The Spanish Review of Financial Economics 10 (2012) 74–83

The Spanish Review of Financial Economics

www.elsevier.es/srfe

Article

From PIN to VPIN: An introduction to order flow toxicity☆

David Abad[a,*], José Yagüe[b]

[a] Universidad de Alicante, Dpto. Economía Financiera y Contabilidad, Facultad de Ciencias Económicas y Empresariales, Ctra. San Vicente, s/n, 03690 San Vicente del Raspeig (Alicante), Spain

[b] Universidad de Murcia, Dpto. Organización de Empresas y Finanzas, Facultad de Economía y Empresa, Campus de Espinardo, s/n, 30100 Murcia, Spain

ARTICLE INFO

Article history:
Received 9 August 2012
Accepted 30 October 2012
Available online 20 November 2012

JEL classification:
G12
G14
C58
D53

Keywords:
PIN
VPIN
Order flow toxicity
High frequency trading (HFT)
Adverse selection
Probability of informed trading
Market microstructure

ABSTRACT

As an update of the well-known PIN measure, Easley et al. (2012a) have developed a new measure of order flow toxicity called Volume-Synchronized Probability of Informed Trading or VPIN. Order flow toxicity makes reference to adverse selection risk but applied to the world of high frequency trading (HFT). We provide a detailed description of the VPIN estimation procedure paying special attention to the main innovations introduced and the key variables of this novel tool. By using a sample of stocks listed on the Spanish market, we compare VPIN to PIN. Although VPIN metric is conceived for the HFT environment, our results suggest that certain VPIN specifications provide proxies for adverse selection risk similar to those obtained by the PIN model. Thus, we consider that the key variable in the VPIN procedure is the number of buckets used and that VPIN can be a helpful device which is not exclusively applicable to the HFT world.

1. Introduction

The 2010 Flash Crash is without a doubt the shortest event in the recent history of financial markets to merit so much attention and generate so much controversy among practitioners and academics. On May 6th 2010 the Dow Jones Industrial Average plunged about 1000 points – or about 9% – only to recover those losses within minutes.[1] Although the ultimate cause of the Flash Crash is still under discussion (e.g., Kirilenko et al., 2011; Madhavan, 2012) it is generally accepted that this event was the result of a new trading paradigm emanating from legislative changes in the US ("Regulation National Market System" of 2005, or "Reg NMS") and Europe ("Markets in Financial Instruments Directive" of 2007, or "MiFID") and prompted by substantial technological advances in computation and communication. The new legislative environment fostered both greater competition and market fragmentation while technological advances made high-speed trading technically possible at and between different trading venues. As a result, the world of high frequency trading (HFT) has appeared as a new reality in current markets that is progressively outshining traditional or low frequency trading (LFT).[2]

A number of studies indicate that HFT is playing a crucial role in liquidity supply activity in current markets. Hasbrouck and Saar (2012), by analyzing low-latency activity (i.e., trading strategies that respond to market events in the millisecond environment) find that it improves traditional market quality measures such as the liquidity in the limit order book. Similarly, Brogaard et al. (2012)

☆ This paper is inspired by the comments that David Abad made about a preliminary version of Easley et al. (2012a) presented at the Workshop "*High Frequency Trading: Financial and Regulatory Implications*" held in Madrid, October 2011. David Abad appreciates helpful comments from Maureen O'Hara and Marcos López de Padro. David Abad acknowledges financial support from the *Ministerio de Ciencia e Innovación* through grants ECO2010-18567 and ECO2011-29751. José Yagüe acknowledges financial support from *Fundación Caja Murcia*. The authors also thank Roberto Pascual for his constructive comments, as well as Zheng Junyan for the help in programming of PIN estimation.

* Corresponding author.
E-mail addresses: goliat@ua.es (D. Abad), ppyague@um.es (J. Yagüe).

[1] The 2010 Flash Crash is also known as 'The Crash of 2:45' or just simply, 'the Flash Crash'.

[2] Easley et al. (2012c) provide a detailed description of this new paradigm and how HFT exploits LFT's structural weaknesses.

2173-1268/$ – see front matter
http://dx.doi.org/10.1016/j.srfe.2012.10.002

find evidence of HFT benefitting price efficiency and the provision of liquidity at stressful times such as the most volatile days and before and after macroeconomic news announcements. Nevertheless, in the HFT environment the liquidity provision activity and its associated risks acquire a new dimension. Thus, Easley et al. (2012a) introduce the concept of "order flow toxicity" to represent adverse selection risk in the HFT context. In the authors' words "order flow is regarded as toxic when it adversely selects market makers who may be unaware that they are providing liquidity at a loss" (p. 1458). Thus, in this case, adverse selection must be understood not only as a problem of asymmetric information but also as a wider notion that may encompass other risks related to liquidity provision. When order flows are essentially balanced, high frequency market makers have the potential to earn razor thin margins on massive numbers of trades. When order flows become unbalanced, however, market makers face the prospect of losses due to adverse selection. These market makers' estimates of the time-varying toxicity level now becomes a crucial factor in determining their participation. If they believe that toxicity is high, they will liquidate their positions and leave the market. To measure "order flow toxicity" Easley et al. (2012a) present the Volume Synchronized Probability of Informed Trading or VPIN metric, a new procedure to estimate the probability of informed trading based on volume imbalance and trade intensity.

VPIN is inspired by the well-known PIN model of Easley et al. (1996), henceforth EKOP (1996). The PIN is a consolidated model to measure the presence of informed traders that has been widely adopted to address a variety of issues in the empirical financial literature, among others: information content of the time between trades (Easley et al., 1997a), trade size (Easley et al., 1997b), analyst coverage (Easley et al., 1998), electronic market order flow (Brown et al., 1999), stock splits (Easley et al., 2001), dealer vs. auction markets (Heidle and Huang, 2002), asset pricing (Easley et al., 2002; Aslan et al., 2011), non-anonymous vs. anonymous trading systems (Gramming et al., 2001), market reaction to public and private information (Vega, 2006), corporate investment decision (Ascioglu et al., 2008; Chen et al., 2007), block ownership (Brockman and Yan, 2009), and market anomalies (Kang, 2010; Chen and Zhao, 2012). However, the PIN is not extent from criticism. First, there is a growing debate as to the appropriateness of PIN in measuring information-based trading (Aktas et al., 2007; Duarte and Young, 2009; Easley et al., 2010; Akay et al., 2012). Second, several papers show that the PIN estimations could suffer several biases for different reasons such as trade misclassification (Boehmer et al., 2007), boundary solutions or the floating-point exception, especially in very active stocks (Easley et al., 2010; Lin and Ke, 2011; Yan and Zhang, 2012), and propose different solutions to mitigate such biases.

PIN and VPIN models require trading volume classified as buy or sell and are based on the notion that order imbalances signal the presence of adverse selection risk. However, the VPIN approach has some practical advantages over the PIN methodology that make it particularly attractive for both practitioners and researchers. The main advantage is that VPIN does not require the estimation of non-observable parameters using optimization or numerical methods thereby avoiding all the associated computational problems and biases. In addition, VPIN allows the capturing of risk variations at intraday level while the original PIN model does not.

In a series of related papers Easley et al. (2011a, 2011b, 2012a) present the VPIN as a useful tool for different market participants. Easley et al. (2011a) show the VPIN of the e-mini S&P500 futures contract achieving its maximum level around the Flash Crash. Higher levels of toxicity force HF market makers to liquidate their positions and leave the market offering a plausible explanation of the Flash Crash. The authors recommend that regulators use VPIN as a warning tool that could herald the implementation of regulatory actions to forestall crashes.[3] Easley et al. (2012a) also show that VPIN has forecasting power over volatility (toxicity-induced) and could become valuable as a risk management tool for market making activity. It can be also useful for trading strategies based on volatility arbitrage and for brokers who look for best time of execution. Easley et al. (2011b) 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. On the other hand, Andersen and Bondarenko (2011) put forward several criticisms questioning the predictive power of VPIN. In particular, the authors document that VPIN is a poor predictor of short run volatility with a limited predictive power emanating from the mechanical relation to the underlying trading intensity. Andersen and Bondarenko's analysis provoked a speedy response from Easley et al. (2012d) who basically point to the confusion in the methodology they use, the analysis they perform and the conclusions they draw.

Using a selected sample of 15 Spanish stocks, the main objective of this paper is to offer a detailed description of the VPIN estimation procedure, its key variables, and its usefulness in an attempt to gain a better understanding of this novel tool. Departing from the PIN model, we document the main innovations introduced in this updated version of the probability of informed trading and we analyze the compatibility of both models. To the best of our knowledge, this is the first study to apply VPIN methodology to a sample of European stocks.[4] Although the relevance of HFT in the Spanish Stock Exchange has not yet been formally measured, mostly because of data availability problems, informal conversations with regulators corroborate the interest of HF traders in the most active stocks listed on the Spanish market.

Our results suggest that certain VPIN specifications provide proxies for adverse selection risk similar to those obtained by the PIN model. In this sense, we consider that the key variable in the VPIN procedure is the number of buckets used, so estimations of VPIN using one bucket are quite similar to those obtained by the PIN model. We conclude that VPIN is, in the main, a straightforward way to measure adverse selection but not exclusively for the high frequency environment.

The paper is organized as follows: Section 2 briefly reviews the PIN model. Section 3 focuses on VPIN putting special emphasis on the main innovations it incorporates and its computational procedure. Section 4 describes the Spanish stock market and the sample employed. Section 5 compares PIN to VPIN aggregated values. Section 6 concludes.

2. PIN model (EKOP 1996)

The probability of information-based trading (PIN) is a measure of the information asymmetry between informed and uninformed trades that builds on the theoretical work of Easley and O'Hara (1987, 1992). The original PIN model was introduced by Easley et al. (1996). Since then, various empirical papers have implemented, adapted, and improved the PIN approach (Easley et al., 1997a,b, 1998, 2008). The PIN measure is not directly observable but is a function of the theoretical parameters of a microstructure model that have to be estimated by numerical maximization of a likelihood function.

The model views trading as a game between liquidity providers and traders (position takers) that is repeated over trading days.

[3] Bethel et al. (2012) confirm that VPIN could have given a strong signal ahead of the Flash Crash event on May 2010 and it can be use for a fully-fledged early warning system for unusual market conditions.

[4] Up to now, VPIN has been mainly applied to high-frequency trading futures contracts.