

Liquidity研究

1. 基本研究思路梳理

我们希望从provider VS taker的视角去衡量市场短期的流动性水平。首先，我们对流动性的定义是市场上的主动交易者所要付出的cost，因此比较合理的方式是用一个类似价差的方式来进行预估。在此基础上，我们需要综合考虑两个信息，一个是当前市场上买卖五档挂单量和价格的分布情况，而不是仅根据买卖一档的spread，综合考虑盘面上所体现的流动性水平；另一个是当前市场的交易水平，即市场上当前在吃单的交易者规模有多大，从而实现normalization，衡量一定交易量下的流动性水平。最后，需要考虑其泛化性和边界情况的处理，观察市场操纵行为或者突发事件对代理变量的影响，研究更好处理这种情形的方式。

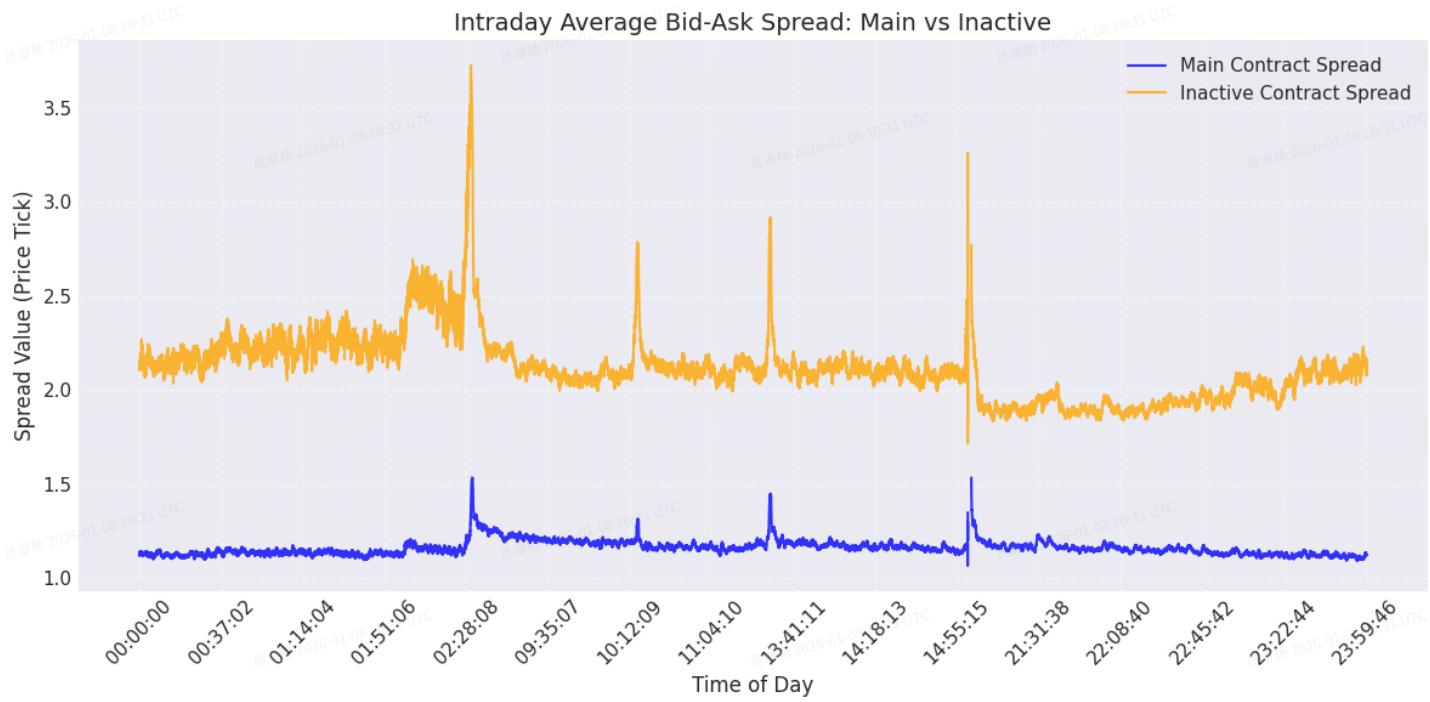
2. 数据处理：

研究中数据包括主力合约和非主力合约共计35个交易日的市场数据，我们首先将主力合约和非主力合约分别拼接，再用timestamp进行merge，便于在同一时间点上比较对于主力合约和非主力合约的流动性指标计算差异，来评估我们的流动性度量是否合理。同时，我们计算了未来5分钟的vwap减去当前midprice的差值并取绝对值，作为一个对未来实际交易中波动性的衡量指标，用我们设计的流动性指标和未来波动性指标求相关性系数，作为评估流动性指标效率的简单研究。

2. 价差作为代理变量的初步研究：

2.1 一档买卖价差

最简单的流动性指标是买卖价差，即卖一价直接减去买一价，作为当前市场上成交一手的成本。我们将过去120个tick的价差滚动取均值作为基准，测试表明主力合约和非主力合约之间价差在同一时间点上的均值存在显著差异，说明价差也可以作为衡量流动性的简单指标。



同样，对价差和未来收益率绝对值计算相关性，发现简单价差和收益率绝对值存在比较显著的正相关性。

| 指标类型 | 和主力合约相关性 | 和非主力合约相关性 |
|------|----------|-----------|
| 简单价差 | 0.273 | 0.158 |

然而，简单价差忽略了5档挂单价格和挂单量的信息，不能充分评估订单簿上反映的流动性情况。同时，在流动性相对稳定的前提下，简单价差基本不变，难以反映更细致的流动性变化情况。

2.2 深度加权价差

结合挂单量信息变化的基本思路是对买卖双方各自五档的挂单量和挂单价格求一个加权均价，表示买卖双方流动性的平均价格，做差得到订单簿深度加权价差。计算公式如下：

1. 卖方加权均价 (Weighted Ask Price):

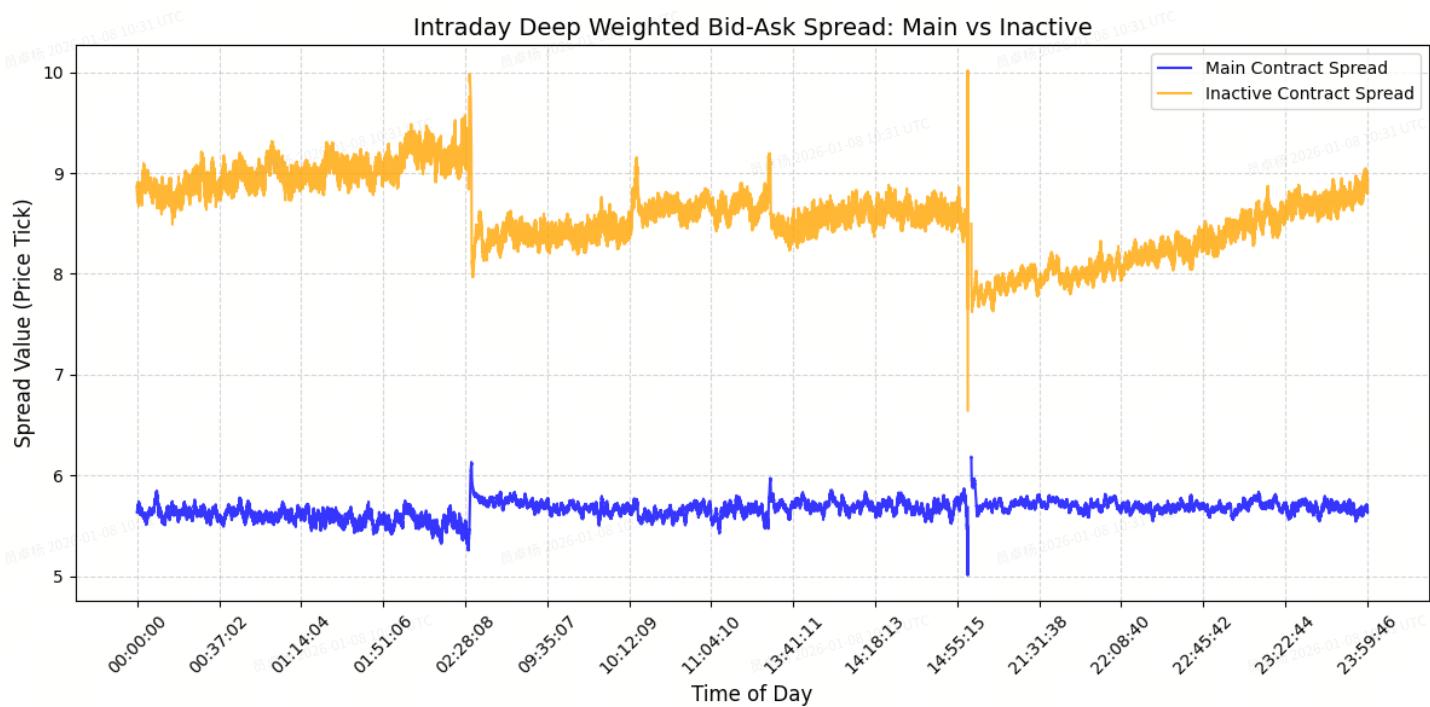
$$\bar{P}_{ask} = \frac{\sum_{i=1}^5 AP_i \cdot AV_i}{\sum_{i=1}^5 AV_i}$$

2. 买方加权均价 (Weighted Bid Price):

$$\bar{P}_{bid} = \frac{\sum_{i=1}^5 BP_i \cdot BV_i}{\sum_{i=1}^5 BV_i}$$

3. 深度加权价差 (Depth Weighted Spread):

$$Spread_{depth} = \bar{P}_{ask} - \bar{P}_{bid}$$



基于深度的价差同样可以明显反映出主力合约和非主力合约的流动性分层，用于评估当前时刻的流动性。可以发现主力合约由于流动性较好，其深度价差通常集中在5-6之间，即大多数时刻每一个档位上都有挂单，深度价差取决于订单簿上的挂单量分布，而非主力合约上可以看出随着开盘时间推进，品种活跃度逐渐降低，挂单变得更加稀疏，因此深度价差拉大。

然而，测试表明深度价差和未来5分钟收益绝对值之间的相关性很低，尤其是主力合约上相关性只有0.03。

| 指标类型 | 和主力合约 相关性 | 和非主力合 约相关性 |
|------|--------------|---------------|
| 深度价差 | 0.036 | 0.119 |

主要原因有以下几点：

1. 忽视流动性需求：

深度价差默认taker要将五档订单全部进行交易，预设了一个很大幅度的流动性吞噬，忽略了当前市场对订单流的需求并没有这么大。例如，流动性充裕时假设挂单量有500手，但市场可能只需要50手，用500手挂单去衡量深度价差显然是不合适的。

2. 大订单影响：

五档深度价差非常容易受到远端大笔挂单的影响，五档上挂大单并不意味着流动性不好，但是会导致深度价差拉大。

3. 对密集订单的流动性区分不明显：

深度价差对主力合约内部流动性的区分程度很低，因为主力合约上挂单非常密集，基本上每档都有挂单，求加权价差很难体现出日内的流动性区分程度。

2.3 基于模拟订单流的深度价差计算

基于之前的实验，我们发现一个好的流动性指标必须结合当前的流动性需求，综合考虑订单供需双方的博弈情况。衡量订单流需求最直观的代理变量就是过去一段时间的交易量。因此，我们选择计算过去30s合约的成交量，并据此计算一个虚拟订单流，用于预估其对当前市场订单簿的冲击成本。

基于该订单流，我们让其对买卖双方分别进行模拟吃单。从1档开始按成本从低到高进行交易，直到该订单流耗尽为止，对买卖双方计算该订单的成交均价，做差得到基于当前市场订单需求的深度价差。显然，模拟订单流的大小是指标计算的关键，设置的模拟订单量阈值过小，其无法击穿订单簿1-2档，几乎等价于一档价差；阈值过大，其击穿全部五档订单流，则基本等于五档深度价差。实验中，我们选择分别用0.5s、1s、3s和5s的平均交易量作为当前市场冲击强度，暂时假设买卖双方需求水平一致，即对交易量除以2，作为双方的模拟订单流。对于击穿订单簿的情况，我们暂时假设其在上方2个tick成交，但一般我们不希望其经常击穿订单簿。

1. 模拟交易量 (Target Volume):

我们选取时间窗口 τ (0.5s, 3s, 5s) 的平均成交量 \bar{V}_τ 作为市场需求强度，假设买卖需求均分：

$$Q_{target} = \frac{\bar{V}_\tau}{2}$$

2. 模拟成交均价 (Simulated Execution Price):

以卖方为例，模拟买入 Q_{target} 的过程。令 v_i 为第 i 档实际吃单量，\$R\\$ 为击穿 5 档后的剩余量， $delta$ 为最小变动价位：

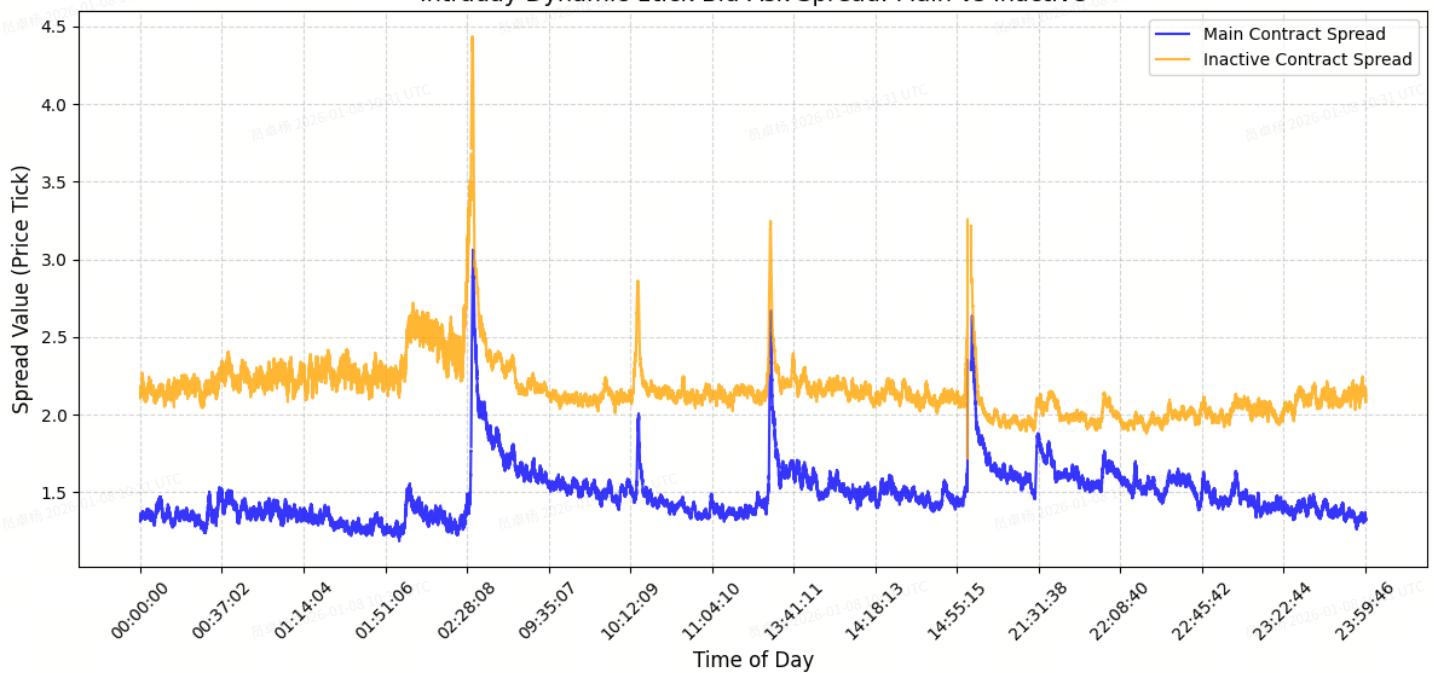
$$\bar{P}_{ask} = \frac{\sum_{i=1}^5 (v_i \cdot AP_i) + R \cdot (AP_5 + 2\delta)}{Q_{target}}$$

同理计算买方均价 \bar{P}_{bid} 。

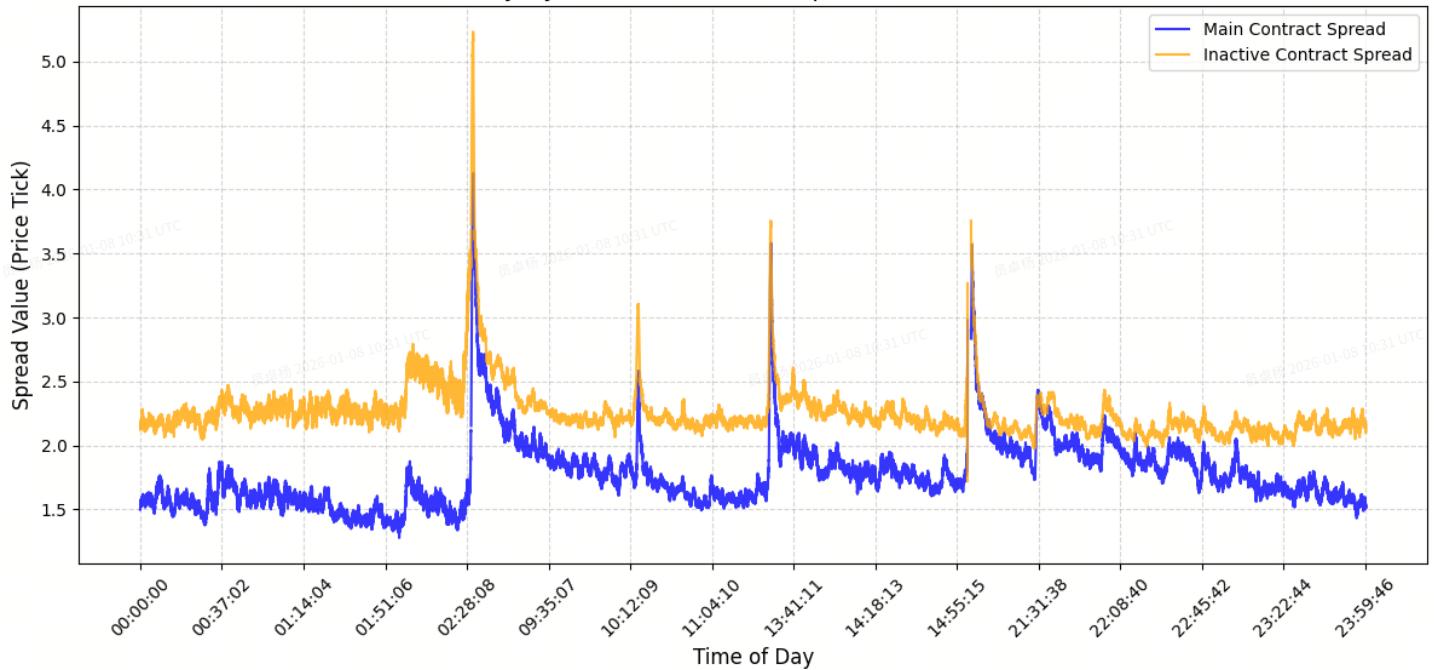
3. 最终模拟订单流深度价差 (Impact Spread):

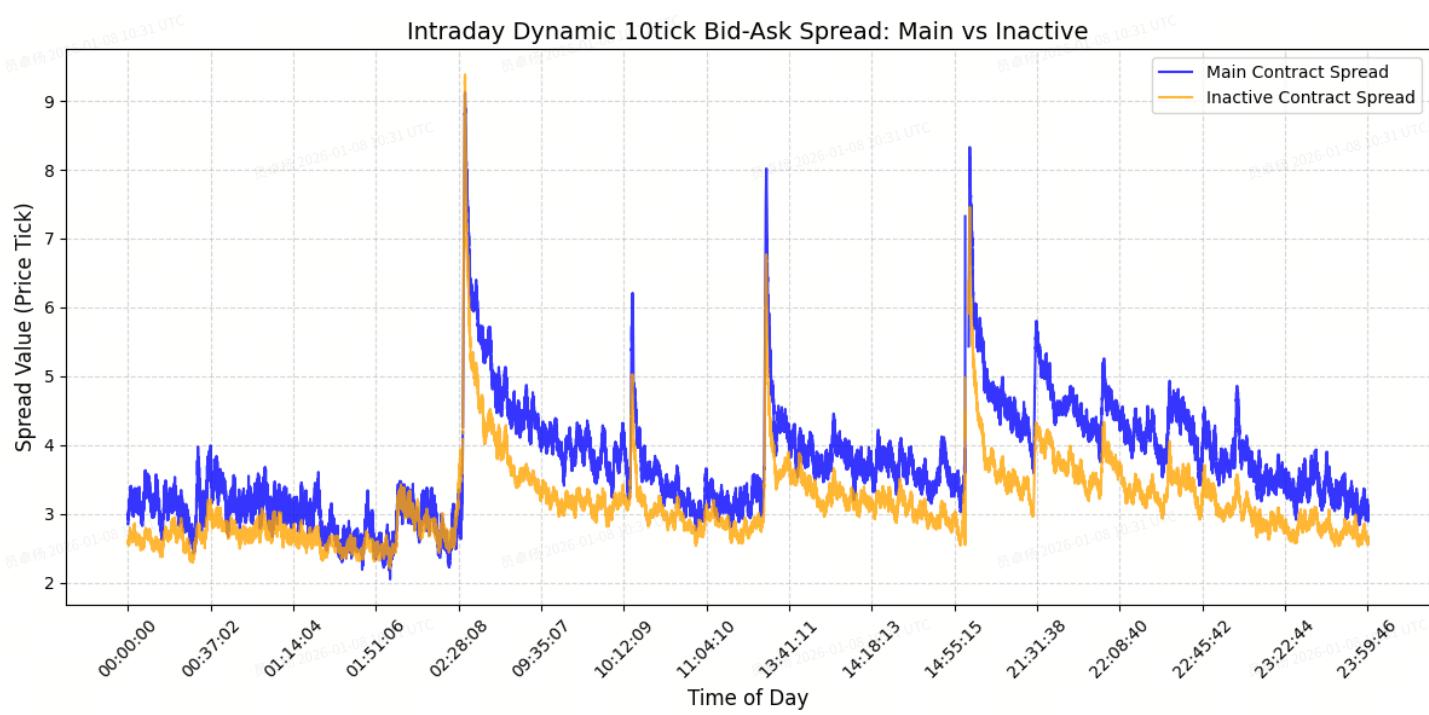
$$Spread_{impact} = \bar{P}_{ask} - \bar{P}_{bid}$$

Intraday Dynamic 1tick Bid-Ask Spread: Main vs Inactive



Intraday Dynamic 2tick Bid-Ask Spread: Main vs Inactive





| 模拟订单选择 | 和主力合约相关性 | 和非主力合约相关性 |
|--------|----------|-----------|
| 1tick | 0.318 | 0.209 |
| 2tick | 0.322 | 0.247 |
| 6tick | 0.321 | 0.267 |
| 10tick | 0.315 | 0.268 |

测试发现，在合约内部，用模拟订单流进行流动性评估的有效性很高，同未来收益率绝对值的相关性明显高于简单一档价差和深度价差。然而，该指标难以对主力合约和非主力合约之间的流动性差异进行有效的刻画。测试发现，主要问题在于主力合约和非主力合约之间的交易量差异过大，主力合约过去30s成交量的中位数是737，非主力合约则是32，相差超过20倍，因此如果简单地对非主力合约也直接通过未来若干秒换算模拟订单流，那么非主力合约的订单流大多数时候小于1，这一指标会退化成简单1档价差，导致无法区分主力合约和非主力合约。所以，我们加入一个动态调整逻辑，如果上一个交易日的模拟订单流中位数小于2，我们就将其乘上一个调整系数为5，即用5s的平均交易量作为模拟订单流大小。对主力和非主力合约进行不同程度的scaling之后，其流动性可以显示出比较明显的分层，同时在同一品种内部，能够比较鲜明地显现出其在不同时间段上的变化差异，即开盘、尾盘特定时段的脉冲效应以及随后流动性逐渐收敛的过程，和未来收益波动的相关性也可以达到0.25-0.3以上，表明相较于简单的一档价差和深度价差，从结果上，基于模拟订单流的深度价差指标更能够反映市场当前的流动性博弈。

可以发现，在日内基于模拟订单流的流动性成本在开盘之后是逐渐下降的，并在尾盘会出现上升，这符合我们认为的市场逻辑，即开盘初期市场具有较大的分歧度和不确定性，做市商为降低逆向选择成本也会拉大价差，导致市场交易的实际流动性成本增大，随着时间推进市场分歧减小，波动下降，做市商也会降低挂单范围。然而在之前计算的一档价差和深度加权价差指标在开盘后是逐渐增大

的，这是因为他们只考虑了当前盘面上的挂单水平，而没有考虑到市场的真实博弈，因此简单地认为价差变大因此流动性下降并不符合市场上流动性博弈的逻辑。

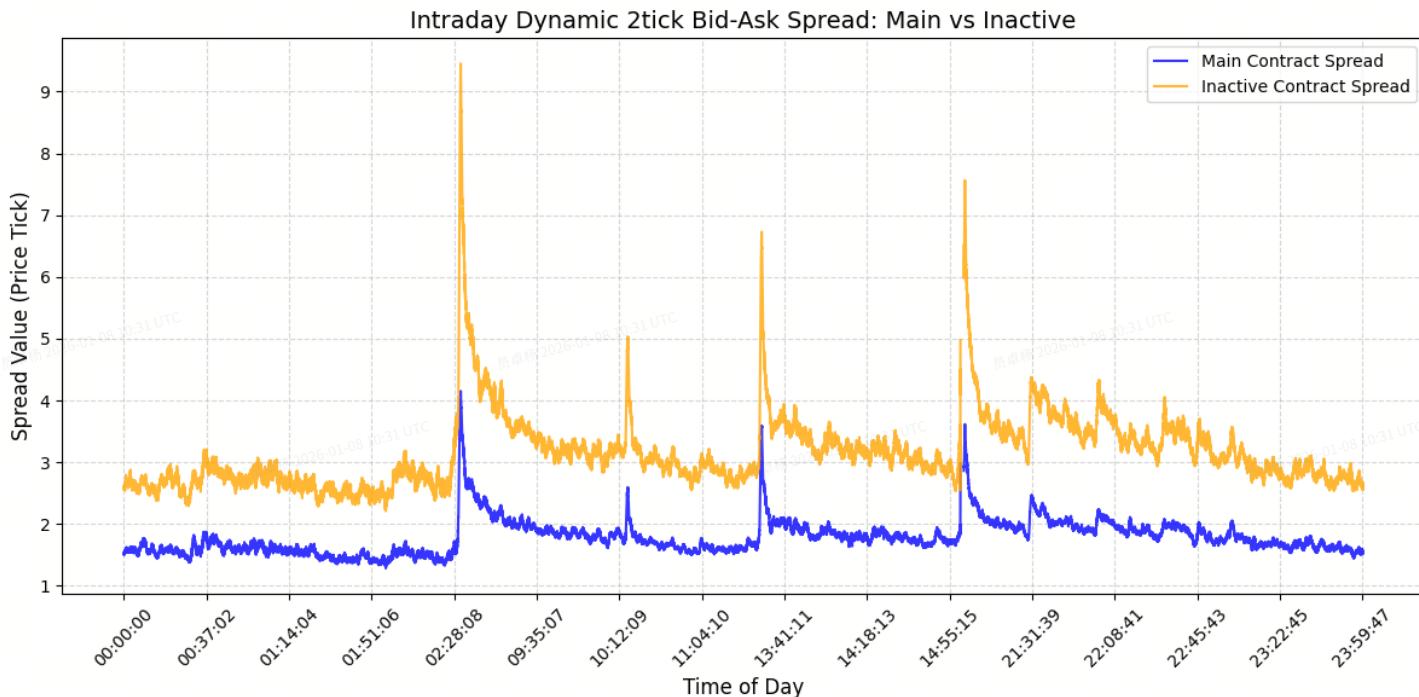
MarketMaker视角：

虽然盘面价差变大，但是市场波动率和分歧度下降，应当主动调低价差否则实际上很难成交，符合交易逻辑；

Taker视角：

如果希望在短时间内完成大量交易，那么必然会砸穿当前的订单流，应该直接考虑简单的深度加权价差估算作为成交成本；

如果希望在未来一段时间内用接近vwap的价格分批完成交易，那么我们并不主导市场，而是市场上全部流动性需求者的一部分，应当充分考虑流动性供需的博弈情况，用模拟订单流价差来评估流动性。



| 模拟订单选择 | 和主力合约相关性 | 和非主力合约相关性 |
|--------|----------|-----------|
| 动态阈值调整 | 0.318 | 0.265 |

3.对于模拟订单流价差指标的改进

基于之前的计算，我们基于模拟订单流价差得到了一个相对更好的流动性评估指标。然而，这一指标在细节处理上存在一些问题。首先，对于订单簿被击穿的情况，研究简单用2个tick来评估击穿后的交易成本，但这一做法并不严谨，可以结合一些订单簿斜率相关的情况尝试进行更精细的估测。其次，我们在研究中简单地认为买方和卖方的订单簿需求是对称的，但这在实际交易中并不符合逻辑，

我们可以尝试对之前的交易量进行更精细的划分，尝试建模过去一段时间内买卖双方各自的订单流需求量，计算各自的潜在市场冲击成本，进一步可以给出对市场流动性偏度的评估指标。同时，研究中没有对大订单及其存续情况进行精细区分，很多大订单不能被简单认为是流动性的提供者，需要进一步测试和区分。最后，模型此前的scaling是简单地对非主力合约乘上一个系数，但对于更广泛的品种，需要研究更合适的scaling方案。

3.1 模拟订单流击穿订单簿的边界情况处理

首先计算当前 5 档深度的价格跨度，作为基础波动率的代理。若价格相同，则使用最小一跳作为保底：

$$W = \max(|P_5 - P_1|, \delta)$$

计算剩余量相对于前 5 档总深度的比例，用于衡量冲击强度：

$$R = \frac{V_{rem}}{\max(V_{filled}, 1)}$$

使用平方根法则计算预期的额外滑点，并施加安全上限截断：

$$S_{extra} = \min\left(W \cdot \sqrt{R} \cdot \alpha, S_{max}\right)$$

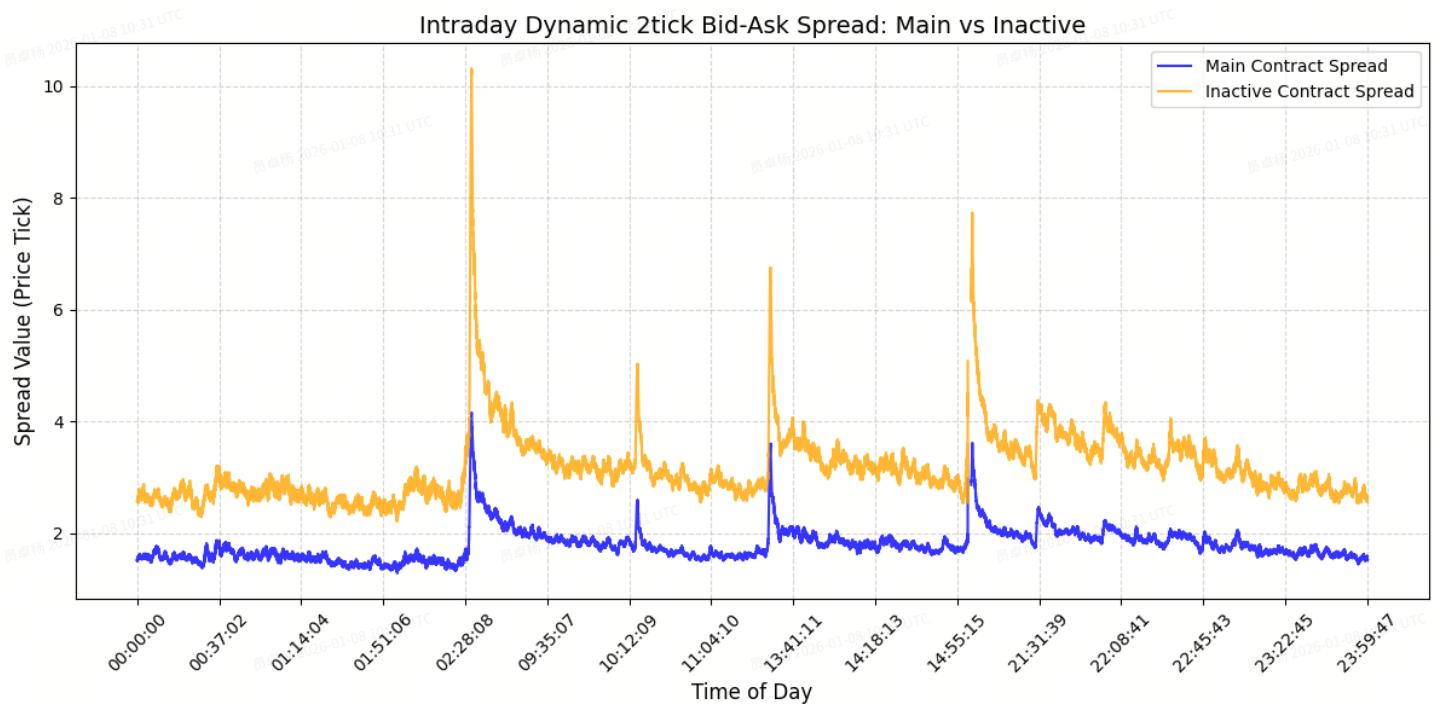
假设剩余订单的平均成交位置为第 5 档价格加上基础跳价和滑点的一半：

$$P_{penalty} = P_5 \pm \left(\delta + \frac{S_{extra}}{2}\right)$$

将前 5 档的实际成交与击穿部分的惩罚价格结合，得到最终的模拟成交均价：

$$\bar{P}_{final} = \frac{\sum_{i=1}^5 (P_i \cdot V_{i,executed}) + V_{rem} \cdot P_{penalty}}{V_{target}}$$

在以上计算中，我们对于击穿订单簿的时刻根据当前订单簿买五到买一的价差和挂单量分布，基于线性假设对击穿订单簿后的额外滑点进行了估计，用平方根和安全上限来约束滑点程度，避免流动性估计过高。通过以上修正，模型对于更高波动率时刻的流动性水平实现了相较于深度加权和简单阈值更好地估计，当前流动性计算指标的分布更符合观察直觉，并且非主力合约上的流动性水平与未来收益率相关性也有略微提升。



| 模拟订单选择 | 和主力合约相关性 | 和非主力合约相关性 |
|--------|----------|-----------|
| 动态阈值调整 | 0.318 | 0.2652 |

3.2 对于不同合约的模拟订单流大小Scaling方法设计

我们之前的处理中，直接对非主力合约的模拟订单流乘以固定的系数。我们现在希望以主力合约为基准，通过计算过去一个交易日主力和非主力合约的比例的交易量比例实现这个推导。对于每个交易日，回看上一个交易日主力合约的总成交量和非主力合约的总成交量，可以计算得到一个比值作为乘数。对于每一个品种，以主力合约作为基准，对非主力合约的模拟订单流乘上这个乘数实现归一化，保证在同一段时间内，对于主力合约和非主力合约，以相同规模的交易量作为流动性的参考，更有效地建立主力合约和非主力合约之间的流动性对比关系，同时保证每个交易日内非主力合约流动性大小的相对关系不受影响。经过调整后，可以发现两个合约之间的流动性水平呈现非常显著的分层性质，同时也保留了原有的日内流动性水平变化结构。可能的问题是在换算到同一基准后，开盘时段非主力合约的流动性冲击过高，导致显著击穿五档订单簿，可能导致部分时刻对模拟订单流冲击成本的估计不够准确。

成交量调整系数计算 (基于上一交易日)

$$\gamma_d = \frac{V_{main,d-1}}{V_{sub,d-1}}$$

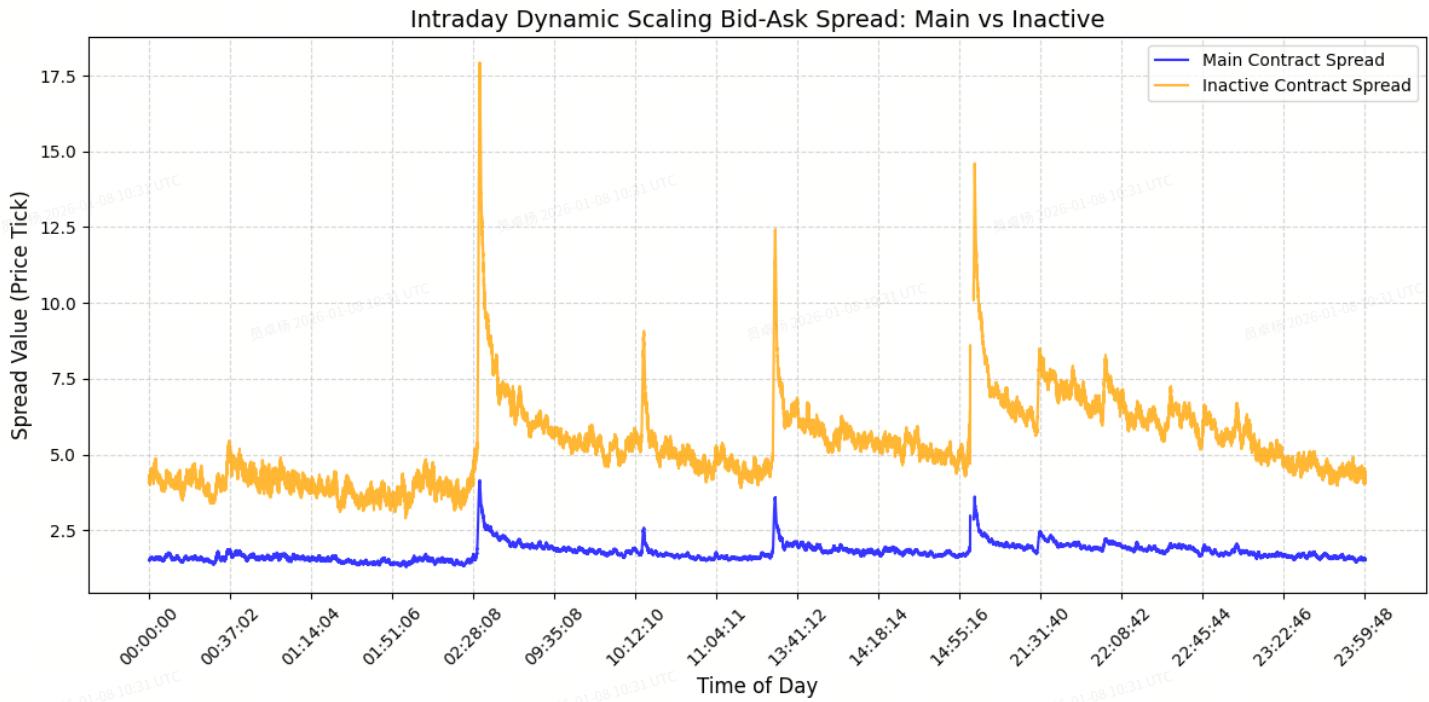
模拟订单流归一化

$$Q_{norm,t} = \begin{cases} Q_{raw,t} \cdot \gamma_d & \text{for Non-Main Contract (非主力合约)} \\ Q_{raw,t} & \text{for Main Contract (主力合约)} \end{cases}$$

此时的非主力合约流动性指标的意义：

如果希望在非主力合约上完成与主力合约相近规模的交易，应当付出多高的流动性成本

这样做标准化的优势在于可以更方便地对同一期货品种内不同合约的流动性进行比较，同时方便进行调整，只需要判断当前哪个是主力合约即可。随着时间推进，近月合约接近交割，近月和远月合约之间交易量之比逐渐收敛，两者的流动性评估也会逐渐收敛到相对应的水平，不需要人为调整。



3.3 对于五档内超大单的修复设计

研究中发现在部分时刻存在不合理大单的情况，例如在2-5档上会出现超过500手的挂单，或者在某个档位挂单量超过1分钟平均成交量。这种大订单可能并没有真实在市场上提供流动性，例如挂单几秒即被撤回，但是在价差计算时通常会被超大单截断，导致流动性评估不准确。因此，我们尝试对大单进行识别，并探究如何修复以更好描述市场流动性。

首先，结合绝对数值和过去一段时间的交易量给出超大单的识别标准：

$$I_{LargeOrder} = \begin{cases} 1 & \text{if } V_{raw} > 500 \\ 1 & \text{if } 100 \leq V_{raw} \leq 500 \text{ and } V_{raw} > 2 \cdot Q_{dynamic} \\ 0 & \text{if } V_{raw} < 100 \end{cases}$$

首先对于小于100手的订单，我们通常不认为他是一个异常值；对于100-500手的订单，如果超过了过去30s交易量的2倍，意味着可能出现了异常大单；如果超过500手，我们认为存在异常大单。

对于大单的处理有以下几种逻辑：

- 1.直接进行截断处理。这种做法对大单的干扰较低，但是在我么计算价差的过程中几乎不会对冲击成本的计算产生影响；
- 2.将大单所在位置的挂单量用相邻挂单或其他四档的中位数填充；

3. 直接认为此处没有挂单，但这并不符合正常的市场逻辑。

因此，我们尝试对异常大单所在位置用订单簿中位数填充，来尝试修复。然而，实际测试表明对大单的识别处理，对于模拟成交价差的计算几乎不产生任何影响，这主要是因为买卖一档上的异常挂单相对较少，而其他档位上的大订单对于价差计算影响较小，同时研究中对过去价差取平均实现了进一步平滑，几乎不会影响结果。因此，暂时不需要考虑大单异常值处理。

4. 结合买卖双方方向对模拟订单流价差的改进和流动性偏度计算

当前的模拟价差计算建立在一个假设上，即同一时刻买卖双方的流动性需求是相同的。但是这个假设在实际交易中可能并不成立。因此，可以尝试对买卖双方分别计算过去一段时间的订单需求，计算各自的模拟订单成交价格来计算价差，能更好地反映双方的流动性冲击成本，便于测算流动性偏度。然而，期货市场不存在逐笔成交记录，因此只能通过tick-level数据来估算双方的成交结果。

4.1 基于Lee-Ready算法分配交易量方向

学术上通过tick数据推断买卖双方的成交需求的最常见方法是Lee-Ready法，通过比较最新成交价与上一时刻的买一/卖一价，以及价格变动方向，将每一笔成交量的增量（Volume Delta）拆分为主动买入量（Active Buy）和主动卖出量（Active Sell）：

1. 增量成交量：

$$\Delta V_t = \max(0, V_{cum,t} - V_{cum,t-1})$$

2. 交易方向推断 (D_t)：

令 A_{t-1}, B_{t-1} 为上一时刻卖一和买一价， P_t 为当前成交价。

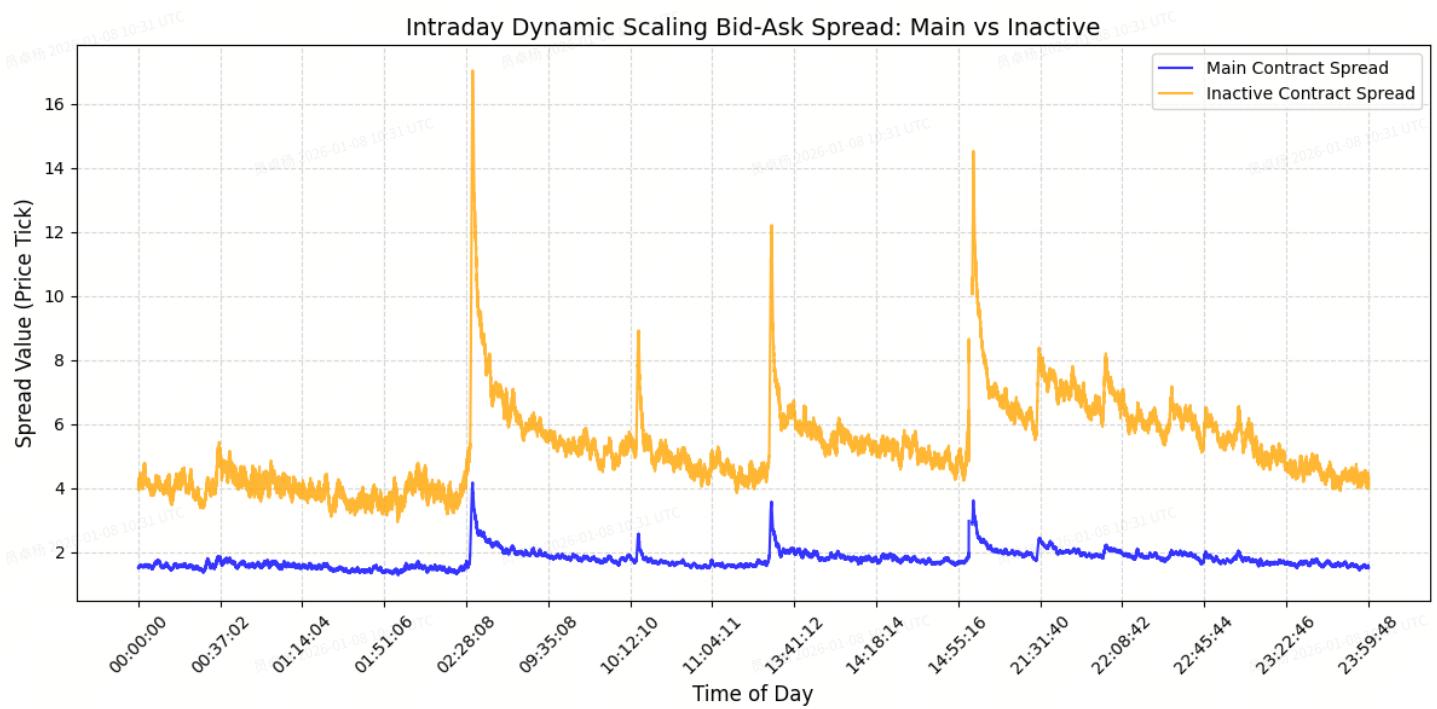
$$D_t = \begin{cases} 1 & \text{if } P_t \geq A_{t-1} \\ 0 & \text{if } P_t \leq B_{t-1} \\ 1 & \text{if } A_{t-1} > P_t > B_{t-1} \text{ and } P_t > P_{t-1} \\ 0 & \text{if } A_{t-1} > P_t > B_{t-1} \text{ and } P_t < P_{t-1} \\ 0.5 & \text{otherwise} \end{cases}$$

3. 流动性需求分解：

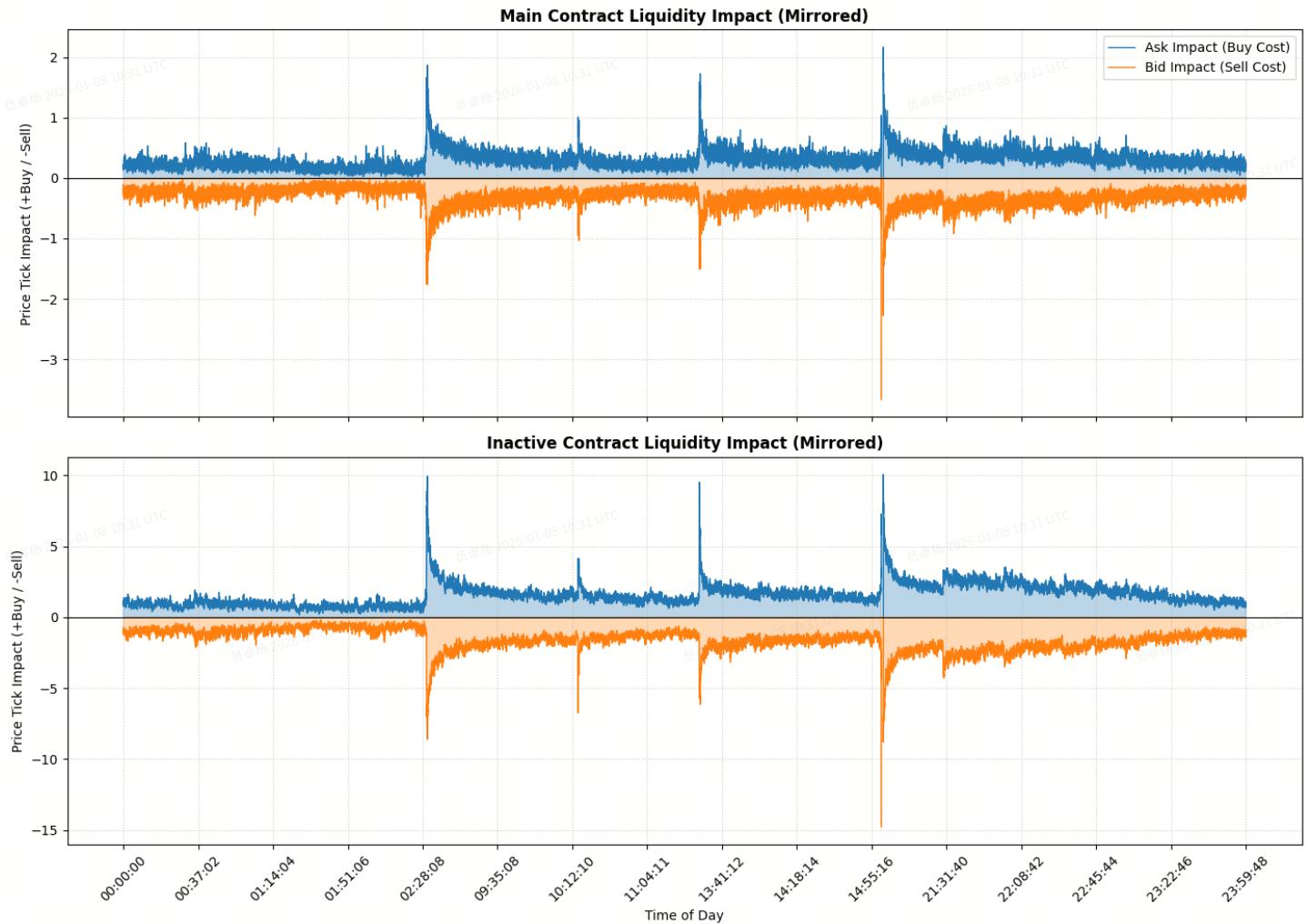
$$V_{active_buy,t} = \Delta V_t \cdot D_t$$

$$V_{active_sell,t} = \Delta V_t \cdot (1 - D_t)$$

这样，我们将过去一段时间的成交量分拆成卖方和买方双方的成交量，之后滚动计算过去30s买卖双方各自的成交量，并换算出平均1s的流动性需求量。之后，在买卖双方各自通过与此前相同的方法计算其成交成本并计算价差，作为流动性的估计指标。



从整体的价差分布来看，将流动性分方向统计对于模拟价差的影响并不大，和未来收益率绝对值的相关系数也基本一致，这主要是因为在市场方向并不十分明确的前提下，30s内买卖双方总成交量的差并不明显，因此分方向计算基本只对边界情况产生影响。



对双边模拟订单流的成交价格各自计算价差，测试表明双方的价差基本上是对称的，在统计上基本不存在时间上的差异。因此整体来看，分边计算价差对于总价差流动性指标的影响并不大，因此我

们主要需要关注其对于双方流动性偏度的影响。

4.2 基于交易量分边统计的流动性偏度计算

对于流动性偏度的计量，更好的方式是从结果反推，计算未来5分钟vwap与当前中间价的差，作为当前时刻买卖一方实际流动性的衡量依据，一方的流动性越弱，价格就越容易发生对应方向的移动，因此移动幅度应当是同向的，而双方的差更加直观的反映出哪一边的流动性更加薄弱，流动性更博弱的一方容易被击穿。因此，可以用类似单因子测试的方式来简单评估流动性偏度的有效性。

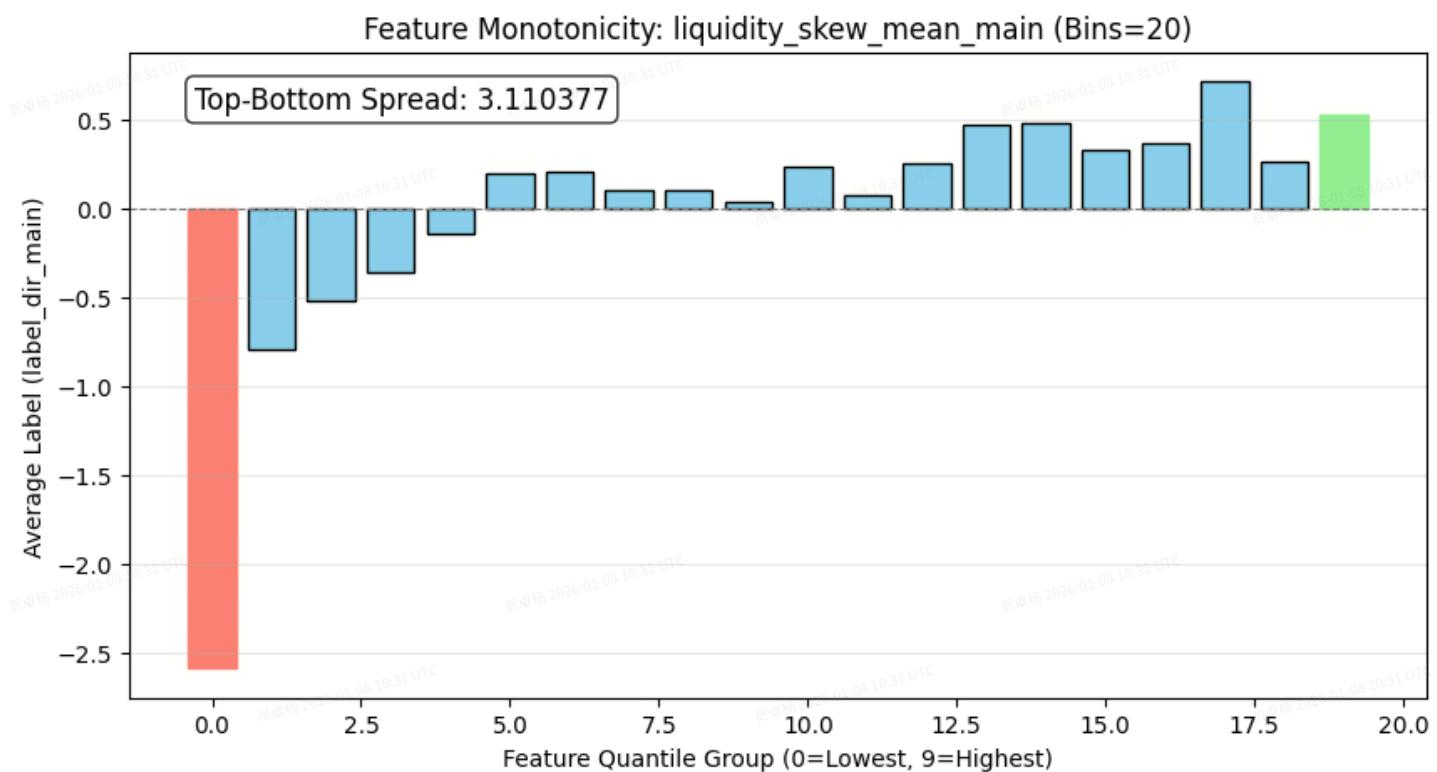
流动性偏度 (Liquidity Skewness) 公式：

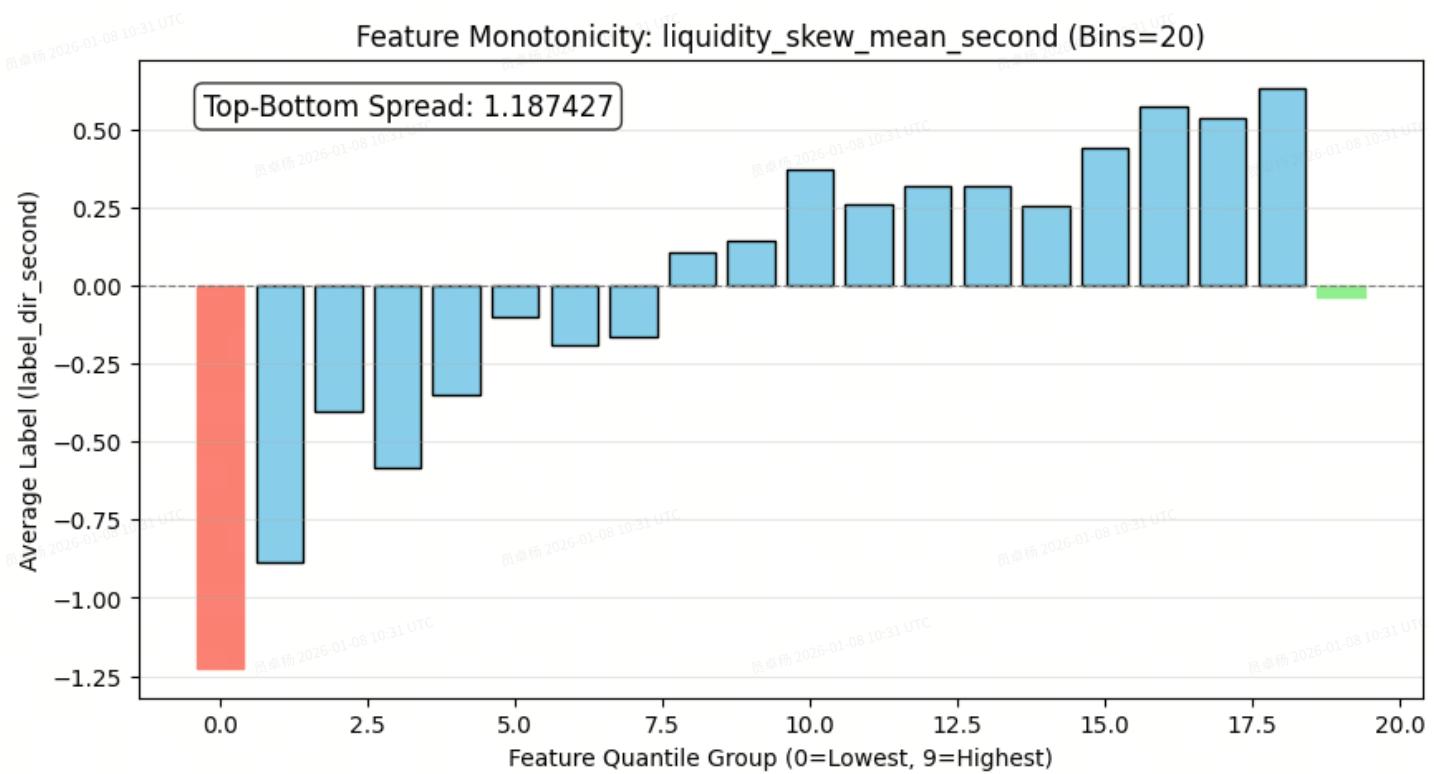
1. 单边冲击成本: $I_{ask,t} = \bar{P}_{ask,t} - A_{1,t}$ $I_{bid,t} = B_{1,t} - \bar{P}_{bid,t}$

2. 瞬时流动性偏度: $\text{Skew}_t = I_{ask,t} - I_{bid,t}$

3. 滚动平均偏度 ($N = 120$): $\overline{\text{Skew}}_t = \frac{1}{N} \sum_{i=0}^{N-1} \text{Skew}_{t-i}$

我们以未来5分钟vwap对于当前时刻偏离程度作为度量，进行20组分组测试，测试表明流动性测试多空分组比较明显，测试期内，空头端即买方流动性薄弱的时候更容易被击穿。





测试中主力合约的分层比非主力合约更明显，价差更大，这主要是主力合约交易活跃度更高的结果。

为建立一个比较基准，我们将当前分方向计算的流动性偏度指标，和之前双方模拟交易量相等的流动性偏度指标的分层测试结果进行对比，测试表明分方向计算的流动性偏度分层性质明显更好，如果简单地认为双方流动性需求一致，则计算得到的流动性偏度几乎完全无法反映真正的流动性供需和博弈情况。因此可以认为分方向统计流动性需求，能够更好的反映两边的流动性需求水平，从而更好地估计当前的流动性偏度。

| 流动性统计 | 主力合约分组价差 | 非主力合约分组价差 |
|-------|----------|-----------|
| 分方向 | 3.11 | 1.19 |
| 不分方向 | 0.86 | -0.31 |

4.3 基于BVC算法的交易量分配和流动性偏度计算实验

之前的交易量分配算法主要是基于价格变化方向实现的，并且是将交易量进行完全分配的。但这一做法的问题在于，一个tick的价格向一侧移动不代表交易完全成交在一边。如果完全将一个时刻的成交量分配在一边，在部分特殊时间段内会导致成交量几乎集中于一侧，一方流动性成本过高而另一方很低，使得流动性的度量不够平滑。为进行改进，我们尝试引入BVC (Bulk Volume Classification) 算法，根据价格变化的剧烈程度依概率进行成交量分配，使其更加平滑。

$$Z_t = \frac{P_t - P_{t-1}}{\sigma_{\Delta P}}$$

我们假设价格的变化率服从正态分布，之后通过正态分布 CDF 计算该笔成交量属于主动买入的概率：

$$\alpha_t = \Phi(Z_t) = \int_{-\infty}^{Z_t} \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx$$

最后，根据概率将总成交量拆解为买方和卖方分量：

$$V_{buy,t} = V_t \times \alpha_t$$

$$V_{sell,t} = V_t \times (1 - \alpha_t)$$

测试发现改为BVC算法以后，对整体流动性价差指标的分布基本没有影响，但是流动性偏度的分组显著性要低于直接对交易量进行分边处理，这可能是由于概率分配交易量带来的信息损失导致的结果。

同样，我们还可以通过midprice对之前订单簿的偏离程度，等比例地进行成交量分配。

计算基准价格相对于 Bid_{t-1} 的偏移程度，并用价差宽度标准化：

$$\lambda_t = \frac{P_t - B_{t-1}}{A_{t-1} - B_{t-1}}$$

将系数限制在 $[0, 1]$ 区间内，以处理价格击穿买一或卖一的情况：

$$\alpha_t = \max(0, \min(1, \lambda_t))$$

成交量分配：

$$V_{buy,t} = V_{total,t} \cdot \alpha_t$$

$$V_{sell,t} = V_{total,t} \cdot (1 - \alpha_t)$$

4.4 结合当前时刻交易基准价格的改进

当前对成交量的双边分配可以分为两个部分，首先是确定当前时刻基准价格，其次是根据基准价格的变化方向来分配成交量。

基准价格的选择共有3种方式：使用中间价，使用last，以及当前tick的成交均价。中间价相对更稳定，但是可能会忽略订单簿被击穿前的价格变化；last表示最后一笔成交价格，更直观反映最新的价格，但是last变化过于频繁，其变动并不说明整个tick的资金流向；成交均价可以相对准确地反映当前tick的交易方向，但在成交不活跃的品种上作用有限。

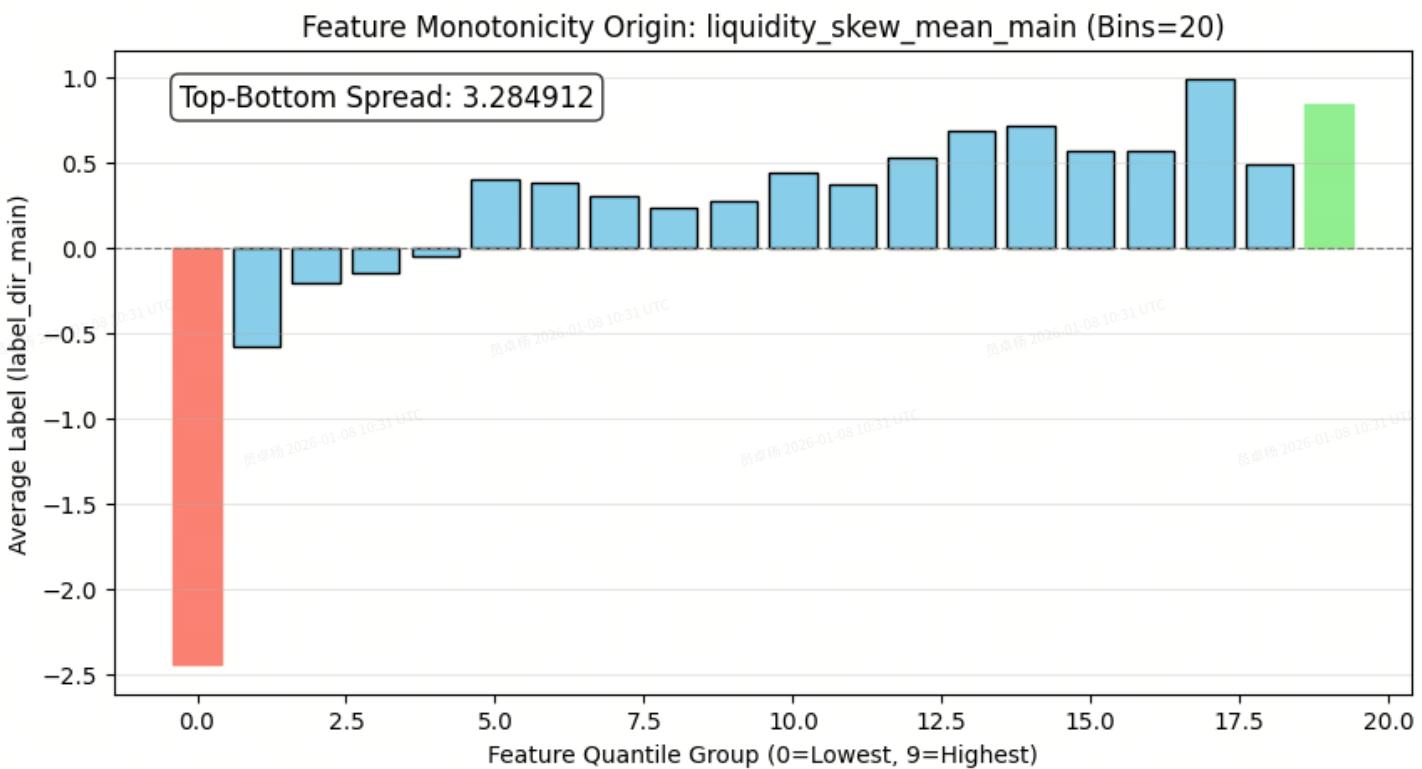
成交量分配方法除了此前提到的Lee-Ready和BVC以外，也可以通过其在订单簿中的位置进行等比例分配，同样有3种方式。之前实验表明成交量分配方法对流动性指标影响不大，主要影响流动性偏度，因此我们对以上的情况进行组合，观察其流动性偏度的分层显著性。

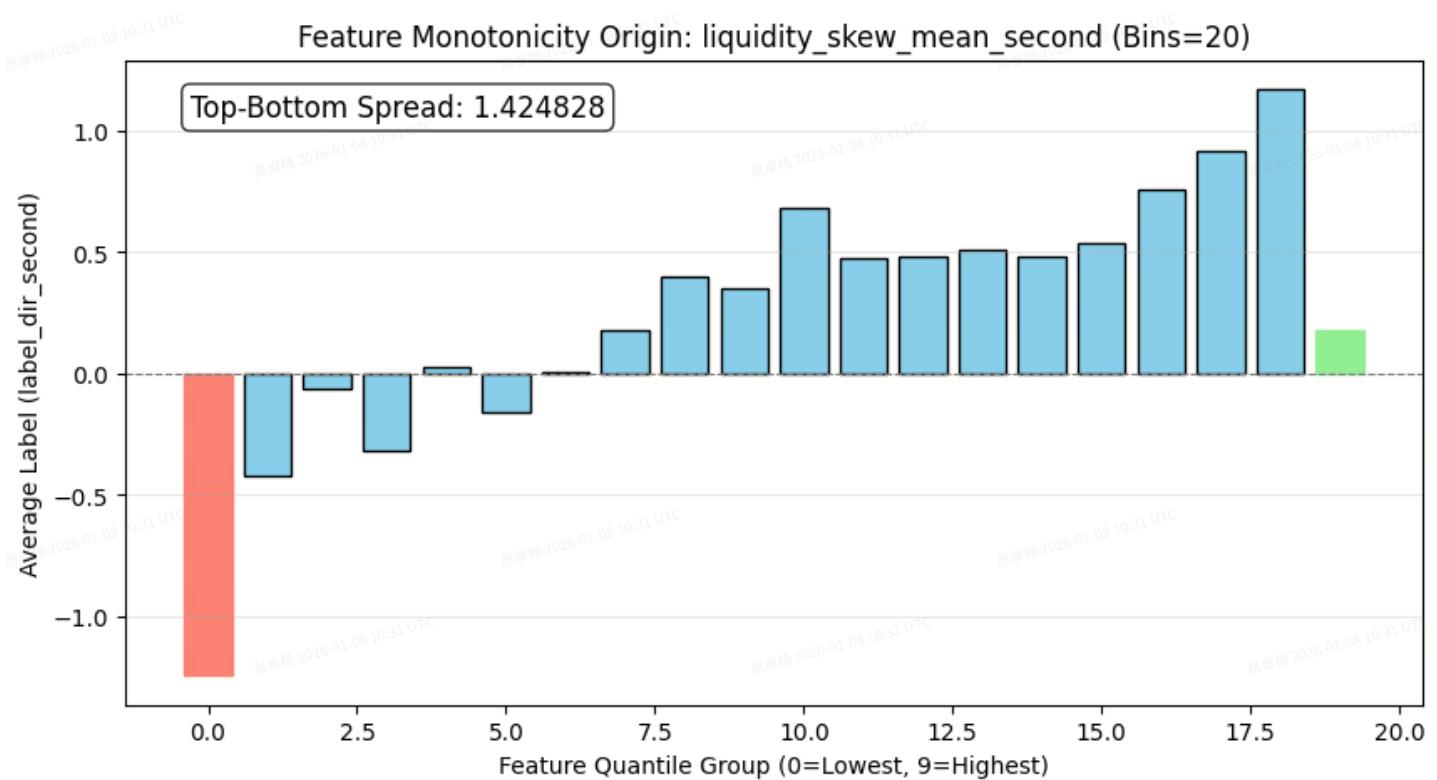
| 成交量分配算法 | 成交量分配 基准价格 | 流动性偏度 主力合约20 档分组价差 | 流动性偏度非主 力合约20档分组 价差 |
|-----------|---------------|--------------------------|---------------------------|
| Lee-Ready | mid_price | 3.11 | 1.19 |

| | | | |
|-----------|-----------|------|-------|
| Lee-Ready | last | 2.00 | 0.455 |
| Lee-Ready | vwap | 3.29 | 1.04 |
| BVC | mid_price | 1.86 | 0.58 |
| BVC | last | 2.78 | 0.66 |
| BVC | vwap | 2.6 | 1.05 |
| 位置比例分配 | mid_price | 3.28 | 1.4 |
| 位置比例分配 | last | 1.97 | 0.60 |
| 位置比例分配 | vwap | 1.66 | 0.69 |

通过测试可以发现，用last作为基准价格显然不是一个衡量流动性偏度的有效方案，因为last中包含的噪声太大。vwap和midprice相比，此前的直觉上会感觉vwap更符合一般意义上对于博弈价格的定义，但实践表明使用midprice的效果更加稳定。其核心原因可能在于成交不活跃的时段上vwap的意义有限，而白银期货五档挂单通常比较稀疏，价格变化很快，挂单的变化可能先于成交价格变化，中间价移动代表订单簿的上移或者下移，对于成交方向的判断准确度更高。

而在算法上，BVC算法过成交量分配平滑导致信息损失，由于白银期货价格移动很快，非常容易击穿上一tick的买卖一档价格，因此在大多时候用lee-ready和位置比例分配计算结果是一致的，而在价格移动没有基础击穿一档的时刻，按照位置偏移比例分配交易量更加符合直觉判断，测试结果也更有效。因此，我们选择采用位移比例分配算法，结合midprice的变化来实现交易量的分配。





4.5 基于订单流补充速度的流动性修正

之前对于流动性偏度的计算已经考虑了当前时刻对于订单的需求以及当前挂单的情况，但是没有考虑订单流的补充速率。在实际交易中，提供流动性的除了当前订单簿上可见的挂单之外，还有潜在地会在未来一段时间内补充挂单的交易者。为使计算的流动性指标更符合真实情景，我们选择计算过去一段时间内买卖双方各自的订单流补充情况来评估当前订单流的恢复速度，对模拟订单流的交易量进行一定修正，更好的刻画流动性供需双方的博弈结果。

瞬时补充流计算

补充流等于“盘口深度的净变化”加上“被消耗掉的成交量”，此处的计算参考了OFI的计算方法：

$$R_{ask,t} = V_{active,t} + \begin{cases} D_t - D_{t-1} & \text{if } P_t = P_{t-1} \text{ (价格不变, 看深度差)} \\ D_t & \text{if } P_t < P_{t-1} \text{ (价格下跌/变优, 新挂单全额补充)} \\ -D_{t-1} & \text{if } P_t > P_{t-1} \text{ (价格上涨/变差, 原挂单全额撤退)} \end{cases}$$

滚动弹性系数

计算过去 w 个时刻（例如 60 ticks）的平均补充强度，并归一化到 $[-\lambda, \lambda]$ 区间：

$$\rho_t = \lambda \cdot \tanh \left(\frac{\frac{1}{w} \sum_{i=0}^{w-1} R_{t-i}}{S} \right)$$

其中， S 为过去 60 个 tick 中每个 tick 成交量均值的 5 倍，即平均 2.5s 的成交量，用于控制弹性系数的平滑程度。

有效订单流修正

根据弹性系数修正原始模拟流量。

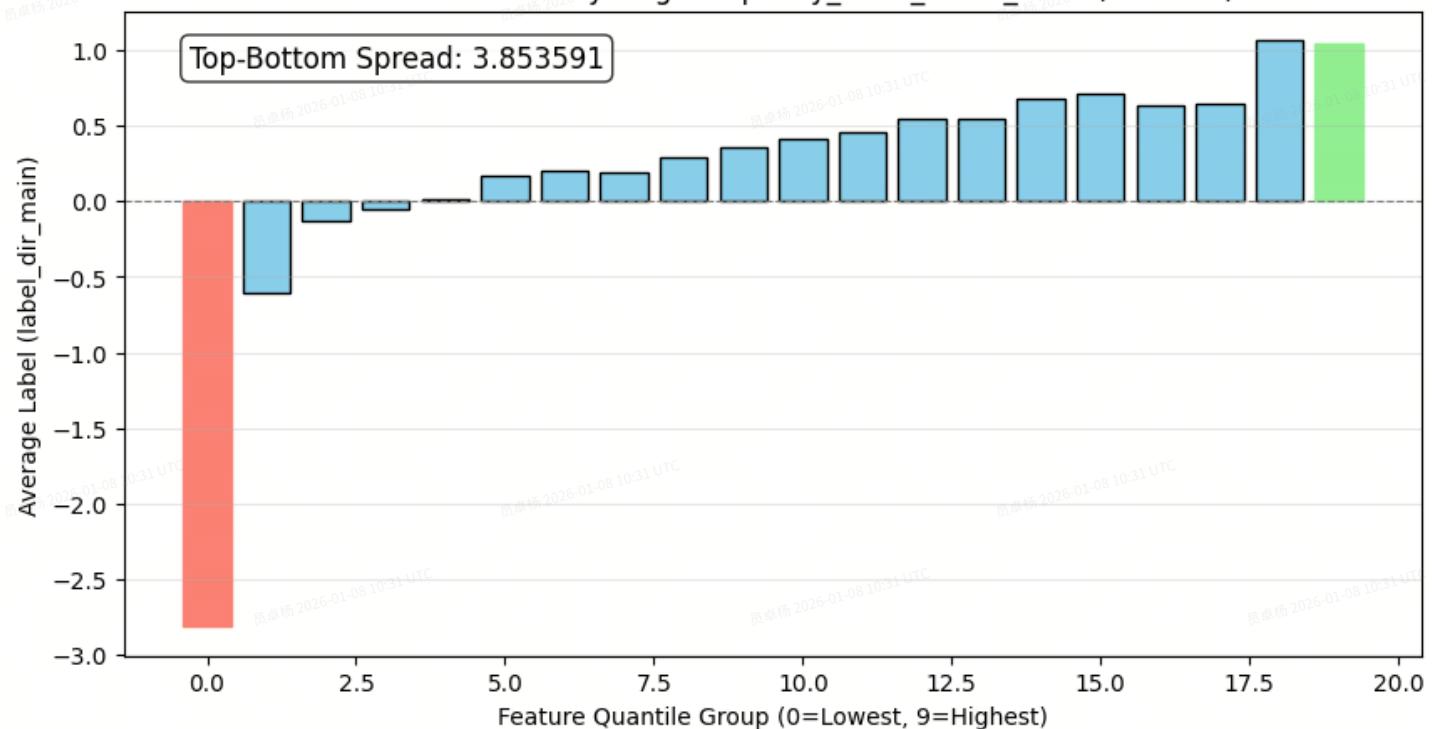
- 若订单流补充速度为正 ($\rho > 0$)，有效流量减少 (冲击力减弱)。
- 若订单流补充速度为负 ($\rho < 0$)，有效流量增加 (冲击力放大)。

$$Q_{eff} = Q_{raw} \cdot (1 - \rho_t)$$

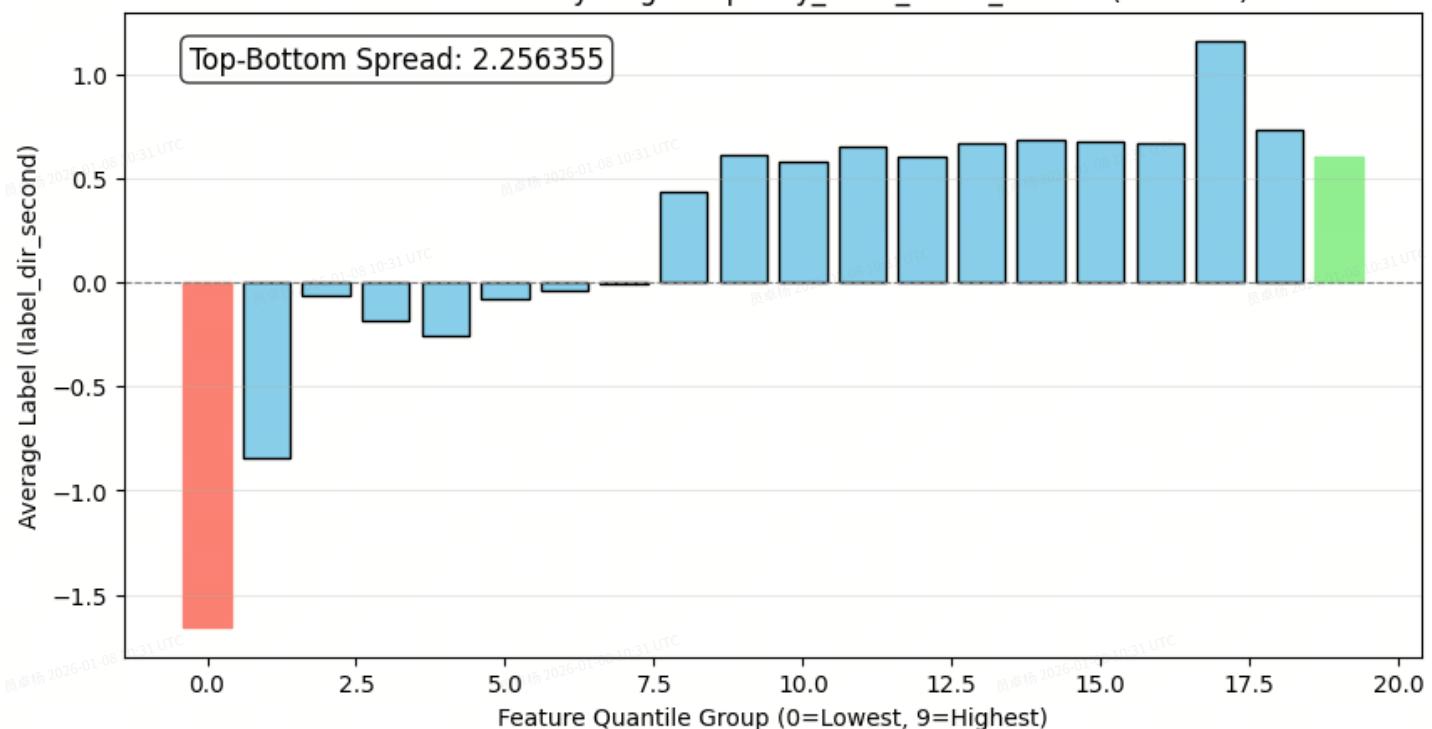
| 订单簿档数 | 弹性系数 | 主力合约价差 | 非主力合约价差 |
|-------|------|--------|---------|
| L1 | 0.3 | 3.31 | 1.50 |
| L1 | 0.5 | 3.286 | 1.51 |
| L1 | 0.7 | 3.283 | 1.58 |
| L5 | 0.3 | 3.55 | 1.99 |
| L5 | 0.5 | 3.75 | 2.29 |
| L5 | 0.7 | 3.85 | 2.26 |
| 原始 | 原始 | 3.28 | 1.4 |

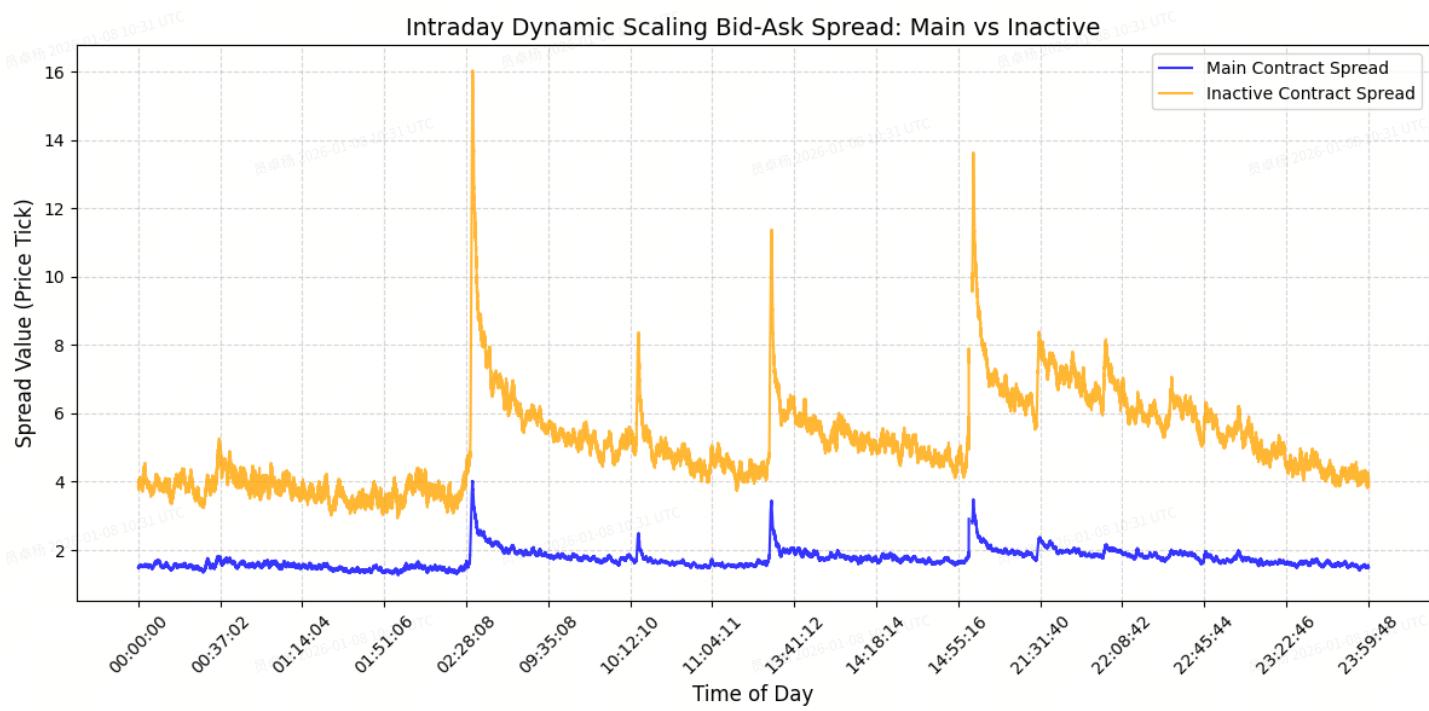
从结果上看，用五档挂单量的变化计算订单簿补充速率时，经过订单流补充修正后的流动性偏度指标明显能够更好的刻画买卖双方流动性的差异，因此对未来一段时间内价格突破的方向有更好的提示作用。从分组图像看，流动性偏度指标在经过订单流补充速率的修正后，其多头端的分组情况有明显改善，即对于卖方的流动性薄弱的情况可以进行更好的刻画，此前多头分组弱的主要原因可能在于，当基于过去成交量分组计算的流动性指标意识到卖方的流动性需求远大于供给时，可能已经到达上涨趋势的尾声，订单流补充速度明显加快，因此卖方的流动性有了很好的改进，不太容易被突破。经过流动性恢复速度的修正后，流动性指标对于流动性恢复更为敏感，在流动性加速恢复时流动性指标会更加平缓，更符合现实中的流动性博弈情况，因此能够更好的衡量买卖双方流动性的相对强弱。

Feature Monotonicity Origin: liquidity_skew_mean_main (Bins=20)



Feature Monotonicity Origin: liquidity_skew_mean_second (Bins=20)





5. 流动性指标聚合

我们之前计算得到了每一个 tick 对应的流动性和偏度指标。此前我们选择将最近 120 个 tick 的流动性指标取平均，作为当前时刻聚合得到的流动性指标。当前流动性指标相对完善后，我们希望测试怎样更有效地对每个 tick 的流动性指标进行聚合。对于简单地滚动平均，可以选择不同聚合时间窗口。同时，我们可以尝试用其他方法进行聚合。如果希望突出当前时刻的影响，可以用指数滑动加权；如果希望度量每个 tick 上对订单整体的冲击成本，可以用成交量进行加权；如果希望尽可能对不同时刻进行平滑，可以用波动率加权。

6. 后续研究计划

6.1 对超大挂单的建模和精细化处理

6.2 不同品种的流动性对比和研究

6.3 流动性和偏度指标在模型中的效果测试