International Journal of Social Science and Economic Research

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

THE FORECAST OF THE PROBABILITY OF INFORMED TRADING ON THE FLASH CRASH—BASED ON THE CSI 300 INDEX FUTURES AND SPOT

Zhu Xiaoyu*

*College of Economic and Management,
Nanjing University of Aeronautics and Astronautic, Nanjing 210016, China

ABSTRACT

The paper is intended to explore the impact of information asymmetry on market volatility in high frequency world by providing an examination concerning the probability of informed trading across the related market. I research on the relationship between informed trading and market volatility in spot market, futures market and cross-market scenarios based on the VPIN toxicity metric model. I find that probabilities of informed trading in the CSI300 stock index futures market and spot market in 2015 stood at 0.30 and 0.33 respectively, slightly higher than the previous level. This indicates the existence of index volatility uncertainties. I also find that the futures market's probability of informed trading negatively relates to the spot market liquidity over the following 4 minutes and positively to the spot market volatility over the following 4 minutes. It serves as an early warning of joint crash across futures and spot markets. The futures market's probability of informed trading is an efficient indicator of toxicity-induced illiquidity.

Keywords: Informed trading; market micro-structure; high frequency trading.

I. INTRODUCTION

Flash crashes emerge increasingly frequently varying across financial markets and over time. Chinese stock index futures intraday dropped over 9% on August 24, 2015. A plummet of S&P500 index by 560 points in 4 minutes triggered a flash crash on May 6, 2010. Bitcoin's price fell 14% in a short period of 30 minutes following a 'flash crash' on exchange Bitfinex on August 19, 2015. Flash crashes quicken the index deviation by implementing that the volatility of security price occurs in a matter of minutes. Consequently, systematic risk exposure disseminated across multiple related financial markets varies the market microstructure by affecting market participants' confidence and behavior. High frequency trading (HFT) is argued

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

that it implies negative implications on market microstructure. However, HFT toxicity level depends on strategies adopted by high frequency traders. HFT strategies vary, ranging from market making to pernicious algorithm. There is a general agreement that market making HFT always enhances market liquidity by providing limit orders. When they supply more liquidity to markets, informational efficiency is enhanced by tightening bid-ask spreads (see Brogaard, Hendershott, and Riordan, 2013; Carrion, 2013). Menkveld (2013) provided the empirical evidence that the participation of HFTs reduce spreads for Dutch stocks trading on Chi-X Europe. However, unlike traditional market making, the high frequency market makers employing a cross-market strategy making it akin to statistical arbitrage are on only one side of the book in each stock, and there is no commitment to continuously providing liquidity (see Virtu Financial, Inc., 2014). They provide empirical evidence on high frequency market making that their trading can induce market instability by periodic illiquidity (see Kirilenko, Kyle, Samadi, and Tuzun, 2011; Madhavan, 2013).

Information asymmetry level of financial internal market induces the financial crashes. Informed traders adversely select market makers who may be unaware they are providing liquidity at a loss. Then market crash tend to occur after the sufficient loss of liquidity providers is accumulated leading this type of investors to quit the market. Pernicious trading gambits operated by high frequency algorithm induce dramatic market volatility as well as information inefficiency through increased information asymmetry (see Ye, Yao, and Jiading, 2013). Pernicious HFT trigger flash crashes by accumulating unbalance orders between buyers and sellers. Researchers have proposed different theories and models to explain the underlying mechanism about the relationship between informed trading and financial flash crashes. ELO (2011) proposed the VPIN toxicity metric model based on the original PIN estimation approach, in which VPIN metric could predict short-term fluctuations timely when a significant reduction occurs on underlying asset index. Order flow imbalance implies a negative shock on market participants' behavior as well as investment performance. Information such as public news drives the price deviation by order imbalance and quotation spread. Bernile.G, Hu.J and Tang.Y (2016) provide the empirical evidence of how to serve early warning of collapse and argue that the E-mini Standard & Poor's 500 futures' abnormal order imbalances direction is consistent with the expected response to public news. Zhou.Q (2015) proposes the empirical evidence that the probability of informed trading relates negatively to market liquidity and positively relates to volatility at Chinese stock index futures market in 2015. Therefore, market A's information asymmetry tends to induce market A's flash crash.

Based on the price co-integration and information infection theory, the paper proposes that market A's information asymmetry tend to induce market B's flash crash when the two markets' indices show signs of co-integration. Lux (1995) proposes the contagion model, in which a selforganizing infective process induces most investors to overvalue their asset before market crashes occurs. HF traders operate the arbitrage strategy by taking advantage of price deviation across the market. HFT enhances the information transmission between the related markets by providing the one side order tied to the underlying assets traded in several markets. Berman (2014) highlights the case of the GLD SPDR, an equity ETF linked to gold. The HFT market maker would be quoting both in GLD and in Gold futures to employ arbitrage. But there are 13 other exchange traded products tied to gold that requires placing bids and asks across all of these 91 potential pairs, in the gold futures as well as the cash market. International systematic risk exposure varies across countries and through time. Pukthuanthong and Roll (2009) focus is measuring time-varying integration. Berger and Pukthuanthong (2012) develop a time-varying measure of systemic risk within international equity markets, in which aggregate systematic risk exposure is high across multiple markets when joint crashes imply a negative shock spreading internationally. If a shock occurs when multiple markets share a high risk exposure to a common factor, then these multiple markets will experience simultaneous market declines. They argue that the probability of a worldwide financial crash is highest when many countries share a high exposure to the world market factor. In other words, market A's information asymmetry factor is a precursor to market B's flash collapse.

Our study mainly concentrates on the information asymmetry across the related financial market. Market A's information asymmetry not only induces its own flash crash, but also contributes Market B's information asymmetry and consequently triggers Market B's flash crash. I research the lead-lag relationship between the futures and spot markets by information infection. Since the futures market is more sensitive to public information and reacts faster than spot market, index futures prices are generally ahead of the index spot prices about few minutes varying cross countries and through time. Futures market lies in a dominant position of price discovery. I select Market A's information asymmetry indicators to predict Market B's flash crash possibility. I propose futures market's probability of informed trading is an efficient indicator of toxicity-induced liquidity of the spot market. I also evaluate the relationship between indicator and volatility across the market. I explore the inherent link between futures markets' order toxicity and the performance of spot market together.

II. METHODOLOGY

The Internal Mechanism of Imbalance Orders, Liquidity and Volatility

The paper proposes the following assumptions. Trading orders are unlimited. Orders arrive according to Poisson arrival process. The difference between buy and sell orders provided by traders during a unit timescale is m. The system order-processing capacity is estimated by

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

processed volumes per timescale represented by n. Transactions follow the rules of price priority and time priority. When the transaction matching system is exposed to an unstable state, in which m>n, trading processing capacity is exceeded by more and more arriving orders hysteresis paired. The system is at steady state in which m<n. The probability of system processing i^{th} order at any t timescale is $\Pr ob_i$. There is a relative balance between unprocessed orders and processed orders in the system. The system state transition probability from 0 to 1 is $m \Pr ob_0$ and state transition probability from 1 to 0 is $n \Pr ob_1$. When system is in the state 0, the equilibrium equation is that $m \Pr ob_0 = n \Pr ob_1$, then $\Pr ob_0 = \frac{n}{m} \Pr ob_1$. The equilibrium equation could be extended to any state where i is greater than 1. State transition probability from i to i+1 turns out to be $m \Pr ob_i + n \Pr ob_{i+1}$ while state transition probability from i to i-1 turns out to be $n \Pr ob_i + m \Pr ob_{i-1}$. The equilibrium equation is derived as $m \Pr ob_{i-1} + n \Pr ob_{i+1} = (m+n) \Pr ob_i$

$$\Pr ob_i = \left(\frac{m}{n}\right)^i \Pr ob_0, i = 1,2,3... \quad (1)$$

Assuming $k = \frac{m}{n} < 1,$

Then $\Pr ob_i = k^i \Pr ob_0, i = 1,2,3...$ (2)

The k in equation 2 is the ratio of average order quantities to average volumes measuring transaction matching system's order processing intensity. K converges on 1 illustrating the higher the system processing intensity, the greater the market transaction volume. Under the equilibrium condition, the probability of the order quantity reaching i in the system is $\Pr ob_i$, as

$$\sum_{i=0}^{\infty} \Pr ob_i = 1,$$

$$\Pr ob_0 + \sum_{i=1}^{\infty} k^i \Pr ob_0 = 1 \quad (3)$$

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

$$\Pr ob_0 = (1+\sum_{i=1}^{\infty} k^i)^{-1} = (\sum_{i=0}^{\infty} k^i)^{-1} = (\frac{1}{1-k})^{-1} = 1-k$$

$$\Pr ob_i = (1-k)k^i, i = 1,2,3... \tag{4}$$

The average order processing quantities in transaction matching system equilibrium state is Q, then, $Q_{bal} = \sum_{i=0}^{\infty} i \Pr ob_i = \sum_{i=0}^{\infty} i(1-k)k^i = \sum_{i=1}^{\infty} k^i = \frac{k}{1-k} = \frac{m}{n-m}$ (5)

When good news A reaches the market with the probability of δ, orders issued by informed trader obey the Poisson distribution is gradually added to market unprocessed orders queue. Then, order quantities raise to m+λ per unit timescale. $Q_{bal}^{*} = \frac{m+\lambda}{n-m-\lambda} > Q_{bal}$, it indicates that good news enhance the trading volume.

Assume $p = f(Q_{bal}), \frac{dp}{dQ_{bal}} > 0,$ (P represents trading price). Therefore, increased trading volumes Q_{bal} stimulate underlying asset trading price promotion. When bad news B reaches the market with the probability of 1-δ, $Q_{bal}^{*} < Q_{bal}$,. Decreased trading volumes Q_{bal} decline underlying asset trading price.

The proof provides the direct mechanism of how information asymmetric affects the underlying asset price. Adverse selection risks simulated by information asymmetric are operated by informed traders. When market makers traded with informed traders, they did not realize that providing liquidity would take excess loss without risk compensation. The direction of market price changes erodes the interests of uninformed traders. The profitability of liquidity providers primarily depends on bid-ask spread. Therefore, the larger the proportion of informed trading, the more liquidity providers would choose to quit the market leading internal market illiquidity risk highly enhanced. Market illiquidity expands market makers' bid-ask spread. Consequently, I propose Assumption 1 that probability of informed trading relates negatively to liquidity in the financial internal market.

Rational investors bear excess risk with requirements for excess return compensation, so holding the excess supply of the underlying asset will raise their expected excess returns. Private information reaching the market drives the underlying asset fluctuation. Uninformed traders cannot detect the same amount of information as informed traders from market price. So they tend to require more excess returns to make up for excess risk. Illiquidity not only widens bid-

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

ask spread but also lead to violent price fluctuations. Therefore, the price fluctuations can effectively reflect the heterogeneous information held by investors. Consequently, I propose Assumption 2 that the probability of informed trading relates positively to volatility in the financial internal market.

The Internal Mechanism of Cointegration Effect

Information spillover effect across the market has been researched extensively. The efficient market hypothesis states that spot and futures prices of a commodity should cointegrate with a unit slope on futures prices. The futures contracts are effective hedging tools to reduce the underlying asset risk exposure. The futures price is affected by the expected spot price of the underlying asset. The price discovery process induces a dynamic relationship between the futures and the spot price.

$$r_{1,t} = \mu_l + \sum_{i=1}^{p} \alpha_{1,i} r_{1,t-i} + \sum_{j=1}^{q} \beta_{1,j} r_{l,2-j} + \sum_{m=1}^{r} \gamma_{1,m} r_{l,t-m} + \sum_{n=1}^{s} \varphi_{l,n} r_{4,t-n} + \mathrm{K}_1 (P_{2,t-1} - P_{1,t-1})$$

$$\varepsilon_t = T(\varepsilon_{i,t}) \left| \Omega_{i,t} \right. \sim t(0, \Sigma_i) \qquad (6)$$

With the widespread acceptance of EMH theory, the short-run one-for-one relationship between spot and futures prices it postulates has verified empirically (see Figuerola-Ferretti and Gonzalo, 2010, Westerlund and Narayan, 2013). Information in the futures market prices is the first to be effectively reflected, followed by the spot price, which indicates that futures market price guides the spot market price. The futures market price can be used as a leading indicator to spot market price by reflecting market effectiveness. As a result, information asymmetry in the leading market may have an impact on price volatility in the lagged market. Thus, I propose that futures market's probability of informed trading relates negatively to spot market next term's liquidity as Assumption 3, and futures market's probability of informed trading relates positively to spot market next term's volatility as Assumption 4.

Measurement of Informed Trading

ELO (2012) present a new procedure to estimate flow toxicity based on volume imbalance, in which trade intensity is named as volume-synchronized probability of informed trading (VPIN). This approach does not require the intermediate numerical estimation of parameters and overcomes the difficulties of estimating multiple parameters assumed in previous models such as PIN (see ELO,2012). It provides a time varying way to measure the toxicity of order flow at

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

high-frequency environment. ELO (2012) calculate buy and sell volumes using one-minute time bars and Let,

$$V_\tau^B = \sum_{i=t(\tau-1)+1}^{t(\tau)} V_i \times Z\left(\frac{P_i - P_{i-1}}{\sigma_{\Delta P}}\right)$$
$$V_\tau^S = \sum_{i=t(\tau-1)+1}^{t(\tau)} V_i \left[1 - Z\left(\frac{P_i - P_{i-1}}{\sigma_{\Delta P}}\right)\right] = V - V_\tau^B \tag{7}$$

Z is the CDF of the standard normal distribution as an estimator of the standard derivation of price changes between time bars. The procedure splits the volume in a time bar equally between buy and sell volume. The important information will be given greater weights by the VPIN model as it induces an increase in the volume of transactions and baskets per unit time. The expected imbalance trading is $E[|V_\tau^B - V_\tau^S|] \approx \alpha\mu$. The VPIN flow toxicity metric is shown below.

$$VPIN = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} = \frac{\alpha\mu}{V} \approx \frac{\sum_{\tau=1}^{n} |V_\tau^S - V_\tau^B|}{nV} \tag{8}$$

III. DATA AND BACKGROUND

I select the database of China's futures market obtained from Resset information Co., Ltd. The paper selects CSI300 index futures contract and CSI300 index under high frequency environment for a sample period spanning from January 2, 2015 to December 31, 2015. The high frequency interval is 1 minute. Data processing applies SAS9.4 and matlab 2015(a). Variables are logarithmically processed in need.

The Distribution of VPIN

Based on the VPIN toxicity metric model (see Easley.D,2012), the paper retrieved data from WIND dataset that contains CSI 300 and CSI 300 index dominant futures contracts for a sample period spanning from January 2, 2015 to December 31, 2015. I calculate 56280 VPIN metrics of CSI 300 index futures and 58210 VPIN metrics of CSI 300 respectively, along with the standard deviation, skewness, kurtosis and JB statistics of VPIN metric serials (See table 1).

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

TABLE 1: Statistical results of all variables in the CSI300 futures market

Statistics	Observations	Mean	SD	Max	Min	Skewness	Kurtosis	ADF
$VPIN^F$	56280	0.30	0.08	0.73	0.00	0.16	3.45	-3.84
$VPIN^S$	58210	0.33	0.05	0.66	0.07	0.17	3.06	-9.43
$Liquidity^F$	56280	0.94	2.35	7.50	-6.72	0.19	3.76	-8.27
$Volatility^F$	56280	5.79	3.68	9.20	0.70	1.69	5.74	-5.80
$Volume^F$	56280	2.07	0.73	4.46	-2.16	0.08	3.41	-11.90
ORT^F	56280	1.87	2.27	6.09	-3.40	-0.08	4.13	-3.59
$Liquidity^S$	58210	0.42	1.19	12.06	3.56	1.71	5.58	-19.35
$Volatility^S$	58210	6.02	2.90	4.67	0.88	1.03	5.51	-14.16
$Volume^S$	58210	1.77	0.68	10.09	5.96	0.16	3.97	-17.02
ORT^S	58210	2.67	1.87	9.75	2.98	-0.05	4.04	-13.02

Notes: Orders turnover rate represents order to trade ratio, $ORT_T = \ln(VOLUME_T / ORDERS_T)$. In 2015, probabilities of informed trading stood at 0.30 in the CSI 300 stock index futures market and 0.33 in spot market respectively, slightly higher than in the previous year.

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

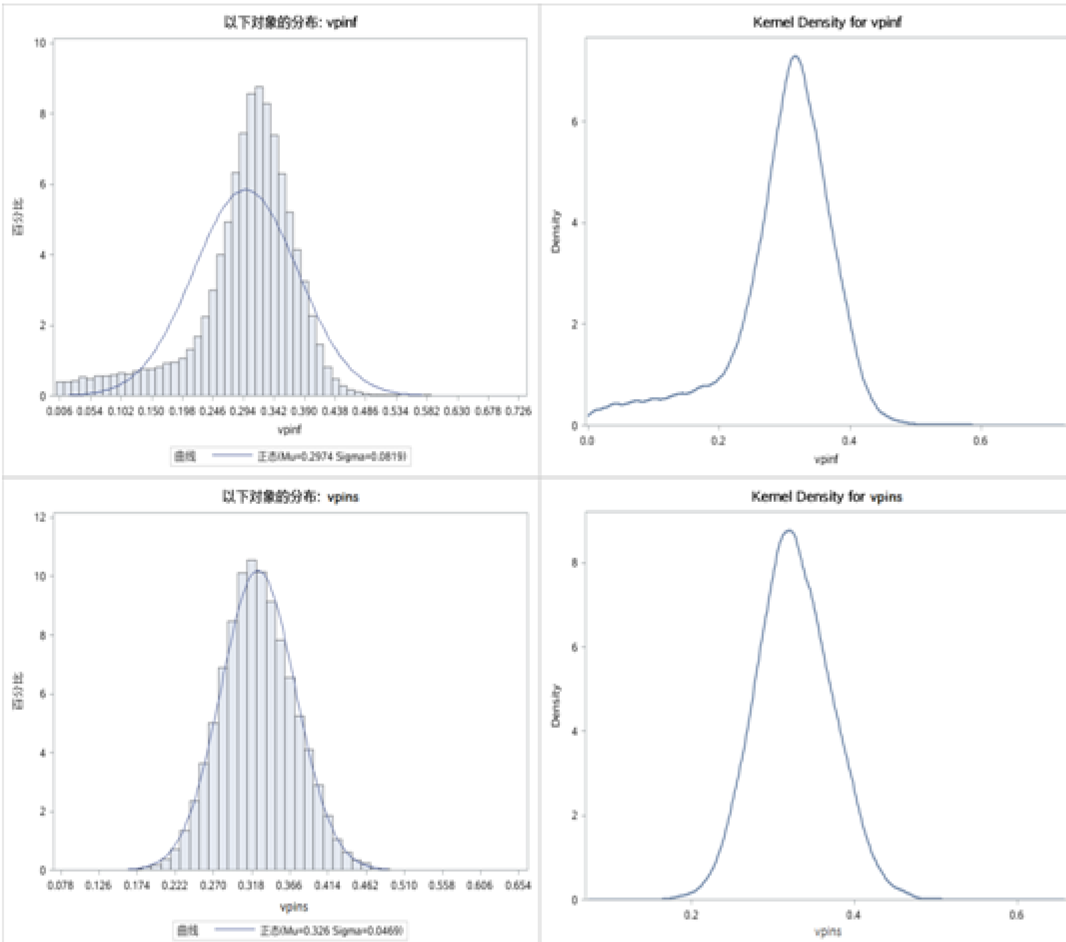

Fig.1-2. Distribution of CSI 300 futures VPIN, KDE of CSI 300 futures VPIN,

Fig.3-4. Distribution of CSI 300 VPIN, KDE of CSI 300 VPIN

The probability of informed trading measured by VPIN metric obeys a normal distribution featuring high peak and fat tail in CSI300 index futures market (see Table 1) and it densely distributes around 0.30. The 99th quartile of futures market's VPIN metric value is 0.44. While the VPIN metric increases over 0.44, the lack of supply willingness leads to significant decreased liquidity which likely induces the market collapse. The CSI300 Index VPIN metric obeys normal distribution approximately with the mean locating around 0.33, which is slightly higher than in previous years where the figure stood at 0.29 in 2014. (See Zhou.Q, Liu.W, 2014) The 99th quartile of index spot market's VPIN metric value is 0.438. The mean of spot market's VPIN metric is higher than that in futures market. The standard deviation of spot market's VPIN metric is lower than that in futures market, suggesting that the probability of informed trading in the

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

spot market is more stable and lower than that in the futures market.

The Distribution of Liquidity

Amihud (2002) constructs an illiquidity indicator that divides the depth and the width of liquidity. The paper draws on this indicator to measure the market liquidity in high frequency environment (see Amihud. Y , Hameed. A, Kang.W, Zhang. H, 2015). The liquidity depth is represented by trading volumes in the set time interval while the liquidity width is represented by the absolute value of bid-ask spread at the set time interval. The liquidity indicator is shown below,

$$LIQUIDITY = \ln\left(\frac{V}{INTERVAL\_SECONDS \times |P_\tau^E - P_\tau^B|}\right) \tag{9}$$

In Equation 9, P_τ^E is the transaction price at the end time of period T, P_τ^B is the transaction price at the beginning time of period T.

The Distribution of Volatility

A large body of empirical literature has developed aimed at assessing to predict market volatility. The extended realized volatility allows for market microstructure frictions in the observed high-frequency returns. In addition to traditional realized volatility measures and the underlying sampling high frequencies, the forecasting performance of realized volatility measures is also designed to mitigate the impact of the microstructure noise. (see Andersen, T., Bollerslev, T., Meddahi, N., 2011). The Realized volatility proposes the realized measurements into the traditional GARCH model Realized GARCH (1, 1) model estimates fluctuation values obtained from the conditional variance through historical variability. The paper adopts this approach to estimate volatility under high frequency environment. The approach process is shown below.

$$\begin{aligned}
&r_t = \ln(P_T^E) - \ln(P_{T-1}^E) \\
&r_t = c + \phi_1 r_{t-1} + \varepsilon_t \\
&\varepsilon_t = \sigma_t z_t \\
&z_t \sim N(0,1) \\
&\log \sigma_t^2 = \omega + \beta \sigma_{t-1}^2 + \gamma \log \chi_{t-1} \\
&\log \chi_t = \xi + \varphi \log \sigma_t^2 + \delta(z_t) + u_t, u_t \sim i.i.d.N(0, \sigma_u^2)
\end{aligned} \tag{10}$$

P_T^E is the transaction price at the end of period T, $\alpha, \beta, \omega, \gamma, \xi, \varphi, \delta$ are the parameters based on the maximum likelihood estimation.

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

The Distribution of Other Variables

All the variables in the model are statutory according to ADF test (see Table 1). The distributions of liquidity, volatility and volume are positively skewed while the turnover ratio is negatively skewed. A look at the means and standard deviations of these variables leads to the conclusion that liquidity in futures market is higher than the spot market, while volatility in the spot market is higher than the futures market.

IV. EMPIRICAL SPECIFICATION AND ESTIMATION

The Prediction in the Index Futures Markets

Market makers can refer VPIN metrics to measure the toxicity of order flow at high-frequency environment. The sudden growth of informed trading induces liquidity providers to bear excess loss without risk compensation. The unexpected loss affects market makers' profits, consequently, market makers form overly pessimistic beliefs about expected investment environment because they vary judgment from positive domain to negative domain. Same amount of loss at different domains leads individuals to react differently in which people tend to react more pessimistic to low outcomes in the negative domain relative to the positive domain (see Kuhnen.C, 2015). The market liquidity tends to be at the critical situation when market makers provide rare orders because of refusing to enter the market. In this section, the paper seeks to address the linkage between informed trading and futures liquidity. Model 1 is a multiple regression characterizing the relationship between probability of informed trading, liquidity, trading volume, order to trade ratio and next term liquidity of stock index futures. It illustrates the prediction of illiquidity induced by informed trading in the futures market.

$$LIQUIDITY_{T+1} = \alpha_0 + \beta_1 VPIN_T + \beta_2 VOLATILITY_T + \beta_3 LIQUIDITY_T + \beta_4 VOLUME_T + \beta_5 ORDER\_TO\_TRADE_T + \xi_{T+1} \tag{11}$$

In Model 1 $LIQUDITY_{T+1}$ represents futures market liquidity at time T+1, $VPIN_T$ represents the probability of informed trading in futures markets. $VOLATILITY_T$ represents futures market's volatility at time T, $LIQUIDITY_T$ represents futures market liquidity at time T. $VOLUME_T$ represents the trading volume of futures market at time T. $ORDER\_TO\_TRADE_T$ represents turnover ratio orders of futures markets at time T.

Model 2 is a multiple regression characterizing the relationship between probability of informed trading, liquidity, trading volume, order to trade ratio and next term volatility of stock index futures. It illustrates the prediction of flash crash induced by informed trading in the futures market.

$$VOLATILITY_{T+1} = \alpha_0 + \beta_1 VPIN_T + \beta_2 VOLATILITY_T + \beta_3 LIQUIDITY_T + \beta_4 VOLUME_T + \beta_5 ORDER\_TO\_TRADE_T + \xi_{T+1} \quad (12)$$

In Model 2 $VOLATILITY_{T+1}$ represents volatility of futures market at time T+1. The remaining variables are the same with model 1.

I draw the conclusion that the stock index liquidity relates negatively to probability of informed trading while stock index volatility relates positively to probability of informed trading in futures market at the significance of 1% (see Table 7). Holding the other independent variables constant, for every 1% of change in the probability of informed trading, the futures market's liquidity at next period would change -1.02% while volatility at next period would change 0.90%. The stock index futures' liquidity at next period relates positively to such variables as current liquidity, trading volume and turnover rate at the significance at least of 5%. It illustrates that the greater the current market volume, the faster the turnover, the better liquidity at next period market (see table 8). The stock index futures volatility at the next period relates negatively to current liquidity, trading volume and turnover rate in futures market at the significance of 5%. It illustrates that an increased stock index futures trading volume, more liquidity and faster trading turnover at current term induce more stable index. Consequently, the probability of informed trading estimated by VPIN metric reflects the futures market's information asymmetry as well as systematic risk exposure. It could efficiently predict liquidity and volatility of stock index futures at the next period.

The Prediction in the Index Spot Markets

Model 3 is a multiple regression characterizing the relationship between probability of informed trading, liquidity, trading volume, order to trade ratio and next term liquidity in the index spot market. It illustrates the prediction of illiquidity induced by informed trading in the index spot market.

$$LIQUIDITY^S_{T+1} = \alpha_0 + \beta_1 VPIN^S_T + \beta_2 VOLATILITY^S_T + \beta_3 LIQUIDITY^S_T + \beta_4 VOLUME^S_T + \beta_5 ORDER\_TO\_TRADE^S_T + \xi_{T+1} \quad (13)$$

In Model 3 $LIQUIDITY^S_{T+1}$ represents spot market liquidity at time T+1. $VPIN^S_T$ represents the probability of informed trading of spot market at time T. $VOLATILITY^S_T$ represents spot market

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

volatility at time T. $LIQUIDITY_T^S$ represents spot market liquidity at time T. $VOLUME_T^S$ illustrates the trading spot market volume at time T.

Model 4 is a multiple regression characterizing the relationship between probability of informed trading, liquidity, trading volume, order to trade ratio and next term volatility in the index spot market. It illustrates the prediction of flash crash induced by informed trading in the index spot market.

$$VOLATILITY_{T+1}^S = \alpha_0 + \beta_1 VPIN_T^S + \beta_2 VOLATILITY_T^S + \beta_3 LIQUIDITY_T^S + \beta_4 VOLUME_T^S + \beta_5 ORDER\_TO\_TRADE_T^S + \xi_{T+1}$$

In Model 4 $VOLATILITY_{T+1}^S$ represents spot market volatility at time T+1. The remaining variables are the same with model 3.

I draw the conclusion that liquidity relates negatively to probability of informed trading while volatility relates positively to probability of informed trading in stock index spot market at the significance of 1% (see Table 7). Holding the other independent variables constant, for every 1% of change in the probability of informed trading, the futures market's liquidity at next period would change -0.39% while volatility at next period would change 0.46%. The stock index liquidity at next period relates positively to such variables as current liquidity, trading volume and turnover rate at the significance at least of 5%. It illustrates that the greater the current market volume, the faster the turnover, the better liquidity at next period market (see table 8). The stock index volatility at the next period relates negatively to current liquidity, trading volume and turnover rate in spot market at the significance of 5%. It illustrates that an increased stock index trading volume, more liquidity and faster trading turnover at current term induce more stable index. Consequently, the probability of informed trading estimated by VPIN metric reflects index spot market's information asymmetry as well as systematic risk exposure. It could efficiently predict spot market's the next period of liquidity and serves as an early warning of spot market's collapse.

Guide Relations between Futures and Spot Market

The paper adopts the Johansen co-integration tests, VECM model and impulse response to analyze the co-integration between CSI300 stock index futures and stock index. It also evaluates how accurately stock index futures price can predict spot market price.

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

TABLE 2: SAMPLE STATISTIC DESCRIPTIONS

Statistics	Observations	Mean	Max	Min	Median	SD	Skewness	Kurtosis
CSI 300	52702	3954.3	5379.5	2952.7	3791.3	584.3	0.7	0.5
CSI 300 futures	52702	3902.3	5389.6	2755.2	3683.0	623.9	0.7	0.5
CSI 300 returns	52702	0.0	0.1	-0.1	0.0	0.0	-4.0	345.4
CSI 300 futures returns	52702	0.0	0.1	0.0	0.0	0.0	-1.6	319.8

Notes: the mean of CSI300 (at roughly 3954.3) is higher than CSI300 futures, while its standard deviation is lower than the CSI futures.

The co-integration test bases on 52702 high frequency data at one-minute interval collected from WIND HFT dataset of CSI 300 index and the CSI 300 stock index futures for a sample period spanning from December 31, 2013 to December 31, 2014. I analyze the CSI 300 index return series noted as RS and the CSI 300 Index futures return series noted as RF. It illustrates that the average returns of CSI 300 stock index and futures are close to zero, with distributions both characterized by high kurtosis, fat tail and negative skewness. However, the CSI300 futures variance is higher than the CSI300 significantly (see Table 2).

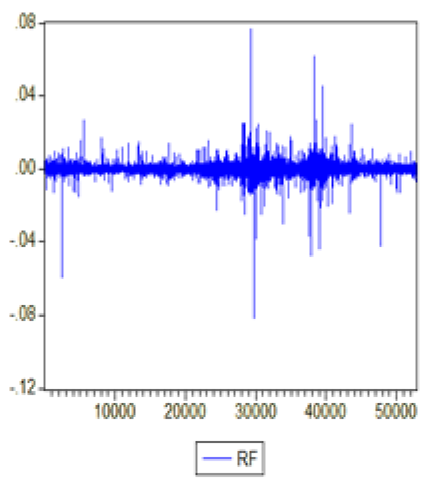

Fig. 5: CSI300 Stock index futures return series

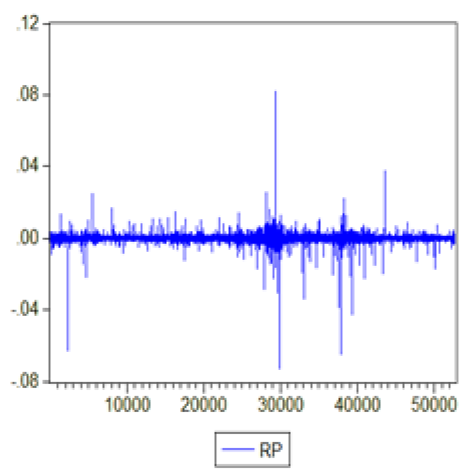

Fig.6: CSI300 Stock index return series

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

TABLE 3: Autocorrelation test results

Statistics	CSI300 index futures return	CSI300 index return
LB(-10)	26.7122	84.6911
LB(-20)	46.4147	54.8788

Notes: LB (-10) and LB (-20) are the Q statistic of logarithmic yield with lag 10 and lag 20 respectively.

TABLE 4: Stationary test results

Statistics	Price series ADF	Return series ADF
CSI300 Stock index return	-1.231(0.6635)	-107.86(0.0001)***
CSI300 index futures return	-1.3569(0.6052)	-230.87(0.0001)***

Notes: *** denotes rejection at the 1% significance level. First-order differential of price series are stationary series, return series is stationary series at 1% significance level.

Autocorrelation test results illustrate that the autocorrelation exists in the CSI300 stock index futures price and the CSI300 index at the significance of 1%. It indicates that the CSI 300 stock index futures and spot price series are non-stationary. The first-order differential of price series and return series are stationary at the significance of 1% (see Table 4).

TABLE 5: Co-integration TESTS and causality test

Johansen test				Granger test			
H_0	Trace	λ_{trace}	P-value	H_0	Observations	F	P-value
R =0	0.0005	30.41	0.000	CSI300 does not Granger Cause CSI300 futures	52696	15.01	0.001
R =1	0.0000	1.67	0.196	CSI300 futures does not Granger Cause CSI300	52696	16.73	0.000

According to the results of Johansen test in which the price includes error items but does not include trend items, the P-value of λ_{trace} below 0.01 illustrates that more than zero cointegration relationship exists at the significance of 1%. However, the P-value of λ_{trace} above 0.01 illustrates one cointegration relationship exists at the significance of 1%. Granger test results indicate long-term cointegration relationship exists between CSI 300 stock index futures price and index series

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

(see table 5).

TABLE 6: Price Guidance coefficient CSI 300 index futures and spot estimates

	CSI300 stock index	CSI300 index futures		CSI300 stock index	CSI300 stock index futures
$\alpha_{s,-1}$	0.186**	0.054***	$\beta_{f,-1}$	0.185***	-0.027***
	(33.716)	(6.4528)		(51.426)	(-4.952)
$\alpha_{s,-2}$	-0.101***	-0.059***	$\beta_{f,-2}$	0.110***	0.006
	(-20.565)	(-7.886)		(30.672)	(0.205)
$\alpha_{s,-3}$	-0.070***	-0.025	$\beta_{f,-3}$	0.0834***	0.002
	(-11.241)	(-0.162)		(12.653)	(0.490)
$\alpha_{s,-4}$	-0.056***	-0.034	$\beta_{f,-4}$	0.037***	0.008
	(-8.669)	(-0.148)		(3.173)	(0.750)
$\alpha_{s,-5}$	-0.012***	0.009	$\beta_{f,-5}$	0.006	0.041
	(-4.641)	(0.922)		(0.544)	(0.622)
$\alpha_{s,-6}$	-0.039**	0.007	$\beta_{f,-6}$	0.013	0.014
	(-2.589)	(0.715)		(0.315)	(0.341)
μ	0.002	0.002**	γ	0.001	-0.004
	(0.524)	(2.731)		(0.049)	(-0.108)

Notes： **, *** denotes rejection at the 5% and 1% significance level separately. Futures on the spot from the impact point of view, continuous lags of 1 to 4 orders are statistically significant at the significance of 1%.

In accordance with AIC guidelines, the optimal price sequence lag order for parameters p, q and r is 4. The VECM test indicates that at the significance of 1%, CSI 300 stock index futures price and spot price are mutually influenced at the lag order of 4. The influence coefficient of stock index futures on spot market and *t* statistic are greater than those of the spot market on futures market illustrating stock index futures price has a greater leading impact on spot index price (see Table 6). In terms of the impact of futures on the spot, continuous lags β_f of 1 to 4 orders are statistically at the significance of 1%. In terms of the impact of spot on the futures, continuous

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

lags $\alpha_{s,}$ of 1 to 2 orders are statistically significant. It presents that the CSI300 stock index futures prices lead spot prices by about 4 minutes. Error correction coefficient γ of CSI300 stock index futures and spot price series are -0.0036 and 0.0011 respectively. It illustrates that the error correction term has a negative regulatory role on stock index futures price series while has a positive regulatory role on the stock price series.

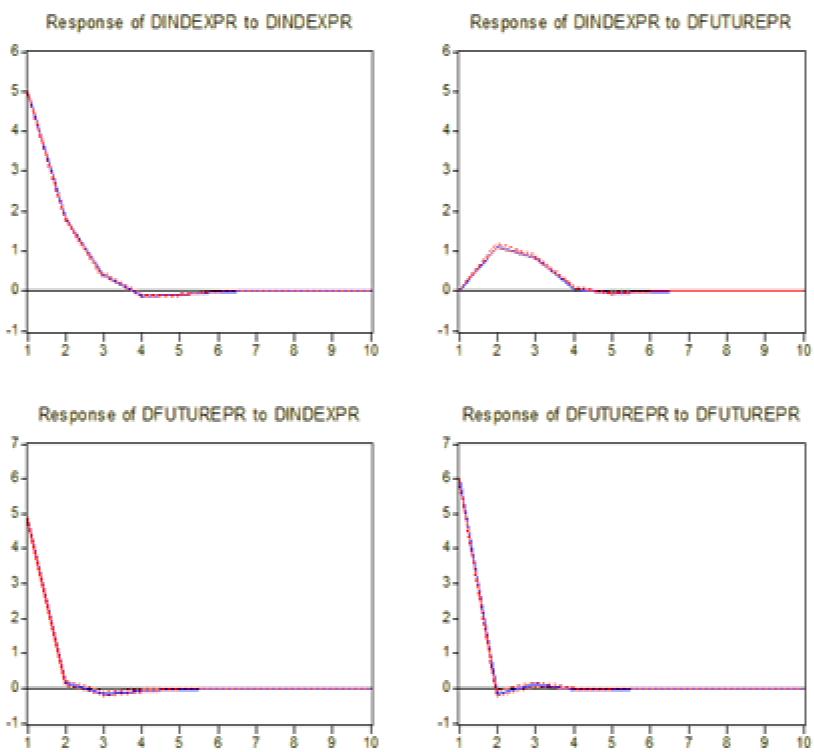

Fig. 7: CSI 300 index futures and spot the impulse response

The results illustrate that the response of CSI 300 stock index futures tend to be stationary in one minutes after disturbance. The stock index futures market is quickly responsive. Futures price is sufficiently effective to reflect the information within one minute (see Fig.7). CSI 300 Index, on the other hand, is less responsive, requiring 4 minutes before the impact dies out. The VECM test results illustrate that a bidirectional price guide relation exists between the futures and spot markets at the significance of 1%. The impact of futures prices on the spot market lasts about 4 minutes while the impact of the spot price on the futures market lasts about 2 minutes. Stock index adjustments for price volatility in the stock index futures market is slow. Stock index prices discovery lasts longer than 4 minutes and is lagging behind the futures market. When the spot index market price is interfered by exogenous factors, stock index futures price completes clearing price in about 1 minute by an immediate adjustment. When the spot index futures market price interfered by exogenous factors, stock index price has a slower adjustment. The index achieves price discovery by adjusting itself within 2 to 4 minutes. The ability of the stock index futures market to absorb external interference performs well. Stock index futures price at current can immediately reflect most of the market information. Consequently, CSI300 stock index futures market leads the spot market for about 4 minutes.

The Prediction Cross the Market

Model 5 is a multiple regression characterizing the relationship between index futures market's probability of informed trading, index spot market's liquidity, index spot market's trading volume, index spot market's order to trade ratio and next four term spot market's liquidity across the related markets. It illustrates the prediction of illiquidity induced by index futures market's informed trading in the index spot market.

$$LIQUIDITY^S_{T+4} = \alpha_0 + \beta_1 VPIN^F_T + \beta_2 VOLATILITY^S_T + \beta_3 LIQUIDITY^S_T + \beta_4 VOLUME^S_T + \beta_5 ORDER\_TO\_TRADE^S_T + \xi_{T+4} \quad (15)$$

In Model 5, $LIQUIDITY^S_{T+4}$ represents spot market's liquidity at time T+4, $VPIN^F_T$ represents the probability of informed trading in the futures markets.

Model 6 is a multiple regression characterizing the relationship between index futures market's probability of informed trading, index spot market's liquidity, index spot market's trading volume, index spot market's order to trade ratio and next four term spot market's volatility across the related markets. It illustrates the prediction of flash crash induced by index futures market's informed trading in the index spot market.

$$VOLATILITY^S_{T+4} = \alpha_0 + \beta_1 VPIN^F_T + \beta_2 VOLATILITY^S_T + \beta_3 LIQUIDITY^S_T + \beta_4 VOLUME^S_T + \beta_5 ORDER\_TO\_TRADE^S_T + \xi_{T+4} \quad (16)$$

In Model 6, $VOLATILITY^S_{T+4}$ represents the spot market volatility at time T+4. The remaining variables are the same with Model 5.

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

TABLE 7: VPIN across market prediction

Model	1	2	3	4	5	6
Dependent variable	Liquidity $_{t+1}$	Volatilty $_{t+1}$	Liquidity $^{s}_{t+1}$	Volatilty $^{s}_{t+1}$	Liquidity $^{s}_{t+4}$	Volatilty $^{s}_{t+4}$
Interception	-3.11*** (-15.02)	-1.54*** (-7.26)	-0.09 (-0.61)	-1.53*** (-19.24)	-0.09*** (-7.58)	-1.56 (-0.90)
$VPIN_t$	-1.02*** (-12.75)	0.90*** (11.25)	-	-	-0.83** (-2.38)	1.45*** (18.13)
Liquidity $_t$	18.52** (2.88)	-15.43** (-2.57)	-	-	-	-
Volatilty $_t$	-14.85*** (-4.04)	12.11** (3.29)	-	-	-	-
$Volume_t$	3.80*** (5.21)	-6.73*** (-9.22)	-	-	-	-
ORT_t	19.41*** (8.55)	-24.99*** (-11.01)	-	-	-	-
$VPIN^s_t$	-	-	-0.39*** (-8.98)	0.46*** (9.50)	-	-
Liquidity ^s_t	-	-	0.12 (0.05)	-0.31 (-0.13)	-4.37** (-2.67)	6.29*** (5.28)
Volatilty ^s_t	-	-	-17.88*** (-6.17)	15.28*** (5.27)	-12.95*** (-4.47)	16.97*** (5.85)
Volume ^s_t	-	-	6.43*** (9.46)	-5.19*** (-7.63)	4.83*** (7.10)	-5.05*** (-7.43)
ORT^s_t	-	-	4.57** (2.44)	-2.66* (-1.42)	3.61* (1.93)	-2.05 (-1.10)
Observations	56135	56135	58212	58212	49781	49781
Adjust R^2	0.2343	0.2928	0.1845	0.3064	0.1893	0.3019

Notes: Model 1 to model 6 correspond equation 11 to equation 16. The stock index liquidity in spot market relates negatively to stock index probability of informed trading in spot market at the significance of 1%. Probability of informed trading changes 1% and spot market's liquidity over the following 4 periods would change -0.83%. Probability of informed trading changes 1% and spot market's volatility over the following 4 periods would change 1.45%.

Transaction information asymmetry induces adverse selection, consequently affecting market liquidity. Different types of information induce different degrees of the underlying asset price volatility. The informed traders operate adverse selection strategies leaving uninformed traders to

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

bear unexpected risk. Adversely selections implemented by informed traders induce market makers unaware of the unexpected loss while providing liquidity. Then market crash tend to occur after the sufficient loss of liquidity providers is accumulated leading this type of investors to quit the market. When the good or bad news arrive the market, informed traders take the lead in the transaction due to information superiority, far ahead of the market price discovery inducing the benefit from prioritized submitted orders. Dramatic volatility of market liquidity induces price fluctuations reflected by the toxicity of order flow. I draw the conclusion that the stock index liquidity in spot market relates negatively to futures market's probability of informed trading at the significance of 5% while stock index volatility in spot market relates positively to futures market's probability of informed trading at the significance of 1% (see Table 7). Holding the other independent variables constant, for every 1% of change in futures market's probability of informed trading, the spot market's liquidity at next four periods would change -0.83% while volatility at next four periods would change 1.45%. The stock index liquidity at next four periods relates positively to such variables as current liquidity, trading volume and turnover rate at the significance of 1%. It illustrates that the greater the current market volume, the faster the turnover, the better liquidity at next four periods market (see table 8). The stock index futures volatility at the next four periods relates negatively to current liquidity, trading volume and turnover rate in futures market at the significance of 5%. It illustrates that an increased stock index futures trading volume, more liquidity and faster trading turnover at current term induce more stable index. Consequently, the probability of informed trading estimated by VPIN metric reflects the futures market's information asymmetry as well as systematic risk exposure. It could efficiently predict liquidity and volatility of stock index between the stock index spot market and futures market 4 minutes in advance.

TABLE 8: Granger test of VPIN to market liquidity and volatility

H_0	Observations	F-Statistic	P-value
$VPIN_t$ does not Granger Cause liquidity $_{t+1}$	56133	230.900	1.E-100
liquidity $_{t+1}$ does not Granger Cause $VPIN_t$	56133	10089.1	0.00000
$VPIN_t$ does not Granger Cause volatility $_{t+1}$	56133	817.431	0.00000
volatility $_{t+1}$ does not Granger Cause $VPIN_t$	56133	174.596	8.3E-79
$VPIN^s_t$ does not Granger Cause liquidity $^s{}_{t+1}$	58210	1883.41	0.00000
liquidity $^s{}_{t+1}$ does not Granger Cause $VPIN^s_t$	58210	186.712	1.5E-81

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

$VPIN^s_t$ does not Granger Cause volatility $^s_{t+1}$	58210	215.708	1.4E-93
volatility $^s_{t+1}$ does not Granger Cause $VPIN^s_t$	58210	265.416	5.2E-108
$VPIN_t$ does not Granger Cause liquidity $^s_{t+4}$	49779	25.4627	8.9E-12
liquidity $^s_{t+4}$ does not Granger Cause $VPIN_t$	49779	0.17395	0.04034
$VPIN_t$ does not Granger Cause volatility $^s_{t+4}$	49779	14.9935	3.1E-07
volatility $^s_{t+4}$ does not Granger Cause $VPIN_t$	49779	12.0655	5.8E-06

It illustrates that futures markets' VPIN metric and the spot markets' next period of liquidity exists mutually reinforcing relationship (see table 8). VPIN metric remains at a high level during the flash crash until the index back to the ordinary level. The probability of informed trading in the futures market could predict the spot markets' next period of liquidity. Granger test results illustrate the futures markets' VPIN metric and the spot markets' next period of liquidity exist reciprocal causation. Consequently, the probability of informed trading in the futures market could predict the spot markets' next period volatility. I find that the futures market's probability of informed trading relates negatively to the spot market liquidity over the following 4 minutes and positively to the spot market volatility over the following 4 minutes. It serves as an early warning of joint crash across futures and spot markets. The futures market's probability of informed trading is an efficient indicator of toxicity-induced illiquidity. From the perspective of the futures market, it predicts changes in the spot market liquidity and volatility. Information acts an important factor to order transactions. Since market makers are passive traders by providing limit orders, it is volume rather than time as the operative metric. When uninformed traders undertake the loss beyond risk affordability, they would withdraw the order and quote. Supply of securities is higher than the demand and market liquidity is at the low level. Persistently high VPIN metric levels have implications for market maker behavior, which in turn fuels price volatility.

The paper studies how informed trading of the futures market influences the spot market. Co-integration test indicates that CSI300 stock index futures lead spot market by 4 minutes. Futures prices market information firstly discovers, and then affects the spot market price. Consequently, the dramatic increase of the futures market's informed trading induces futures market's liquidity deteriorates, stock index futures price volatility occurs. Stock index futures price linkage spot prices volatility. When VPIN metric varies beyond 95th percentile, regulators could take preventative actions to supervise potential price manipulation. VPIN metric is a risk management tool of cross market arbitrage to slow down market volatility. Thus, in terms of cross-market, the probability of informed trading in futures market relates negatively to the liquidity and positively

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

to the volatility of spot market 4 minutes in advance illustrating that futures market's probability of informed trading may predict flash crash in spot market. Consequently, the probability of informed trading measures the effectiveness of information asymmetry as well as systematic risk exposure across the related market. The paper proposes market regulators can view the index futures VPIN metric detected beforehand liquidity crisis in the market as a reference.

V. SUMMERY AND CONCLUSIONS

Information asymmetry is accumulated to induce serious market fluctuations always accompanied by market volatility. Technology has dramatically changed the nature of market microstructure. Flash crashes depressed investor sentiment leading to net capital continuous outflow four weeks after a stock collapse in July 2015 of Chinese stock market (see China Statistical Office reports). I examine the probability of informed trading in futures and spot markets separately based on the VPIN toxicity metric model. I research on the relationship between informed trading and market volatility within spot market, futures market and cross-markets based on the VPIN toxicity metric model. I find that probabilities of informed trading in the CSI300 stock index futures market and spot market in 2015 stood at 0.30 and 0.33 respectively, slightly higher than the previous level. This indicates the existence of index volatility uncertainties.

Secondly, in stock index spot and futures market, the probability of informed trading in the market relates negatively to the liquidity while positively to the volatility. It is indicated that an increased probability of informed trading adversely affects the smooth operation of the market, the ability to predict futures market volatility. VPIN metric tends to enhance beyond the 95th percentile when the underlying asset prices suddenly rises or falls in minutes. In terms of the spot market, the probability of informed trading may be used to predict flash crash. Then, the co-integration test shows that the CSI300 stock index futures market lead the spot market by about 4 minutes. In terms of cross-market, I find that the futures market's probability of informed trading relates negatively to the spot market liquidity over the following 4 minutes while positively to the spot market volatility over the following 4 minutes. The futures market's probability of informed trading is an efficient precursor of toxicity-induced illiquidity of the spot market. Consequently, the futures market's probability of informed trading is an efficient precursor of flash crash of the spot market. The indicator of index futures market's information asymmetry compared with indicator of index spot market, can effectively predict the possibility of flash crash in index spot market 4 minutes in advance.

Finally, the paper has shown that the VPIN metric can effectively signal the toxicity-induced volatility. VPIN metric estimates the order imbalance of each transaction by measuring the

ISSN: 2455-8834

Volume:04, Issue:02 "February 2019"

proportion of the market maker in informed trading effectively portrayed the level of order flow toxicity. More liquidity providers tend to quit the market in the influence of the larger proportion of informed trading leading to augmented internal market illiquidity risk. Illiquidity not only expand bid-ask spread but also lead to violent price fluctuations. VPIN metrics help market makers detect potential liquidity risk and serve as an early warning of joint crash across futures and spot markets. Market A's information asymmetry factor is a precursor to correlated market B's flash collapse.

REFERENCES

Andersen, T., Bollerslev, T., Meddahi, N., 2011. Realized volatility forecasting and market microstructure noise. Journal of Econometrics, vol. 160, no. 1, pp, 220-234.

Amihud, Y., Hameed, A., Kang, W., Zhang, H., 2015. The illiquidity premium: International evidence. Journal of Financial Economics, vol. 117, no. 2, pp. 350-368.

Berger. D., Pukthuanthong.K.2012. Market fragility and international market crashes. Journal of Financial Economics, 105, pp. 565–580.

Brogaard, J. Hendershott, T.J., Riordan, R., 2013. High frequency trading and price discovery. Review of Financial Studies 27 (8), 2267–2306.

Berman, G., 2014. What drives the complexity and speed of our markets? US Securities and Exchange Commission Speeches, April 15. www.sec.gov/speeches.

Bernile, G., Hu, J., Tang, Y., 2016. Can information be locked up? Informed trading ahead of macro-news announcements. The Journal of Financial Economics, vol. 121, no.3, pp. 496-520.

Carrion, A., 2013. Very fast money: high-frequency trading on Nasdaq. Journal of Financial Markets, 680–711.

Easley, D., López de Prado M, O'Hara M. 2011. The Micro-structure of the 'Flash Crash': Flow Toxicity, Liquidity Crashes and the Probability of Informed Trading. The Journal of Portfolio Management, vol. 37, no.2, pp. 118-128.

Easley, D., López de Prado M, O'Hara M. 2012. "Flow toxicity and liquidity in a high- frequency world," Review of Financial Studies, vol. 25, no.5, pp. 1457-1493.

Figuerola-Ferretti, J., Gonzalo, 2010. Modeling and measuring price discovery in commodity markets. J.Econometrics, 158, pp. 95–107.

J. Westerlund, P. Narayan, 2013. Testing the efficient market hypothesis in conditionally heteroskedastic futures markets. J. Futures Markets, 33, pp. 1024–1045.Kirilenko A, Kyle A P, Samadhi M, 2010. The Flash Crash: The Impact of High-Frequency Trading on an E lectronic Market. SSRN Working paper.

Kuhnen.C. 2015. Asymmetric learning from financial information. The Journal of Finance, 5, pp. 2029-2062.

Lux T., 1995. Herd behavior, bubbles and crashes. The Economic Journal, pp. 881-896.

Madhavan, A., 2013. Exchange-traded funds, market structure, and the flash crash. Financial Analysts Journal 68, 20–35.

Pukthuanthong, K., Roll, R., 2009. Global market integration: a better way to measure it and its application. Journal of Financial Economics 94, 214–232.

Virtu Financial, Inc., 2014. Form S-1. Filed with the Securities and Exchange Commission. March 10.

Ye, M., Yao, C., Jiading, G., 2013. The Externalities of high-frequency trading. SSRN working paper. University of Illinois.

Zhou.Q, Zhu.Y, 2015. Probability of Informed Trading Market, Liquidity and Volatility with Evidence from China stock index futures market. Journal of Financial Research, vol. 419, no.5, pp. 132-147.

