Johnson School Research Paper Series #10-2011

The Exchange of Flow Toxicity

David Easley—Cornell University
Marcos Mailoc Lopez de Prado—Tudor Investment Corp.; RCC at Harvard
Maureen O'Hara—Cornell University

January 2011

Electronic copy available at: https://ssrn.com/abstract=1748633

THE EXCHANGE OF FLOW TOXICITY

David Easley[1]
dae3@cornell.edu

Marcos M. López de Prado[2]
marcos.lopezdeprado@tudor.com

Maureen O'Hara[3]
mo19@cornell.edu

January 17, 2011

ABSTRACT

Flow toxicity can be measured in terms of the probability that a liquidity provider is adversely selected by informed traders. In previous papers we introduced the concept of Volume-synchronized Probability of Informed Trading (the VPIN* metric), and provided a robust estimation procedure. In this study, we discuss the asymmetric impact that an incorrect estimation of the VPIN metric has on a market maker's performance. This asymmetry may be part of the explanation for the evaporation of liquidity witnessed on May 6th 2010. To mitigate that undesirable behavior, we present the specifications of a VPIN contract, which could be used to hedge against the risk of higher than expected levels of toxicity, as well as to monitor such risk. Among other applications, it would also work as an execution benchmark, and a price discovery mechanism, since it allows for the externalization of market participants' views of future toxicity.

Keywords: Liquidity provision, flow toxicity, market microstructure, probability of informed trading, VPIN.

JEL codes: C51, C53, G10, G12, G14.

[1] Scarborough Professor and Donald C. Opatrny Chair Department of Economics, Cornell University.
[2] Head of High Frequency Futures, Tudor Investment Corp., and a Postdoctoral Fellow of RCC at Harvard University.
[3] Robert W. Purcell Professor of Finance, Cornell University.

Electronic copy available at: https://ssrn.com/abstract=1748633

Order flow is deemed to be *toxic* when it adversely selects liquidity providers. Earlier papers have discussed the merits of the Probability of Informed Trading (*PIN*) as a measure of flow toxicity. In Easley, López de Prado and O'Hara (2010a), we presented a modified estimation procedure for PIN that applies to high frequency trading. This modified procedure uses trade time rather than clock time to produce a measure of flow toxicity that we called the VPIN metric. In Easley, López de Prado and O'Hara (2010b), we presented evidence that the VPIN metric anticipated the events of May 6th 2010 (also known as the '*flash crash*') with as much as a 2 hour lead. In contrast, the VIX index was a weaker forecaster of the crash, and performed erratically during the incident. This difference in performance of the two measures motivated an analysis of the properties of the VPIN metric's time series, with the conclusion that an increase in the VPIN does foreshadow an increase in volatility. In our earlier work, we also introduced the idea of a futures contract with the VPIN metric as underlying (henceforth called the *FVPIN* contract*), which could be used to monitor the probability of a crash and serve as a risk management tool.

This paper develops more fully the rationale for the FVPIN contract and describes its construction. The paper is organized as follows. We first present the range of levels at which a market maker is willing to contribute liquidity, and shows that it is a function of PIN. Because PIN is not perfectly known, estimation errors are to be expected. We then show the asymmetric impact that a wrong estimation of PIN has on a liquidity provider's performance. This underscores the need for a procedure by which the market reaches a consensus level for PIN that can be used by all market makers. The mechanism we propose takes the form of a futures contract (the FVPIN contract). We then discuss a few examples of how this contract could fill an important role for market makers, execution brokers, risk managers and regulators. We also provide the specifics of how this contract could be implemented in practice.

THE BID-ASK SPREAD AS A FUNCTION OF PIN

As shown by Easley, Kiefer, O'Hara and Paperman (1996) and Easley, Kiefer and O'Hara (1997a, 1997b), liquidity providers who are subject to adverse selection set trading bands with a positive bid-ask spread to protect against the risk of trading with better informed traders. In a simple version of that analysis, in which good and bad events are assumed to be equally likely, this spread is

$$A - B = \left(\overline{S} - \underline{S}\right)\frac{\alpha\mu}{\alpha\mu + 2\varepsilon} = \left(\overline{S} - \underline{S}\right)PIN \tag{1}$$

where:

- α is the probability associated with an information event.
- the information event results in either a low outcome $\underline{S}$ or a high outcome $\overline{S}$.
- μ is the rate of arrival of informed trades.
- ε is the rate of arrival of uninformed trades.
- A is the ask level offered by a market maker.
- B is the bid level offer by a market maker.
- PIN is the fraction of expected order flow that arises from informed trade.

Electronic copy available at: https://ssrn.com/abstract=1748633

These bounds reflect a balance between the gains the market maker expects to make trading with uninformed traders (the benign order flow) and the losses she expects from trading with informed traders (the toxic flow). The more toxic a flow is, measured in terms of the *Probability of Informed Trading* (PIN), the wider the optimal bid-ask spread.

THE MARKET MAKER'S 'ASYMMETRIC PAYOFF DILEMMA'

The analysis above assumes that liquidity providers know the toxicity of the order flow. It's more realistic to assume that their beliefs about flow toxicity (the PIN) may not be correct. In this section, we assume that liquidity providers have common, but possibly incorrect beliefs about PIN, and we let PIN* be their estimate of PIN. Because we assume that liquidity providers are identical, we analyze the market from the point of view of a representative liquidity provider.

We first consider the impact that an inaccurate estimate of PIN (our PIN^*) has on the representative liquidity provider's expected performance. Because trade with the uninformed is generally offsetting (i.e. the uninformed are equally likely to buy and sell), the market maker expects to capture a gain $(A^* - B^*)$ on average from each trade with an uninformed counterparty. Unfortunately, by the end of the day our representative market maker has also lost an average of $\frac{A^* - B^* - (\overline{S} - \underline{S})}{2}$ on each contract exchanged with informed traders (i.e., toxic flow). With the uninformed trading at a rate of 2ε and the informed at a rate $\alpha\mu$, this leaves the market maker with an expected daily profit (denoted PnL) of

$$PnL = \left(A^* - B^*\right)\varepsilon + \frac{1}{2}\alpha\mu\left[\left(A^* - B^* - \left(\overline{S} - \underline{S}\right)\right)\right] \tag{2}$$

The expected order flow is $V = \alpha\mu + 2\varepsilon$. Because $A^* - B^* = \left(\overline{S} - \underline{S}\right)PIN^*$ and $V \cdot PIN = \alpha\mu$, simple calculation shows that this expected profit is

$$PnL = V\frac{\overline{S} - \underline{S}}{2}\left(PIN^* - PIN\right) \tag{3}$$

So clearly the market maker's expected profit will depend upon the accuracy of her PIN calculation as well as on the volume the market maker transacts.

To this point in the analysis we have treated the expected total order flow as exogenous and, most importantly, independent of the trading band set by our representative liquidity provider. In reality, the total order flow will depend on the size of the spread the liquidity provider demands. For simplicity here we take a reduced form approach and assume that the expected order flow is a linear function of the bid-ask spread relative to the difference between high and low outcomes for the asset

$$V = \kappa\left(1 - \frac{A^* - B^*}{\overline{S} - \underline{S}}\right) \tag{4}$$

where κ is a constant. The intuition for this formulation is that larger spreads discourage trade, while larger potential gains from trade (the difference between high and low outcomes) induces trade.[4]

Our assumption about the responsiveness of total order flow to spread implies that the representative market maker's expected daily profit will be

$$PnL = \frac{(\overline{S} - \underline{S})\kappa}{2}(1 - PIN^*)(PIN^* - PIN) \tag{5}$$

This expression reveals the dilemma faced by the representative market maker. If the market maker underestimates the order toxicity she is facing ($PIN^* < PIN$), then she will experience expected losses on every trade, and she will have greater volume due to her setting a more aggressive spread. Conversely if she overestimates the toxicity ($PIN^* > PIN$), every trade yields an expected profit but the volume of trade is reduced.[5]

The analysis above assumes a particular functional form for the response of volume to PIN^* but a qualitatively similar expression would result from any reasonable relation between volume and the bid-ask spread.[6] The most important conclusion from this analysis is that the market maker's expected payoff is asymmetric in her estimate of PIN^*. If she underestimates PIN she will have a large order flow but an expected loss on every trade. Alternatively, if she overestimates PIN she has an expected profit on each trade, but a reduced order flow.

The following numerical example illustrates these relationships.

DAILY PnL AS A RESULT OF (PIN*, PIN)

(PIN*,PIN)	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	0.00	-50.00	-100.00	-150.00	-200.00	-250.00	-300.00	-350.00	-400.00	-450.00	-500.00
0.1	45.00	0.00	-45.00	-90.00	-135.00	-180.00	-225.00	-270.00	-315.00	-360.00	-405.00
0.2	80.00	40.00	0.00	-40.00	-80.00	-120.00	-160.00	-200.00	-240.00	-280.00	-320.00
0.3	105.00	70.00	35.00	0.00	-35.00	-70.00	-105.00	-140.00	-175.00	-210.00	-245.00
0.4	120.00	90.00	60.00	30.00	0.00	-30.00	-60.00	-90.00	-120.00	-150.00	-180.00
0.5	125.00	100.00	75.00	50.00	25.00	0.00	-25.00	-50.00	-75.00	-100.00	-125.00
0.6	120.00	100.00	80.00	60.00	40.00	20.00	0.00	-20.00	-40.00	-60.00	-80.00
0.7	105.00	90.00	75.00	60.00	45.00	30.00	15.00	0.00	-15.00	-30.00	-45.00
0.8	80.00	70.00	60.00	50.00	40.00	30.00	20.00	10.00	0.00	-10.00	-20.00
0.9	45.00	40.00	35.00	30.00	25.00	20.00	15.00	10.00	5.00	0.00	-5.00
1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Table 1 – Daily PnL as a result of various estimation errors of flow toxicity

[4] The ratio of the bid-ask spread to the difference between high and low outcomes for the asset is by the pricing equation (1) equal to the estimate of PIN, PIN^*.

[5] This discussion refers to the total profit or loss of market makers, here captured by our representative market maker. Any individual market maker is one of many and so her own trading bands have little or no effect on the total profit or loss of market makers. For this reason our representative market maker does not take into account the dependence of total profit or loss on the market's trading bands.

[6] A more important assumption is that we have not taken into account the possibility that market makers may have differing, incorrect beliefs about PIN. In this case, the pricing game that market makers play would be potentially quite complex. One issue that arises is what a market maker with the best quote can infer about her own beliefs from the fact that she has posted the best quote. If market makers simply have differing beliefs this may tell her nothing other than the fact that others disagree. Alternatively, if their differing beliefs are generated from some common structure then she should take the "winners curse" of having the most optimistic beliefs into account in setting her trading bands.

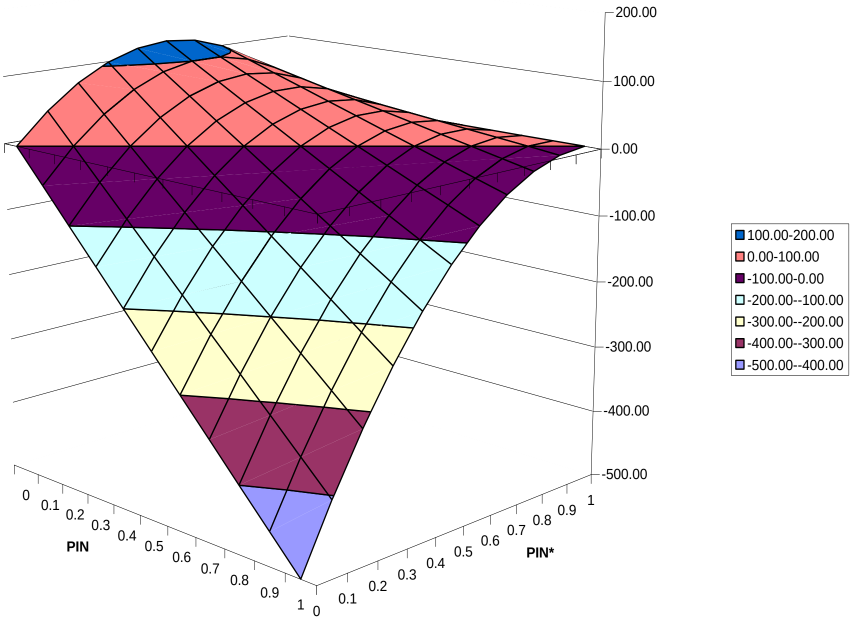

Figure 1 – Market Maker's daily PnL for $\overline{S} - \underline{S} = 1$ *and* $\kappa = 1{,}000$

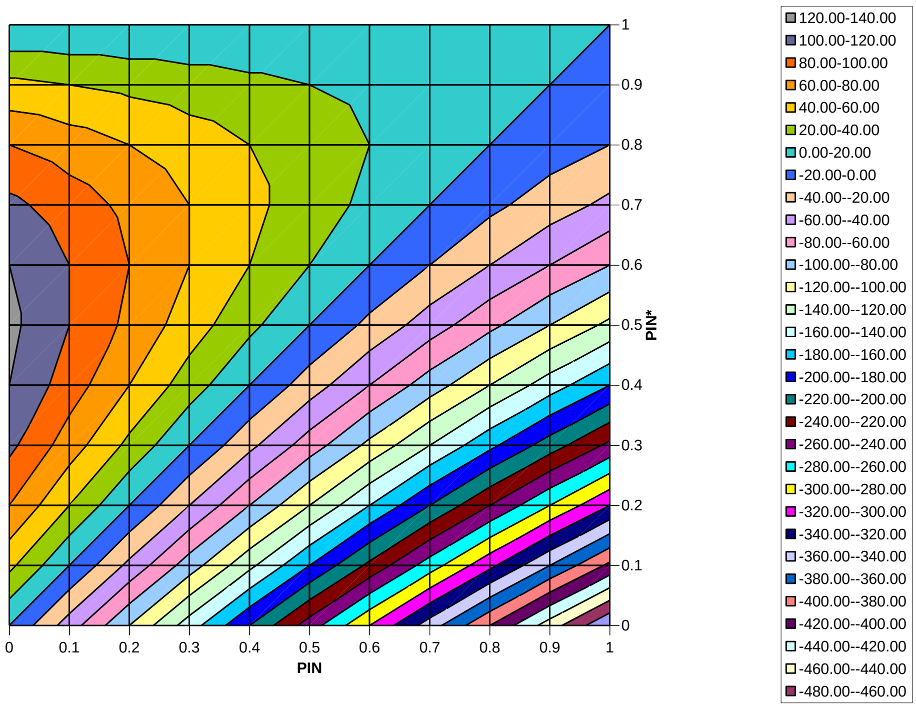

Figure 2 – Market Maker's daily PnL for $\overline{S} - \underline{S} = 1$ *and* $\kappa = 1{,}000$ *(heat map)*

An intuitive way of understanding this asymmetric payoff is to think of market makers as sellers of options to be adversely selected (i.e. contingent recipients of toxic flow) at a premium of $\frac{A^* - B^*}{2}$.[7] If market makers underestimate PIN the premium they charge for the option they sell is too small; in this case they lose (in expectation)

[7] This notion of the spread as an option premium was first suggested by Copeland and Galia (1983) in their analysis of the effects of information on the bid-ask spread.

on each trade. Those who underestimate the true value of *PIN* will experience mounting losses and may be forced to withdraw from the market. This theoretical behaviour finds a practical example in the evaporation of liquidity witnessed on May 6th 2010.[8]

The representative market maker's profit is a jointly concave function of the true value PIN and the estimate of PIN, PIN*. This concavity arises because mis-estimation negatively affects the market maker, and it sets the stage for why the market maker would prefer to insure against this estimation risk. To see why this is so, suppose that over time the true value of PIN is stochastic and the estimate PIN* is also stochastic. Further, suppose that when the true toxicity is PIN, the estimate of PIN is $PIN^{*}=PIN+\omega$, where ω is random and uncorrelated with PIN. Denote the mean of PIN by E[PIN], the mean of the error in the estimate of PIN by E[ω], and the variance the error term by VAR[ω]>0. Over time, as PIN varies and the estimate of PIN also varies, the representative market maker's average profit will be the expectation of profit

$$E[PnL] = E\left[\frac{(\overline{S}-\underline{S})\kappa}{2}(1-PIN^{*})(PIN^{*}-PIN)\right] = \frac{(\overline{S}-\underline{S})\kappa}{2}\left[E[\omega](1-E[PIN])-VAR(\omega)-E[\omega]^{2}\right] \tag{6}$$

Note that if the mean of the error in the estimate of PIN, E[ω], is 0 then expected profit is negative. This arises because the profit function is concave and it is zero when the error is zero.

Negative long run average profit is not a sustainable situation and, therefore, we would expect that over time the representative market maker would have an overestimate of PIN. That is, the mean of the distribution of the estimate PIN* must be greater than the mean value of PIN in order for the market maker's expected profit to be non-negative. So the market will act as if the average value of toxicity is greater than it actually is, which in turn means that the market will provide less liquidity than if the estimation risk could be avoided. Consequently, if there were a contract that allowed liquidity providers to insure against variation in PIN, the market would act according to the actual PIN rather than the overestimate of PIN induced by the concavity of the profit function.

[8] Cf. Easley, López de Prado and O'Hara (2010b).

THE FVPIN CONTRACT

In Easley, López de Prado and O'Hara (2010a, 2010b) we provided an improved estimate of PIN which we called VPIN. This improvement estimates PIN in trade time rather than in clock time, but otherwise it has the same interpretation as PIN. In this section we use VPIN in place of PIN to describe the futures contract we suggest. Let's first define $fVPIN_t = -Ln(VPIN_t)$, which is bounded $[0,\infty)$ as $VPIN_t \in [0,1]$.

Consider a futures contract called *FVPIN*, which forecasts the value of the *fVPIN* index at expiration (*T*). The contract expires at the end of each session, when it is cash-settled. Note that a long position on this contract is a wager on $VPIN_T < VPIN_t$, i.e. it sells protection against toxicity of the order flow. A buyer of the contract at time $t \in [0,T]$ at price $FVPIN_t$ is entitled to receive from the seller a sum determined as a linear function of $fVPIN_T - FVPIN_t$ if positive; and is obliged to pay that sum to the seller if negative.

Liquidity providers need to come up with a correct *ex-ante* estimate of PIN in order to identify the optimal liquidity contribution levels. A large deviation between their *ex-ante* and the *ex-post* PINs will result in suboptimal performance from excess inventory or insufficient portfolio turnover. Let us suppose that a market maker expects to receive a level of toxicity consistent with a value $VPIN^*$ over the interval $[t,T]$. She can either

1. Provide liquidity at levels determined by $VPIN^*$.
2. *Sell the contract at* $FVPIN_t$ and provide liquidity at levels consistent with e^{-FVPIN_t}. The exact amount of *FVPIN* to be sold will depend on how sensitive her market making strategy is to changes in toxicity.
3. Start with option 1 and turn to option 2 as her inventory grows beyond tolerance levels.

POTENTIAL USES OF THE FVPIN CONTRACT

The securitization of PIN would serve multiple functions for a variety of market participants. Examples are:

1. A critical application is to provide a ***mechanism*** by which all market participants reach a market-consensus on the prevailing toxicity levels, and allow for a transfer of risks associated with it. Liquidity providers could use this contract for insurance, but other market participants may wish to speculate using the contract. For example: An investor who has completed the execution of a large order and therefore has a reasonable expectation that the market is about to become more balanced, will be interested in *buying the FVPIN contract*.
2. As we argued earlier, it provides a ***risk management tool*** for market makers. One of the advantages of hedging with the *FVPIN* contract is that it will allow market makers to continue providing liquidity, even if toxicity exceeds the levels originally expected. This could help to dampen the kind of liquidity evaporation we witnessed on May 6th 2010. Example: *"Buy VPIN hedge if inventory grows over an outright risk of $1m 1-StDev/day"*.
3. ***A regulators' tool***: The FVPIN contract can help to monitor the level of 'pain' that is being inflicted on market makers on a particular day. Since the VPIN

metric provides a signal of developing adverse liquidity conditions in the market, it may be useful to base circuit-breakers on the FVPIN contract in addition to basing them on prices. Example: *"Order a temporary market halt if the VPIN contract goes over the 90% cumulative probability threshold"*.

4. In an earlier paper, we presented evidence that the VPIN metric Granger-causes volatility, and provided an explanation for the mechanism by which greater toxicity anticipates a rise in volatility. This makes the FVPIN contract a desirable ***security for the volatility arbitrage business***. Example: *"Buy the FVPIN contract, sell VIX"*.
5. Brokers could use the VPIN metric as an ***indication for the correct timing of execution***. In the presence of toxic flow, it is easy to get passive orders filled on the wrong side, and hard to get them filled on the right side. When clients commission brokers to execute their orders on the side that is being adversely selected, they could set a limit to the level of toxicity acceptable, as registered by the VPIN metric. Example: *"Buy 1,000 e-Mini S&P 500 around the level 1120 as long as the VPIN metric stays below 37%"*.

CONTRACT SPECIFICATIONS

Below is the proposed definition for the *FVPIN* contract.

Name	Volume-based Probability of Informed Trading Index Futures (the FVPIN contract).
Underlying	Estimates of the VPIN metric index, as computed in stochastic time (matching volume arrival) by the exchange, based on the methodology presented in Easley, López de Prado and O'Hara (2010a).
Contract size	\$5,000 times the value of $-Ln(VPIN)$
Contract expirations	Daily, at end of session.
Trading hours	Same as the contract on which order flow the VPIN metric is measured (e.g., E-mini S&P 500 futures).
Dollar value per tick	0.01 points, or \$50 per contract.
Crossing	1 contract.
Settlement	The contract is cash-settled on a daily basis against the value of $-Ln(VPIN)$ that prevailed at the end of the session.

The reason for daily cash settlement is to replicate market makers' intraday seasonality in risk-taking. The timing of their portfolio's risk mimics the seasonality of intraday trading activity. As liquidity providers start and end the day with reduced inventory, it makes sense to provide a daily settlement for the risks associated with their activity.

CONCLUSIONS

In this paper we provide a detailed description and motivation for a futures contract with an underlying based on the VPIN contract. We believe that the FVPIN contract will help to reduce the bid-ask spread and introduce a price discovery mechanism that is currently missing in the market. It may also add efficiency and liquidity to the markets, by making available a risk management tool for market makers and regulators. Prior research shows that an increase in the VPIN metric foreshadows volatility, which makes this contract also useful in the volatility arbitrage business. Finally, clients could set limits to the toxicity they are willing to accept in the execution of their passive orders, thus communicating to their brokers what the VPIN metric levels they are allowed to work with.

Easley, López de Prado and O'Hara (2010b) argues that a contributing factor to the May 6^{th} 2010 '*flash crash*' was the evaporation of liquidity resulting from the substantial losses sustained by market makers in the period leading up to the crash. We believe that the VPIN contract would have made a difference that day by allowing market makers to observe and condition on the level of toxicity they were (inadvertently) experiencing, and/or allowing for a transfer of the risk associated with it.

BIBLIOGRAPHY

- Copeland, T. and D. Galai, (1983): *"Information effects on the bid-ask spread,"* Journal of Finance, 38(5).
- Easley, D., N. Kiefer, M. O'Hara and J. Paperman (1996): *"Liquidity, Information, and Infrequently Traded Stocks"*, Journal of Finance, September.
- Easley, D., N. Kiefer and M. O'Hara (1997a): *"The Information Content of the Trading Process"*, Journal of Empirical Finance, No. 4.
- Easley, D., N. Kiefer and M. O'Hara (1997b): *"One Day in the Life of a Very Common Stock"*, Review of Financial Studies, Fall.
- Easley, D., M. López de Prado and M. O'Hara (2010a): *"Measuring Flow Toxicity in a High Frequency World"*, SSRN Working paper, http://ssrn.com/abstract=1695596
- Easley, D., M. López de Prado and M. O'Hara (2010b): *"The Microstructure of the Flash Crash: Flow toxicity, liquidity crashes and the Probability of Informed Trading"*, Journal of Portfolio Management *(forthcoming, Winter 2011).*

DISCLAIMER

The views expressed in this paper are those of the authors and not necessarily reflect those of Tudor Investment Corporation. No investment decision or particular course of action is recommended by this paper.