Market Microstructure: A Review of models

Rahman Doostian
Assistant Professor Department of Accounting, Khorramabad Branch, Islamic Azad University, Khorramabad, Iran
Rahman_doostian@yahoo.com

Omid Farhad touski
Assistant Professor Department of Accounting, Khorramabad Branch, Islamic Azad University, Khorramabad, Iran
(Corresponding author)
farhadi_omid58@yahoo.com

Submit: 30/04/2022 Accept: 30/10/2022

ABSTRACT

In recent decades, the development of capital market microstructure theory has led to a broad understanding of market performance, market organizational structure, transaction costs, and asset prices. Certainly one of the most important goals of microstructure modeling is to understand and describe the quality of markets. Define market microstructure as the process by which investors' latent demands are ultimately translated into prices and volumes. Define market microstructure as the study of trading mechanisms and regulations used to accomplish a trade. definition of market microstructure which is the study of the intermediation and the institutions of exchange. The main purpose of this article is to review the most important microstructure models of the market. Defines market microstructure as the study of the process and outcomes of exchanging assets under explicit trading rules. The main body of market microstructure theory consists of inventory-based models and information-based models. This article focuses on information-based models. Studying open microstructure models can help market participants understand the pricing process and the impact of information on pricing. The study of market microstructure theory and models leads to a deep understanding of the performance and organizational structure of the market, transaction costs, asset prices, and an understanding and description of market quality.

Keywords:

Information asymmetry, Probability of informed trading, Market Microstructure.

1. Introduction

In recent decades, the development of capital market microstructure theory has led to a broad understanding of market performance, market organizational structure, transaction costs, and asset prices. Market microstructure is the study of the process and outcomes of exchanging assets under explicit trading rules. While much of economics is concerned with the trading of assets, market microstructure research focuses on the interaction between the mechanics of the trading process and its outcomes, with the specific goal of understanding how actual markets and market intermediaries behave. This focus allows researchers to pose applied questions regarding the performance of specific market structures, as well as more theoretical queries into the nature of price adjustment. Certainly one of the most important goals of microstructure modeling is to understand and describe the quality of markets. O’Hara (1995) Defines market microstructure as "the study of the process and outcomes of exchanging assets under explicit trading rules”. Spulber (1996) has provided a broader definition of market microstructure which is the study of the intermediation and the institutions of exchange. Madhavan (2000) Define market microstructure as the process by which investors' latent demands are ultimately translated into prices and volumes. Asmar and Ahmad (2011) Define market microstructure as the study of trading mechanisms and regulations used to accomplish a trade. *Trading mechanisms* refer to the methods of trading securities. The mechanisms are determined by several dimensions including *market type, price discovery, order forms and degree of transparency*. *Market Regulations* on the other hand refer to the rules of trading securities defined by securities market to control various aspects of trading process, such as the rules of order priority, tick size and spread, listing, trading unit, price thresholds, trading status, short selling and off-market trading (Aigbovo & Isibor, 2017). Price and volume dynamics resulting from orders can help to understand the behavior of investors as well as the amount of information they use in trading. The traditional view of financial market theories is that price is formed in the process of supply and demand. However, the reality is that factors such as information asymmetry and the difference in trading time (and other cases) can upset the price. These factors create inventory-based and information-based models. In the first generation of microstructure models, the market maker seeks to increase the bid-ask spread price for inventory price risk by providing market liquidity. The second generation of models is related to the asymmetry between market participants and the creation of price gaps to compensate for the costs of adverse selection. The basis of market microstructure models is based on the order-oriented model (order flow shifts prices) developed by Kyle (1985). This model was later developed by Glosten and Milgrom (1985), who consider a trader-centric market, according to which differences in information between traders can lead to the cost of information for traders. Most microstructure models based on information are divided into three main parts. The first section includes models that study the price effect of information (Hasbrouck, 1991a, 1991b; Madhavan & Smidt, 1991). The second group of models used certain criteria such as the bid-ask spread price (Bagehot, 1971; Jaffe & Winkler, 1976; McInish & Wood, 1992), the volume and size of the deal (Keim & Madhavan, 1995, 1997), firm size (Hasbrouck, 1991b), number of transactions (Jones, Kaul, & Lipson, 1994) and ratio of insiders (Jones et al., 1994). The third generation provides sequential trading models that describe the trading process and estimate the probability of informed trading. The first numerical perspective formula for estimating the probability of a conscious trade was developed by (Easley, Kiefer, & O'hara, 1996; Easley & O'hara, 1992) by estimating the maximum likelihood probability. And a set of subsequent models has also been developed.

One of the most important issues in the capital market microstructure is the relationship between informed and uninformed traders. Researchers believe that informed traders have an information advantage and use this advantage to trade to benefit uninformed traders. A common feature of many theoretical models of market microstructure is the existence of an expert who deals with two types of traders: informed traders and liquidity (uninformed) (Bagehot, 1971; Easley, Hvidkjaer, & O'hara, 2002; Easley, Kiefer, & O'hara, 1996, 1997; Easley, Kiefer, O'hara, & Paperman, 1996; Easley & O'hara, 1987, 1992, 2004; Easley, O'hara, & Saar, 2001; Glosten & Milgrom, 1985; Kyle, 1985).

Traditional microstructure theory provides two major directions to explain price setting behavior: *inventory models and asymmetric information based* *models.* The branch of inventory models investigates the uncertainty in order flow and the inventory risk and optimization problem of liquidity suppliers under possible risk aversion and The main result of inventory models is that the market maker sets quotes in such a way that the costs of order processing and inventory maintenance are covered while the asymmetric information based models, model market dynamics and adjustment processes of prices using insights from the theory of asymmetric information and adverse selection. The two main approaches in asymmetric information models are *sequential trade models* and *strategic trade models*. In addition to the asymmetric information based models there is also the *synthetic model* that incorporates both the adverse selection and inventory/order handling cost (Aigbovo & Isibor, 2017). Seminal papers in the field of inventory-based models include Garman (1976), Stoll (1978), Amihud and Mendelson (1980), and (Ho & Stoll, 1981, 1983), Roll (1984), Hasbrouck (1991a), (Huang & Stoll, 1994, 1997), and Madhavan, Richardson, and Roomans (1997) among others. Two main classes of asymmetric information models are the *sequential* trade models and the strategic models. Seminal papers in the field of *sequential* trade models include Copeland and Galai (1983); Glosten and Milgrom (1985); Easley and O'hara (1987); O'Hara (2003); Easley et al. (1997) and Easley et al. (2002) among others and Seminal papers in the field of *strategic trade models* include Kyle (1985); Admati and Pfleiderer (1988); Foster and Viswanathan (1996), among others.

The main focus of this article is on the capital market microstructure. And the purpose of this article is a comprehensive overview of asymmetric information-based capital market microstructure models. Models that form the core of capital market microstructure theory. The rest of the paper is structure as follows. In section two we review the market microstructure models. Section three provides the summary and conclusion.

2. Market Microstructure models

2.1 The Rational Expectations Equilibrium Model

Competitive models of rational expectations are in fact the basis of market microstructure theory. The components of modern market microstructure models are called rational expectations models. The rational expectations model is consistent with standard models (Admati, 1985; Grossman & Stiglitz, 1980; Hellwig, 1980). The basis of rational expectation models has been developed based on various economic phenomena: acquisition of information in financial markets (Verrecchia, 1982), performance of information markets (Admati & Pfleiderer, 1986, 1987, 1990), learning multi-asset information (Admati, 1985), internal trading (Leland, 1992), supply and return on assets (Breon-Drish, 2015), and price volatility (Gao, Song, & Wang, 2013). The basic model of rational expectations follows the original one-period framework-based model (Grossman & Stiglitz, 1980).

The building blocks of the modern market microstructure models are so called Rational Expectation Models. These are static competitive models. The models are formulated as follows:

Period: Two period's $t = 0, 1$.

Asset: one risky asset pays a normal distributed dividend d in period $1, d \sim N\left(\overline{d}, \sigma^2\right)$ and it is traded at $t = 0$ at a price *p*. The supply of the risky asset is random $S > 0$.

Riskless rate: The riskless rate between periods 0 and 1 is *r*.

Traders: There are N agents in the market, with CARA (constant absolute risk aversion) utility and the risk aversion coefficient is α. Among them, N_I are informed and N_U are uninformed, where $N_I + N_U = N$. The informed observe a signal of the risky asset $S = d + \varepsilon$, ε is independent of d and $\varepsilon \sim N(0, \sigma_\varepsilon^2)$. The uninformed observe no signal.

Steps in the analysis of this model: (1) derive distribution of asset's payoff conditional on public and private information. (2) Find demand schedule of traders by utility maximization. (3) Find equilibrium price by equating aggregate supply and demand.

Equilibrium price reveals some of I's private information If U is smart (rational), he will take this information into account in his decisions.

Consider an uninformed agent with wealth W_0, if at $t = 0$ she buys x shares, her period 1 wealth is:

$$W = (W_0 - xp)(1 + r) + xd$$

And her expected utility is:

$$-E\ exp(-\alpha W)$$

This form of utility (CARA) has the good property that if $y \sim N(\mu, \sigma^2)$, then

$$E\ exp(y) = exp\left(\mu + \frac{1}{2}\sigma^2\right) \tag{3}$$

The expected utility is therefore:

$$-exp\left[-\alpha\left((W_0 - xp)(1+r) + x\overline{d} - \frac{\alpha}{2}\sigma^2 x^2\right)\right] \tag{4}$$

Thus, the uninformed can be viewed as maximizing:

$$max_x\left((W_0 - xp)(1+r) + x\overline{d} - \frac{\alpha}{2}\sigma^2 x^2\right) \tag{5}$$

This is a concave function and by solving F.O.C, we have the uninformed demand:

$$x_U(p) = \frac{\overline{d} - (1+r)p}{\alpha\sigma^2} \tag{6}$$

Here the price of the risky asset enters through budget constraint and the initial wealth is not in this demand.
The informed face a conditional probability, suppose (x, y) are jointly normal, then conditional on x, y is normal and

$$E(y|x) = E(y) + \big(x - E(x)\big)\frac{cov(x,y)}{\sigma_x^2} \tag{7}$$

$$\sigma_{y/x}^2 = \sigma_y^2 - \frac{cov(x,y)^2}{\sigma_x^2} \tag{8}$$

So in this model, we have

$$E(d|s) = \overline{d} + \big(s - \overline{d}\big)\frac{\sigma^2}{\sigma^2 + \sigma_\epsilon^2} \tag{9}$$

$$\sigma_{d/s}^2 = \frac{\sigma^2\sigma_\epsilon^2}{\sigma^2 + \sigma_\epsilon^2} \tag{10}$$

Hence, for the informed, the expected utility is:

$$-exp\left[-\alpha\left((W_0 - xp)(1+r) + xE(d|s) - \frac{\alpha}{2}\sigma_{d/s}^2 x^2\right)\right] \tag{11}$$

And therefore

$$x_I(p) = \frac{E(d|s) - (1+r)p}{\alpha\sigma_{d/s}^2} \tag{12}$$

By equating demand with supply, $N_I x_I(p) + N_U x_U(p) = S$ we solve for price:

$$p = \frac{\overline{d}}{1+r} + N_I k\left(1 - \frac{\sigma_{d/s}^2}{\sigma^2}\right)\big(s - \overline{d}\big) - S\alpha k\sigma_{d/s}^2 \tag{13}$$

$$k = \frac{1}{(1+r)\left(N_I + N_U\frac{\sigma_{d/s}^2}{\sigma^2}\right)} \tag{14}$$

In this model, the price fully reveals the signal of the informed (price is linear in signal). For the uninformed, they could use price to infer informed signal.

2.2 The Kyle [1985] Model

A seminal strategic model is studied in Kyle (1985). The Kyle model is a model of a batch-auction market, in which market makers see the order imbalance an each auction date. And market makers compete to fill the order imbalance, and matching orders are executed at market clearing prices. Unlike the sequential trade model, the strategic informed agent could trade at multiple times. Kyle develops the optimal trading behavior for the informed trader and shows that the agent will trade on his information only gradually, rather than exploit it to the maximum extent possible.
In another market microstructure context that consider when trader has superior information to derived the security price or how the prices adjust to the full information value is introduced in Kyle (1985). This model proposes that the new equilibrium price partially reflects the new full information with three types of market participants: market makers, noise traders, and informed traders. Kyle proposed a single period model which at the value of an asset is a random variable: $S \sim N(P_0, \sigma_0^2)$, at the end of period asset value of $\overline{S} \sim N(P_0, \Sigma_0)$.

There are also noise or uninformed traders who trade for exogenous reasons and submit a market order for $\overline{U}$ quantity, where $\overline{U} \sim N(0, \sigma_u^2)$. Also, Kyle assumes that at the beginning of trading period informed trader can get information of the price of the asset from the value of S, and he knows the value of the asset at the end of the period is equal to $\overline{S}$ but does not know the quantity $\overline{U}$; however, an informed trader chooses to submit a market order to maximize their profit at x quantity. At this point, the market maker observes the net order flow $y = u + x$ and sets a price p, however, the market maker cannot distinguish which part of the order comes from noise or informed traders. At the equilibrium, the market maker earns a zero expected profit, so he takes a position $-(u+x)$ to clear the market and earn $-\big(\overline{S} - P\big)(u+x)$. The market maker sets p following the function of $(u+x)$, thus $P(x+u) = E\big(\overline{S}\big|u+x\big)$.

The informed trader chooses to submit quantity x that depends on $\overline{S}$ to maximize his profit $\overline{\pi}$ at the end of a period. Thus, the market maker sets the equilibrium price as:

$$MAX_x E(\overline{\pi}|\overline{S}) = MAX_x E\left(\left(\overline{S} - P(x+\overline{u})\right)x\middle|\overline{S}\right)$$

This model assumes at the equilibrium that the market maker price is a linear function given by the posterior expectation;

$P(y) = \mu + \lambda y.$

So, the maximize profit of the informed traders

$$E(\pi) = MAX_x E\left(\left(\overline{S} - P(x+\overline{u})\right)x\middle|\overline{S} = S\right)$$

$$= MAX_x E\left((S - \mu - \lambda\,(x-\overline{u}))x\middle|S\right)$$

$$= MAX_x (S - \mu - \lambda\,\mathrm{x})x,$$

When the last step follows from the fact that $E(\overline{u}) = 0$.

Then, maximizing the expected profit, the solution to informed trader for optimal trade is

$$x = argmax_x E(\pi) = \frac{S-u}{2\lambda} = \alpha + \beta v$$

$$where\ \alpha = -\frac{u}{2\lambda} and\ \beta = \frac{1}{2\lambda}, if\ \lambda > 0.$$

To solve the linear market maker pricing and trade order parameters, we assume that the market maker sets the market price in order to earn zero profits at $P = E(\overline{S}|u+x)$. The expected order flow is $E(y) = \overline{u} + x = \overline{u} + \alpha + \beta\overline{S}$, where S and y are jointly normally distributed, that is, $E(\overline{S}|y)$. Therefore, the market maker should follow the maximum likelihood estimator to optimal pricing rule that equals to $\mu + \lambda y$ where μ and y minimize, $min\, E\left(\left(\overline{S} - P(y)\right)^2\right) = min_{\mu,\lambda} E\left(\left(\overline{S} - \mu - \lambda y\right)^2\right) = min_{\mu,\lambda} E\left(\left(\overline{S} - \mu - \lambda\left(\overline{u} + \alpha + \beta\overline{S}\right)\right)^2\right) = min_{\mu,\lambda} E\left(\left(\overline{S}(1-\alpha\beta) - \alpha\overline{u} - \mu - \lambda\alpha\right)^2\right)$

Following assumptions $E(S) = P_0, E((S-P_0)^2) = \Sigma_0, E(u) = 0, E(u^2) = \sigma_u^2$ and $E(uS) = 0,$ then $min_{\mu,\lambda} E(1-\lambda\beta)^2(\Sigma_0 + P_0^2) + (\mu + \lambda\sigma)^2 + \lambda^2\sigma_u^2 - 2(\mu + \lambda\alpha)(1-\lambda\beta)P_0$.

The first condition respect to μ and λ are $\mu = \lambda\alpha + P_0(1-\lambda\beta)$, thus

$$\lambda = \frac{\beta\Sigma_0}{\beta^2\Sigma_0 + \sigma_u^2}$$

$$Use\ \alpha = -\frac{\mu}{2\lambda} and\ \beta = \frac{1}{2\lambda}, we\ have\ \mu = P_0\ and\ \lambda = \frac{\sqrt{\Sigma_0}}{2\sigma_u}$$

At equilibrium, the market maker trade price is (15)

$$P = P_0 + \frac{\sqrt{\Sigma_0}}{2\sigma_u}(\overline{u} + \overline{x})$$

Where the informed order is

$$x = \frac{(\overline{S} - P_0)\sigma_u}{\sqrt{\Sigma_0}}$$

From equation 21, we can see that the informed order is greater or more active in the magnitude provided by the volatility of the order from uninformed traderσ_u. Substituted 21 into 20, we get $P = P_0 + \frac{\sqrt{\Sigma_0}\overline{u}}{2\sigma_u} + \frac{(\overline{S}+P_0)}{2}$. Thus, only one half of private information $\frac{1}{2}v^2$ is reflected to p, therefore the equilibrium price is not fully reveled by informed trader's information. The expected profit of the informed trader, unconditional on knowing the value of $\overline{S}$ at the beginning of trading period is (16)

$$E(\overline{\pi}) = \frac{\sigma_u(\overline{S} - P_0)^2}{2\sqrt{\Sigma_0}} \quad (17)$$

Since the market maker sets the trade price in condition to earn zero profit, the expected gain for informed traders is the expected loss from noise traders, not the market maker. As the expected profit from informed order is a linear in noise volatility, this can be assumed that informed trader hide their order with the orders from uninformed trader to hide the position. Consider the illiquidity parameter $\lambda = \frac{\sqrt{\Sigma_0}}{2\sigma_u}$ which presents the value that the market maker raises the price when the net order flow $y = u + x$ increases by one unit. Therefore, the $\lambda y = \sqrt{\Sigma_0}\frac{y}{2\sigma_u}$ is liquidity risk scaled by volatility of security, and $\frac{y}{\sigma_u}$ is similar to the percentage of volume. Hence, the amount of order that raises the price by 1 dollar equals $\frac{1}{\lambda}$ which is measured by the market depth or market liquidity. Intuitively, the greater number of noise traders, the greater profits of informed traders that gain from the loss of uniformed traders. However, with a greater number of uninformed traders, an individual loss is less.

(18)

2.3 The Glosten and Milgrom [1985] Model

In Glosten-Milgrom model, orders arrive and are executed by a market maker individually. The arrival rates of *informed* and *uninformed* trader are determined exogenously. Informed traders trade when chosen by this mechanism as if they have no future opportunities to trade. In other words, when trade is profitable, they trade as much as possible.

Consider one security valued at $V \in \{V_h, V_l\}$, with $Pr(V_l) = \delta$. The value is revealed at the end of trade. There are two types of traders: the informed *I* and the uninformed *U*, the proportion of informed traders among the population is μ. The market maker posts bid and ask quotes, *B* and *A*. A trader is randomly drawn from the population. If the trader is informed, he buys if $V = V_h$, sells if $V = V_l$. If the trader is uninformed, she buys or sells randomly with equal probability. The market maker does not know the types of the trader. A buy is a purchase by the trader at the dealer's asks price, *a*; a sell is a trading at the bid, *B*. We assume that the competition among dealers drives the expected profit to zero. The market maker's inference given that the first trade is a buy or a sell can be summarized by his posterior belief about the low outcome.

Let $p_k(buy)$, or $p_k(sell)$ $k = 1, 2, \ldots$ denote the probability of a low outcome given the k_{th} trade is a buy (or a sell). p_0 Is the unconditional probability being a low outcome, which is δ. Let B_k denote k_{th} order is buy, S_k denote k_{th} order is sell. Then the market maker's posterior belief of a low outcome after the first trade is buy is,

$$P_1(buy) = Pr(V_l|B_1) = \frac{Pr(V_l, buy)}{Pr(buy)} = \frac{\delta(1-\mu)}{1+\mu(1-2\delta)}$$

And dealer's expectation of the value given first buy order is $E(V|B_1) = Pr(V_l|buy)V_l + \left(1 - Pr\left((V_l)\middle|buy\right)\right)V_h$. If competition drives the expected profit to zero, then the posted "ask price" is the dealer's expected value.

$$A = E(V|B_1) = \frac{\delta(1-\mu)V_l + (1-\delta)(1+\mu)V_h}{1-\mu(1-2\delta)}$$

The bid price is similar, followed by a sell to the dealer. The dealer saw the first trader is a sell order and post the bid price.

$$P_1(sell) = Pr(V_l|s_1) = \frac{Pr(V_l, sell)}{Pr(sell)} = \frac{\delta(1+\mu)}{1-\mu(1-2\delta)}$$

$$B = E(V|S_1) = \frac{\delta(1+\mu)V_l + (1-\delta)(1-\mu)V_h}{1+\mu(1-2\delta)}$$

The bid-ask spread is:

$$S = A - B = \frac{4(1-\delta)\delta(V_h - V_l)\mu}{1-\mu^2(1-2\delta)^2}$$

The dealer updates his belief and posts new quotes on each trade sequentially. This process repeats for $k = 1, 2,$ this updating procedure could be expressed in general forms since all probabilities in the event trees are constant except $p_k(0)$.

$$P_k\left(buy|P_{k-1}(0)\right) = \frac{P_{k-1}(0)(1-\mu)}{1+\mu\left(1-2P_{k-1}(0)\right)}$$

$$P_k\left(sell|P_{k-1}(0)\right) = \frac{P_{k-1}(0)(1+\mu)}{1-\mu\left(1-2P_{k-1}(0)\right)}$$

It can be shown that $p_k\left(buy\ l p_{k-1}(buy), p_{k-2}(sell)\right) = p_k\left(buy\ l p_{k-1}(sell), p_{k-2}(buy)\right)$, for all *k*. The arrival sequence of the buy or sell orders does not matter. Therefore the proportion of buy or sell orders is deterministic to the outcome.

The conditional expectation of the ask can be decomposed as

$$A = E(V|buy) = E(V|U, buy)Pr(U|buy) + E(V|I, buy)Pr(I|buy)$$

Rearranging terms gives

$$\left(A - E(V|U, buy)\right)Pr(U|buy) = -\left(A - E(V|I, buy)\right)Pr(I|buy) \quad (23)$$

In this model, the economic interpretation for equation (31) is that the gain from an uninformed trader on the left side is equal to the loss to the informed trader on the right side (subject to zero profit expectation for the market maker). There is net wealth transfer from the uninformed to the informed.

Although the trader is independently drawn from both population for order execution, one subset of the population (the informed) always trade in the same direction. The result is that orders are serially correlated. (24)

One important economic justification of G-M model is trades update the price. For any security at k_{th} given trade, a buy order on the $(k+1)_{th}$ trade will make a upward revision in the conditional probability of a high outcome, and consequently increase both ask and bid quotes and drive trading price upward. In contrast, a sell order will drive price downward. The trade price impact is a particular useful empirical implication.

In the Roll model, we denote $\{q_t\}$ as the trade direction variable (+1 buy, -1 sell) with equal probability. In the G-M model, the order flow has no equal probability attributes to asymmetric information processed by difference traders, the informed traders always trade in the direction of his knowledge.

The asymmetric information in the G-M model is μ, the proportion of the informed trader in the population. In equation (29) and (27), the asymmetric information parameter μ is positively related to $p_k(sell)$, and the bid-asked spread. The justification behind is when the market have more informed traders, a sell order will be more likely submitted by an informed trader instead of a uninformed, the probability of a low outcome after sell is high; similarly, the probability of a high outcome given buy order is also high. In consequence, the dealer will post wider bid-ask spread in response to the change of posterior beliefs. These results suggest use of the bid-ask spread or the impact of an order has on subsequent prices as proxies for the asymmetric information. We have more discussions in the empirical study.

The limitation of G-M model is the informed traders are drawn randomly by the market mechanism. When she is selected, she will trade once and the maximum (one unit of order). There are no trading strategies for the informed trader to maximize her profit. The order execution timing and order sizes are two important aspects to the informed in empirical work while remain unaddressed in G-M model.

2.4 The Easley et al. [1996] Model - *Probability of Informed Trading (PIN)*

In their model, Easley, Kiefer, O'hara, et al. (1996) assume that individuals trade a risky asset and money over $i = 1, \ldots, N$ days, where time is indexed by t for each trading day and is considered continuous. The market makers quotes the bid and ask prices at which investors buy and sell securities and is considered to be risk-neutral, so the prices are the expected value of the stock, conditional on the information the dealer has when making the transaction. In the current methodology, there are two types of traders-informed traders who benefit from signals that give the true value of a stock and uninformed traders who receive no signals on the future movement of a stock's price. Both groups of traders enter the market following independent Poisson processes at any minute during the trading day; however, in the case of informed traders, they receive good-news signals that encourage them to enter the market. When such signals reveal themselves, the traders buy the stock. Easley, Kiefer, O'hara, et al. (1996) define information events α which can occur both with probability δ for good-news events and $1 - \delta$ for bad-news events whenever the nature of the news determines from the sample at the beginning of each day whether an event with informational impact on the fundamental value of the asset will appear. In this scenario, $(V_i)_{i=1}^{N}$ represents a random variable which gives us the fundamental value of the asset at the end of every day in the sample. The value of the asset on a day with good news is given by the random variable denoted by V_i, while $\underline{V_i}$ denotes the value of the asset on a bad-news day. If the value of the asset on a day with no information is denoted V_i^*, we will have the inequality $\underline{V_i} < V_i^* < \overline{V}_i$ The model is developed using simple binomial logic. The event generating news at the beginning of each day can be of either a good-news or bad-news type, and the appearance of informed traders that are competitive and risk-neutral does not depend on the nature of the news. The arrival of the news to one trader at a certain moment in time and that trader's actions in the market follow a Poisson process with the arrival rate denoted by μ. We note that all arrival processes are assumed to be independent in the Easley, Kiefer, O'hara, et al. (1996) framework. On days when good-news events are generated (through an independent Poisson process), the arrival rates are given by $\varepsilon + \mu$ for buy orders and ε for sell orders. On days when bad-news events are generated, the arrival rates are given by μ for buy orders and $\varepsilon + \mu$ for sell orders. If, at the beginning of the tree, there is no news-generating event, then only uninformed traders take part in the process, with an arrival rate equal to ε. At the end of the day, the market maker has complete information on each actor in the market and quotes the true value of the stock. The market maker has information on the probabilities for the above events and on the information arriving in the market. The occurrence of such events is unknown. Easley, Kiefer, O'hara, et al. (1996) assume that the dealer is Bayesian in the sense that his information is being updated with the arrival of new trade orders. The information is treated independently across days-therefore, each day is treated as a different observation in computing the probability of informed trading. Based on this fact, if we denote $P(t) = (P_n(t), P_b(t), P_g(t))$ as the market maker's prior beliefs regarding information events at the beginning of each day, when $t = 0$, we will have: $P(0) = (1-\alpha, \alpha\delta, \alpha(1-\delta))$.

If we denote $P(t|S_t)$ as the market maker's updated belief vector that takes into account the history of trades and quotes prior to time *t*, by using the Bayes rule the posterior probability of no news at time *t*, if an order to sell arrives at *t*, is:

$$P_n(t|S_t) = \frac{P_n(t)\varepsilon}{\varepsilon + P_b(t)\mu} \quad (32)$$

In a similar way, the probability of bad news will be given by:

$$P_b(t|S_t) = \frac{P_b(t)(\varepsilon+\mu)}{\varepsilon + P_b(t)\mu} \quad (33)$$

And the probability of good news is:

$$P_g(t|S_t) = \frac{P_g(t)\varepsilon}{\varepsilon + P_b(t)\mu} \quad (34)$$

If we take into account (1), (2) and (3) and the zero-profit hypothesis, the expected bid-price denoted *b*(*t*) at any time *t* on day *I*, is:

$$b(t) = \frac{P_n(t)\varepsilon V_i^* + P_b(t)(\varepsilon+\mu)\underline{V_i} + P_g(t)\varepsilon\overline{V}_i}{\varepsilon + P_b(t)\mu} \quad (35)$$

Based on a similar calculation, the ask price at time *t* is:

$$a(t) = \frac{P_n(t)\varepsilon V_i^* + P_b(t)\varepsilon\underline{V_i} + P_g(\varepsilon+\mu)\varepsilon\overline{V}_i}{\varepsilon + P_b(t)\mu} \quad (36)$$

The expected value of the asset, $E(V_i|t)$, based on the values that we set at the beginning of this section is a function that depends on each probability that we computed in the first three equations. So, $E(V_i|t)$ is:

$$E(V_i|t) = P_n(t)V_i^* + P_b(t)\underline{V_i} + P_g(t)\overline{V}_i \quad (37)$$

Based on relation (6), the values of the bid and ask prices that market makers calculated based on prior information until time t on day i are given by:

$$b(t) = E(V_i|t) - \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)}\left(E(V_i|t) - \underline{V_i}\right)$$

And

$$a(t) = E(V_i|t) - \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)}\left(\overline{V}_i - E(V_i|t)\right)$$

Let $\sum(t) = a(t) + b(t)$ be the spread at time t. Then, in order to identify the factors that are influencing the spread, we can write $\sum(t)$ as:

$$\sum(t) = \frac{\mu P_g(t)}{\varepsilon + \mu P_g(t)}\left(\overline{V}_i - E(V_i|t)\right) + \frac{\mu P_b(t)}{\varepsilon + \mu P_b(t)}\left(E(V_i|t) - \underline{V_i}\right)$$

According to Easley, Kiefer, O'hara, et al. (1996), the spread at time t represents information based probability multiplied by the expected loss to informed buyers plus a symmetric term for sells. So the probability of informed trading represents the sum of the aforementioned probabilities, explicitly:

$$PI(t) = \frac{\mu(1-P_n(t))}{\mu(1-P_n(t)) + 2\varepsilon} \quad (41)$$

When the market opens, at t = 0, if we assume that good and bad news occurs with the same probability, then the spread can be computed as:

$$\sum(0) = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon}\left(\overline{V}_i - \underline{V_i}\right) \quad (42)$$

In what follows, we will give an overview of the analytical and empirical implementation of the above model.

On a day with a bad-news event, the observed sequence of buy and sell trades has the following probability:

$$P(B,S) = e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}e^{-(\varepsilon+\mu)}\frac{((\mu+\varepsilon)T)^S}{S!} \quad (43)$$

On a day with no information-revealing events, the probability becomes:

$$P(B,S) = e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} \quad (44)$$

On a day with a good-news event, the probability is:

$$P(B,S) = e^{-(\mu+\varepsilon)T}\frac{((\mu+\varepsilon)T)^B}{B!}e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} \quad (45)$$

We write the likelihood of trading activity, which is independent across days:

$$L((B,S|\theta)) = (1-\alpha)e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} + \alpha\delta e^{-\varepsilon T}\frac{(\varepsilon T)^B}{B!}e^{-(\varepsilon+\mu)}\frac{((\varepsilon+\mu)T)^S}{S!} + \alpha(1-\delta)e^{-(\varepsilon+\mu)}\frac{((\varepsilon+\mu)T)^B}{B!}e^{-\varepsilon T}\frac{(\varepsilon T)^S}{S!} \quad (46)$$

The parameter space is given by $\theta = \{\alpha, \delta, \varepsilon, \mu\}$.
The likelihood of observing the data *(i=1to h)* am given by the product of daily likelihoods:

$$L(M|H) = \prod_{i=1}^{h} L(\theta|B_i, S_i) \quad (47)$$

The above function is rearranged as per Aktas, De Bodt, Declerck, and Van Oppens (2007) and Easley, Engle, O'Hara, and Wu (2008). Subsequently, the probability of uninformed trading is the unconditional probability that traders buy or sell assets at any point in time t. The higher this probability is, the higher the risk uninformed traders face in their actions of buying or selling stocks.
We write the probability of informed trading as:

$$PIN_t = \frac{\alpha\mu}{\alpha\mu + 2\varepsilon} \quad (48)$$

2.5 The Easley et al. [2012] Model – *Volume - Synchronized Probability of Informed Trading (VPIN)*

The main link between Probability of Informed Trading (PIN) and Volume-Synchronized Probability of Informed Trading (VPIN) was created in 2008 by Easley et al. (2008) and in 2010 the VPIN model was introduced by Easley, de Prado, and O'Hara (2010). The study of Volume-Synchronized Probability of Informed Trading (VPIN) presents the impact of high frequency trading (HFT) on order flows. Easley, López de Prado, and O'Hara (2012) Introduce the concept of "order flow toxicity" to represent the adverse selection risk in HFT context. They state that the market makers might not be aware that they provide liquidity at a loss, and order flow is toxic when it has adverse selection on these market makers.

To measure order flow toxicity, Easley et al. (2012) impute order imbalances through a monotone function of the absolute price changes to gauge the probability of information-based trading on the basis of the volume imbalance and the trade intensity, and use the BV-VPIN (bulk volume classification procedure BV-VPIN) metric to forecast the market volatility induced by toxicity. The inner algorithm is that market makers face the prospect of losses due to adverse selection when order flows become imbalanced. Hence, the estimates of time-varying toxicity level become a crucial factor in determining the participation of market makers. If they believe that toxicity is high, they will liquidate their positions and leave the market.
VPIN model is actually PIN estimation in high frequency conditions. This method is used on a time-volume basis. In this method, the time-volume range and bulk volume must be specified. Since all bulk volume are the same size, V,

$$\frac{1}{n}\sum_{\tau=1}^{n}(V_\tau^B + V_\tau^S) = V = \alpha\mu + 2\varepsilon \quad (49)$$

In this regard, n is the number of volume bucket used to estimate the VPIN. In this method, volume bucket need to be divided into buy volume and sale volume. Instead of using *Tick-rule, Lee-Ready* or other transaction classification techniques, they use a new method of volume classification. In the name of classification, they offer volume bucket. In this way, first the volume is classified in bucket and then part of the volume is classified as buy volume and the rest as sale volume. In a volume bucket, the amount of volume classified as buy is equal to:

$$V_\tau^B = \sum_{i=t(\tau-1)+1}^{t(\tau)} V_i Z\left(\frac{S_i - S_{i-1}}{\sigma\Delta s}\right) \quad (50)$$

In this regard, $t(\tau)$ is an indicator of the last bar (volume or time) in the volume bucket T, V_τ^B represents the volume of buy (transaction with Ask), $V_i$represents the total volume in each bulk, Z represents the standard normal distribution and $\sigma\Delta s$ Indicates the standard deviation of price changes between bars (volume or time). Since all bulk have the same amount of volume V,

$$V_\tau^S = \sum_{i=t(\tau-1)+1}^{t\tau} V_i \left(1 - Z\left(\frac{S_i - S_{i-1}}{\sigma\Delta s}\right)\right) = V - V_\tau^B \tag{51}$$

Since $E[|V^S - V^B|] \approx \alpha\mu$ and $PIN = \frac{\alpha\mu}{\alpha\mu+2\varepsilon}$ so we can say that VPIN is an approximation of PIN.

$$VPIN = \frac{\sum_{\tau=1}^{n}|V_\tau^S - V_\tau^B|}{nV} \tag{52}$$

A brief overview of the three steps of calculating VPIN (Abad, Cutillas-Gomariz, Sánchez-Ballesta, & Yagüe, 2018):

(1) Time bars: The original procedure begins with trade aggregation in time bars. Bar size is the first key variable of the VPIN computation process. Easley et al. (2012) initially use 1- minute time bars. In each time bar, trades are aggregated by adding the volume of all the trades in the bar (if any) and by computing the price change for this period of time. Afterwards, and in order to take into account trade size, the sample is "expanded" by repeating each bar price change as many times as the volume in the bar. Thus, the original raw sample became a sample of one-unit trades, each of them associated with the price change of the corresponding bar.

(2) Volume buckets, bulk classification and order imbalance: Volume bucket is the second essential variable in VPIN metric. Volume buckets represent pieces of homogeneous information content that are used to compute order imbalances. In Easley et al. (2012) volume bucket size (VBS) is calculated by dividing the average daily volume (in shares) by 50, which is the number of buckets they initially consider. Therefore, if we depart from the average daily volume, it is the number of buckets which fully determine VBS. Buckets are filled by adding the volume in consecutive time bars until completing the VBS. If the volume of the last time bar needed to complete a bucket is for a size greater than required, the excess size is given to the next bucket. In general, a volume bucket needs a certain number of time bars to be completed although it is also possible that the volume in a time bar could be enough to fill one (or more) volume buckets. At the same time of bucket completion, time bar volume is classified as buyer-or seller-initiated in probabilistic terms. Normal distribution is employed labeling as "buy" the volume that results from multiplying the volume bar by the value of the normal distribution evaluated in the standardized price change $Z(\Delta P|\sigma_{\Delta P})$. To standardize, we divide the corresponding price change by the standard deviation of all price changes for the whole sample. Analogously, we categorize as "sell" the volume that results from multiplying the volume bar by the complementary of the normal distribution for the buy side. $1 - Z(\Delta P|\sigma_{\Delta P})$. Order imbalance (OI) is then computed for each bucket by simply obtaining the absolute value of the difference between buy volume and sell volume in the assigned time bars.

(3) VPIN and sample length: Finally, in the last step we obtain VPIN values. To do that, it is necessary to define a new variable: sample length (n). This variable establishes the number of the buckets with which VPIN is computed. where VPIN is simply the average of order imbalances in the sample length, that is, the result of dividing the sum of order imbalances for all the buckets in the sample length (proxy of the expected trade imbalance) by the product of volume bucket size (VBS) multiplied by the sample length (n) (proxy for the expected total number of trades). VPIN metric is updated after each volume bucket in a rolling-window process. For example, if the sample length is 50, when bucket #51 is filled, we drop bucket #1 and we calculate the new VPIN based on buckets #2 to #51. (Easley et al. (2012)) first consider sample length equal to the number of buckets (50).

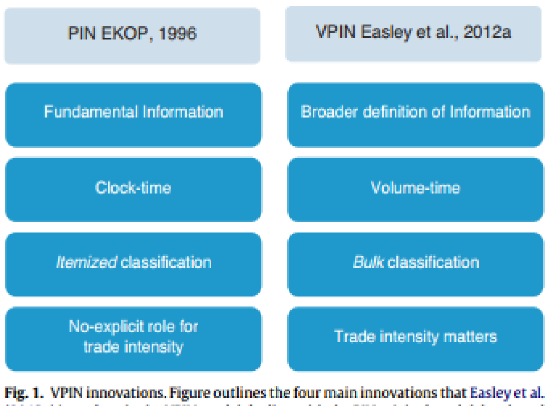

Fig. 1. VPIN innovations. Figure outlines the four main innovations that Easley et al. (2012a) introduce in the VPIN model dealing with the PIN original model developed by EKOP (1996).

3. Conclusion

This is a summary of the existing models related to the market microstructure. Traditional microstructure theory provides two major directions to explain price setting behavior: *inventory models and asymmetric information based models.* The inventory models investigates the uncertainty in order flow and the inventory risk and optimization problem of liquidity suppliers under possible risk aversion while the asymmetric information based models, model market dynamics and adjustment processes of prices using insights from the theory of asymmetric information and adverse selection. The two main approaches in asymmetric information models are *sequential trade models* and *strategic trade models*. In addition to the asymmetric information based models there is also the *synthetic model* that incorporates both the adverse selection and inventory/order handling cost. According to the above study, market microstructure models show the process of pricing by buyers and sellers. These models show the process by which the actual trading process affects the formation of price and trading volume in the market. Microstructural models differ from traditional financial models in recognizing that legitimate information about corporate principles may be unequally distributed and interpreted differently among market participants. Therefore, it can no longer be assumed that prices immediately reflect information, even if all participants are reasonable. The microstructure literature argues that both the risk of information resulting from asymmetric information and the difference in liquidity over time and between firms affect the price of long-term equilibrium in the market. We hope this study will help participants understand the pricing process and how information risk has led to information asymmetry and how the liquidity difference over time and between firms affects the long-run equilibrium price in the stock market. In fact, market microstructure studies are important in order to provide approaches to assist investors in designing investment strategies and stock market stakeholders and policy makers in order to formulate rules and trading mechanisms. It is a safe prediction that in the years to come, the trading environment for Iran equities will undergo further significant change. These changes, and developments in other markets, will provide ample opportunities for microstructure researchers to make further progress in our understanding of the effect of market structure on market quality.

References

1) Abad, D., Cutillas-Gomariz, M. F., Sánchez-Ballesta, J. P., & Yagüe, J. (2018). Real earnings management and information asymmetry in the equity market. *European Accounting Review, 27*(2), 209-235.
2) Admati, A. R. (1985). A noisy rational expectations equilibrium for multi-asset securities markets. *Econometrica: Journal of the Econometric Society*, 629-657.
3) Admati, A. R., & Pfleiderer, P. (1986). A monopolistic market for information. *Journal of Economic Theory, 39*(2), 400-438.
4) Admati, A. R., & Pfleiderer, P. (1987). Viable allocations of information in financial markets. *Journal of Economic Theory, 43*(1), 76-115.
5) Admati, A. R., & Pfleiderer, P. (1988). A theory of intraday patterns: Volume and price variability. *The review of financial studies, 1*(1), 3-40.
6) Admati, A. R., & Pfleiderer, P. (1990). Direct and indirect sale of information. *Econometrica: Journal of the Econometric Society*, 901-928.
7) Aigbovo, O., & Isibor, B. (2017). Market Microstructure: A Review of Literature. *Research Journal of Financial Sustainable Reporting, 2*(2), 116-143.
8) Aktas, N., De Bodt, E., Declerck, F., & Van Oppens, H. (2007). The PIN anomaly around M&A announcements. *Journal of Financial Markets, 10*(2), 169-191.
9) Amihud, Y., & Mendelson, H. (1980). Dealership market: Market-making with inventory. *Journal of financial economics, 8*(1), 31-53.
10) Asmar, M., & Ahmad, Z. (2011). Market microstructure: The components of black-box. *International Journal of Economics and Finance, 3*(1), 152-159.
11) Bagehot, W. (1971). The only game in town. *Financial Analysts Journal, 27*(2), 12-14.
12) Breon-Drish, B. (2015). On existence and uniqueness of equilibrium in a class of noisy rational expectations models. *The Review of Economic Studies, 82*(3), 868-921.
13) Copeland, T. E., & Galai, D. (1983). Information effects on the bid-ask spread. *the Journal of Finance, 38*(5), 1457-1469.

14) Easley, D., de Prado, M. L., & O'Hara, M. (2010). Measuring flow toxicity in a high frequency world. *Unpublished Working paper, Cornell University and Tudor Investment Corp*.
15) Easley, D., Engle, R. F., O'Hara, M., & Wu, L. (2008). Time-varying arrival rates of informed and uninformed trades. *Journal of Financial Econometrics*, *6*(2), 171-207.
16) Easley, D., Hvidkjaer, S., & O'hara, M. (2002). Is information risk a determinant of asset returns? *The journal of finance*, *57*(5), 2185-2221.
17) Easley, D., Kiefer, N. M., & O'hara, M. (1996). Cream-skimming or profit-sharing? The curious role of purchased order flow. *The Journal of Finance, 51*(3), 811-833.
18) Easley, D., Kiefer, N. M., & O'Hara, M. (1997). One day in the life of a very common stock. *The Review of Financial Studies, 10*(3), 805-835.
19) Easley, D., Kiefer, N. M., O'hara, M., & Paperman, J. B. (1996). Liquidity, information, and infrequently traded stocks. *The Journal of Finance, 51*(4), 1405-1436.
20) Easley, D., López de Prado, M. M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high-frequency world. *The Review of Financial Studies, 25*(5), 1457-1493.
21) Easley, D., & O'hara, M. (1987). Price, trade size, and information in securities markets. *Journal of Financial economics, 19*(1), 69-90.
22) Easley, D., & O'hara, M. (1992). Time and the process of security price adjustment. *The Journal of finance, 47*(2), 577-605.
23) Easley, D., & O'hara, M. (2004). Information and the cost of capital. *The journal of finance, 59*(4), 1553-1583.
24) Easley, D., O'hara, M., & Saar, G. (2001). How stock splits affect trading: A microstructure approach. *Journal of Financial and Quantitative analysis*, *36*(1), 25-51.
25) Foster, F. D., & Viswanathan, S. (1996). Strategic trading when agents forecast the forecasts of others. *The Journal of Finance, 51*(4), 1437-1478.
26) Gao, F., Song, F., & Wang, J. (2013). Rational expectations equilibrium with uncertain proportion of informed traders. *Journal of Financial Markets, 16*(3), 387-413.
27) Garman, M. B. (1976). Market microstructure. *Journal of financial Economics, 3*(3), 257-275.
28) Glosten, L. R., & Milgrom, P. R. (1985). Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*, *14*(1), 71-100.
29) Grossman, S. J., & Stiglitz, J. E. (1980). On the impossibility of informationally efficient markets. *The American economic review*, *70*(3), 393-408.
30) Hasbrouck, J. (1991a). Measuring the information content of stock trades. *The Journal of Finance, 46*(1), 179-207.
31) Hasbrouck, J. (1991b). The summary informativeness of stock trades: An econometric analysis. *The Review of Financial Studies, 4*(3), 571-595.
32) Hellwig, M. F. (1980). On the aggregation of information in competitive markets. *Journal of economic theory, 22*(3), 477-498.
33) Ho, T. S., & Stoll, H. R. (1981). Optimal dealer pricing under transactions and return uncertainty. *Journal of Financial economics, 9*(1), 47-73.
34) Ho, T. S., & Stoll, H. R. (1983). The dynamics of dealer markets under competition. *The Journal of finance*, *38*(4), 1053-1074.
35) Huang, R. D., & Stoll, H. R. (1994). Market microstructure and stock return predictions. *The Review of Financial Studies, 7*(1), 179-213.
36) Huang, R. D., & Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *The Review of Financial Studies, 10*(4), 995-1034.
37) Jaffe, J. F., & Winkler, R. L. (1976). Optimal speculation against an efficient market. *The Journal of Finance*, *31*(1), 49-61.
38) Jones, C. M., Kaul, G., & Lipson, M. L. (1994). Information, trading, and volatility. *Journal of Financial Economics, 36*(1), 127-154.
39) Keim, D. B., & Madhavan, A. (1995). Anatomy of the trading process empirical evidence on the behavior of institutional traders. *Journal of Financial Economics*, *37*(3), 371-398.
40) Keim, D. B., & Madhavan, A. (1997). Transactions costs and investment style: an inter-exchange analysis of institutional equity trades. *Journal of Financial Economics, 46*(3), 265-292.
41) Kyle, A. S. (1985). Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*, 1315-1335.

42) Leland, H. E. (1992). Insider trading: Should it be prohibited? *Journal of Political Economy, 100*(4), 859-887.
43) Madhavan, A. (2000). Market microstructure: A survey. *Journal of financial markets, 3*(3), 205-258.
44) Madhavan, A., Richardson, M., & Roomans, M. (1997). Why do security prices change? A transaction-level analysis of NYSE stocks. *The Review of Financial Studies, 10*(4), 1035-1064.
45) Madhavan, A., & Smidt, S. (1991). A Bayesian model of intraday specialist pricing. *Journal of Financial Economics, 30*(1), 99-134.
46) McInish, T. H., & Wood, R. A. (1992). An analysis of intraday patterns in bid/ask spreads for NYSE stocks. *the Journal of Finance, 47*(2), 753-764.
47) O'Hara, M. (2003). Presidential address: Liquidity and price discovery. *The journal of Finance, 58*(4), 1335-1354.
48) O’Hara, M. (1995). Market Microstructure TheoryBlackwell. *Cambridge, MA.*
49) Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *The Journal of finance, 39*(4), 1127-1139.
50) Spulber, D. F. (1996). Market microstructure and intermediation. *Journal of Economic perspectives, 10*(3), 135-152.
51) Stoll, H. R. (1978). The supply of dealer services in securities markets. *The Journal of Finance, 33*(4), 1133-1151.
52) Verrecchia, R. E. (1982). Information acquisition in a noisy rational expectations economy. *Econometrica: Journal of the Econometric Society*, 1415-1430.