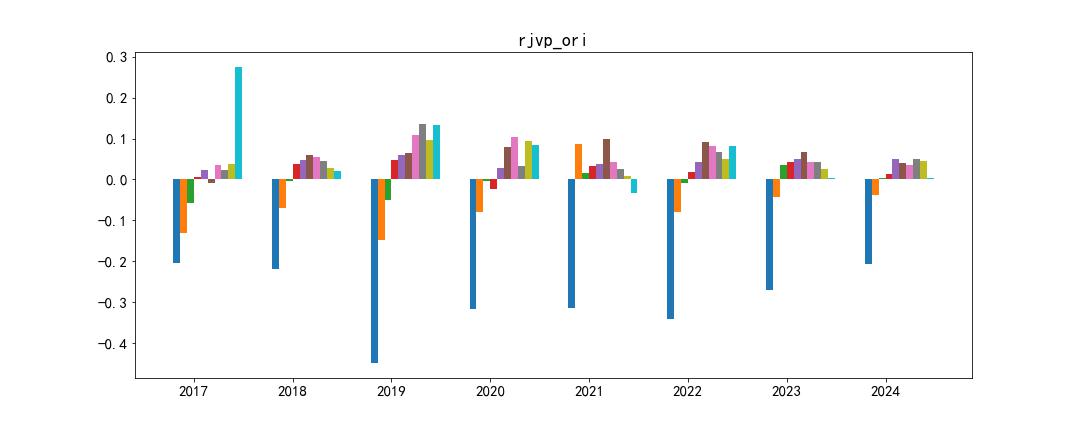
1. 已实现跳跃波动类因子

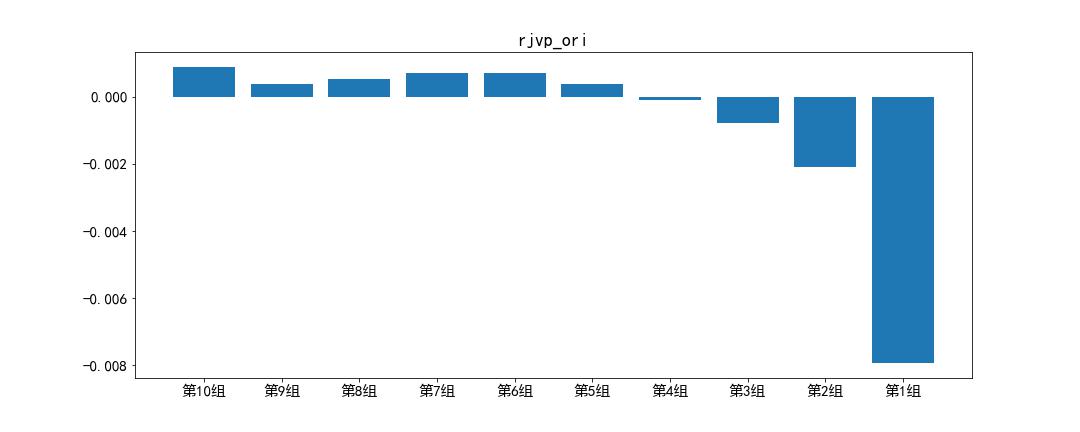
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 因子名称 | t值绝对值平均值 | IC平均值 | IRIC | IC>0概率 |
| srsjv | 4.426 | -0.038 | -0.445 | 29.72% |
| srljv | 3.822 | -0.050 | -0.726 | 20.93% |
| srjv | 4.992 | -0.066 | -0.699 | 21.96% |
| sj | 5.150 | -0.069 | -0.694 | 21.71% |
| rsjvp | 5.272 | -0.049 | -0.577 | 27.91% |
| rsjvn | 4.933 | -0.013 | -0.142 | 42.38% |
| rsjv | 4.679 | -0.029 | -0.348 | 34.11% |
| rljvp | 4.564 | -0.055 | -0.779 | 19.64% |
| rljvn | 3.180 | -0.003 | -0.043 | 45.99% |
| rljv | 4.941 | -0.057 | -0.724 | 21.45% |
| rjvp | 5.959 | -0.076 | -0.783 | 19.90% |
| rjvn | 5.220 | -0.022 | -0.213 | 36.95% |
| rjv | 5.750 | -0.064 | -0.651 | 23.00% |
| r\_srsjv | 4.281 | -0.022 | -0.260 | 38.24% |
| r\_srljv | 3.541 | -0.035 | -0.498 | 27.39% |
| r\_srjv | 4.963 | -0.044 | -0.435 | 31.01% |
| r\_sj | 5.289 | -0.048 | -0.450 | 30.23% |
| r\_rsjvp | 4.717 | 0.020 | 0.190 | 57.88% |
| r\_rsjvn | 5.034 | 0.051 | 0.477 | 67.70% |
| r\_rsjv | 5.828 | 0.048 | 0.384 | 63.57% |
| r\_rljvp | 3.785 | -0.030 | -0.409 | 33.07% |
| r\_rljvn | 2.754 | 0.013 | 0.203 | 55.56% |
| r\_rljv | 3.527 | -0.016 | -0.229 | 42.12% |
| r\_rjvp | 5.580 | -0.019 | -0.158 | 41.34% |
| r\_rjvn | 5.219 | 0.047 | 0.409 | 63.82% |
| r\_rjv | 5.995 | 0.019 | 0.145 | 54.01% |

（1）RJVP因子分10组回测结果



（2）RJVP因子10分组分年表现



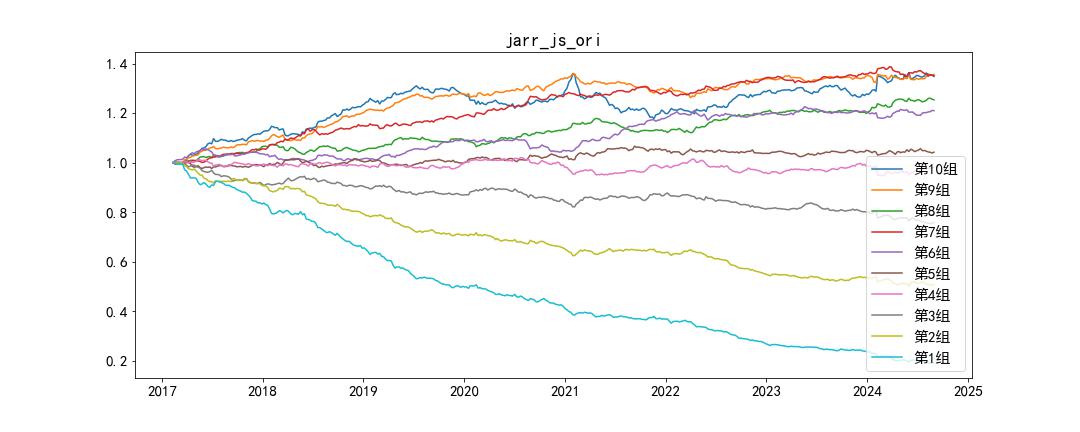


以回测区间内效果较好的因子已实现上行跳跃波动（RJVP）为例，周度换仓的RankIC达-0.076，年化RankICIR为-5.65，IC胜率（<0的比例）达80%，样本内分层单调性较好，且体现出显著的空头负向超额获取能力。

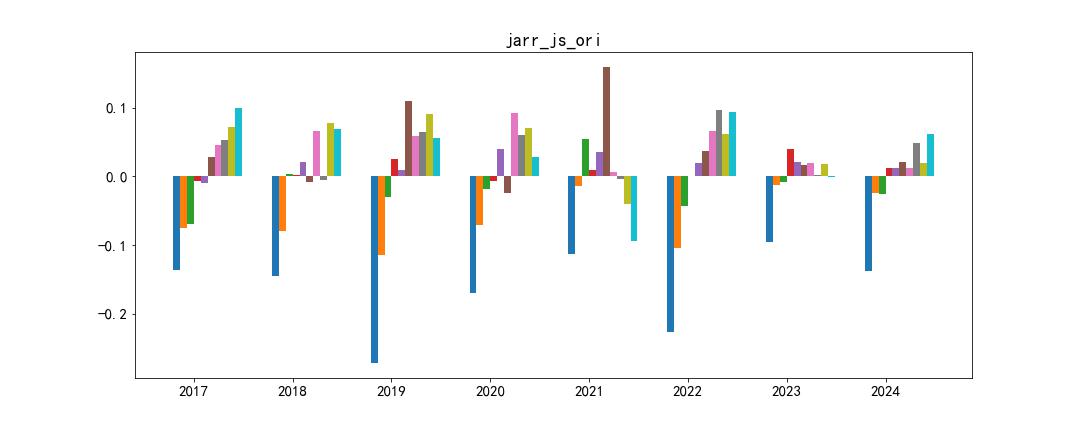
2.跳跃到达率类因子

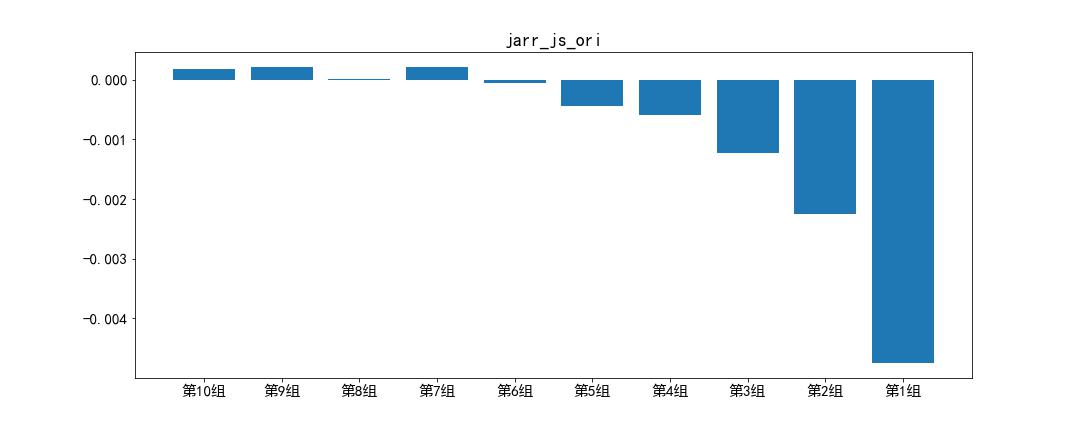
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 因子名称 | t值绝对值平均值 | IC平均值 | IRIC | IC>0概率 |
| jt\_avg\_js | 3.155 | -0.036 | -0.476 | 29.53% |
| jt\_avg\_bns | 5.759 | 0.012 | 0.097 | 54.40% |
| jarr\_pos\_js | 4.340 | -0.045 | -0.535 | 29.79% |
| jarr\_pos\_bns | 3.868 | -0.006 | -0.076 | 45.08% |
| jarr\_neg\_js | 2.371 | -0.016 | -0.352 | 32.64% |
| jarr\_neg\_bns | 3.792 | 0.018 | 0.220 | 56.99% |
| jarr\_js | 4.258 | -0.046 | -0.568 | 27.20% |
| jarr\_bns | 4.824 | 0.008 | 0.077 | 53.63% |

（1）JARR\_JS因子分10组回测结果



(2) JARR\_JS 10组分年回测结果



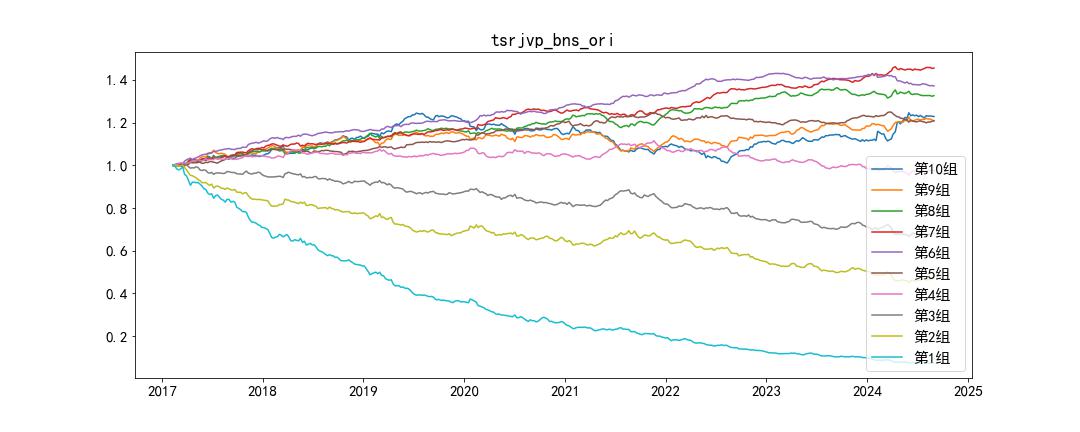


以跳跃到达率类效果相对突出的因子过去20天跳跃到达率（jarr\_js）为例，回测区间内该因子的周度RankIC为-0.046，年化RankICIR为-4.10，IC胜率（<0的比例）约73%，样本内分层单调性一般，空头端贡献显著的负向超额收益。

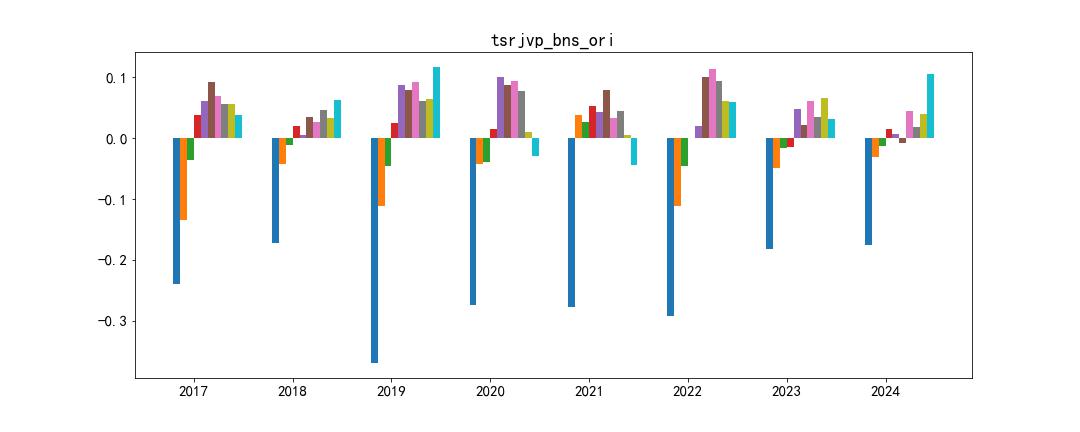
3.跳跃识别与跳跃波动结合类因子

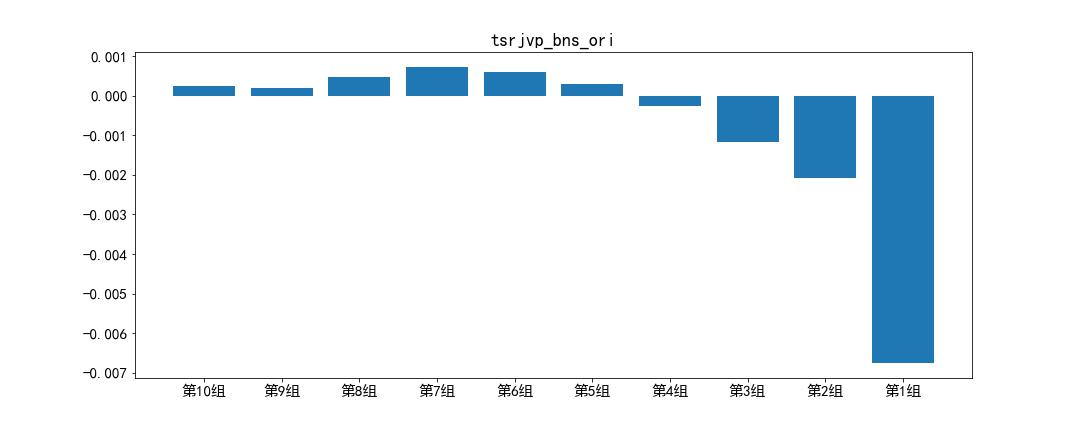
|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 因子名称 | t值绝对值平均值 | IC平均值 | IRIC | IC>0概率 |
| tsrjvp\_js | 6.679 | -0.076 | -0.642 | 24.09% |
| tsrjvp\_bns | 6.361 | -0.078 | -0.708 | 22.28% |
| tsrjv\_js | 2.652 | -0.043 | -0.597 | 26.17% |
| tsrjv\_bns | 4.093 | -0.058 | -0.746 | 21.76% |

（1）tsrjvp\_bns因子分10组回测结果



(2) tsrjvp\_bns 10组分年回测结果





以跳跃识别与跳跃波动结合类表现较好的因子tsrjvp\_Tbns为例，回测区间内该因子的周度RankIC为-0.078，年化RankICIR为-5.10，IC胜率（<0的比例）约77%，样本内分层单调性欠佳，仍然体现了突出的空头负向超额获取能力。

六、小结

本文参考现有的文献和研报资料，汇总了股票价格跳跃的主要研究架构，以及直接使用跳跃波动和跳跃检验统计量相关的因子定义，并对所构建的单因子效果进行了简单的周度测试。

在下篇中，将进一步介绍跳跃识别与跳跃收益结合、使用跳跃信息对振幅进行改进，从而构建的因子，并采用LGBM模型对两篇文章涉及的跳跃因子，进行简单合成，提高因子整体表现。

参考资料：

1. 《20220831-广发证券-金融工程专题-基于股价跳跃模型的因子研究-高频数据因子研究系列九》

2.《20221209-广发证券-金融工程专题-再谈股价跳跃因子研究-多因子Alpha系列报告之(四十六)》

3.《20220922-方正证券-多因子选股系列研究之六：个股股价跳跃及其对振幅因子的改进》

4.《20240725-招商证券-金融工程研究——如何识别股价的跳跃？》