

Cardiff Economics Working Papers



Working Paper No. E2013/7

The dynamics of trading duration, volume and price
volatility – a vector MEM model

Yongdeng Xu

April 2013

Cardiff Business School
Aberconway Building
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
business.cardiff.ac.uk

This working paper is produced for discussion purpose only. These working papers are expected to be published in due course, in revised form, and should not be quoted or cited without the author's written permission.

Cardiff Economics Working Papers are available online from:

econpapers.repec.org/paper/cdfwpaper/ and

business.cardiff.ac.uk/research/academic-sections/economics/working-papers

Enquiries: EconWP@cardiff.ac.uk

The dynamics of trading duration, volume and price volatility – a vector MEM model

Yongdeng Xu*

12/12/2012

Abstract

We propose a general form of vector Multiplicative Error Model (MEM) for the dynamics of duration, volume and price volatility. The vector MEM relaxes the two restrictions often imposed by previous empirical work in market microstructure research, by allowing interdependence among the variables and relaxing weak exogeneity restrictions. We further propose a multivariate lognormal distribution for the vector MEM. The model is applied to the trade and quote data from the New York Stock Exchange (NYSE). The empirical results show that the vector MEM captures the dynamics of the trivariate system successfully. We find that times of greater activity or trades with larger size coincide with a higher number of informed traders present in the market. But we highlight that it is unexpected component of trading duration or trading volume that carry the information content. Moreover, our empirical results also suggest a significant feedback effect from price process to trading intensity, while the persistent quote changes and transient quote changes affect trading intensity in different direction, confirming Hasbrouck (1988,1991).

JEL Classification: C15, C32, C52

Keywords: Vector MEM, ACD, GARCH, intraday trading process, duration, volume, volatility

* Economic Section, Cardiff Business School, Cardiff University, CF10 3EU, UK. Email: xuy16@cardiff.ac.uk.

1. Introduction

Microstructure theory generally indicates that trading duration¹ and trading volume convey information with respect to fundamental asset prices, and reflect the behaviour of financial market participants.² Since French and Roll (1986) have found evidence that price volatility is caused by private information that affects prices when informed investors trade, the empirical studies on trade and price processes have been based on increasingly on the analysis of the dynamics of trading duration, volume and price volatility. However, prior research on this issue is based on a recursive framework, in which the trade and price processes are independent of each other.

In this paper, we extend the recently developed recursive framework of Engle (2000) and Manganelli (2005) for high frequency data to a vector MEM model in which the trading duration, volume and price volatility are involved simultaneously and are interdependent. We further propose a multivariate lognormal for the distribution of the vector model, which allows the innovation terms to be correlated contemporaneously. In addition, maximum likelihood is proposed as a suitable estimation strategy. In this way, we can build a system that incorporates various causal and feedback effects among these variables. We also construct impulse response functions that show how the price reacts to a perturbation of its long-run equilibrium. The method is applied to a trade and quote dataset of the NYSE, and the model is estimated using a sample of ten stocks.

Our empirical results are generally consistent with the previous findings in the empirical microstructure literature (see, for example, Dufour and Engle (2000), Engle (2000) and Manganelli (2005)). But our work is novel in two ways. First, we find that duration and duration shocks have a significant impact on price volatility, while only the unexpected components of volume are considered to carry information content

¹ Duration is defined as the time that elapses between two consecutive transactions.

² In general, duration is considered to reflect the trading strategy of informed traders or is an indicator of liquidity (Easley and O'Hara 1992), while volume is viewed as an important determinant of the strength of a market move and reflects information about changes in investors' expectations (Harris and Raviv, 1993).

with respect to price. This generally suggests that it is the unexpected components of trading characteristics rather than the trading variables themselves that carry information content with respect to fundamental asset prices. In addition, impulse response analysis shows that shocks to duration or volume are incorporated appropriately into the price within one trading day for frequently traded stocks, but this takes up to one week for infrequently traded stocks. Second, our empirical results suggest that volatility has a negative impact on trading intensity, while volatility shock has a positive impact on trading intensity. We explain this by considering the persistent quote change (volatility) to be motivated by information based reason, and transient quote change (volatility shock) to be motivated by inventory based reason. The results confirm Hasbrouck (1988,1991)'s prediction that persistent quote changes (volatility) reduce trading intensity and transient quote changes increase trading intensity.

The remainder of this paper is organized as follows. Section 2 reviews the relevant literature; the theoretical and empirical work on the relationship of duration, volume and volatility are reviewed in this section. Section 3 outlines the empirical motivation and describes the model and methodology used in the analysis. Section 4 introduces the high frequency data and empirical results. Section 5 concludes the paper.

2. Literature Review

Theoretically, the market microstructure literature explains trading activity using two types of model: information based and inventory based models. Specifically, trading occurs either for information motivated or liquidity motivated reasons. Accordingly, predictions of the relations between duration, volume and price volatility differ. In empirical analysis, the operation of the market is customarily undertaken by using time-series, high-frequency data. The dynamics of such positive-valued variables is generally modelled by a type of autoregressive conditional duration (ACD) model (Engle and Russell, 1998). In this section, we first review the relevant market microstructure theory and its prediction of the relations between duration, volume and volatility. And then the ACD model of the relevant empirical findings on these relationships is reviewed.

2.1 Review of market microstructure theory

In the information-based model, three types of traders are assumed: informed traders; uninformed traders; and market makers. Informed traders are usually defined as corporate officers with private information, while uninformed traders are liquidity motivated and simply behave according to their current information. Market makers are also assumed to be uninformed. Obviously, the different traders have asymmetric information. Informed traders hope to obtain profits from their information so, on average, the market makers lose out to the informed traders. Market makers are specialists and can access information by reading the signals in the market, such as trading intensity and volume, and can thus recoup any losses as uninformed traders. Their activities are covered by the sequential trade model (Diamond and Verrecchia, 1987; Glosten and Milgrom, 1985) and the strategic trade model ((Admati and Pfleiderer, 1988; Easley and O'Hara, 1992; Kyle, 1985).)

In the sequential trade framework, the market maker and market participants behave competitively. Trades take place sequentially, with only one trader allowed to transact at any given point in time. Informed traders would like to trade as much (and as often) as possible. So the market maker would quickly (perhaps instantly) adjust prices to reflect this information. It is obvious that trading volume is positively (perhaps contemporaneously) correlated with price volatility. The strategic model allows the agents to act strategically. For example, in order to make full use of their private information, the informed traders may conceal their trading type by timing their trades carefully or choosing their trade sizes (Easley and O'Hara, 1992; Kyle, 1985). Uninformed traders may also learn by observing the actions of informed traders. In particular, Admati and Pfleiderer (1988) distinguish two types of uninformed traders in addition to informed traders: non-discretionary traders are similar to liquidity traders in the previous model; while discretionary traders, while uninformed, trade strategically. Discretionary traders choose the timing of their trades. They usually select the same period of transaction in an attempt to minimize adverse selection costs, and informed traders follow the pattern introduced by discretionary traders.

In inventory based models, the trading process is effectively motivated by the market makers desire to keep their inventory position at some specific level. Based on their inventory position and uncertainty about order flow, dealers alter their bid and

ask prices to elicit the desired imbalance of buy and sell orders thereby moderating deviations in order flow. The dealer's action in the market is simply independent of information. It only depends on trading costs, the dealer's previous position and net demand to the dealer (Ho and Stoll, 1981; O'Hara and Oldfield, 1986).

These types of model generally induce patterns of various trade characteristics, such as timing, price and volume. These factors contain information and reflect trade behaviour in the market.

2.2 Prediction from market microstructure theory

Among the key variables considered, the timing of the trade plays an important role. It is ignored initially, and incorporated explicitly into market microstructure models by Diamond and Verrecchia (1987) and Easley and O'Hara (1992).

Diamond and Verrecchia (1987) use a rational expectations model with short-sale constraints. The informed traders' actions are driven by the arrival of private information, while uninformed traders are assumed to trade for reasons unrelated to the arrival of such information. If the news is bad, informed traders will wish to sell (or, alternatively, to short-sell if they do not own the stock). Given short-sale constraints, there may be no trade. Therefore, long durations are associated with bad news and should lead an adjustment of the prices and hence to increase the return volatility. This is summarized as 'No trade means bad news'.

Easley and O'Hara (1992) provide a different explanation for the role of time. Informed traders only trade when there is new information (whether good or bad) arriving in the market. So variations in trading intensity are closely related to the change in the participation rate of informed traders. It follows that short trade duration is a signal that informed traders are participating in the market. Consequently, the market maker adjusts his/her prices to reflect the increased risk of trading with informed traders, which reveals a higher volatility and wider bid-ask spreads in the market. To summarize, 'No trade means no news'. In the strategic trading assumption, the informed trader may choose to segment large volume trades into a greater number of smaller-volume, information-based trades, and hence conceal their type and make full use of private information. It follows that both trading intensity and trading volume may provide information concerning the behaviour of market participants.

A relationship between duration and volatility is also explained by the model of Admati and Pfleiderer (1988). It is assumed that frequent trading is associated with liquidity traders, and therefore low trading means that liquidity (discretionary) traders are inactive, which leaves a high proportion of informed traders in the market. This again translates into quick price adjustment and hence high volatility.

Goodhart and O'Hara (1997)) examine the price effect of trade. Traders may learn over time from the information-based model, and adjust their speed of trading in reaction to this. For example, a large change in a market maker's mid-quote price may be a signal to the informed traders that their private information has been revealed to the market makers, assuming that no new signal has been released subsequently. This means that private information is no longer superior, and therefore the incentive to trade disappears, which decreases trading intensity. However, from the inventory model perspective, large quote changes would immediately attract opposite-side traders, thus increasing trading intensity. In addition, when uninformed traders behave strategically (O'Hara, 1995), it becomes more complex, since the uninformed will increase the probability they attach to the risk of informed trading when they observe large absolute returns or large trading volume. Consequently, they will reduce the overall trading intensity. Hasbrouck (1988,1991) explains the two effects using the short-run and long-run characteristics of trading behaviour. The private information is persistent and long-lived; the persistent quote change is related to private information, and should have a negative impact on trading intensity. The inventory level in stationary and inventory control is inherently a transient concern, the transient quote change is related to inventory control, and has a positive impact on trading intensity.

Table 1 summarizes the related market microstructure literature and its predictions.

Table 1: Summary of the related market microstructure literature

Model		Authors and year	Main feature	Predictions
Information-based model	Sequential trade model	Glosten and Milgrom (1985)	All agents act competitively	Volume is positive correlated with volatility
		Diamond and Verrecchia (1987)	Short sale constraints Incorporating time	No trade means bad news (duration is correlated positively with volatility)
	Strategic trade model	Kyle (1985)	Informed traders act strategically Long-lived information	
		Easley and O'Hara (1992)	Incorporating time	No trade means no news (duration is correlated negatively with volatility)
		Admati and Pfleiderer (1988) Parlour (1998)	Uninformed traders also act strategically Short-lived information Rational expectations	Trade intensity increases, the informativeness of trades decreases Large quote change is a risk of informed trading; liquidity traders may leave or slow down trading activity
Inventory-based model	Ho and Stoll (1981) O'Hara and Oldfield (1986) Hasbrouck (1991)		Market makers use price to balance their inventory	Large quote changes attract opposite-side traders, thus increasing trading intensity

2.3 Empirical studies

Empirical investigation of market microstructure predictions is subject to the availability of high-frequency transaction data. Statistically speaking, high-frequency data are realizations of point processes; that is, the arrival of the observations is random. This, jointly with other unique features of financial data (long memory; strong skewness; and kurtosis) implies that new methods and new econometric models are needed. It was first addressed, by Engle and Russell (1998) in the context of an ACD model for the dynamics of transaction time. It represents the time duration as product of a (autoregressive) scale factor and non-negative valued random process.

In the ACD framework, the trade characteristics associated with time are incorporated and modelled simultaneously, so that the market microstructure predictions can be evaluated at the transaction level. Among them, Engle (2000) proposed a recursive framework to represent the dynamics of duration and volatility. The joint density of duration and volatility is expressed as the product of the marginal density of the duration times and the conditional density of the volatility, given the duration. The result provides evidence of the bad-news effect of long durations, which is the reverse of the Diamond and Verrecchia (1987) result. The recursive framework of Engle (2000) reduces the complexity of the model, since each process is estimated separately, and used widely by later empirical works. Engle and Sun (2007) model the joint density of the duration and the tick-by-tick returns within a recursive framework. They build an econometric model for estimating the volatility of the unobserved efficient price change. Using this model, it is easy to forecast the volatility of returns over an arbitrary time interval through simulation using all the observations available. Taylor (2004) models future market trading duration using various augmentations of the basic ACD model, and confirms that bid–ask spread and transaction volume have a significant impact on the subsequent trading intensity.

Manganelli (2005) notes that other high-frequency data (trading volume, bid–ask spread) share similar characteristics to duration (for example, they are positive-valued and persistently clustered over time), so that their dynamics can be represented using the same autoregressive process. He incorporates the trading volume into Engle (2000)’s model and develops a framework to model jointly duration, volume and price volatility. Following Engle (2000), the joint distribution of duration, volume and

volatility is decomposed into the product of the marginal distribution of duration; the marginal distribution of volume, given duration; and the conditional distribution of volatility, given duration and volume. Further assumptions of weak exogeneity are made, such as that the three processes are independent so they can be estimated separately. Manganelli (2005) studies the causal and feedback effects among the three variables and found that times of greater activity coincided with a larger fraction of informed traders being present in the market. However, his empirical results suggest that lagged volatility increases trading intensity, which is in contrast to Easley and O'Hara (1992), but confirms the inventory based model predictions that large returns attract opposite side traders and increase trading intensity.

Grammig and Wellner (2002) noticed that duration and volatility might be interdependent. They have extended Engle (2000)'s recursive model by formulating an interdependent intraday duration and volatility model. In this model, conditional volatility and intraday duration evolve simultaneously. The conditional volatility is formulated as a generalized autoregressive conditional heteroskedasticity (GARCH) process, with time-varying parameters that are functions of the expected interday duration. Their empirical results show that lagged volatility significantly reduces transaction intensity, which is consistent with Easley and O'Hara (1992). Hautsch (2008) analyses the return volatility, trade size and trading duration under the Multivariate Error Model (MEM)³ framework. Rather than using transaction data, Hautsch (2008) uses the cumulated five-minute data and focuses on the study of the underlying common factors that jointly drive the trading processes. He finds that the common factor captures most causal relations and cross-dependencies between the individual variables. The existence of common factors is also an indicator of the interdependence of the three processes.

In addition to the ACD framework, the vector autoregressive (VAR) model is used in the study of high frequency data. For example, Bowe et al. (2009) used a trivariate VAR model to analyse the interrelationship between trading volume, duration and price volatility, which is similar to Dufour and Engle (2000). But it is also similar to the recursive model and assumes that trade and price processes are cross-independent. Using the data from an emerging futures market, they find that

³ MEM is an extension of ACD model.

duration is affected positively by volatility, which is consistent with Diamond and Verrecchia (1987).

To summarize the empirical studies, the recursive frameworks are generally adopted for the analysis of high frequency data, but this is challenged by some empirical evidence. The empirical results with respect to the relations of trade and price process as are partially contradictory and there is no uniform conclusion at present.

3. Methodology

In this section, we first specify the dynamics of duration volume and price volatility according to the Engle (2000) and Manganelli (2005) recursive framework and discuss the statistic and economic concerns with this framework. We then extend the recursive framework of Engle (2000) and Manganelli (2005) to a vector specification in which trading duration, volume and price volatility evolve simultaneously and are interdependent.

3.1 Duration, volume and price volatility --- a recursive framework

Define $\{d_t, v_t, r_t\}, t=1, \dots, T$ as the three-dimensional time series associated with intraday trading duration, trading volume and the return process, respectively. In particular, duration is defined as the time elapsing between consecutive trades, volume is the trade size associated with each transaction and return is measured as the mid-quote change. The trivariate trading process - duration, volume and return volatility - can be modelled as follows:

$$\{d_t, v_t, r_t\} \sim f(d_t, v_t, r_t | \mathcal{F}_{t-1}; \theta) \quad (1)$$

where \mathcal{F}_{t-1} denotes the information available up to period $t-1$, and θ is a vector incorporating the parameters of interest.

In the recursive model (Manganelli, 2005), the joint distribution is decomposed into the product of three components: marginal density of durations, the conditional density of volumes given durations and the conditional density of the return volatility given durations and volumes. Specially,

$$\{d_t, v_t, r_t\} \sim g(d_t | \mathcal{F}_{t-1}; \theta_d) \cdot h(v_t | d_t, \mathcal{F}_{t-1}; \theta_v) \cdot k(r_t | d_t, v_t, \mathcal{F}_{t-1}; \theta_r). \quad (2)$$

For the dynamics of such a non-negative valued financial point process, Engle and Russell (1998) first propose an ACD specification for financial duration. They model duration as the product of its conditional expectation and the non-negative supported innovation term,

$$d_t = \psi_t(\theta_d; \mathcal{F}_{t-1})u_t, \quad u_t \sim i.i.d.(1, \sigma_u^2). \quad (3)$$

The ACD model is further characterized by the assumptions that the conditional duration ψ_t follows a GARCH-type process and the innovations are independently and identically distributed. The base (1,1) specification of ψ_t is:

$$\psi_t = \omega + \alpha d_{t-1} + \beta \psi_{t-1}. \quad (4)$$

The logarithmic version is also specified (Bauwens and Giot, 2000) to ensure positivity of the conditional duration,

$$\log \psi_t = \omega + \alpha \log d_{t-1} + \beta \log \psi_{t-1}. \quad (5)$$

To close the model, the parametric density function for the innovations is needed. Engle and Russell (1998) initially consider the exponential and Weibull distribution, which is extended later by Grammig and Maurer (2000), Allen et al. (2009) and Xu (2011a), offering more flexible density and hazard functions.

Following the ACD model, Manganelli (2005) considers similar specifications for volume and volatility. Then the trivariate system has the following specifications:

$$\begin{aligned} d_t &= \psi_t(\theta_d; \mathcal{F}_{t-1})u_t, \quad u_t \sim i.i.d.(1, \sigma_u^2) \\ v_t &= \varphi_t(\theta_v; d_t, \mathcal{F}_{t-1})\eta_t, \quad \eta_t \sim i.i.d.(1, \sigma_\eta^2) \\ \hat{r}_t &= \sqrt{h_t(\theta_r; d_t, v_t, \mathcal{F}_{t-1})}\xi_t, \quad \xi_t \sim i.i.d.(0,1) \\ \text{or } \hat{r}_t^2 &= h_t(\theta_r; d_t, v_t, \mathcal{F}_{t-1})\xi_t, \quad \xi_t \sim i.i.d.(1, \sigma_\xi^2) \end{aligned} \quad (6)$$

where \hat{r}_t^2 is the proxy for volatility⁴, (ψ_t, ϕ_t, h_t) are the conditional expectations of duration, volume and volatility, respectively, and $\theta = (\theta_1, \theta_2, \dots, \theta_s)$ is a vector of s parameters of interest. Manganelli (2005) considers the univariate exponential distribution for the innovations in this specification.

To capture the causal and feedback effect among these variables, he specifies the following first order autoregressive conditional model:

⁴ In order to obtain a price change sequence which is free of the bid-ask bounce that affects price, we follow Ghysels, et al. (1998) and \hat{r}_t is obtained as the residuals of an ARMA(1,1) process of return series. See also in Hautsch (2008). One advantage of using \hat{r}_t is that it avoids the problem of exact zero values in r_t .

$$\begin{aligned}
\psi_t &= w_1 + (a_{11}d_{t-1} + a_{12}v_{t-1} + a_{13}\hat{r}_{t-1}^2) + (b_{11}\psi_{t-1} + b_{12}\phi_{t-1} + b_{13}h_{t-1}), \\
\phi_t &= w_2 + (a_{21}d_{t-1} + a_{22}v_{t-1} + a_{23}\hat{r}_{t-1}^2) + (b_{21}\psi_{t-1} + b_{22}\phi_{t-1} + b_{23}h_{t-1}) + a_0^{12}d_t, \\
h_t &= w_3 + (a_{31}d_{t-1} + a_{32}v_{t-1} + a_{33}\hat{r}_{t-1}^2) + (b_{31}\psi_{t-1} + b_{32}\phi_{t-1} + b_{33}h_{t-1}) + a_0^{13}d_t + a_0^{23}v_t.
\end{aligned} \tag{7}$$

Under the restrictions of weak exogeneity ($b_{ij} = 0$ for $i = j$) and independence of the innovations terms, the three components are estimated separately. This approach is generally adopted in the existing empirical literature (see, for example, Engle (2000), Dufour and Engle (2000), Manganelli (2005) and Engle and Sun (2007)).

3.2 Econometric concerns

Following Manganelli (2005), there are two concerns regarding the recursive model. First, it assumes that the specific processes are independent. To incorporate the contemporaneous information, Manganelli (2005) specifies causality from duration to volume and from duration and volume to price volatility. However, modelling the distribution of price as being conditional on duration and volume is just one strategy to obtain their joint distribution. As pointed out by Engle and Sun (2007), it is also possible to go from the price process and model duration conditional on its contemporaneous return. Theoretically, variation in duration and variation in the price process would be related to the same news events or the underlying information process. Empirical studies also address this issue. For example, Hautsch (2008) finds the existence of a common unobserved component that jointly drives the dynamics of the trade and price processes. This common component explains most of the causality between the trade and the price processes, even if the contemporaneous effect of the trade variable on the price variable is controlled. We tested this restriction in our previous paper (Xu, 2011b) and show the existence of cross-dependence between the trading and price process. Therefore, the advisable approach is to allow the innovation terms to be contemporaneous correlated, and specify a vector form for the dynamics of the trivariate system.

Second, Manganelli (2005) assumes weak exogeneity, which means the conditional expectation of one variable is a function only of its own past conditional expectation, while the past conditional expectations of other variables are not taken into consideration. This strategy has been adopted by most empirical microstructure papers (see, for example, Dufour and Engle (2000)). However, we argue that this assumption is too restrictive. When studying the price impact of trade, various

specifications of duration and volume should be considered. For example, trade innovation is an exclusive a manifestation of the private information of the informed trader. Engle (2000) and Wuensche et al. (2007) argue that it is the unexpected components of the trade process that carry informational content with respect to the fundamental asset price, since price change is unpredictable. And the same happens for the feedback effects from price to trading intensity. For example, Grammig and Wellner (2002) find that expected volatility and volatility shocks have significant effects on trading intensity. Manganelli (2005) conducts a robustness test on this restriction. Specifically, he regresses the residuals of the three equations against past conditional expectations of other variables. The results indicate that the coefficients of past expected variables are almost never significant, and thus the recursive model is correctly specified. However, the robustness check might be misleading, since the dynamics of expected variables have been distorted when estimating and predicting the expected variables using recursive models. It is also shown by Grammig and Maurer (2000) in a simulation study that the misspecification of the conditional mean has severe consequences for the expectation of conditional duration.

We therefore extend the recursive model into a vector form, by allowing the three processes to be interdependent and relaxing weak exogeneity.

3.3 Vector MEM

Let $x_t = (d_t, v_t, r_t)'$, $\mu_t = (\psi_t, \phi_t, h_t)'$ and $\varepsilon_t = (u_t, \eta_t, \xi_t)'$. Following Cipollini et al. (2007); Engle (2002), we write this system of equations as a trivariate vector multiplicative error model (MEM). The three-dimensional vector MEM for duration, volume and volatility is:

$$x_t = \mu_t \odot \varepsilon_t \quad (8)$$

where \odot denotes the Hadmard (element by element) product; the components of ε_t are process-specific innovation terms which are assumed to be cross-dependent; and ε_t has a mean vector I with all components unity and a general variance-covariance matrix Σ , i.e., $\varepsilon_t | \mathcal{F}_{t-1} \sim D(I, \Sigma)$. The multivariate specification for μ_t is:

$$\mu_t = \omega + \sum_{l=1}^p A_l x_{t-l} + \sum_{l=1}^q B_l \mu_{t-l} + A_0 z_t \quad (9)$$

where z_t is a vector of predetermined variables.

We do not specify recursively the contemporaneous relationship from duration to volume and from duration and volume to volatility (Manganelli, 2005). However, we allow the innovation terms to be contemporaneously correlated. By this specification, the conditional expectation of one variable is a function not only of its own past conditional expectation, but also of past conditional expectations of other variables. The two restrictions imposed by the recursive model are released.

The mean equation is further extended to be a logarithmic version to ensure the positivity of the individual processes without imposing additional parameter restrictions.

$$\ln(\mu_t) = \omega + \sum_{l=1}^p A_l \ln(x_{t-l}) + \sum_{l=1}^q B_l \ln(\mu_{t-l}) + A_0 \ln(z_t) \quad (10)$$

The first two moment conditions of the vector MEM are given by:

$$\begin{aligned} E(x_t | \mathcal{F}_{t-1}) &= \mu_t \\ \text{Var}(x_t | \mathcal{F}_{t-1}) &= \mu_t \mu_t' \odot \Sigma = \text{diag}(\mu_t) \Sigma \text{diag}(\mu_t) \end{aligned} \quad (11)$$

which is a positive defined matrix by construction, as emphasized by Engle (2002).

3.4 Specification of \mathcal{E}_t

A completely parametric formulation of the vector MEM requires a full specification of the conditional distribution of \mathcal{E}_t . In the ACD literature, Engle and Russell (1998) initially consider the exponential and Weibull distribution for the error \mathcal{E}_t , which is extended later by Grammig and Maurer (2000) to be a Burr distribution, by Lunde (1999) to be a generalized gamma distribution, and recently by Allen et al. (2009) and Xu (2011a) to be a lognormal distribution. Figure 1 plots the comparison of density functions implied by these parametric distributions. It can be seen that only the exponential distribution implies a monotonically decreasing density function, while the others imply hump shaped density functions. Xu (2011a) tests the specification of the duration distributions, and finds that the lognormal ACD model is superior to the Exponential ACD and Weibull ACD models, while its performance is similar to the Burr or Generalized Gamma ACD models. It is well known that price volatility is typically lognormally distributed, while Andersen et al. (2001) and Cizeau et al. (1997), among others, also showed that the lognormal distribution fitted the

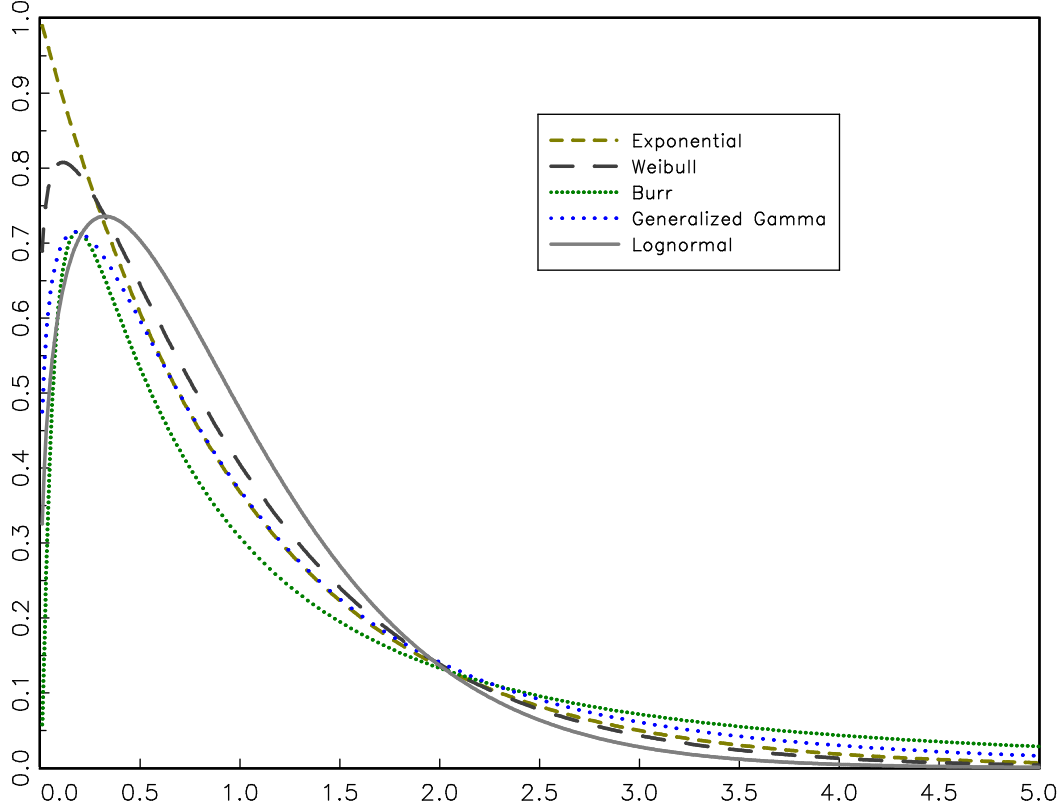


Figure 1: A comparison of parametric density

realized volatility distribution very well.

So we propose to use the multivariate lognormal distribution for the MEM. Indeed, the multivariate lognormal distribution seems to be the only feasible choice in the specification of vector MEM. It has a closed form conditional density function, so that ML estimation can be conducted. Cipollini et al. (2007) consider appropriate multivariate gamma versions but find that the only useful version admits only positive correlation, which is too restrictive. The multivariate lognormal distribution admits both positive and negative correlations. Moreover, Allen et al. (2008) prove that the lognormal distribution is sufficiently flexible to provide a good approximation to a wide range of non-negative distributions, and is also sufficiently accurate so as not to induce unnecessary numerical difficulties.

Assume ε_t follows a multivariate lognormal distribution such that $\varepsilon_t | \Omega_t \sim \ln N(\nu, D)$ ⁵. The density function is:

⁵ where $\nu_i = -\frac{1}{2}d_{ii}$ to guarantee that $E(\varepsilon_t | \mathcal{F}_{t-1}) = I$

$$f(\varepsilon_t | \mathcal{F}_{t-1}, D) = (2\pi)^{-K/2} |D|^{-1/2} \left(\prod_{i=1}^K \varepsilon_{i,t}^{-1} \exp \left(-\frac{1}{2} (\ln \varepsilon_t - \nu)' D^{-1} (\ln \varepsilon_t - \nu) \right) \right) \quad (12)$$

where $\varepsilon_t > 0$. The conditional density of x_t is then:

$$f(x_t | \mathcal{F}_{t-1}, \theta) = (2\pi)^{-K/2} |D|^{-1/2} \prod_{i=1}^K x_{i,t}^{-1} \exp \left(-\frac{1}{2} (\ln x_t - \ln \mu_t - \nu)' D^{-1} (\ln x_t - \ln \mu_t - \nu) \right). \quad (13)$$

The log likelihood of the model is then:

$$l = \sum_{t=1}^T l_t = \sum_{t=1}^T \ln f(x_t | \mathcal{F}_{t-1}, \theta) \quad (14)$$

where

$$l_t = \ln f(x_t | \mathcal{F}_{t-1}, \theta) = -\frac{K}{2} \ln(2\pi) - \frac{1}{2} |D| - \sum_{i=1}^K \ln(x_{i,t}) - \frac{1}{2} (\ln x_t - \ln \mu_t - \nu)' D^{-1} (\ln x_t - \ln \mu_t - \nu) \quad (15)$$

The first and second moments of the multivariate lognormal distribution are given by:

$$\begin{aligned} \tau &= E(\varepsilon) = (\tau_1, \tau_2, \dots, \tau_k)' & \tau_i &= e^{v_i + \frac{1}{2}d_{ii}} = 1 \\ \Sigma &= E(\varepsilon - \tau)(\varepsilon - \tau)' = \sigma_{ij} & \sigma_{ij} &= e^{(v_i + v_j + \frac{d_{ii} + d_{jj}}{2})} (e^{d_{ij}} - 1) = e^{d_{ij}} - 1 \\ \rho_{ij} &= \frac{e^{d_{ij}} - 1}{\sqrt{(e^{d_{ii}} - 1)(e^{d_{jj}} - 1)}} \end{aligned}$$

where $v_i = -\frac{1}{2}d_{ii}$ and d_{ij} is the ij th element of D . It is clear that if $(\ln \varepsilon_1, \ln \varepsilon_2, \dots, \ln \varepsilon_k)$ are independent, then $(\varepsilon_1, \varepsilon_2, \dots, \varepsilon_k)$ are also independent and vice versa. The multivariate lognormal distribution allows both positive and negative correlation, which is much more flexible than the multivariate gamma distribution (Cipollini et al., 2007).

The lognormal belongs to the exponential family. The parameters are still consistently estimated, even if the chosen density is wrong. The asymptotic distribution of the QML estimator differs from that of the ML estimator. The variance-covariance matrix is not the inverse of the Fisher information. It has the so-called ‘sandwich’ form.

$$\sqrt{N}(\hat{\theta}_{QML} - \theta) \rightarrow N(0, I^{-1}(\hat{\theta})J(\hat{\theta})I^{-1}(\hat{\theta})) \quad (16)$$

where $I(\hat{\theta}) = -E \left[\frac{\partial^2 \ln L(\hat{\theta}; x)}{\partial \hat{\theta} \partial \hat{\theta}'} \right]$, $J(\hat{\theta}) = E \left[\frac{\partial \ln L(\hat{\theta}; x)}{\partial \hat{\theta}} \left(\frac{\partial \ln L(\hat{\theta}; x)}{\partial \hat{\theta}} \right)' \right]$ are,

respectively, the components of the empirical average Hessian and the empirical average outer product of the gradients evaluated at the estimates $\hat{\theta}$.

3.5 Impulse response function

Following the vector MEM, we can derive the impulse response functions. We concentrate on the first order model and exclude the predetermined variables.

$$\begin{aligned} x_t &= \mu_t \odot \varepsilon_t, \\ \ln \mu_t &= \omega + A \ln x_{t-1} + B \ln \mu_{t-1}. \end{aligned} \quad (17)$$

In the impulse response, we work on the impulse of $\nu_0 = \ln \varepsilon_t$ on the natural logarithmic of the interested variable $\ln x_t$. The impulse responses function of the model (17) for $t > 0$ is⁶:

$$\frac{\partial \ln x_t}{\partial \nu_0} = \Phi_t \quad (18)$$

where $\Phi_t = (A + B)^{t-1}(A - B)$, $\Phi_0 = I$.

This process can be rewritten in such a way that the residuals of different equations are uncorrelated. For this purpose, we choose a decomposition of the white noise covariance matrix $\Sigma_\nu = W \Sigma_\tau W'$, where Σ_τ is a diagonal matrix with positive diagonal elements and W is a lower triangular matrix with unit diagonal. Thus,

$$\ln x_t = \bar{\omega} + \sum_{i=0}^{\infty} \Theta_i \tau_{t-i}, \quad \Theta_i = \Phi_i W^{-1}. \quad (19)$$

Then the impulse response function of the model (17) for $t > 0$ is:

$$\frac{\partial \ln x_t}{\partial \tau_0} = \Theta_t. \quad (20)$$

The standard errors for the impulse response are computed as followings. Let

$$\begin{aligned} \theta_{(p \times 1)} &= [\theta_d', \theta_v', \theta_r']' \quad \text{and} \quad \kappa_t \equiv \text{vec}(\Phi_t(\theta)) \quad . \quad \text{If} \quad \sqrt{T}(\hat{\theta} - \theta) \xrightarrow{a} N(0, Q) \quad , \quad \text{then} \\ \sqrt{T}(\hat{\kappa}_t - \kappa_t) &\xrightarrow{a} N(0, G_t Q G_t'), \quad \text{where} \quad G_t = \frac{\partial \kappa_t}{\partial \theta'}. \end{aligned}$$

⁶ See Appendix 3: Proofs of impulse response function

3.6 Vector ARMA representation

One of the advantages of using the lognormal distribution for the vector MEM model is that it has an equivalent Vector ARMA specification with an innovation that follows a multivariate Gaussian distribution.

From the following log vector MEM model,

$$x_t = \mu_t \odot \varepsilon_t, \quad (21)$$

$$\ln(\mu_t) = \omega + \sum_{l=1}^p A_l \ln(x_{t-l}) + \sum_{l=1}^q B_l \ln(\mu_{t-l}) + A_0 \ln(z_t). \quad (22)$$

If we take logs of (21), we obtain

$$\ln(x_t) = \ln(\mu_t) + \ln(\varepsilon_t) = c + \ln(\mu_t) + e_t \quad (23)$$

where $e_t | \mathcal{F}_{t-1} \sim iid N(0, \Pi)$.

Then,

$$\ln(\mu_t) = \ln(x_t) - c - e_t, \quad (24)$$

$$\sum_{l=1}^q B_l \ln(\mu_{t-l}) = \sum_{l=1}^q B_l \ln(x_{t-l}) + \sum_{l=1}^q B_l c - \sum_{l=1}^q B_l e_{t-l}. \quad (25)$$

Substituting $\ln(\mu_t)$ and $\sum_{l=1}^q B_l \ln(\mu_{t-l})$ into Equation (22), it follows that

$$\ln(x_t) = \bar{c} + \left(\sum_{l=1}^p A_l + \sum_{l=1}^q B_l \right) \ln(x_{t-l}) + e_t - \sum_{l=1}^q B_l e_{t-l} + A_0 \ln(z_t) \quad (26)$$

where $\bar{c} = c + \omega - \sum_{l=1}^q B_l c$.

It is interesting that the vector MEM model is equivalent to a VARMA specification. In particular, it provides a good way to adopt the VARMA inference⁷ to make inferences in the vector MEM model.

4. Empirical analysis

4.1 Data

We use the data from the Trades and Quotes (TAQ) dataset at NYSE. The TAQ data consists of two parts: the first reports the trade data, while the second lists the

⁷ See Appendix 2: Inference of VARMA Models

Table 2: Stocks used in the analysis

A. Frequently traded		B. Infrequently traded	
TRN	Trinity Industries	ABG	Group ABG
R	Ryder System Inc.	OFG	Oriental Finl Grp Hold Co.
ARG	Airgas Inc.	LSB	LSB Industries Inc.
FMO	Federal-Mogul Corp.	HTD	Huntingdon Life S.G.
VTs	VERITAS DGC INC	HUN	Hunt Corp.

quote data (bid and ask data) posted by the market maker. The data were kindly provided by Manganelli (2005). He constructed 10 deciles of stocks covering the period from Jan 1,1998 to June 30, 1999, on the basis of the 1997 total number of trades of all stocks quoted on the NYSE. We randomly selected 5 stocks from the eighth decile (frequently traded stocks) and 5 from the second decile (infrequently traded stocks) covering the period from Jan 1,1998 to June 30, 1999. The tickers and names of the ten stocks are reported in Table 2:

Before the analysis began, we adopted Manganelli (2005)'s strategy to prepare the data. First, all trades before 9:30 am or after 4:00 pm were discarded. Second, durations over night were computed as if the overnight periods did not exist. For example, the time elapsing between 15:59:50 and 9:30:05 of the following day is only 15 seconds. We keep overnight duration because our samples for infrequently traded stocks are very small. Eliminating this duration would cause the loss of important data for these stocks. Third, all transaction data with zero duration are eliminated. These transactions are treated as one single transaction, and the related volumes are summed. Fourth, to deal with the impact of dividend payments and trading halts, we simply deleted the first observation whose price incorporated the dividend payment or a trading halt. Fifth, to adjust the data for stock splits, we simply multiplied the price and volume by the stock split ratio. Sixth, the price of each transaction is calculated as the average of the prevailing bid and ask quote. To obtain the prevailing quotes, we use the 5 second rule used by Lee and Ready (1991) which links each trade to the quote posted at least 5 seconds before , since the quotes can be posted more quickly than trades are recorded. This procedure is standard in microstructure studies. Seventh, the returns were computed as the difference of the log of the prices. To obtain a return sequence that is free of the bid-ask bounce that affects prices (see Campbell et al., 1997, chapter 3), we follow Ghysels et al. (1998) in using the residuals of an ARMA(1,1) model estimated on the return data.

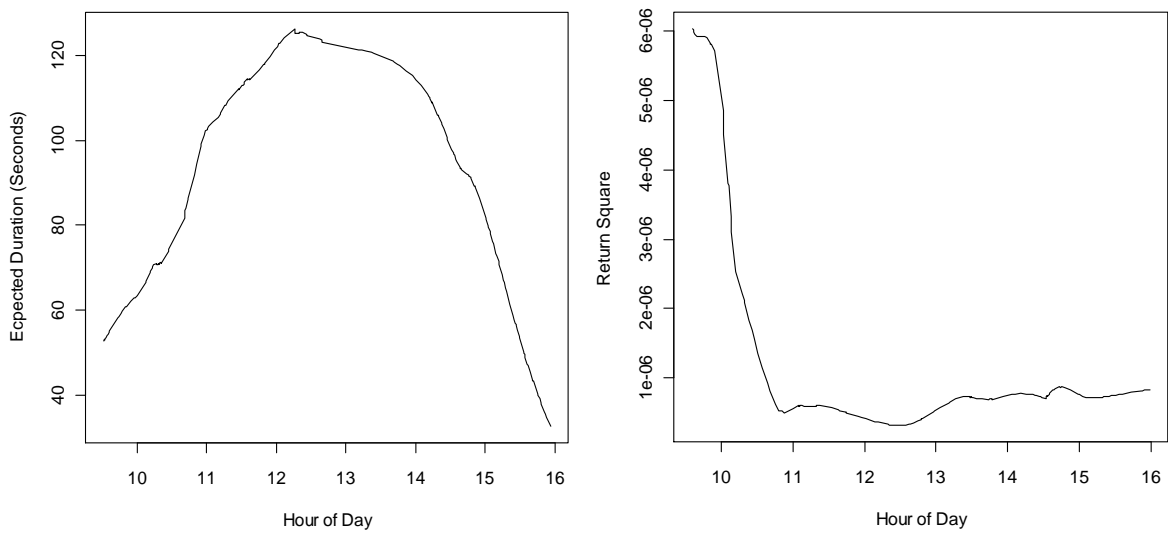


Figure 2: Nonparametric estimate of daily pattern of transaction duration and return square.

The second issue to be addressed prior to the analysis concerns the intraday pattern in the data. It is well known that duration, volume and volatility exhibit strong intraday periodic components, with a high trading activity at the beginning and end of the day. To adjust for this, we make use of a method used by Engle (2000). We regress the durations, volumes and returns squares on a piecewise cubic spline with knots at 9:30, 10:00, 11:00, 12:00, 13:00, 14:00, 15:00, 15:30 and 16:00. The original series are then divided by the spline forecast to obtain the adjusted series. Figure 2 depicts the nonparametric estimate of daily pattern of duration and return square for one typical stock ARG. Generally, less frequently traded stocks do not exhibit any regular intraday pattern. More frequently traded stocks typically show the inverted U pattern for duration, the L pattern for return squares, and no regular pattern for volume.

Table 3 presents some summary statistics for the ten stocks. For the frequently traded stocks, the number of observations range from 33,850 to 69,720 in the sample period, and the average trading duration ranges from 87 seconds to 259 seconds. For the infrequently traded stocks, the number of observation ranges from 2,074 to 7,212 in the sample period, with the average trading duration ranging from 1,215 seconds to 4,215 seconds. The trading volume does not show any difference between frequently traded stocks and infrequently traded stocks. The number of trading volumes ranges

Table 3: Summary statistics for the 10 stocks

	Obs	Mean		LB(20)			MLB(20)
		Duration	Volume	Duration	Volume	Variance	
TRN	55582	157.86	1369.43	3780.09	1383.35	3769.80	12744.02
GAS	101332	86.54	3118.12	5951.85	2338.08	4073.09	19049.05
TCB	55208	158.94	1855.20	4171.36	2644.11	2925.82	14716.45
R	69702	125.67	2492.98	14072.3	7276.91	23685.7	58049.96
ARG	33850	259.2	1280.70	3780.09	1383.35	3769.80	12744.02
ABG	2074	4214.88	5259.05	120.28	225.07	146.00	760.08
OFG	7212	1214.58	833.86	523.16	1343.43	738.09	3557.98
LSB	2962	2962.19	1971.61	481.41	435.69	523.58	2110.88
HNN	5887	1483.73	1070.02	2431.00	660.60	788.81	4564.73
JNS	3949	2215.94	2748.60	268.52	682.92	297.01	1571.99

Notes: LB(20) denotes Ljung–Box statistics for order 20. MLB(20) denotes multivariate Ljung–Box statistics.

from 833 to 5,295. The multivariate Ljung–Box statistics, computed according to Hosking (1980) and is given by

$$MLB(s) := n(n+2) \sum_{j=1}^s \frac{1}{n-j} \text{trace}(\hat{C}'_j \hat{C}_0^{-1} \hat{C}'_j \hat{C}_0^{-1}) \sim \chi^2(ks) \quad (27)$$

where k denotes the dimension of the process (in this case $k=3$), s is the number of lags taken into account, and \hat{C}_j is the j th residual autocovariance matrix. It is apparent that duration, volume and volatility show strong serial autocorrelations, and this is particularly true for high frequency traded stocks. The large multivariate Ljung-Box statistics in the table indicate that the trivariate system reveals strong dynamic and contemporaneous dependencies. These indicators suggest the use of vector form MEM.

We also depict the non-parametric density and parametric densities implied by the exponential and lognormal distributions.⁸ Figure 3 reports the comparison of parametric and non-parametric densities for one typical stock LSB. It can be seen that the lognormal distribution fits with the true density very well for the duration data. This result is consistent with Xu (2011a). For volume data, we are surprised to find the lognormal distribution has the best performance. And the raw data fluctuates closely around the lognormal distribution. Even for volatility, the lognormal

⁸ See Xu (2011a) and Grammig and Maurer (2000) for the discussion of parametric and non-parametric density.

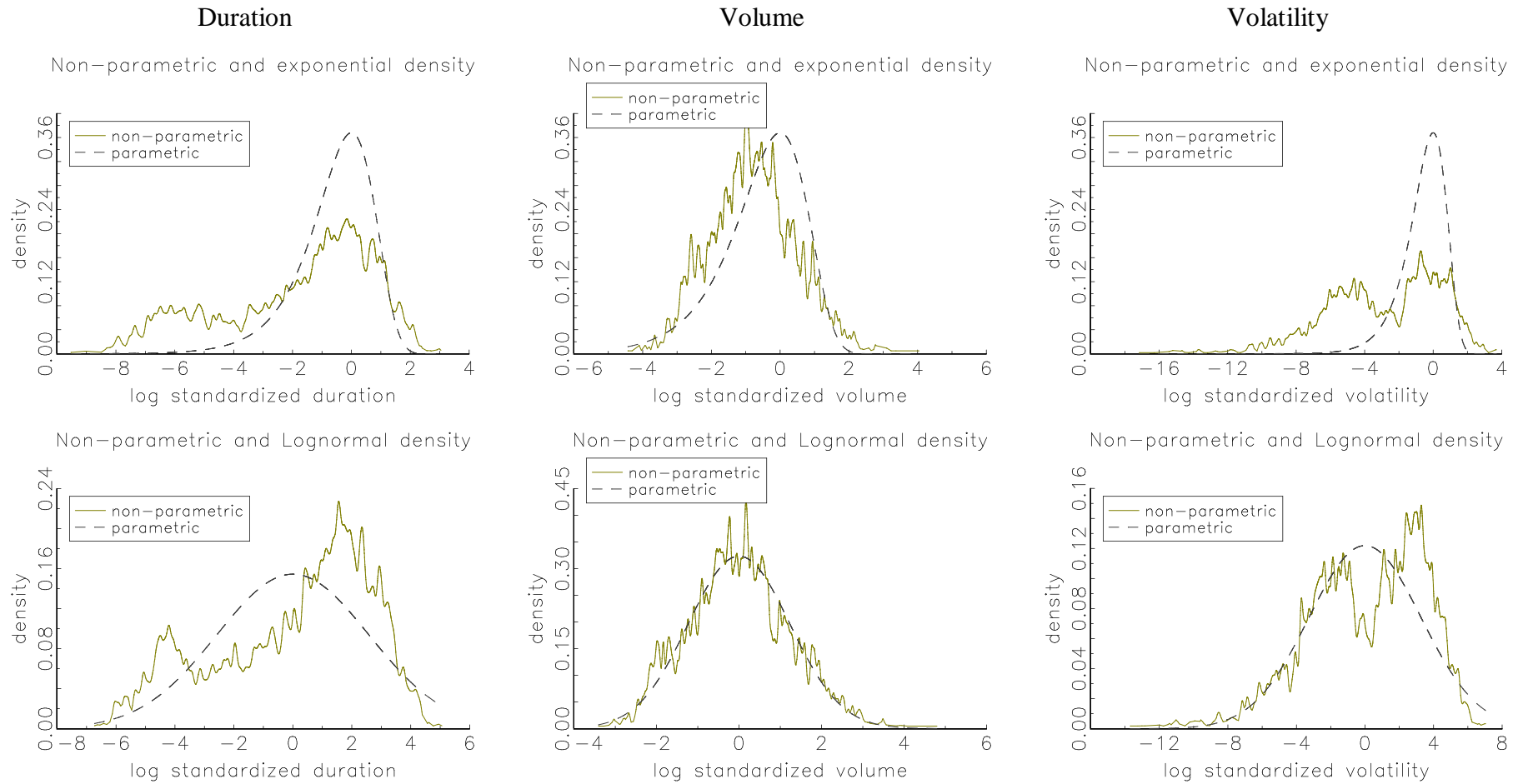


Figure 3: A comparison of parametric density and non-parametric densities--LSB

distribution also performs well. For brevity, the other 9 stocks have are not been reported for brevity, but these findings are robust across the stocks.

The data we use in this paper strongly support the multivariate lognormal MEM model for the dynamics of duration, volume and price volatility.

4.2 Empirical model

In the empirical analysis, we are interested in the causal and feedback effects among the variables. In contrast to the previous recursive model, we allow trade duration, volume and innovations of these variables to affect price volatility and vice versa: the volatility and volatility shocks are allowed to affect trading intensity. So we specify and estimate the following vector MEM:

$$\begin{aligned}
 x_t &= \mu_t \odot \varepsilon_t, \quad \varepsilon_t | \mathcal{F}_{t-1} \sim D(I, \Sigma) \\
 \ln \mu_t &= \omega + A \ln x_{t-1} + B \ln \mu_{t-1} + C \ln \frac{x_{t-1}}{\mu_{t-1}}
 \end{aligned} \tag{28}$$

where B is a diagonal matrix and C is a matrix where the diagonal elements are zero. Then, a_{31} (a_{32}) measures the impact of duration (volume) on price volatility, c_{31} (c_{32}) measures the impact of duration (volume) shocks on price volatility, a_{13} measures the impact of volatility on trading intensity and c_{13} measures the impact of volatility shocks on trading intensity. The estimation results and various diagnostics for the five frequently traded stocks are reported in Table 4 and results for the five infrequently traded stocks are reported in

Note: LL denotes Log likelihood function. BIC denotes Bayes Information Criterion. LB denotes Ljung-Box statistics of flitted residuals and MLB denotes. multivariate Ljung-Box statistic. The Ljung-Box statistics are computed based on 20 lags. Critical values $\chi^2(6)_{0.05}=12.59$, $\chi^2(6)_{0.01}=16.81$

Table 5.

Considering the diagnostic statistics of the model, these suggest that the vector MEM improves the dynamic properties of the model significantly, as we can see from the sharp drop in the Ljung-Box statistics. This is particularly true for the volatility process. Moreover, the vector MEM reduces the multivariate Ljung-Box statistic significantly, indicating that the vector MEM does a good job in capturing the multivariate dynamics and interdependencies between the individual processes. For frequently traded stocks, the dynamics of the system are still not captured completely by the model. But this is commonly the case with such large time series (see, for example, Engle (2000)). For infrequently traded stocks, the dynamics of the system are captured completely by the vector MEM.

Table 4: Estimation results and diagnostics: frequently traded stocks.

	ARG	TRN	TCB	GAS	R
α_{11}	0.060***	0.089	0.055***	0.062	0.064
α_{12}	0.107**	0.228***	0.122***	0.216***	0.168
α_{13}	0.012***	0.007***	0.025***	0.009	0.018
α_{21}	-0.067**	0.113**	-0.003**	0.121	0.018
α_{22}	0.125	0.124	0.098	0.118***	0.125
α_{23}	-0.009	-0.004	-0.009***	-0.007***	-0.011**
α_{31}	-0.337***	-0.204***	-0.387***	-0.065	0.071***
α_{32}	-0.109	0.219***	0.371***	-0.019	-0.434
α_{33}	0.389***	0.241***	0.278***	0.316***	0.195
b_{11}	0.939***	0.730***	0.942***	0.724***	0.912***
b_{22}	0.508***	0.638***	0.695***	0.706***	0.606**
b_{33}	0.239***	0.301***	0.075***	0.246***	0.629***
c_{12}	-0.171***	-0.331***	-0.202***	-0.338***	-0.265
c_{13}	-0.023***	-0.017***	-0.035***	-0.014***	-0.032
c_{21}	0.064	-0.127**	-0.013***	-0.132	-0.031
c_{23}	0.004	0.005**	0.008***	0.007***	0.009***
c_{31}	0.015	-0.084***	0.015***	-0.170	-0.414***
c_{32}	0.629***	0.353***	0.260***	0.474***	0.882**
LR test ⁹					
$H_0: c_{ij} = 0, i \neq j$	240	519	345	408	1587
Diagnostics					
<i>LL</i>	-221998	-358758	-365788	-260314	-407773
<i>BIC</i>	444247	717780	731838	520883	815812
<i>MLB</i>	565.8***	991.6**	1018***	684.3***	1086
<i>LB_d</i>	101.4***	36.23**	104.0***	52.82***	60.37***
<i>LB_v</i>	95.97***	184.1***	182.3***	83.37***	355.9***
<i>LB_ŕ²</i>	174.3***	308.7***	457.0***	219.9***	70.30***

Note: LL denotes Log likelihood function. BIC denotes Bayes Information Criterion. LB denotes Ljung-Box statistics of flitted residuals and MLB denotes. multivariate Ljung-Box statistic. The Ljung-Box statistics are computed based on 20 lags. Critical values $\chi^2(6)_{0.05}=12.59$, $\chi^2(6)_{0.01}=16.81$

⁹ We estimate five different vector MEMs for comparison. The results have not reported for brevity. LR test is based on the likelihood values of restricted and unrestricted models.

Table 5: Estimation results and diagnostics: infrequently traded stocks.

	ABG	HTD	LSB	HUN	FEP
α_{11}	0.042***	0.019**	0.032	0.056***	0.056***
α_{12}	-0.079***	0.015***	0.049	-0.002	0.016
α_{13}	-0.021***	0.018***	-0.004	0.013**	0.085
α_{21}	0.019***	-0.564***	-0.083	-0.005	-0.051
α_{22}	0.231***	0.198	0.166	0.133**	0.185***
α_{23}	-0.079***	-0.090**	-0.009	-0.025***	-0.155
α_{31}	-0.023***	-0.446	-0.145***	0.032	-0.062
α_{32}	-0.599***	-0.408***	-0.140	-0.374	-0.134
α_{33}	-0.079***	-0.090**	-0.009	-0.025***	-0.155
b_{11}	0.912***	0.980***	0.967***	0.932***	0.910***
b_{22}	0.366***	0.290***	0.569	0.708***	0.643***
b_{33}	0.682***	0.665***	0.522***	0.67***	0.318**
c_{12}	0.108***	-0.061***	-0.118	-0.013***	-0.103***
c_{13}	0.024***	-0.038***	0.006	-0.028**	-0.094
c_{21}	-0.034***	0.549***	0.087	0.000	0.042
c_{23}	0.079***	0.095**	0.013	0.028***	0.149***
c_{31}	-0.024***	0.370	0.028	-0.160***	-0.006***
c_{32}	0.919***	0.663***	0.389	0.721**	0.580**
LR test					
$H_0: c_{ij} = 0, i \neq j$	35.1	54.7	10.4	110	34.0
Diagnostics					
LL	-14392	-17116	-19370	-37243	-28574
BIC	28967	34420	38932	74695	57310
MLB	221.3**	191.3	249.3***	167.2	195.7
LB_d	36.47**	32.81**	48.37**	25.04	29.12
LB_v	22.17	17.11	37.52***	19.98	10.38
LB_r^2	27.78	21.57	35.27**	12.73	65.36***

Note: LL denotes Log likelihood function. BIC denotes Bayes Information Criterion. LB denotes Ljung-Box statistics of flitted residuals and MLB denotes. multivariate Ljung-Box statistic. The Ljung-Box statistics are computed based on 20 lags.

Critical values $\chi^2(6)_{0.05}=12.59$, $\chi^2(6)_{0.01}=16.81$

In Manganelli(2005) 's recursive model, the assumption of weak exogeneity is made in the specification of the conditional mean. The past expected variables are assumed not to carry any information ($c_{ij} = 0$). Manganelli (2005) and Dufour and Engle (2000) also conduct robustness tests of this restriction, in which the residuals of the three models are regressed against lagged expected variables. They find that the lagged expected variables are insignificant. However, we find that the most lagged expected variables are significant (c_{ij}) in our vector MEMs. It is particularly true for infrequently traded stocks. The LR tests also suggest that the lagged expected variables are jointly significant in almost all cases. We argue that the robustness checks conducted by Manganelli (2005) and Dufour and Engle (2000) are misleading, since the dynamics of expected variables has been distorted by the marginal model. Therefore, the weak exogeneity assumption is not supported by the empirical data. The lagged expected variables should be incorporated in this trivariate system.

4.3 Empirical results

Looking first at the price volatility (h_t) process. The coefficient of duration (a_{31}) and coefficient of duration shocks (c_{31}) in the volatility equation are negative and significant in most cases. This is consistent with Easley and O'Hara (1992), indicating that trades with short duration or the shocks of trading intensity are related to the arriving of new information, which reveals a higher volatility impact. The implicit application is that market makers will associate the higher trading activity or trading activity that is higher than its expected level as a signal of informed trading.

The volume coefficient (a_{32}) is only significant for 4 out of 10 stocks and the sign is unclear. However, the volume shocks coefficient (c_{32}) are all significant and positive. This implies that the unexpected component of volume rather than the raw volume carry information. Implicitly, market makers will only consider trade size that is larger than its expected level as a signal of private information, and adjust bid-ask price accordingly. The expected large trade size is simply for liquidity reason. The results partly support the prediction from Easley and O'Hara (1987,1992).

This exercise of the price impact of trade is novel in two aspects. First, most empirical market microstructure literature (see, for example,Dufour and Engle (2000) and Manganelli (2005)) uses raw duration (volume) to determine the presence of

informed traders in the market. We highlight that it is the unexpected components of trade that carry information with respect to asset prices. Second, in contrast to Manganelli (2005), our findings are generally robust for less frequently traded stocks. There is no reason why the informed traders should avoid taking advantage of their private information if it is related to infrequently traded stocks.

The strikingly different results, with respect to the feedback effects from the price process to trading intensity, are found in the duration equation. For the frequently traded stocks, the coefficient of volatility (a_{13}) is always positive but significant for 3 out of 5 stocks and the coefficients on volatility innovation (c_{13}) is always negative but significant for 4 out of 5 stocks. Following Hasbrouck (1988,1991), we explain this by considering the persistent quote change (volatility) to be information motivated and transient quote change (volatility shock) to be inventory motivated. Then our results are consistent with microstructure predictions. For example, information motivated large absolute quote changes indicate a risk of informed trading and the liquidity traders may leave or slow down the trading activity to avoid adverse selection (Admati and Pfleiderer, 1988; Easley and O'Hara, 1992), while inventory motivated large quote changes may attract opposite side traders and increase trading intensity. Similar results can be found for infrequently traded stocks, but the effects are less significant.

In the existing empirical microstructure literature, Dufour and Engle (2000) and Manganelli (2005) find that short durations follow large returns, while Grammig and Wellner (2002) find that lagged volatility significantly reduces trade intensity. Our findings enhance the existing literature by incorporating both of these effects in one model.

4.4 Impulse Response Analysis

From the estimates of the MEM in equation (28), we generate the impulse responses which trace the effect of a one-time shock to one of the innovations on the future values of the endogenous variables. The impulse response function for two representative stocks TRN and ABG are plotted in Figure 4 and Figure 5. This gives the effects of a variation on the forecast up to the 10th trade. Since the impulse-response functions are plotted in transaction time, they are not directly comparable among different stocks. We use the Manganelli (2005) method to

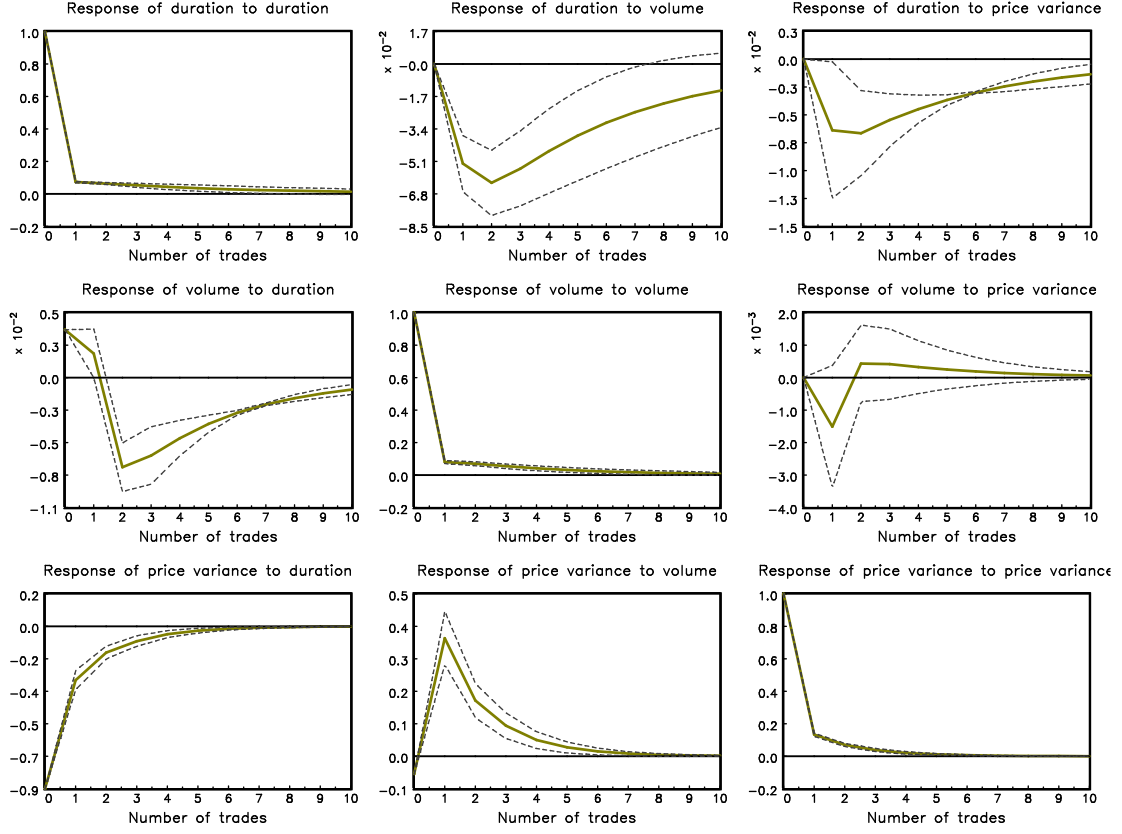


Figure 4: Impulse response function for TRN

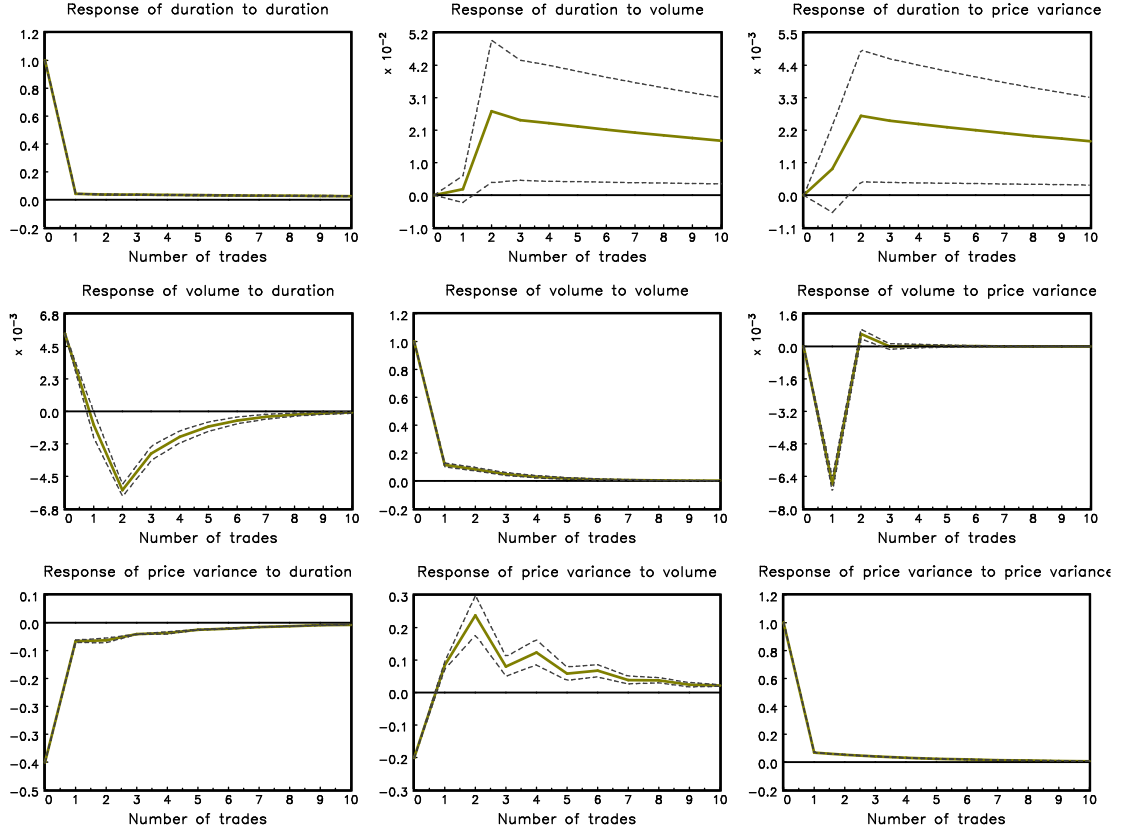


Figure 5: Impulse response function for ABG

approximate the calendar time the system takes to return to its long-run equilibrium, by multiplying the number of transactions by their average duration. The average duration per trade of the two representative stocks is 158 seconds for TRN and 4215 seconds for ABG. This implies, for example, that a shock to the duration of TRN is absorbed by the expected duration after about 27 trades, or, on average, after 1.2 hours. In the case of ABG, the same shock is absorbed after 54 transactions, which corresponds, on average, to a period of 63.3 hours. Similar results hold for the other impulse-responses, indicating that the more traded the stock, the faster the market returns to its full information equilibrium after an initial perturbation. In particular, this is consistent with the (plausible) assumption that the more frequently traded the stock the higher the number of informed traders.

Table 6 summarizes the results for the other stocks, confirming that the price volatility of frequently traded stocks converges much faster to its long-run equilibrium¹⁰ after an initial perturbation. In general, for frequently traded stocks, the new information is implicitly incorporated in the price within one trading day, while it takes up to a week for the new information to be included into the price for infrequently traded stocks. Overall, the effect is to suggest that the market is reasonably efficient. This result, in contrast to Kyle (1985), confirms Admati and Pfleiderer (1988) and Holden and Subrahmanyam (1992)'s finding that information is short lived. For example, Holden and Subrahmanyam (1992) show that with multiple informed traders there will be more aggressive trading in the early periods, causing more information to be revealed earlier in the process.

Table 6: Time (in hours) it takes to absorb shocks to the long term equilibrium variances

	ARG	TRN	TCB	GAS	R
Shock to duration	2.5	1.2	0.8	1.7	3.1
Shock to volume	2.5	1.2	0.8	1.7	3.1
Shock to price volatility	2.4	1.1	0.7	1.7	3.0
	ABG	HTD	LSB	HUN	FEP
Shock to duration	63.3	59.0	37.8	29.7	7.7
Shock to volume	69.1	61.9	38.7	31.3	8.9
Shock to price volatility	63.3	59.0	38.7	29.3	7.2

¹⁰ The threshold at which the shock producing the impulse-response is assumed to be absorbed is at $1e-7$ for shocks. That is, Table 7 reports the time it takes for the impulse-response of the variance to fall below $1e-7$.

5. Conclusion

In this paper, we extend the recursive framework of Engle (2000) and Manganello (2005) for the transaction data to a vector MEM, in which trading duration, volume and price volatility are interdependent. We further propose a multivariate lognormal for the distribution of the vector MEM, which allows the innovations terms to be contemporaneously correlated. In this way, we can build a system that incorporates various causal and feedback effects among these variables. The method is applied to the trade and quote dataset of the NYSE and the model is estimated using a sample of 10 stocks. The empirical findings are summarized as follows:

- (1) The diagnostic statistics show that the vector MEM improves the dynamic properties of the model significantly. Moreover, the lagged (un)expected variables are widely significant in the MEM model, challenging the weak exogeneity assumptions used in the empirical market microstructure literature.
- (2) We find a significant price impact of trade. However, we highlight the effect of unexpected components of trading characteristics. Both duration and duration shocks carry price information, while only unexpected volume carries most of the volume related information content.
- (3) We also find significant feedback effects, with volatility and volatility shocks affecting duration in different directions. This finding confirms Hasbrouck (1988,1991)'s prediction that persistent quote changes are driven by private information, which decreases trading intensity, while the transient quote changes are motivated by inventory control, which would attract opposite side traders and increase trading intensity. However, this effect is only robust for frequently traded stocks.
- (4) With the impulse response, we find that the new information is implicitly incorporated in to the price within one trading day for frequently traded stocks, and it takes up to one week for infrequently traded stocks.

With respect to further research, the methodology used in this paper can easily be extended to model any non-negative valued variables. An interesting extension is to

model financial volatilities. For example, there are different measures of volatility, but no individual one appears to be a sufficient measure on its own. One possibility is to consider absolute daily returns, daily high-low range and daily realized volatility in the vector MEM for forecasting volatility (see Engle and Gallo (2006)). A second example, the multivariate GARCH model is usually used in modelling dynamics interactions among volatilities in different markets. But it is hindered by parametric limitations. However, one can model directly the volatility proxy (i.e. daily range) for each market and insert other markets' volatility in the expression of its conditional expectations in the vector MEM. This is a very promising possibility, since there is no parametric limitation.

Appendices:

Appendix 1: Lognormal distribution

[A] Univariate lognormal distribution

A lognormally-distributed random variable is a random variable whose logarithm is normally-distributed. Consider a standard lognormally-distributed x , whose logarithmic transformation $y = \log(x)$ is normally-distributed with mean μ and standard deviation σ . The probability density function for a lognormal distribution is given by,

$$f(x | \sigma) = \frac{1}{\sqrt{2\pi\sigma^2} x} \exp\left[-\frac{(\log x - \mu)^2}{2\sigma^2}\right], \quad x \geq 0$$

As noted, for example, in Hines and Montgomery (1990) this distribution is skewed with a longer tail to the right of the mean. When μ and σ are known for y , the corresponding mean and variance for x can be found from the following:

$$E(x) = e^{\mu + \frac{1}{2}\sigma^2}$$
$$Var(x) = e^{2\mu + \sigma^2} (e^{\sigma^2} - 1)$$

[B] Multivariate lognormal distribution

Let $y = (y_1, y_2, \dots, y_k)$ be a k -dimensional random variable having multivariate normal distribution with mean v and covariance matrix $D = (d_{ij})$. The probability density function of y is defined as:

$$f_y(y | D) = (2\pi)^{-k/2} |D|^{-1/2} \exp\left(-\frac{1}{2}(y-v)' D^{-1}(y-v)\right)$$

So the variable, $x = \exp(y)$, has a multivariate lognormal distribution. It is defined as $x \sim \ln N(v, D)$. Using the Jacobian transformation, and $y = h(x) = \ln(x)$, the probability density for the multivariate lognormal distribution has the following form:

$$f_x(x | D) = f_y(h(x) | D) \left| \frac{dh}{dx} \right|$$
$$= (2\pi)^{-k/2} |D|^{-1/2} (x_1 x_2 \cdots x_k)^{-1} \exp\left(-\frac{1}{2}(\ln x - v)' D^{-1}(\ln x - v)\right)$$

Law and Kelton (2000) show that the covariance and correlation of the bivariate lognormal variables $x = (x_1, x_2, \dots, x_k)$ are given by:

$$\begin{aligned}\mu &= E(x) = (\mu_1, \mu_2, \dots, \mu_k)' & \mu_i &= e^{v_i + \frac{1}{2}d_{ii}} \\ \Sigma &= E(x - \mu)(x - \mu)' = \sigma_{ij} & \sigma_{ij} &= e^{(v_i + v_j + \frac{d_{ii} + d_{jj}}{2})} (e^{d_{ij}} - 1) \\ \rho_{ij} &= \frac{e^{d_{ij}} - 1}{\sqrt{(e^{d_{ii}} - 1)(e^{d_{jj}} - 1)}}\end{aligned}$$

where d_{ij} is the ij th element of D . It is clear that if y_1, y_2, \dots, y_k are independent, then x_1, x_2, \dots, x_k are also independent and vice versa.

[C] Jacobian transformation

Let $y = (y_1, \dots, y_k)$ be a k -dimensional random variable with probability density function (pdf) $f_y(y): R^k \rightarrow R$

Define some 1:1 differentiable transformation of y into x using $g: R^k \rightarrow R^k$,

$$g(y) = \begin{bmatrix} g_1(y) \\ \vdots \\ g_k(y) \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_k \end{bmatrix} = x$$

with inverse

$$h(x) = \begin{bmatrix} h_1(x) \\ \vdots \\ h_k(x) \end{bmatrix} = \begin{bmatrix} y_1 \\ \vdots \\ y_k \end{bmatrix} = y$$

The pdf of y , the transformed random variable, is

$$f_x(x) = f_y(h(x)) \left| \frac{dh}{dx} \right|$$

where

$$\left| \frac{dh}{dx} \right| = \left| \frac{\partial(h_1, \dots, h_k)}{\partial(x_1, \dots, x_k)} \right| = \begin{vmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_1}{\partial x_2} & \dots & \frac{\partial h_1}{\partial x_k} \\ \frac{\partial h_2}{\partial x_1} & \frac{\partial h_2}{\partial x_2} & \dots & \frac{\partial h_2}{\partial x_k} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_k}{\partial x_1} & \frac{\partial h_k}{\partial x_2} & \dots & \frac{\partial h_k}{\partial x_k} \end{vmatrix}$$

Appendix 2: Inference of VARMA Models

A general K -dimensional, linear, time-invariant, covariance stationary VARMA (p,q) model takes the form¹¹

$$\chi_t = A_1\chi_{t-1} + \dots + A_p\chi_{t-p} + \nu_t + B_1\nu_{t-1} + \dots + B_q\nu_{t-q}$$

Or, in lag operator notation, as

$$A(L)\chi_t = B(L)\nu_t$$

where $A(L) := I_K - A_1L - \dots - A_pL^p$ and $B(L) := I_K + B_1L + \dots + B_qL^q$

[A] Stationarity and invertibility

The VARMA(p,q) system will be stationary if $\det(I_n + A_1\lambda + \dots + A_p\lambda^p) \neq 0$ for $|\lambda| < 1$. The VARMA system will be invertible if and only if $\det(I_n + B_1z + \dots + B_qz^q) \neq 0$ for $|z| < 1$.

[B] Identification

It is well known that the VARMA (p,q) model is generally not identified unless special restrictions are imposed on the coefficient matrices (Lütkepohl, 2005).

[C] The Gaussian quasi-maximum likelihood function

Assume our VARMA(1,1) process is a Gaussian, stationary, and invertible process,

$$\chi_t = \gamma + A_1\chi_{t-1} + \nu_t + B_1\nu_{t-1}$$

Assume also that we have a sample $\chi_1, \chi_2, \dots, \chi_T$ and define

$$\Gamma = \begin{bmatrix} I_3 & 0_3 & \dots & 0 & 0 \\ -A_1 & I_3 & \dots & 0 & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & I_3 & 0 \\ 0 & \dots & -A_1 & I_3 \end{bmatrix}, \quad M = \begin{bmatrix} B_1 & I_3 & \dots & 0 \\ 0 & B_1 & I_3 & \dots & 0 \\ \ddots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & B_1 & I_3 \\ 0 & \dots & B_1 & I_3 \end{bmatrix}$$

where Γ is a $3T \times 3T$ matrix and M is a $3T \times 3(T+1)$ matrix,

¹¹ In principle, it should contain an intercept term. This has not been done here because it is assumed that the mean has been subtracted prior to estimation.

Then we get

$$\Gamma \begin{bmatrix} \chi_1 \\ \vdots \\ \chi_T \end{bmatrix} + \begin{bmatrix} \gamma \\ \vdots \\ \gamma \end{bmatrix} + \begin{bmatrix} -A_1 \chi_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \mathbf{M} \begin{bmatrix} \nu_0 \\ \nu_1 \\ \vdots \\ \nu_T \end{bmatrix}$$

Hence, for given, fixed presample values χ_0

$$\chi = \begin{bmatrix} \chi_1 \\ \vdots \\ \chi_T \end{bmatrix} \sim N(\Gamma^{-1} \mu_0, \Gamma^{-1} \mathbf{M} (I_{T+1} \otimes \Sigma_v) \mathbf{M}' \Gamma'^{-1})$$

where $\mu_0 := \begin{bmatrix} \gamma \\ \vdots \\ \gamma \end{bmatrix} + \begin{bmatrix} -A_1 \chi_0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

Assuming that ν_t is Gaussian white noise, the corresponding likelihood function, conditional on χ_0 , is

$$\begin{aligned} & l(A_1, B_1, \gamma, \Sigma_v | \chi, \chi_0) \\ & \propto \left| \Gamma^{-1} \mathbf{M} (I_{T+1} \otimes \Sigma_v) \mathbf{M}' \Gamma'^{-1} \right|^{-1/2} \\ & \quad \times \exp \left\{ -\frac{1}{2} (\chi - \Gamma^{-1} \mu_0)' \Gamma' [\mathbf{M} (I_{T+1} \otimes \Sigma_v) \mathbf{M}']^{-1} \Gamma (\chi - \Gamma^{-1} \mu_0) \right\} \\ & = \left| \mathbf{M} (I_{T+1} \otimes \Sigma_v) \mathbf{M}' \right|^{-1/2} \\ & \quad \times \exp \left\{ -\frac{1}{2} (\Gamma \chi - \mu_0)' [\mathbf{M} (I_{T+1} \otimes \Sigma_v)]^{-1} (\Gamma \chi - \mu_0) \right\} \end{aligned}$$

where $|\Gamma| = 1$ has been used.

Even if the Gaussian white noise assumption of ν_t is invalid, maximization of the Gaussian log likelihood function can provide consistent estimates of the parameters of this linear representation. This is a quasi-maximum likelihood estimation solution. However, the standard errors have to be adjusted.

Appendix 3: Proofs of impulse response function

Model: $x_t = \mu_t \odot \varepsilon_t$

$$\ln \mu_t = \omega + A \ln x_{t-1} + B \ln \mu_{t-1}$$

Firstly, we Hasbrouck (1988,1991) transform the vector MEM into a VARMA model, by substituting

$$\ln \mu_t \text{ with } \ln x_t - \ln \frac{x_t}{\mu_t}.$$

Then,

$$\begin{aligned} \ln x_t &= \omega + A \ln x_{t-1} + \ln \frac{x_t}{\mu_t} + B \left(\ln x_{t-1} - \ln \frac{x_{t-1}}{\mu_{t-1}} \right) \\ &= \omega + A' \ln x_{t-1} + \ln \varepsilon_t + B' \ln \varepsilon_{t-1} \end{aligned}$$

where $A' = A + B$, and $B' = -B$

The causal and feedback effect are not affected by this transformation. Therefore, it is feasible to assume that $\ln \varepsilon_t$ follows a multivariable Gaussian distribution. Then, (quasi) maximum likelihood estimation can be used to estimate the parameters of VARMA model. Suppose v_t is a multivariable Gaussian distributed random variables, then

$$\ln x_t = \bar{\omega} + A' \ln x_{t-1} + v_t + B' v_{t-1}$$

where $v_t \sim N(0, \Sigma_v)$, $\bar{\omega} = \omega + D + BD$, $D = \text{dia}(\Sigma_v)$

In the impulse response, we work on the impulse of v_t on the $\ln x_t$ in a standard way. Writing the VARMA (1,1) equation as an infinite VAR model:

$$\ln x_t = \bar{\omega} + \sum_{i=0}^{\infty} \Phi_i v_{t-i} = \bar{\omega} + \Phi(L) v_t$$

$$\begin{aligned} \Phi(L) &= (I - A'L)^{-1}(I + B'L) \\ &= (I + A'L + (A'L)^2 + (A'L)^3 + \dots)(I + B'L) \\ &= I + (A' + B')L + A'(A' + B')L^2 + A'^2(A' + B')L^3 + \dots \end{aligned}$$

$$\text{And } \Phi_i = A'^{i-1}(A' + B') = (A + B)^{i-1}(A - B) \quad \Phi_0 = I$$

The impulse response function of the model (17) for $t > 0$ is :

$$\frac{\partial \ln x_t}{\partial v_0} = \Phi_t$$

This process can be rewritten in such a way that the residuals of different equations are uncorrelated. For this purpose, we choose a decomposition of the white noise covariance matrix $\Sigma_v = W\Sigma_\tau W'$, where Σ_τ is a diagonal matrix with positive diagonal elements and W is a lower triangular matrix with unit diagonal. This decomposition is obtained from the Choleski decomposition $\Sigma_v = PP'$ by defining a diagonal matrix D which has the same main diagonal as P and by specifying $W = PD^{-1}$ and $\Sigma_\tau = DD'$

$$\ln x_t = \bar{\omega} + \sum_{i=0}^{\infty} \Theta_i \varepsilon_{t-i} \quad \Theta_i = \Phi W_i^{-1}$$

Then the impulse response function of the model (17) for $t > 0$ is :

$$\frac{\partial \ln x_t}{\partial \tau_0} = \Theta_t$$

References

- Admati AR, Pfleiderer P A theory of intraday patterns: Volume and price variability. *Review of Financial Studies*. 1988; 1; 3-40.
- Allen D, Chan F, McAleer M, Peiris S Finite sample properties of the QMLE for the Log-ACD model: Application to Australian stocks. *Journal of econometrics*. 2008; 147; 163-185.
- Allen D, Lazarov Z, McAleer M, Peiris S Comparison of alternative ACD models via density and interval forecasts: Evidence from the Australian stock market ☆. *Mathematics and Computers in Simulation*. 2009; 79; 2535-2555.
- Andersen TG, Bollerslev T, Diebold FX, Ebens H The distribution of realized stock return volatility. *Journal of financial economics*. 2001; 61; 43-76.
- Bauwens L, Giot P The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks. *Annales d'Economie et de Statistique*. 2000; 117-149.
- Bowe M, Hyde S, McFarlane L The interrelationship between trading volume, duration and futures contract price volatility. *Manchester Business School working paper*. 2009.
- Cipollini F, Engle RF, Gallo GM A model for multivariate non-negative valued processes in financial econometrics. Unpublished paper: Stern School of Business, New York University. 2007.
- Cizeau P, Liu Y, Meyer M, Peng CK, Eugene Stanley H Volatility distribution in the S&P500 stock index. *Physica A: Statistical Mechanics and its Applications*. 1997; 245; 441-445.
- Dacorogna MM An introduction to high-frequency finance. *Academic Pr*; 2001.
- Diamond DW, Verrecchia RE Constraints on short-selling and asset price adjustment to private information. *Journal of financial economics*. 1987; 18; 277-311.
- Dufour A, Engle RF Time and the price impact of a trade. *The Journal of Finance*. 2000; 55; 2467-2498.
- Easley D, O'Hara M Price, trade size, and information in securities markets. *Journal of financial economics*. 1987; 19; 69-90.
- Easley D, O'Hara M Time and the process of security price adjustment. *Journal of Finance*. 1992; 577-605.
- Engle R New frontiers for ARCH models. *Journal of Applied Econometrics*. 2002; 17; 425-446.
- Engle RF The Econometrics of Ultra - high - frequency Data. *Econometrica*. 2000; 68; 1-22.
- Engle RF, Gallo GM A multiple indicators model for volatility using intra-daily data. *Journal of econometrics*. 2006; 131; 3-27.
- Engle RF, Russell JR Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica*. 1998; 1127-1162.
- Engle RF, Sun Z When is noise not noise—a microstructure estimate of realized volatility. Manuscript, Stern School, New York University. 2007; 38.
- French KR, Roll R Stock return variances: The arrival of information and the reaction of traders. *Journal of financial economics*. 1986; 17; 5-26.
- Ghysels E, Gouriéroux C, Jasiak J Stochastic volatility duration models. 1998.
- Glosten LR, Milgrom PR Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. *Journal of financial economics*. 1985; 14; 71-100.
- Goodhart CAE, O'Hara M High frequency data in financial markets: Issues and applications. *Journal of empirical finance*. 1997; 4; 73-114.

-
- Grammig J, Maurer KO Nonmonotonic hazard functions and the autoregressive conditional duration model. *The Econometrics Journal*. 2000; 3; 16-38.
- Grammig J, Wellner M Modeling the interdependence of volatility and inter-transaction duration processes. *Journal of econometrics*. 2002; 106; 369-400.
- Harris M, Raviv A Differences of opinion make a horse race. *Review of Financial Studies*. 1993; 6; 473-506.
- Hasbrouck J Trades, quotes, inventories, and information. *Journal of financial economics*. 1988; 22; 229-252.
- Hasbrouck J Measuring the information content of stock trades. *Journal of Finance*. 1991; 179-207.
- Hautsch N Capturing common components in high-frequency financial time series: A multivariate stochastic multiplicative error model. *Journal of Economic Dynamics and Control*. 2008; 32; 3978-4015.
- Hines WW, Montgomery DC Probability and Statistics in Engineering and Management Science. John Wiley & Sons, Inc. 1990.
- Ho T, Stoll HR Optimal dealer pricing under transactions and return uncertainty. *Journal of financial economics*. 1981; 9; 47-73.
- Holden CW, Subrahmanyam A Long-Lived Private Information and Imperfect Competition. *Journal of Finance*. 1992; 47; 247-270.
- Hosking JRM The multivariate portmanteau statistic. *Journal of the American Statistical Association*. 1980; 75; 602-608.
- Kyle AS Continuous auctions and insider trading. *Econometrica: Journal of the Econometric Society*. 1985; 1315-1335.
- Lütkepohl H New introduction to multiple time series analysis. 2005.
- Law AM, Kelton WD Simulation Modeling and Analysis. Mc Graw Hill. 2000.
- Lee CMC, Ready MJ Inferring trade direction from intraday data. *Journal of Finance*. 1991; 733-746.
- Lunde A A generalized gamma autoregressive conditional duration model. Department of Economics, Politics and Public Administration, Aalborg University, Denmark. 1999.
- Manganelli S Duration, volume and volatility impact of trades. *Journal of Financial Markets*. 2005; 8; 377-399.
- O'Hara M Market microstructure theory. Blackwell; 1995.
- O'Hara M, Oldfield GS The Microeconomics of Market Making. *Journal of Financial and Quantitative Analysis*. 1986; 361-376.
- Taylor N Trading intensity, volatility, and arbitrage activity. *Journal of Banking & Finance*. 2004; 28; 1137-1162.
- Wuensche O, Grammig J, Theissen E Time and the Price Impact of a Trade: A Structural Approach. 2007.
- Xu Y Lognormal ACD model and comparison with alternative specification Working paper, Cardiff Business School, UK. 2011a.
- Xu Y Weak exogeneity in financial point processes. Working paper, Cardiff Business School, UK. 2011b.