

证券研究报告

2011-09-14



A股市场知情交易概率初探

国信证券经济研究所 金融工程部
量化交易策略小组

证券分析师： 秦国文
证券投资咨询执业资格证书编码： S0980511080001
联系人： 张璐楠



国信证券经济研究所

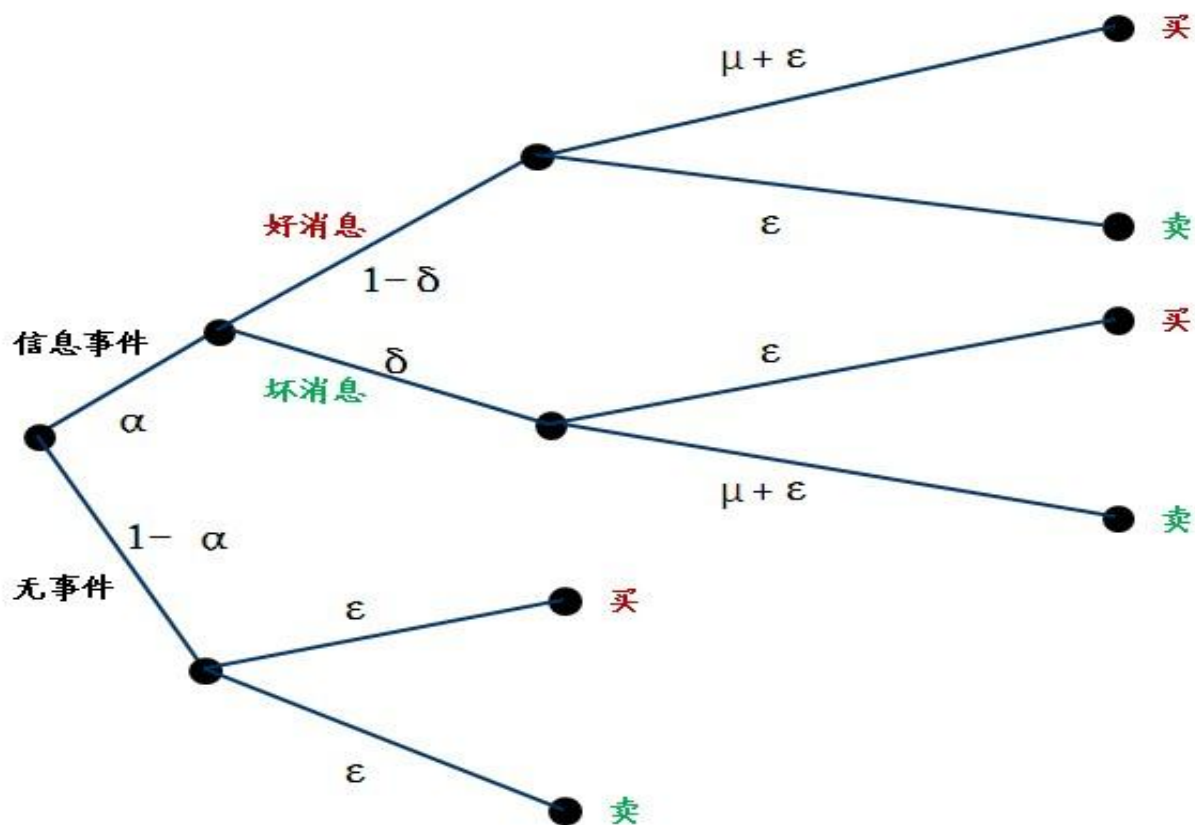
Guosen Securities Economic Research Institute



背景及内容摘要

- 介绍交易指令的逆向选择、交易流毒性和PIN。
- 知情交易概率（PIN）的模型说明
- 新的VPIN算法
(VPIN是一个不依赖于不可观测参数的交易毒性估计模型。VPIN估计值越大，说明市场毒性越高，知情交易者占的比例越高。另外，VPIN更新的速度越快，说明市场上的知情信息越多。)
- 实证检验，通过估计50ETF的VPIN，展示了其对应的CDF函数在大盘崩盘前提前预警提示大盘毒性极高，告诉投资者市场面临大幅波动（在高位一般指大幅调整）的概率迅速增大。
- VPIN对于波动性有较为明显的预测能力，并且它对未来的高绝对收益提供了指示作用。

不确定事件下的交易序列



- 微观结构模型

资料来源：国信证券经济研究所整理

不确定事件下的交易序列

$$P(t) = (P_n(t), P_b(t), P_g(t)) \longrightarrow P(0) = (1-\alpha, \alpha\delta, \alpha(1-\delta))$$

$$E[S_i | t] = P_n(t)S_i^* + P_b(t)\underline{S}_i + P_g(t)\bar{S}_i$$

$$\text{上一刻的标的资产期望值 } S_i^* = \delta\underline{S}_i + (1-\delta)\bar{S}_i$$

$$\text{卖出询价 Bid} \longrightarrow B(t) = E[S_i | t] - \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)} [E[S_i | t] - \underline{S}_i]$$

$$\text{买入询价 Ask} \longrightarrow A(t) = E[S_i | t] - \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)} [\bar{S}_i - E[S_i | t]]$$

传统PIN模型

在t时刻买入卖出询价的价差为：

$$\Sigma(t) = A(t) - B(t)$$

$$\Sigma(t) = \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)} [\bar{S}_i - E[S_i | t]] + \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)} [E[S_i | t] - \underline{S}_i]$$

在时段的初期，好坏事件出现概率相等的时候，即 $1-\delta = \delta$ ；

阶段初始的价差有一个非常简单的形式：

$$\Sigma = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} [\bar{S}_i - \underline{S}_i]$$



PIN模型

这个模型的关键点是初始指令来自知情交易者的概率，知情交易的概率记作PIN。可以直观地看出在一段时间内的第一笔交易它是“知情的”概率由下式给出：

$$PIN = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon}$$

分子是基于有信息而下达指令单的发生概率

分母是所有交易指令单的发生概率

上述的公式描述了：流动性提供者需要正确地估计他们的PIN以发现进入市场的最优报价。在PIN的非预期变化后，加入不调节报单价格，流动性提供商将会产生损失。



VPIN与参数估计

计算PIN模型的标准方法是使用极大似然法估计出驱动交易的不可观测的参数 $(\alpha, \mu, \delta, \varepsilon)$ ，然后根据这些参数的估计值推断PIN。

在今年初的相关研究中提出另一种可以直接估计PIN的方法，该方法不需要对不可观测的参数进行中间估计。同时，它还能够在随机时间更新（传统估算PIN的方法并非随机更新，而是通过历史数据定期用MLE做参数优化），与市场上新信息的到达速率相匹配。这种基于成交量的方法，我们记为VPIN，对高频环境下的交易指令流毒性提供了一种简单的度量。

信息与时间的概念

高频交易模型最重要的一方面就是交易并不是在时间上平均分布的。交易以不规则的频率到达，有些交易比其它交易更加重要因为它们能够揭示不同数量的信息。



资料来源：天软资讯，国信证券经济研究所整理

不同产品的信息在不同的时间到达，因此产生了不同的日内成交量周期性。

成交量划分

为了实现基于成交量的取样，我们将交易序列以均匀成交量分组，称为一个“成交量篮子”， V 。一个成交量篮子是成交量加总为 V 的一组交易。

$$V = V_t^B + V_t^S$$

完成篮子的最后一笔交易如比所需要的大，超过的成交量将被划分到下一个篮子，按照这样的成交量篮子取样后我们获取到一个观测值。构建成成交量篮子时我们对**每一笔买入和卖出成交量进行分类**，成交量的方向与交易指令单毒性存在潜在联系，结合总成交量水平和成交量的方向可以指示出存在新信息的可能性。因此，买入成交量占多数说明毒性来自好的信息，反之亦然。我们通过观测买单和卖单之间交易的强度和不平衡性来估计出VPIN。

VPIN计量法

标准的PIN模型仅根据买卖交易的数量来推测潜在信息结构，成交量在模型中没有明显的角色。在我们分析的高频领域，交易数量是有问题的，很多时候看到的交易数量都是经过合成得出的。回到PIN的理论基础，我们真正想要的信息是知情和非知情交易者的交易目的。交易目的和交易数据间的联系是充满噪音的，因为交易目的可以被切分成很多块从而对市场的影响最小化。一个指令单也许会产生很多的执行，知情交易可以以不同的形式出现。基于这些原因，我们将每一个交易视为单位大小的加总。

Easley等人[2008]详细的介绍了
预期的交易失衡 (trade imbalance) :

$$E[|V_t^S - V_t^B|] \approx \alpha \mu$$

所有交易的
预期到达速率:


$$\frac{1}{n} \sum_{t=1}^n (V_t^B + V_t^S) = V = \underbrace{\alpha(1-\delta)(\varepsilon + \mu + \varepsilon)}_{\text{volume from up event}} + \underbrace{\alpha\delta(\mu + \varepsilon + \varepsilon)}_{\text{volume from down event}} + \underbrace{(1-\alpha)(\varepsilon + \varepsilon)}_{\text{volume from no event}} = \alpha\mu + 2\varepsilon$$

VPIN计量法

$$E[|V_t^S - V_t^B|] \approx \alpha\mu \quad \text{分子}$$

$$\frac{1}{n} \sum_{t=1}^n (V_t^B + V_t^S) = V = \underbrace{\alpha(1-\delta)(\varepsilon + \mu + \varepsilon)}_{\text{volume from up event}} +$$

$$\underbrace{\alpha\delta(\mu + \varepsilon + \varepsilon)}_{\text{volume from down event}} + \underbrace{(1-\alpha)(\varepsilon + \varepsilon)}_{\text{volume from no event}} = \alpha\mu + 2\varepsilon \quad \text{分母}$$


$$VPIN = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} = \frac{\alpha\mu}{V} \approx \frac{\sum_{\tau=1}^n |V_{\tau}^S - V_{\tau}^B|}{nV}$$

估计VPIN需要对每一个篮子的成交量V以及用来估计预期交易失衡和密集程度的篮子的数量n进行选择。举例，我们将V设定为每日交易量的五十分之一。如果我们选择n=50，我们将会结合50个篮子计算VPIN，假设一个交易日的交易量和平均交易量相等，这个数值就是日VPIN。我们会展示最优的V和n值可以通过一个最优化的方法找到。

VPIN的更新时间说明

1、我们希望更新VPIN的速度能够模拟信息到达的速率。为了实现这个目的，**我们使用成交量作为信息到达的一个代理变量。**

2、我们希望**每一次VPIN的更新基于可比的信息量。**

在低成交量的区间里，交易也可能非常失衡，但是低成交量时刻很难有新的信息在市场上产生。因此时钟时间更新VPIN将容易错误的利用了信息。



例子：以2009年7月29日的上证50ETF作为例子。这一天的成交量极高，所以我们的方法产生了137个VPIN计量结果，而每天平均是50个估计值。因为我们的样本容量n也是50，2009年7月29日用于估计VPIN的时间跨度仅仅是1个多小时，而之前平均是4小时。



资料来源：天软资讯，国信证券经济研究所整理

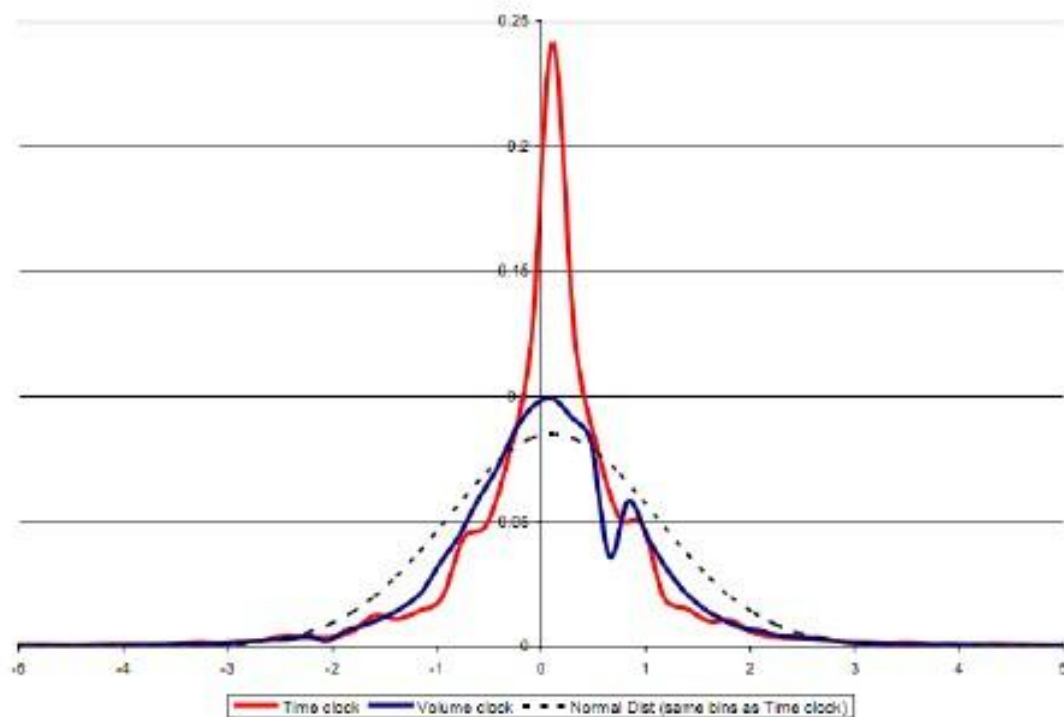
成交量时间 vs 时钟时间

- 2009.1.1至2011.9.1
- 平均每日50 个数据
- 一阶差分后标准化

红色是始终时间

VS

蓝色是成交量时间



资料来源：国信证券经济研究所整理

VPIN实证

1分钟频率取样

50ETF价格 Vs VPIN估计值



资料来源：天软资讯，国信证券经济研究所整理

VPIN实证

30秒频率取样

50ETF价格 Vs VPIN估计值

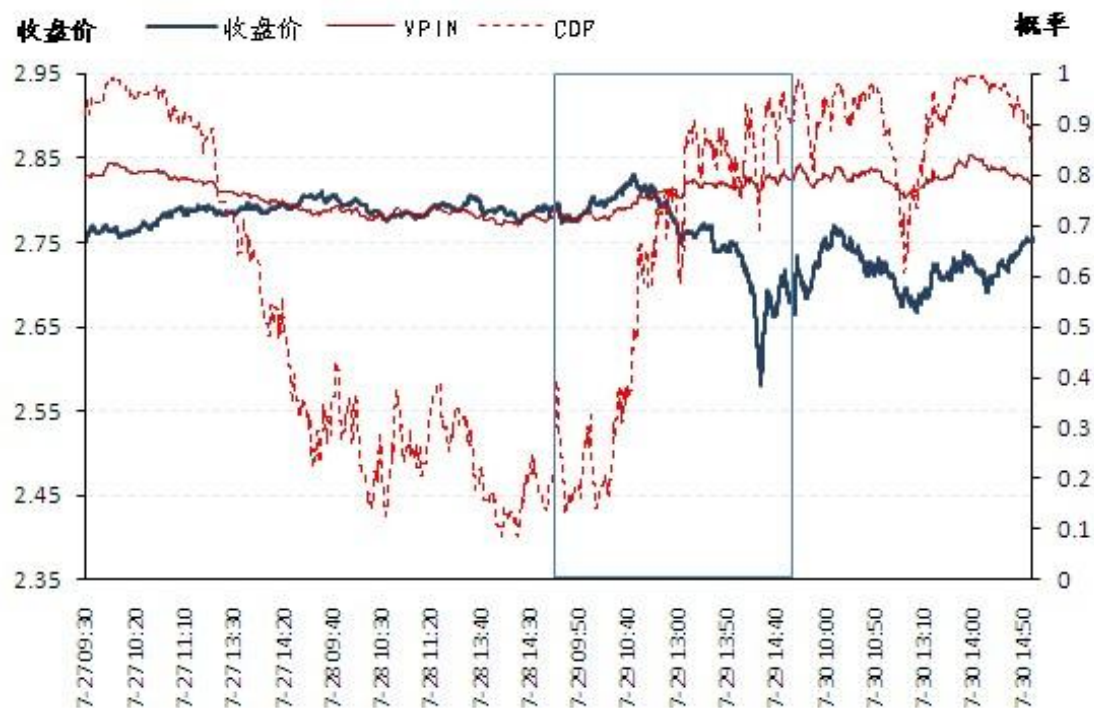


资料来源：天软资讯，国信证券经济研究所整理

VPIN实证

10秒频率取样

50ETF价格 Vs VPIN估计值



资料来源：天软资讯，国信证券经济研究所整理

VPIN实证

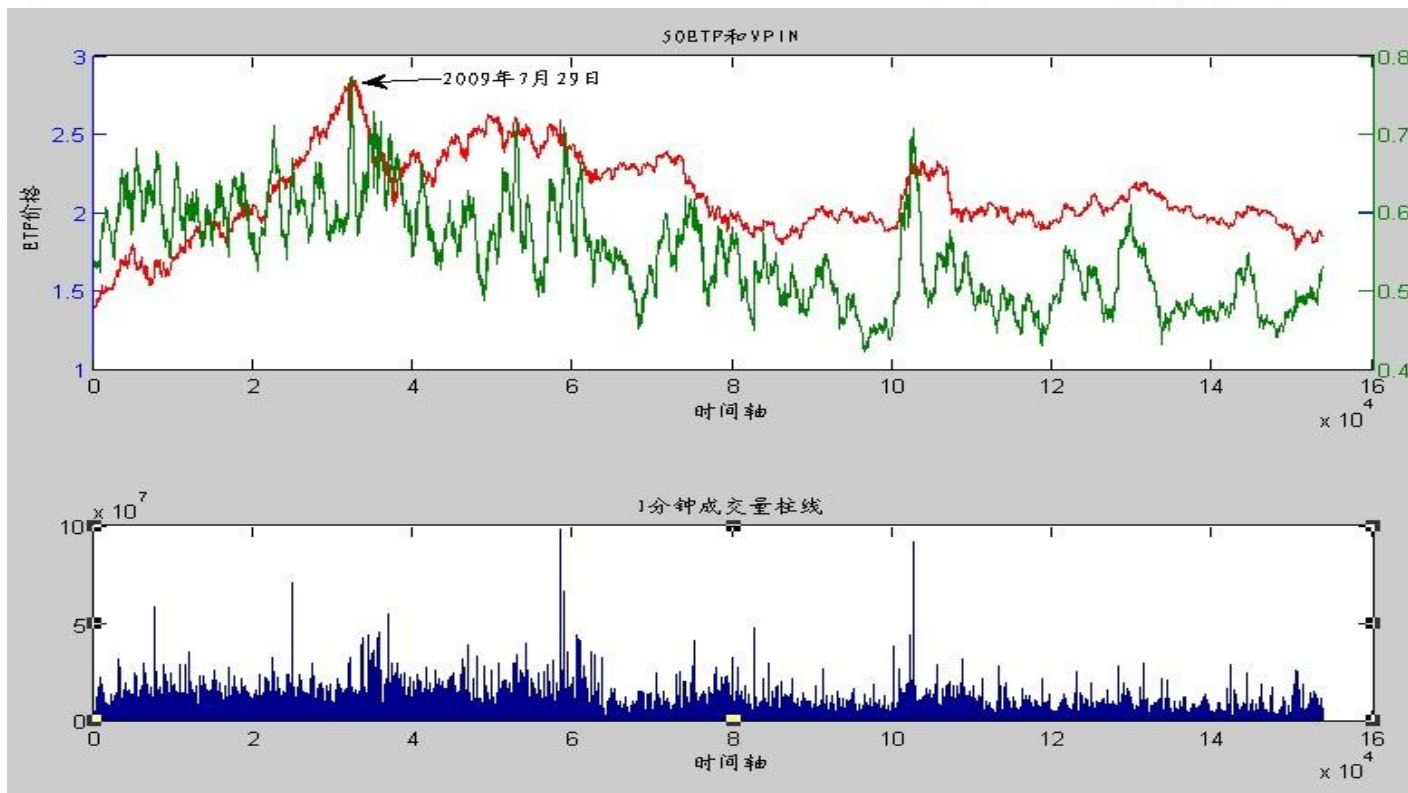


交易的买卖分类错误，
导致计算结果存在误差。

资料来源：天软资讯，国信证券经济研究所整理

ETF使用VPIN的历史回测

50ETF价格 Vs VPIN估计值



资料来源：天软资讯，国信证券经济研究所整理



VPIN与绝对收益的相关系数曲面

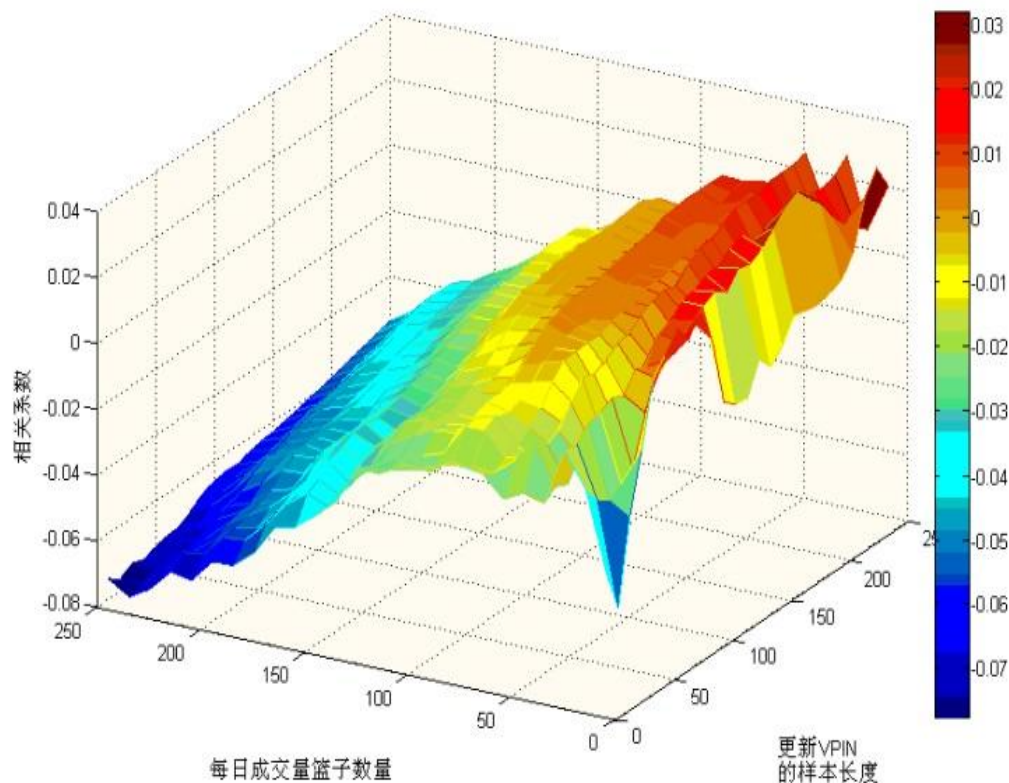
我们开始研究每日成交量篮子数量以及用于构建VPIN的篮子数量的不同的组合对VPIN及波动性预测能力的影响。

我们先用VPIN对数和50ETF的绝对收益之间简单的相关系数:

$$\rho\left(\ln(VPIN_{\tau-1}), \left|\frac{S_{\tau}}{S_{\tau}-1}-1\right|\right)$$

我们使用绝对收益 $\left|\frac{S_{\tau}}{S_{\tau}-1}-1\right|$ 作为价格波动的代理变量。

VPIN与绝对收益的相关系数曲面



对于每天50个成交量篮子，250个样本长度，我们可以从超过80000个观测数据中得到相关系数：

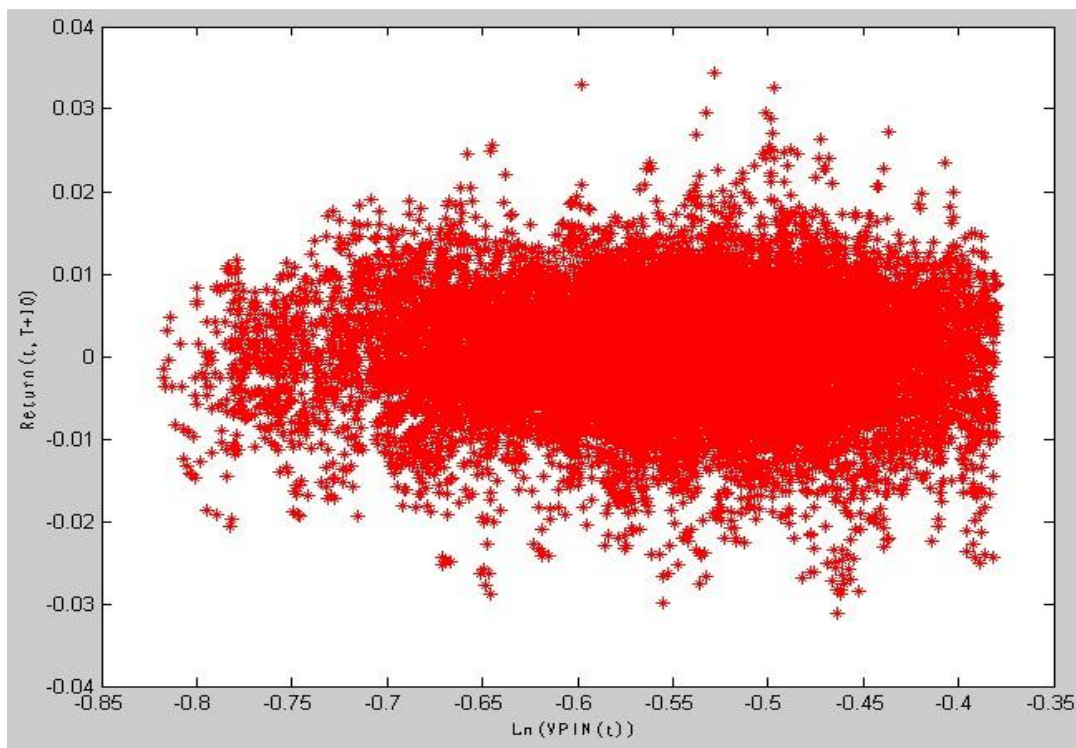
$$\rho\left(\ln(VPIN_{\tau-1}), \left|\frac{P_{\tau}}{P_{\tau-1}} - 1\right|\right) = 0.0248$$

正相关性可以看做VPIN的值越高，标的资产绝对回报越高的表现。

资料来源：天软资讯，国信证券经济研究所整理

VPIN与相对应的50ETF绝对收益

50 ETF收益 $\text{Return}(t, t+10)$ 对应更新10个成交量篮子（每日成交量的1/5）VPIN的关系

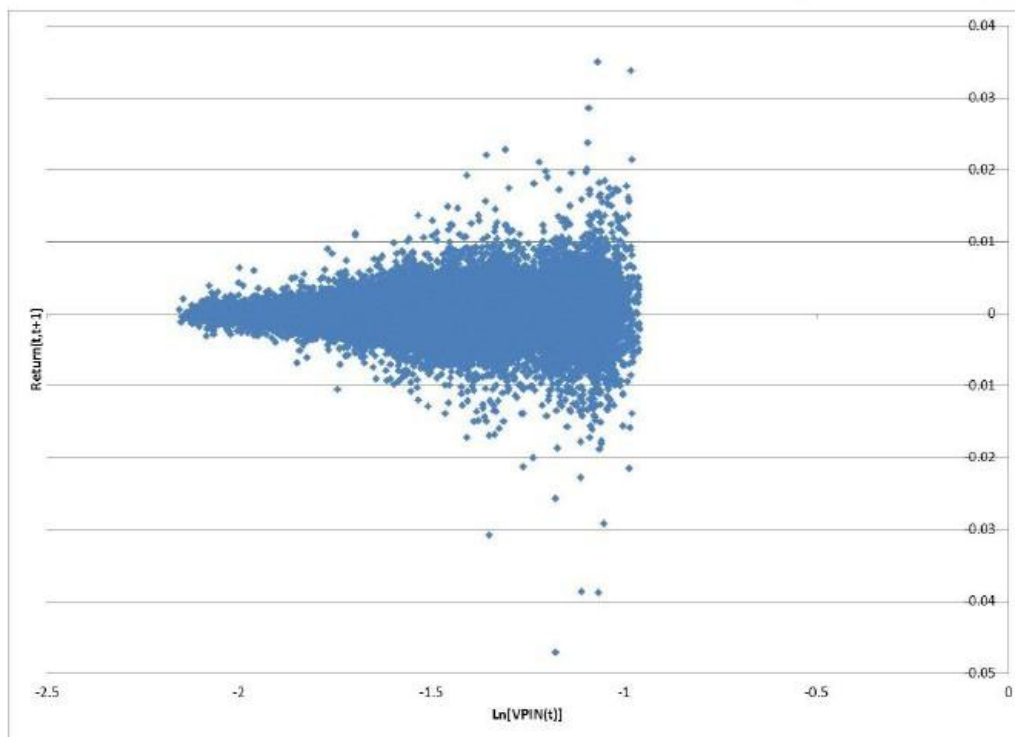


毒性越高所对应的收益绝对值也越高（靠右边的点），相反毒性较低时相应的绝对收益较低（靠左边的点）

资料来源：天软资讯，国信证券经济研究所整理

VPIN与绝对收益 (E-mini S&P500期指)

数据量大，效应明显。毒性越高所对应的收益绝对值也越高。



资料来源：天软资讯，国信证券经济研究所整理

小结

- 首先介绍交易指令的逆向选择、交易流毒性和PIN之间的联系。
- 提出一个新的VPIN算法。
- VPIN相对于之前的估计方法的一个重要的优势就是它的更新频率与成交量契合，这样更新频率可以与信息到达时间同步。
- VPIN是一个不依赖于不可观测参数的交易毒性估计模型。VPIN估计值越大，说明市场毒性越高，知情交易者占的比例越高。另外，VPIN更新的速度越快，说明市场上的知情信息越多。
- 在实证检验中，通过估计50ETF的VPIN，展示了其对应的CDF函数在大盘崩盘前提前预警提示大盘毒性极高，告诉投资者市场面临大幅波动（在高位一般指大幅调整）的概率迅速增大。我们还展示了VPIN对于波动性有较为明显的预测能力，并且它对未来的高绝对收益提供了指示作用。



VPIN应用

- 因为上述的性质，我们认为VPIN可以作为一种高频交易领域的风险管理工具。
- 通过交易指令估计交易毒性，并可通过高的VPIN值提示后续价格将有较高的波动，做市商可以通过VPIN调整做市价差。流动性风险不仅对做市商而言至关重要，对交易员来说，流动性可能会导致价格潜在的大幅波动，适时通过VPIN控制交易执行风险也会是一个不错的选择。
- 后续我们将对不同的交易品种进行VPIN估计的系统性测算，同时进一步挖掘VPIN与标的资产波动性的潜在关系。



期待您的关注和意见！ 谢谢！

风险提示

本报告信息均来源于公开资料，我公司对这些信息的准确性和完整性不作任何保证。报告中的内容和意见仅供参考，并不构成对所述证券买卖的出价或询价。我公司及其雇员对使用本报告及其内容所引发的任何直接或间接损失概不负责。我公司或关联机构可能会持有报告中所提到的公司所发行的证券头寸并进行交易，还可能为这些公司提供或争取提供投资银行业务服务。本报告版权归国信证券所有，未经书面许可任何机构和个人不得以任何形式翻版、复制、刊登。

证券投资咨询业务的说明

证券投资咨询业务是指取得监管部门颁发的相关资格的机构及其咨询人员为证券投资者或客户提供证券投资的相关信息、分析、预测或建议，并直接或间接收取服务费用的活动。

证券研究报告是证券投资咨询业务的一种基本形式，指证券公司、证券投资咨询机构对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向客户发布的行为。



国信证券经济研究所

Guosen Securities Economic Research Institute