**Anderson and Bondarenko Paper Summary:**

This paper explores various types of VPIN measures, discuss about the implementation details of each, and then evaluates the upside and downside of those measures. Author’s key findings are CAPITALIZED within the below summary.

The first VPIN (for what is VPIN, please refer to Jiongjia’s summary) measure they tested is TR-VPIN (Ticker-Rule VPIN). And TR-VPIN, like all other VPINs mentioned in the paper, uses volume bucketing (please refer to Jiongjia’s summary for this). Within each volume bucket, TR-VPIN aggregates ticks from an arbitrary time interval delta together (for example, a popular choice is set delta=60s) and classifies this aggregated “transaction” either as a buy or sell depending on the price change between the very first tick and the last tick. A positive or zero price change would result in buy (+1) classification, and a negative price change would result in a sell (-1) classification. They claim this is a poor measure as the OIs (order imbalance) used to compute TR-VPIN (TR-VPIN is essentially a moving average of OIs with a window size of L, and L is set to 50) are highly correlated with the intensity of the trading patterns, and minimally related with trade classification. (The author spends almost all of the remaining paper to support this claim)

**FINDING 1) The author mentioned the performance of TR-VPIN prior to the flash crash showed substantial jump, but was by no mean exceptional.**

Now, to show you how TR-VPIN works and why it may not work well, the author provided a hypothetical example; this is critical in understanding both the implementation and the shortcomings of TR-VPIN, so instead of summarizing the results, I want to show at least a part of the process. The example is as follows:

Assume we use a volume bucket with bucket size = V = 1000 (ticks), the following is what happened within the volume bucket:

{(9:38:00;10.01;100); (9:39:00;10.02;200); (9:40:00;10.02;200);(9:41:00;10.01;300); (9:42:00;10.01;100); (9:43:00;10.00;100)}, each triplet reads as (time of the transaction; price; volume)

Using this volume bucket, we compute OI = |100+200+200-300-100-100|/1000 = 0. Note the formula implies that each trade classification is weighted by their volume.

Alternatively you can think of the above formula as OI = |Sigma[t=1 to 6](w(t)\*b(t))|, where w(t) = volume(t)/V, and b(t) = +1 or -1.

Now assume nothing changes, but trading is intensified two times as the above case, the volume bucket becomes:

{(9:38:00;10.02;300); (9:39:00;10.01;500); (9:40:00;10.00;200)}

And OI = |300-500-200|/1000 = 0.4

The author continues with hypothetical situation when trading volume is intensified three times, four times, and five times…the results show OI calculated using TR-VPIN can be arbitrary, and depends solely on intensity as oppose to trade classification.

The author continues with an alternative FB-VPIN measure, in FB-VPIN, instead of using a time bar, we use volume bars with contract (tick) size of 200 (an arbitrary number just for demonstration purposes). Then a typical volume bucket may look like:

{(9:39:00;10.02;200); (9:40:00;10.02;200); (9:41:00;10.01;200);

(9:41:00;10.01;200); (9:43:00;10.00;200)}

The author states that FB-VPIN is a preferable approach to TR-VPIN as it is invariant to trade intensities.

The author goes on with a more detailed analysis; perhaps the earlier example was too simplistic, since trading intensity itself may be informative; in such a case TR-VPIN can still be valid. This time, the author constructed a benchmark VPIN called U-VPIN (u for “uninformed”), U-VPIN aims to incorporate only the effect of trading intensity instead of both intensity and trade classification. As the author explains that U-VPIN measures how TR-VPIN is expected to evolve given the variation in trade patterns (i.e. change in intensity), but absent any systematic order flow imbalance. U-VPIN assumes the next aggregated “transaction” has the equal likelihood of being classified as a buy or a sell, so E[b(t)] = 0. The one variable that is critical in constructing U-VPIN is the number of volume bars within a volume bucket, denoted as Q. And the author defines F(Q) = E[OI(Q)], where E[] is the expectation. It is worthy to point out that F(Q) is non-increasing in Q and approaches zero at a rate of Q^(-0.5).

U-VPIN is formally defined as a moving average of E[OI(Q)], with a window of size L, and we set L=50, which is analogous to the case of TR-VPIN, which is a moving average measure of OIs.

**FINDING 2) F(Q) is decreasing in Q. This implies as trade intensified, OI and the associated TR-VPIN will tend to increase as Q declines.**

**FINDING 3) All else equal, the level of OI, and associated TR-VPIN is monotonically related to the length of the time bar. This is consistent with FINDING 2) since longer time bars cause the number of “transactions” within a volume bucket (i.e. Q) to decrease.**

The author next introduces a second benchmark as he claims U-VPIN does not take into account the impact of volume ACROSS the bars. The author generalizes the definition of U-VPIN by saying U-VPIN, which incorporates E[OI] is the L1 norm of signed order imbalance = Sigma(|(w(t)\*b(t))|); and the second benchmark, denoted U2-VPIN will utilize the L2 norm of order imbalance, namely sqrt(Sigma(w(t)^2)). Formally, U2-VPIN is a moving average of L2 norms of order imbalances. Again, L=window size=50. The author noted that by Jensen’s inequality L2 norm of OI will always be > L1 norm of OI = E[OI].

**FINDING 4) All else equal, E[OI], and associated TR-VPIN is increasing in degree of volume heterogeneity across time bars. (My interpretation of FINDIG 4) is that TR-VPIN is positively correlated with volatility)**

Now we have introduced FB-VPIN, TR-VPIN, U-VPIN, U2-VPIN. The author proceeds by presenting empirical findings:

The market involved in the study is the E-mini S&P 500 future’s market, and on average of 2,165,000 contracts are exchanged daily. During regular trading hours, 390 transactions per minute or 6.5 transactions per second , are executed, indicating sufficient liquidity within the market. The author set the volume bucket size to 40,000 (V=40,000), and L = 50. The author again emphasizes the importance of Q (# of aggregated transactions within each volume bin). And he mentioned that there is a high skewness in trade size within each time bar. Typical contract size within each time bar is one or two, whereas a few time bars can entail as much as 40, to a few hundred contracts.

For testing, the author used three different TR-VPIN measures, each with a time bar interval of 10, 60, and 300 seconds. And three FB-VPIN measures, each with a volume bar of size 1, 1000, and 5000. Below are there findings:

**FINDING 5) TR-VPIN is highly correlated with U-VPIN and U2-VPIN, and the correlation increases with time bar length. Hence the trading classification has negligible impact; this finding is consistent with earlier hypothesis that TR-VPIN is BAD (for a lack of better word)**

**FINDING 6) TR-VPIN and FB-VPIN behave very differently, and only TR-VPIN attains an exceptional value on May 6, 2010. (Flash Crash)**

The study also shows there is a negative correlation between FB-VPIN computed from relative small bins and trading volume; and the correlation drops monotonically as the size of the volume bars grow. And very importantly since FB-VPIN is not sensitive to trading intensity, but only to order classification, this negative correlation shows trader indicator (i.e. order classification) are largely unrelated to overall trading volume.

Another highlight is that TR-VPIN displays a positive correlation with VIX, while FB-VPIN displays a negative correlation with VIX. To conclude, the author states:

**FINDING 7) TR-VPIN is strongly correlated with volume, and weakly correlated with VIX. FB-VPIN is weakly, and negatively correlated with volume, while strongly and negatively correlated with VIX.**

The author proceeds to explore how TR-VPIN behaves as a forecast variable relative to other standard volatility predictors, and how TR-VPIN is correlated with future volatility. He performs the experiment by running regressions. The author claims TR-VPIN performs poorly as a volatility forecast. The reason is when he compares TR-VPIN’s performance relative to U-VPIN, TR-VPIN actually shows a lower correlation, suggesting the trade classification employed in TR-VPIN does not provide any incremental value (if not detrimental). Then he used data to show TR-VPIN’s trade classification is actually detrimental to its predictive power of future volatility by citing R^2. The study also finds that VIX actually shows the highest correlation with future volatility (beats both TR-VPIN and U-VPIN).

One striking finding is that FB-VPIN is actually negatively correlated with future volatility, and it is especially so when the volume bar sizes chosen are small. The author went on to discuss how TR-VPIN performs so poorly as even a random walk volatility model (use one-day lagged trading volume) can outperform TR-VPIN by a factor of 4. And his conclusion is that VIX is the most superior predictor of future volatility.

**FINDING 8) TR-VPIN is horrible, its underperformance relative to the U-VPIN measures is evident that trade classification mechanism employed by TR-VPIN is detrimental to its ability to forecast volatility.**

**FINDING 9) If we control trading volume, and thus remove the mechanical correlation, TR-VPIN is negatively related to future return volatility. Again TR-VPIN is BAD, for lack of a better word.**

Next the author tries to refute ELO’s finding and claims tick data actually provide the best opportunity to misclassification. (Note FB-VPIN with a volume bar of size 1 translates to using tick level data) The author provides an intuition behind why FB-VPIN is negatively correlated with VIX.

**FINDING 10) FB-VPIN are negatively correlated with future volatility, especially for small volume bar size, because using finer grained level data, trade indicators (classifications) show less serial correlations. Intuitively, if the market is volatile, then tick level transactions vary more, and FB-VPIN drops as OI decreases => negative correlation with volatility.**

Next the author starts an entirely new VPIN measure, called bulk volume VPIN (or BV-VPIN), originally presented by ELO. BV-VPIN uses a probabilistic trade classification scheme, so b(t)s are no longer binary values, but are instead defined as 2\*Z(delta(P)/std) – 1. Where delta(P) is the change in price, and std is the standard deviation of price changes in-between bars (may it be time or volume). Z is the standard normal cdf. The advantage of this approach is that every bar is classified as partially buy and partially sell. (for more information, you can refer to Jiongjia’s 1st summary).

The authors shows if BV trade indicator process has mean zero, is iid distributed, and is independent of volume weights, then sqrt(E[OI^2]) = L2 norm \* std(b), since std(b) < 1, E[BV-OI^2] < E[TR-OI^2]. Author also points out the L2 norm of OI is inversely related to Q (# of bars within a bucket). Using Z as a CDF for transforming variable serves the only purpose of converting units into [0,1] range, and does not imply any underlying distribution. With simple mathematical manipulation, it can be shown that b(t) = 2\*Z(delta(P)/std(t)\*std(t)/std) – 1, here std(t) is the standard deviation within each time bar, whereas std is the standard deviation cross different time bars. From there, we can see if std(t) increases, the P(|b(t)| > 0.5) increases as well. Therefore, the author suggests that trade intensity and price volatility are primary drivers of the BV-VPIN order imbalance. The author expects BV-VPIN to be more sensitive to price volatility than TR-VPIN. On top of that price volatility and price intensity can have combined effect on BV-VPIN measure, which is observed during the subsequent Flash Crash study.

Next the author tries to determine if BV-VPIN’s trade classification provide incremental value. (based on all the findings so find, you can probably guess the outcome…)

The author provides empirical correlation findings of different BV-VPINs based on variable-length time bars and U-VPIN, U2-VPIN, FB-VPIN, volume, VIX. They observed that BV-VPINs are negatively correlated with FB-VPINs, strongly correlated with volume, and moderately correlated with VIX, and strongly correlated with U-VPIN and U2-VPIN. When regressing BV-VPIN with future volatilities (one step and fifty steps ahead), BV-VPIN’s performances are much better than that of TR-VPIN, achieved an R2 of 11.09% for one step future volatility and 39.26% for fifty steps future volatilities; yet VIX outperform BV-VPIN by obtaining R2s of nearly double the values achieved by BV-VPINs. The author claims there is indeed great correlation between BV-VPIN and future volatility, yet it is not that miraculous as it has already incorporated volume and price volatility into the BV-VPIN OI measures. As a volatility forecasting variable, it significantly underperforms the contemporary realized (historical) volatility (so it underperforms a random walk model).

Next, he revisits the Flash Crash. Here is the idea: the author asserts “okay, perhaps TR-VPIN and BV-VPIN don’t do too well in ‘benign times’, but may be of great importance during turbulent times…let’s revisit the Flash Crash”. The author basically verified their FINDING 1) that TR-VPIN reached moderately high value (0.53), but the value itself does not heed an atypical scenario, and thus does not possess predictive power. You shall take a look of the VPIN figures presented in figure 11 and 12 of the original paper to come up with potentially alternative interpretations

**FINGDING 11) TR-VPIN at any specific point in time on a given trading day is sensitive to the exact sequence of trades recorded prior to that trading day. TR-VPIN can even produce divergent behaviors based on the sample starting values.**

**FINDING 12) SOI (signed OI) is better than OI**

They proceed to show both volume and VIX are better predictors of the flash crash than TR-VPIN and BV-VPIN.

**CONCLUSION: All the different variations of VPIN do not warrant good assessment of order imbalance and predictability of episodic market stress scenarios.**