

情報の非対称性のリアルタイム計測としての VPIN を巡って

大阪大学 大学院経済学研究科
金融・保険教育研究センター 大屋幸輔

1 はじめに

市場の情報非対称性の計測に関しては、これまで数多くの研究が行われてきたが、その一つに Easley, Hvidkjaer and O'Hara (2002) で提案された PIN (Probability of informed trading) がある。この指標は買い手主導による取引と売り手主導による取引の件数から情報の非対称性を計測するもので、直感的にはオーダー・インバランスの程度を表す指標と考えることもできる。すべての投資家が情報を共有し、有する情報に差がない状況とは異なって、市場に私的情報を有する情報投資家と公的情報のみにもとづいて行動を行う投資家がいる場合、市場には情報の非対称性があるという。情報投資家は彼らが有する情報が、価格が上昇することに関するものなら買い注文、下落することに関するものなら売り注文を行い、いずれでもなければ発注しない。一方、公的情報のみにもとづく投資家は、情報投資家の有する情報とは関係なく、平均的に一定数の売買注文を行い、それらが約定している状況を考える。このとき結果として、期待される総約定件数のうち、情報投資家の注文による約定の期待件数の割合を推定したものが PIN とよばれ、市場の情報の非対称性の程度を測る尺度と考えられている。例えば、私的情報にもとづく取引の約定件数が総約定件数に比べて相対的に大きい場合、PIN は (1 に近い) 大きな値をとる。しかしながら PIN の推計方法が前提としている状況は、短い時間に非常に多くの取引が成立する状況とは大きく異なっており、現在の市場をより適切に反映する推計方法が望まれてきた。その中で Easley, Lopez de Prado and O'Hara (2011a) *1, ELO (2011b), ELO (2012a) で提唱されているのが VPIN (Volume-synchronized PIN) である*2。この VPIN が注目を集める一つの理由として、VPIN の挙動を観測することで 2010 年 5 月 6 日に米国市場で発生したフラッシュ・クラッシュのような出来事を事前に予想できると VPIN の提唱者たちが主張していることがあげられる。確かに彼らはその主張の根拠を論文で示しているが、一方でその主張に対し Andersen and Bondarenko (2014a) のように異議を唱える研究にも注目する必要がある*3。

以下、本稿では VPIN の具体例、推計方法を示しながら、VPIN と関連する問題について概説する。

2 長期国債先物市場の例

はじめに例示のため、日本国債先物取引に関連する指標をみていく。図 1 は 2013 年 1 月 4 日から 10 月 31 日の長期国債先物の終値、一日の売買高、そのオプションから計算されたインプライド・ボラティリティ・インデックス*4、そして VPIN の推移を示している*5。

図中の縦破線は先物価格の急落によりサーキット・ブレーカーが発動された 2013 年 4 月 5 日と 5 月 13 日

*1 以降、Easley, Lopez de Prado and O'Hara を ELO と略記する。

*2 Google patents で VPIN という単語を検索すると、関連する特許が米国において出願されている情報を調べることができる。

*3 この論争に関しては *Journal of Financial Markets*, 17, 2104 に ELO (2014), Andersen and Bondarenko (2014b) が掲載されており、合わせて参照されたい。

*4 大阪大学金融・保険教育研究センターにて試算されている CSFI-VXJGB

*5 ボラティリティ・インデックスのリアルタイム版は算出されておらず、日次であるため、他の指標は日次のものをもちいている。VPIN に関しては各日における最大値。

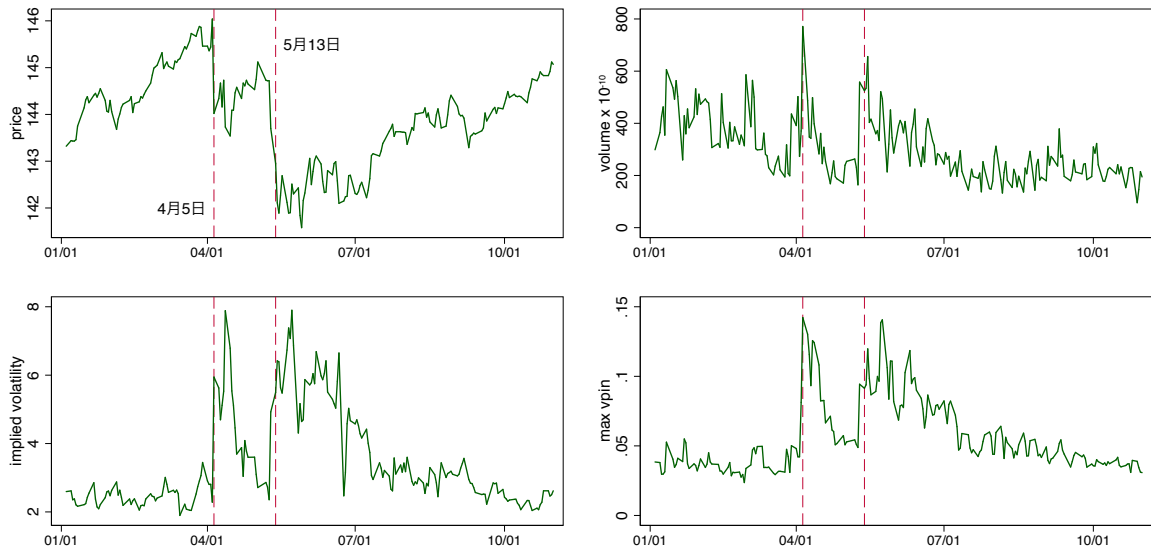


図1:長期国債先物：価格（上段左）、売買高（上段右）、ボラティリティ（下段左）、VPIN（下段右）

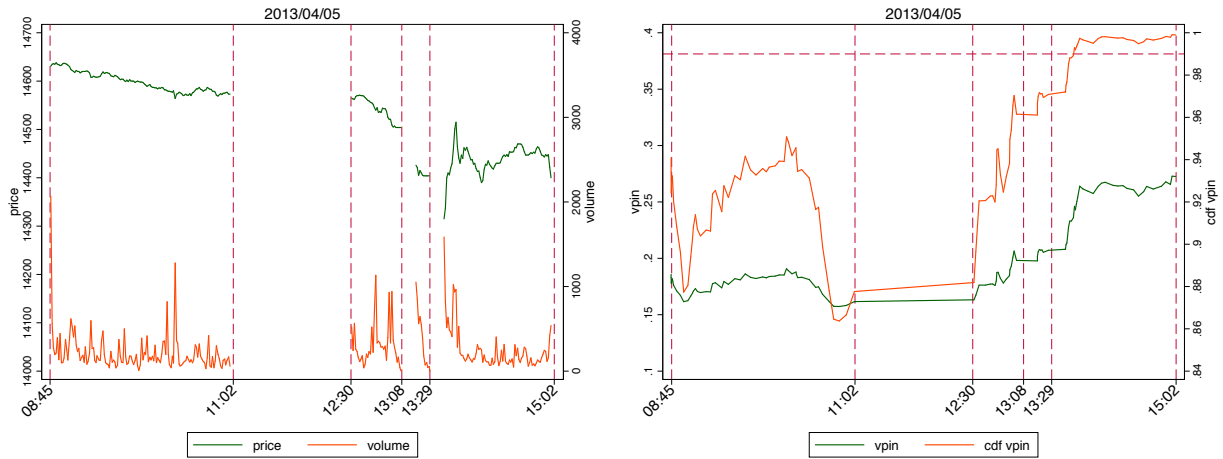


図2: 価格・売買高（左）、VPIN・CDF VPIN（右）

を示している。各指標が日次であるため、状況の変化のスピードが速い市場にあっては、各指標間の因果関係や先行・遅行関係をこの図から明確にとらえることはできないが、売買高、ボラティリティ、VPINともに、長期国債先物価格の急落に関連する動きを示していることは見てとれる*6。本来、VPINは時々刻々と変化する市況をとらえる指標であるため、日次データを示した図1からはその特徴を十分にとらえることはできない。そこで2013年4月5日午後1時8分頃と午後1時29分頃に2度のサーキット・ブレーカーが発動された長期国債先物市場の様相を示したものが次の図2である。左は先物価格（price）の売買高（volume）のグ

*6 ボラティリティについては、他の指標より遅れて反応しているようにも見えるが、他の指標と同じタイミングで変化を示した後に、その後の市場の予想を反映して大きくジャンプしているというのが実情であろう。

ラフである。昼に取引が行われていない時間帯以外に、サーキット・ブレーカー発動時の様相が示されている。これに対して右の図は VPIN と VPIN の累積分布関数 (CDF: cumulative distribution function) が示してある。ELO (2012a) など提唱者たちの一連の研究によれば、VPIN の値自体よりも、VPIN の累積分布関数が 0.99 を越えた時点に注目するといった方法が提案されている。図からは、売買高の大きな変化は価格変動に関連しているが、VPIN やその CDF の方がより明瞭になっていることがわかる。ただしそのような指標の変化が、市場の急変に呼応した結果なのか、急変自体を事前に伝えているものなのかの判断は、この図からは難しい。

実際、VPIN の算出にあたっては、事前に決めておかなければならない値があり、その設定値の違いで VPIN の挙動は異なってくる。VPIN の提唱者たちは、設定値の影響は大きくないと主張しているが、Andersen and Bondarenko (2014a) での論点は、事前に設定しなければならぬ値によって VPIN の形状は異なること、また VPIN やその CDF が必ずしも市場急変の予兆とはなっていないと主張している点にある。

3 VPIN の構成と関連する議論

既に述べたように VPIN はマイクロストラクチャーの理論を背景に、情報投資家による取引高の割合の程度をリアルタイムに測ることを目的としている。以下、極めて単純化した形で説明を行う。

- 売買高 V に関して、買手主導のものを V^B 、売手主導のものを V^S 、すなわち $V = V^B + V^S$ とする。
- 売買高 V のうち、公的情報のみにもとづく投資家による平均的な買いと売り数量はそれぞれ同数の ε とする。
- 情報投資家が情報を得る確率を α 、そして情報投資家はその情報により売り注文か、買い注文かを選択し発注を行う。その平均的な取引数量を μ とする。
- このとき $E[V] = E[V^B + V^S] = \alpha\mu + 2\varepsilon$ 、 $E[|V^B - V^S|] \approx \alpha\mu$
- 観測されたデータから $E[|V^B - V^S|]$ を推定するために、ティックデータを売買高が V になる毎にグループとしてまとめ、第 τ 番目のグループにおける V^B と V^S をそれぞれ V_τ^B 、 V_τ^S 、($\tau = 1, \dots, n$) とする。
- $|V^B - V^S|$ の期待値を対応する標本 $|V_\tau^B - V_\tau^S|$ の算術平均で推定する。

このとき VPIN は次のように定義できる*7。

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

この VPIN の特徴は、観測される市場データの集約方法にある。通常、統計的な分析を行う場合、特定の同一時間間隔、例えば 1 分間や 5 分間といった時間間隔で観測時点を区切る clock time による方法が利用されることが多い。しかし VPIN では高頻度に観測されるデータを一定の売買高毎に区切る方法、いわゆる volume time を採用している。この volume time によって区切られたグループは volume bucket と呼ばれている。

VPIN の構成法の基本的な考え方は上述の通りであるが、ELO (2012a) によれば、実際は下記の 2 段階の手続きによって VPIN はもとめられている。

*7 ELO (2012a) 参照

第一段階：volume bar

大きさ ν の volume time で、観測開始時点からのティックデータを区切る*8。最初のグループの区切りの時点を T_1^V 、次を T_2^V としていくと、volume time により区切られた時点 $T_1^V < \dots < T_i^V < \dots$ が得られる。これらの間隔は一定ではなく、取引が非常に多い時間帯においてはその間隔は狭く、取引量が少ない時間帯では広くなる。さらに、第 i 番目の bar である区間 $[T_{i-1}^V, T_i^V)$ における最後の取引価格を P_i^V とし、そうして得られた価格系列 P_1^V, P_2^V, \dots から価格変化の列 $\Delta P_1^V, \Delta P_2^V, \dots$ を得る。この価格変化の標準偏差を $\sigma_{\Delta P}$ と定義する。

ELO(2011a), ELO (2011b), ELO (2012a) などに掲載されている多くの事例では、volume bar ではなく、clock time による time bar が利用されている。しかし、短い時間間隔の time bar によって区切られた価格による収益率の分布は正規分布や類似する分布とは大きく異なり、ゼロに極端に集中したものとなっていることが知られている。VPIN では、後述するように取引分類（買手主導か売手主導かの分類）において正規分布や類似する分布を利用しており、第一段階目においても、time bar よりも volume bar の利用が望ましい。

第二段階：volume bucket と order flow 分類

高頻度に観測されているティックデータは volume bar としてまとめられるが、さらにそれらを大きさ V の volume time でまとめたものを volume bucket と呼んでいる*9。

第 τ 番目の volume bucket に関して、買手主導と売手主導の取引量 (V_τ^B, V_τ^S) は以下の分類法によって推測される。この方法では、volume bucket に含まれる各 bar での売買高分類を価格変化の大きさに応じて行っている。

Order flow 分類 (bulk classification)

大きさ V の volume bucket の中に、大きさ ν の volume bar が Q 個含まれているものとする。このとき第 τ 番目の volume bucket に関して以下の分類法を bulk classification 法による分類と呼ぶ。

$$V_\tau^B = \sum_{i=Q \times (\tau-1) + 1}^{Q \times \tau} \nu \times \Phi\left(\frac{P_i - P_{i-1}}{\sigma_{\Delta P}}\right), \quad V_\tau^S = V - V_\tau^B \quad (2)$$

また $\Phi(\cdot)$ は標準正規分布の累積分布関数である。

約定した取引が買手主導であるのか、売手主導であるかの判断については配信されるデータに売買フラグなどの情報が付加されていれば、それ自体は困難ではない。しかしその情報から、どの取引が情報投資家からのものなのかを見分けることは容易ではない*10。

VPIN

各 volume bucket での (V_τ^B, V_τ^S) が得られれば、その系列の移動平均を計算することによって、即時的な VPIN の系列をもとめることができる。

$$VPIN_{n+k-1} = \frac{1}{n} \sum_{\tau=n \times (k-1) + 1}^{n \times k} \frac{|V_\tau^B - V_\tau^S|}{V}, \quad k = 1, \dots \quad (3)$$

*8 Andersen and Bondarenko (2014a) ではこの volume bar は volume bin と呼ばれている。他方、clock time で区切られたものは time bar と呼ばれる。

*9 例えば、ELO では標本期間に含まれる各取引日の取引量から1日の平均取引量をもとめ、その50分の1を V と定めている。

*10 その識別が容易でない例については ELO (2014) 48 ページ脚注 1 を参照。

ここで一つの VPIN を計算する際に n 組の (V_{τ}^B, V_{τ}^S) を利用しているが、この n の値は VPIN の系列を特徴付けるものとして重要である。その値が小さいと算出された VPIN の系列は、変化の激しいものとなる一方、値が大きければ滑らかな推移を示す系列となるが、その場合、市場急変に対するシグナルとしてはその鋭敏さを失う可能性がある。

CDF VPIN

市場の情報の非対称性が極度に高まっているとして、アラートを出す目的でこの VPIN を利用する場合、VPIN の値そのものがある特定の値（閾値）を越えた時点に注目する方法が考えられるが、その閾値を客観的に定めることは難しい。そこで ELO (2012a) では、計測された VPIN の経験累積分布関数をもとめ、累積確率が 0.99 となる VPIN の値を閾値として用いることを提案している^{*11}。図 2 の右の上部に CDF VPIN の閾値が横破線として示してあるが、CDF VPIN がその閾値を越えるのは二度目のサーキット・ブレーカーが発動された後であり、市場の急変に対する予兆としては機能していない例となっている。

一般論として、CDF VPIN が累積確率 0.99 を越える回数は、CDF VPIN の算出に利用した VPIN 全体の数の 1% に相当する。市場が平穏な状況を示している期間の VPIN だけで CDF VPIN を算出した場合より、市場が急変している時期の VPIN をその算出に含む方が CDF VPIN が示す閾値は高くなる。VPIN の水準が全般的にトレンドをもっていたり、その水準が、複数の期間で明らかに異なっている状況では、どの期間を CDF VPIN の算出にもちいるか注意が必要である。

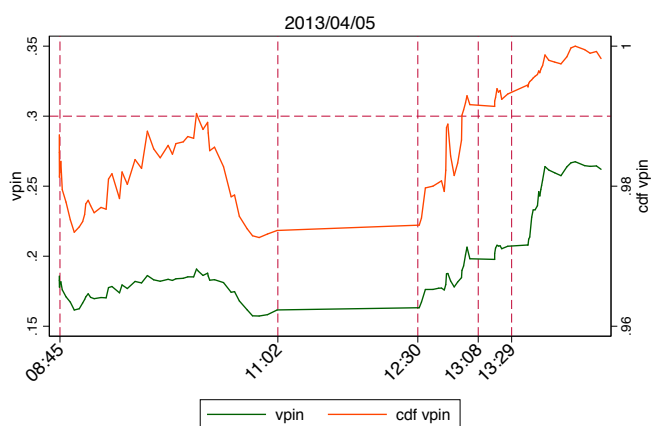


図 3: VPIN と CDF VPIN

図 2 での CDF VPIN は 2013 年 1 月 4 日から 2014 年 2 月 28 日までの VPIN を使って算出されていたのに対して、図 3 では、2013 年 4 月 5 日に焦点をあたるため、2013 年 1 月 4 日から 4 月 5 日午後 2 時 30 分までの VPIN を使っている^{*12}。この場合は、1 度目のサーキット・ブレーカーの前に既に CDF VPIN はその閾値を超えており、市場の急変をあらかじめ知らせるものとなっている。図 1 の VPIN の各日の最大値（下段右）のグラフからもわかるように、2013 年 4 月 5 日の市場急変後しばらくの間、VPIN は高い水準のまま継続している。図 2 の CDF VPIN は、注目している 4 月 5 日以降の高水準の VPIN を、将来時点の

ものであるにもかかわらず利用しており、その高水準の VPIN をもちいていない図 3 での閾値よりも高い閾値になっている。ELO (2012a) では、VPIN を算出している期間全体、フラッシュ・クラッシュ前後で分割した期間と、それぞれ異なる期間で CDF VPIN を算出し、これらに大きな差がないことを確認し、CDF VPIN の算出に対して標本期間の選択は大きな影響を与えないと結論づけている。異なる期間での累積経験分布関数が同じ形状になっているという事実は、彼らが観測対象とした市場において、観測期間のすべてにわたって

^{*11} 例えば、100 個の VPIN があつた場合、それらを大きき順に並べると、小さい方から 99 番目の VPIN の値が累積確率 0.99 を与える VPIN の値となる。

^{*12} 2013 年 4 月 5 日の 2 度目のサーキット・ブレーカーが発動された約 1 時間後までの VPIN を利用している。

VPIN はある程度一定の水準で推移しており、市場が急変する局面で、その水準から乖離し大きな値をつけていることを示している。一方、ここでとりあげている長期国債先物市場の例では、観測期間をとおして、その水準は一定にあるとは言いがたく、先行研究で示されている VPIN, CDF VPIN の算出方法の機械的な適用には注意が必要となる証左となっている。

4 おわりに

市場の情報非対称性をリアルタイムで計測する方法としての VPIN と、それから派生して算出される CDF VPIN は、提唱者らによれば、フラッシュ・クラッシュに対する予兆として機能していることが報告されているが、一方で Andersen and Bondarenko (2014a) で述べられている批判があることも事実である。批判の多くは、time bar を利用して算出された VPIN に対するもの、本稿で説明しなかった Tick rule によるオーダーフローの分類法などに対するものであり、volume bar, bulk classification を利用した VPIN や CDF VPIN に関して、今後も議論・検証の余地が残されている。

現在の市場をとりまく環境において、リアルタイムに市況の急変を事前にとらえる指標の開発が望まれていることは言うまでもない。そのような指標の開発には、科学的・客観的事実の積み重ねが必要であり、今後、市場関係者や研究者の密接な連携によりその目的が達成されることを期待し、本稿の結びにかえる。

参考文献

- Andersen, T. G. and Bondarenko, O. (2014) “VPIN and the Flash Crash,” *Journal of Financial Markets*, **17**, 1-46.
- Andersen, T. G. and Bondarenko, O. (2014) “Reflecting on the VPIN Dispute,” *Journal of Financial Markets*, **17**, 53-64.
- Easley, D., Hvidkjaer, S. and O’Hara, M. (2002) “Is information risk a determinant of asset returns ?,” *Journal of Finance*, **57**, 5, 2185-2221.
- Easley, D., Lopez de Prado, M. and O’Hara, M. (2011a) “The microstructure of the ‘Flash Crash’: Flow toxicity, liquidity crashes and the probability of informed trading,” *Journal of Portfolio Management*, **37**, 2, 118-128.
- Easley, D., Lopez de Prado, M. and O’Hara, M. (2011b) “The Exchange of flow toxicity,” *Journal of Trading*, **6**, 2, 8-13.
- Easley, D., Lopez de Prado, M. and O’Hara, M. (2012a) “Flow Toxicity and Liquidity in a High Frequency World,” *Review of Financial Studies*, **25**, 5, 1457-1493.
- Easley, D., Lopez de Prado, M. and O’Hara, M. (2014) “VPIN and the Flash Crash: A rejoinder,” *Journal of Financial Markets*, **17**, 47-52.

本資料に関する著作権は、株式会社大阪取引所にあります。

本資料の一部又は全部を無断で転用、複製することはできません。

本資料は、デリバティブ商品の取引の勧誘を目的としたものではありません。