CARDIFF BUSINESS SCHOOL
WORKING PAPER SERIES

Cardiff Economics
Working Papers

Woon K. Wong, Dijun Tan and Yixiang Tian

Nonlinear ACD Model and Informed Trading: Evidence from Shanghai Stock Exchange

E2008/8

Cardiff Business School
Cardiff University
Colum Drive
Cardiff CF10 3EU
United Kingdom
t: +44 (0)29 2087 4000
f: +44 (0)29 2087 4419
www.cardiff.ac.uk/carbs

ISSN 1749-6101
April 2008

Nonlinear ACD Model and Informed Trading: Evidence from Shanghai Stock Exchange

Woon K. Wong*

Investment Management Research Unit

Cardiff Business School

Dijun Tan

School of Management

University of Electronic Science and Technology

Yixiang Tian

School of Management

University of Electronic Science and Technology

28 January 2008

*Corresponding author: Aberconway Building, Colum Drive, Cardiff, CF10 3EU, United Kingdom. Tel: +44 29 20875079; Fax: +44 29 20874419; Email: wongwk3@cardiff.ac.uk

Nonlinear ACD Model and Informed Trading:
Evidence from Shanghai Stock Exchange

Abstract

Dufour and Engle (J. Finance (2000) 2467) find evidence of an increased presence of informed traders when the NYSE markets are most active. No such evidence, however, can be found by Manganelli (J. Financial Markets (2005) 377) for the infrequently traded stocks. In this paper, we fit a nonlinear log-ACD model to stocks listed on Shanghai Stock Exchange. When trading volume is high, empirical findings suggest presence of informed trading in both liquid and illiquid stocks. When volume is low, market activity is likely due to liquidity trading. Finally, for the actively traded stocks, our results support the price formation model of Foster and Viswanathan (Rev. Financial Studies (1990) 593).

Keywords: Informed trading, Liquidity trading, Duration, Volume, Volatility

JEL Classification: G11, G14, G15

1. Introduction

Due to the availability of high frequency intraday trade data, there have been increasing empirical interests in the role of duration, time between trades, in conveying information to market participants. The theoretical motivations for the study on the role of time between transactions can be traced back to Diamond and Verrecchia (1987) and Easley and O'Hara (1992). According to Diamond and Verrecchia, long durations are likely to be associated with bad news because informed traders will always trade unless they do not own the stock and short-sale constraints exist. On the other hand, in the model studied by Easley and O'Hara, informed traders can always trade as soon as there is a signal or news. As a result, long durations are likely to be associated with no news.

Generally speaking, informed traders, for fear of newly received information becoming worthless, tend to trade as quickly as possible and as much as possible. However, as pointed out by Easley and O'Hara (1987), informed traders may be recognised by their large volume trading and their profit opportunities would not be maximised. Therefore, informed traders may choose to break up large volume trades, thereby generating a large number of information-based trades, which results in higher trading rates. This analysis is not only consistent with the empirical findings by Chakravarty (2001) that medium-size trades by (informed) institutions cause disproportionately large stock prices changes, but also suggests that the variations in trading rates in Easley and O'Hara (1992) are associated with changing numbers of informed traders.[1] Clearly, above literatures suggest that duration conveys information.

Using Hasbrouck's (1991) vector autoregressive model for prices and trades,

[1] Recently, literature suggests that institutions are relatively more informed; see, e.g., Lee and Radhakrishna (2000) and Anand, Chakravarty and Martell (2005).

Dufour and Engle (2000) study empirically the role played by duration in the process of price formation. Dufour and Engle find that high trading intensity (short duration) is associated with larger price impact of trades and faster price adjustment to trade-related information, suggesting an increased level of information present in the market. Manganelli (2005) presents a framework that models duration, volume and volatility simultaneously, incorporating causal and feedback effects among the variables. Manganelli applies the methodology to two groups of NYSE stocks, classified according to trade intensity or liquidity. Findings similar to that of Dufour and Engle are obtained for the frequently traded stocks. For the infrequently traded stocks, both lagged volumes and squared returns hardly affect the durations at all. Contrary to the findings for the American markets, Cellier (2003) applies Manganelli's model to the Paris Bourse to find significantly positive relationship between duration and past volatility, implying that larger price variations tend to be associated with lower trade intensity. Attributing the results to the different learning process in the purely order-driven Paris Bourse, Cellier claims his findings as evidence that the French stock market is dominated by liquidity trading.

In this paper, we use a nonlinear (piecewise linear) log Autoregressive Conditional Duration (ACD) model to study the relationships among the duration, volume and volatility for the stocks listed on Shanghai Stock Exchange (SHSE). Motivated by literatures indicating that volume could be used as a proxy for information flow, we consider a piecewise linear log-ACD model according to the size of trading volume. While our findings are consistent with those of Dufour and Engle (2000) and Manganelli (2005), they contribute to the literature in the following ways. First, in the case of Manganelli's study, times of greater activity coincide with a higher presence of informed traders *only* for the frequently traded stocks.[2] The results obtained by our nonlinear log-ACD model indicate otherwise. Specifically, when the volume is high, greater trading activity is found to be associated with larger price variations for *both* frequently and infrequently traded stocks. Since there is no reason to exclude informed traders from trading in the less liquid stocks, our finding is more plausible.

Second, Dufour and Engle (2000) reject Admati and Pfleiderer's (1988) model in favour of Foster and Viswanathan's (1990) on the ground that both the price impact of trades and the speed of price adjustment to trade-related information increase as the time duration between trades decreases. This view of Dufour and Engle may be understood by considering the work of Seppi (1997) who associates market liquidity to the *temporary* or *non-informational* price impact of market orders of different sizes. Accordingly, Dufour and Engle interpret large price impacts of trades and fast price adjustment to new information as signs of a market with reduced liquidity, a consequence of an increased presence of informed traders. Strictly speaking, it is difficult to differentiate the two microstructure models of Admati and Pfleiderer and Foster and Viswanathan in that the former also has an increased presence of informed traders (albeit along with uninformed liquidity traders) whose trading would also make price more informative. Fortunately, the empirical work of Foster and Viswanathan (1993) illustrates a way to substantiate the claim made by Dufour and Engle. Foster and Viswanathan postulate that the presence of informed traders would deter discretionary liquidity traders from trading, especially when the public information to be released proxy well the private information. Accordingly, they find for actively traded stocks (thus with plenty of public news), trading volume on Monday is on average lower than other weekdays. The reason is that a large number informed traders with private information

---

[2] All the stocks studied by Dufour and Engle (2000) are also highly active stocks.

accumulated over the weekend are keen to trade to maximize their profits on the first day of trading, thereby discouraging discretionary liquidity traders from trading and thus resulting in a lower volume. Consistent with the model of Foster and Viswanathan, we find for the frequently traded stocks in our sample low duration (or high trading activity) coincides with low trading volume.

Third, we observe that when trading volume is low, market activity on the stocks is essentially liquidity motivated. Our conjecture is consistent with the notion of liquidity as defined by Seppi (1997) and Dufour and Engle (2000).[3] According to them, liquidity can be regarded as a measure of market quality in which trades have a lower impact on prices, and new trade-related information takes longer to be fully incorporated into prices. Therefore, our finding of a positive association between duration and price variation (when the trading volume is low) implies that little new information is impounded on price when the price variation is small. On the other hand, if there is significant new information to be incorporated, price adjustment takes a longer duration to do so.

Finally, our empirical results also suggest that a nonlinear (or piecewise linear) model is preferable to describe the complicated relationship between duration, volume and volatility. This remark is substantiated by two observations. One is the fact that, as noted above, Manganelli's (2005) linear VAR system fails to uncover signs of informed trading in the infrequently traded stocks though there is no reason why informed traders should not exploit their information advantage in the illiquid stocks. The other observation is the (incorrect) inference implied by Cellier's (2003) model estimates for the Paris Bourse. Like the Paris Bourse, Shanghai stock market is also a purely order-driven market. It is interesting to observe that when a linear

[3] See also Grossman and Miller (1988), Harris (1990) and Brennan and Subrahmanyam (1996) for more expositions on the definition of liquidity.

log-ACD model is used, we arrive at the same conclusion as Cellier. However, likelihood ratio tests reject a linear specification, and the inference that high trading activity is due to liquidity trading contradicts both existing theoretical predictions and empirical findings. Therefore, we conclude that a linear relationship fails to describe the complex dynamics of duration, volume and price variation.

The rest of the paper is organised as follows. Section 2 provides information on the institution background of SHSE. Section 3 describes the econometric models whereas Section 4 presents the empirical results. Finally, Section 5 concludes.

2. Institutional background

China has the largest fast growing economy in the world. In US dollar term, the size of its economy stands at \$2.7 trillion in 2006, ranked after US, Japan and Germany. In parallel with the fast growing economy, the combined market capitalization of its two domestic stock exchanges, the Shanghai Stock Exchange (SHSE) and the Shenzhen Stock Exchange (SZSE), have grown to \$3.7 trillion in 2007. In particular, SHSE is one of the most actively traded stock exchanges. By the end of 2004, its 837 listed companies have already reached an annual share turnover of 288.7%.

Market structure wise, SHSE is a purely order-driven market without designated market makers. It runs an electronic automated trading system and opens from Monday to Friday with three sessions: 0915-0925 for call auction, 0930-1130 and 1300-1500 for continuous trading double auction. Only limit orders are allowed in SHSE. Orders are valid for one day and are stored in the limit order book, of which the best five bid and ask prices and the corresponding depths of the book are revealed continuously to public investors. The tick size (minimum price variation unit) is 0.01

RMB while the minimum trading quantities unit is 100 shares (one lot). In the pre-trading call auction, submitted orders are batched for execution, resulting in an equilibrium opening price that maximizes the total trading volume; see also Xu (2000). In the subsequent trading sessions, submitted buy and sell limit orders are matched continuously based on the price and time priority rules. While the matched orders result in a trade, the unmatched orders remain in the order queues in the limit order book, waiting for future executions.

Trading on SHSE is dominated by individual investors: 99.5% of the 68.8 million domestic investor accounts in 2002 are held by individual investors.[4] Short selling is absolutely prohibited in SHSE. Also, to dampen extreme price movements and to provide a cool-off period in the events of overreaction, SHSE currently sets the daily price limit at 10%. Due to the growing importance of China economy and its financial markets, there is an increasing research on China stock markets; see, e.g., Feng and Seasholes (2004), Chan, Menkveld and Yang (2007) and Ng and Wu (2007).

3. Econometric Models

3.1 A linear log-ACD model

The Autoregressive Conditional Duration (ACD) model of Engle and Russell (1998) forms the basis for various models of irregularly spaced transaction data; see, e.g., the Ultra-High-Frequency GARCH model by Engle (2000), the log-ACD model by Bauwens and Giot (2000), the nonlinear ACD model by Zhang, Russell and Tsay (2001), and the stochastic volatility duration models by Ghysels, Gourieroux and Jasiak (2004). The ACD model employs a marked point process to describe the dynamics of transaction duration, which may be written as follows:

---

[4] See the Chinese Securities Depository & Clearing Co. Ltd.

$$x_i = \varphi_i z_i\,, \tag{1}$$

$$\varphi_i = E(x_i \mid \Omega_{i-1}) = \mu + \sum_{j=1}^{p} \kappa_j x_{i-j} + \sum_{j=1}^{q} \eta_j \varphi_{i-j}\,. \tag{2}$$

Here, x_i is the i^{th} duration, and φ_i is the conditional mean of x_i; μ, κ_j and η_j are coefficients; Ω_{i-1} is the information set at the time $i-1$; and $\{z_i\}$ is an *iid* innovation process. Distribution of $\{z_i\}$ can be either Exponential, Weibull or Gama with $E(z_i)=1$ and $Var(z_i)=\delta$. To ensure a positive duration, we consider a simple log-ACD model proposed by Bauwens and Giot (2000) with $p = q = 1$ as given below.

$$x_i = e^{\varphi_i} z_i\,, \tag{3}$$

$$\varphi_i = E(x_i \mid \Omega_{i-1}) = \mu + \kappa \ln(x_{i-1}) + \eta \varphi_{i-1}\,. \tag{4}$$

As mentioned above, Dufour and Engle (2000) find that duration plays an important role in the process of price formation. They discover that as duration decreases, the price impact of trades and the speed of price adjustment to trade-related information increase, suggesting an increased presence of informed traders. Building on the results of Dufour and Engle, we analyze the influence of volume and volatility on duration. Our approach is a log-ACD model augmented with volume and price volatility, so Equation (4) above is replaced by

$$\varphi_i = E(x_i \mid \Omega_{i-1}) = \mu + \kappa \ln(x_{i-1}) + \eta \varphi_{i-1} + \xi Volume_{i-1} + \gamma \left|u_{i-1}\right|, \tag{5}$$

where $Volume_{i-1}$ is the trading volume series and $\left|u_{i-1}\right|$ is the proxy for volatility.[5] Above augmented log-ACD model is identical with the duration equation of the VAR system of duration, volume and volatility proposed by Manganelli (2005) to

---

[5] u_i in (5) is obtained as the residuals of an MA(1) process: $r_i = \mu + \rho u_{i-1} + u_i$. That is, u_i is the residual series after removing the microstructure effect of the original price return series; see Dacorogna, Gencay and Muller (2000). We have also used u_i^2 in place of $\left|u_i\right|$, and essentially similar results are obtained.

study NYSE stocks. According to Admati and Pfleiderer (1988), Dufour and Engle (2000), Manganelli (2005) and others, if the high trading intensity is attributed to informed trading, then price volatility is high.[6] That is, volatility is positively related with trading intensity and negatively associated with duration, so γ in Equation (5) is expected to be negative. Otherwise, if the high trading intensity is related to liquidity trading, γ is expected to be positive.

Similar argument holds for volume. For example, Holden and Subrahmanyam (1992) generalize Kyle (1985) model to incorporate competition among multiple risk-averse insiders and demonstrate that competition among insiders is associated with high trading volume and rapid revelation of private information. Generally speaking, analyses of Easley and O'Hara (1992), O'Hara (1995), and Easley, Kiefer and O'Hara (1997) suggest there is some implied information in the trading volume that may not be reflected in the price process timely. All these studies share the same conclusion that there is a positive relationship between volumes and informed trading. Therefore, the volume coefficient ξ in (5) is expected to be negative in the presence of informed traders.

3.2 A nonlinear log-ACD model

In addition to the above microstructure literatures on the association of volume and informed trading, numerous studies have documented the importance of volume as a proxy for information. For example, Lamoureux and Lastrapes (1990) find that augmenting the variance function with trading volume for an individual stock removes evidence of GARCH effects; Andersen (1996) in a stochastic volatility framework regards the variation in trading volume as the information arrival process. Therefore, since the dynamics of informed trading is likely to differ from those of

[6] In Hasbrouck (1991), the trade-correlated component of variance of changes in the efficient price is regarded as a measure of private information impounded on the market through trading.

liquidity trading, a nonlinear relationship dependent on the level of trading volume is considered. Another motivation for a nonlinear model comes from the empirical results of Manganelli (2005) for the less frequently traded stocks, where most of the volatility coefficients γ's are found to be insignificant. Since there is no theoretical ground to exclude investors with private information to trade on illiquid stocks, we conjecture that the insignificant finding of Manganelli is likely due to the possibility that a linear model fails to uncover the presence of informed trading in a less liquid stock. We thus propose here a simple nonlinear (piecewise linear) log-ACD model to differentiate the relationship between volatility and duration according to the size of trading volume as stated below,

$$\begin{aligned}\varphi_i &= E(x_i \mid \Omega_{i-1}) \\ &= \mu + \kappa \ln(x_{i-1}) + \eta\varphi_{i-1} + \xi Volume_{i-1} + \gamma_L V_{i-1}^L |u_{i-1}| + \gamma_H V_{i-1}^H |u_{i-1}|.\end{aligned} \tag{6}$$

V_i^H is an indicator variable that equals to 1 if $Volume_i \geq Mean(Volume)$, 0 otherwise.[7] The other indicator variable is simply defined as $V_i^L = 1 - V_i^H$. The above nonlinear model is actually a piecewise linear log-ACD model in which the relationship between duration and volatility is captured by γ_H when volume is above average ($V_i^H = 1$, $V_i^L = 0$); when volume is below average ($V_i^L = 1$, $V_i^H = 0$), the relationship is described by γ_L.

Since it is theoretically plausible that (discretionary) liquidity trading also causes concentrated trading (see Admati and Pfleiderer, 1988), the advantage of (6) is to allow for concentrated trading to be caused by informed trading at certain periods of time (say, when volume is high), as well as by liquidity trading at other time intervals (when volume is low). If this hypothesis was correct, our nonlinear log-ACD model would detect presence of informed traders for both liquid and

[7] $Mean(Volume)$ is the mean value of volume over the entire sample.

illiquid stocks.

As it turns out, our empirical results in the next section shows that high trading activities at different volume state do suggest a rather different economic dynamic: short duration at high-volume state implies informed trading whereas at low-volume state, concentrated trading is likely due to liquidity traders. We therefore consider a step further in which a different dynamic also exists between duration and volume according to the size of volume, as described below.

$$\begin{aligned}\varphi_i &= E(x_i \mid \Omega_{i-1}) \\ &= \mu + \kappa \ln(x_{i-1}) + \eta \varphi_{i-1} \\ &\quad + (\xi_L Volume_{i-1} + \gamma_L |u_{i-1}|) \cdot V_{i-1}^{L} + (\xi_H Volume_{i-1} + \gamma_H |u_{i-1}|) \cdot V_{i-1}^{H}.\end{aligned} \tag{7}$$

Generally speaking, high trading volume is associated with rapid revelation of private information. However, Foster and Viswanathan (1990, 1993) claim that the presence of informed traders could also deter discretionary liquidity traders from trading and thus resulting in a relatively lower volume. Our nonlinear log-ACD model above enables us to formally test the claim made by Foster and Viswanathan. For highly active stocks with plenty of news coverage, the Foster and Viswanathan's model predicts a positive ξ_H; in all other cases, ξ's should be negative.

4. Empirical results

4.1 Data

We consider 10 stocks listed on the SHSE and extract their transaction data from the CSMAR database.[8] To select the 10 stocks, we first classify all the stocks listed on the SHSE into large, medium and small groups according to their market value, and five stocks with the highest market value in the large and small groups are

[8] CSMAR stands for China Securities Market and is a registered trade mark of GTA Information Technology, Co. Ltd.

selected. For our analysis, we use price of trades, time stamp of trades, size (volume) of trades, and bid-ask quotes. Our sample period begins on 1 September 2003 and ends on 30 June 2005. As is noted before, in each trading day, there are four trading hours in two sessions of continuous trading, from 0930 to 1130 and from 1300 to 1500, with a noon-break in between. Similar to Engle (2000), the effective duration is defined as the time interval between two trades with a price change (trades with the same price are aggregated). The first trade in both the morning as well as the afternoon sessions is deleted. Descriptive statistics of the 10 stocks are shown in Table 1. Basically, durations and spread are smaller for large and actively traded stocks.

< Insert Table 1: Sample stocks >

Similar to microstructure variables such as spread and volume, duration has a strong intraday periodicity; see, e.g., Engle and Russell (1998), Andersen and Bollerslev (1997) and Martens (2001). We apply the smoothing method of Engle and Russell to remove the intraday periodicity of duration and volume series. Taking the duration series Dur_i as an example, the smoothing method is,

$$x_i = Dur_i / s(t_i)\,, \tag{8}$$

$$s(t_i) = \sum_{j=1}^{N} D_j (\hat{a}_j + \hat{b}_j (t_i - T_j) + \hat{c}_j (t_i - T_j)^2 + \hat{d}_j (t_i - T_j)^3)\,. \tag{9}$$

Here, Dur_i is the i^{th} duration, $s(t_i)$ is the periodic factor, x_i is the adjusted duration, N is the number of sample sections in each trading day, and T_j is the corresponding specified time point of the sample section. Since each section lasts for half an hour, $N = 8$ and T_j ($j = 1,...,8$) refers to 0930, 1000, . . . , 1430, 1500. Finally, D_j is a dummy variable that attributes each duration to a specified section ($D_j = 1$ if the i^{th} duration takes place in the j^{th} section, $D_j = 0$ otherwise), and t_i

is the time at which the i^{th} duration is taking place. The estimators of $\hat{a}_j$, $\hat{b}_j$, $\hat{c}_j$, $\hat{d}_j$ can be obtained from regression of equation (9), and the fitted $s(t_i)$ is used for duration periodicity adjustment. The empirical intraday patterns of duration and volume are found to be qualitatively similar to those of Engle (2000) and other literature on intraday seasonality.

4.2 Duration and trading activity

Here, we provide statistics on volume and spread for a large stock 600019 and a small stock 600063 in order to preliminarily assess the role played by duration in the process of price formation. We first consider the scenarios of high and low volume. Then for each observation i, duration x_i is sorted into short-medium-long groups and price volatility $|u_i|$ is sorted into small-medium-large groups. Relevant statistics only for the short and long duration groups as well as the small and large size groups are reported in Table 2 below.

< Insert Table 2: Duration and trading activity >

The figures in Panel A are average number of shares per unit time (in second) transacted between two trades that result in a price change. It is clear that short duration in SHSE coincides with high trading intensity, regardless of whether it is in a high or low volume state. Moreover, as can be seen from Panel B, the average total volume statistics reveal that, despite their short time span, short durations account for a significant portion of trading volume. Panel C provides figures on the spread, defined as asks minus bids quotes. Consistent with existing literature, when duration is short and price is volatile, trading is especially active and spread is large. Spread can be decomposed into two parts, asymmetry cost and inventory cost; see, e.g., Madahavan, Richardson and Roomans (1997). Higher spread is thus often regarded as higher asymmetry cost, which implies a higher likelihood of presence of informed traders.

Finally, R is defined as the ratio of large-$|u_i|$ to small-$|u_i|$ figures. R measures the relative increase in trading intensity when price becomes volatile. So we can see that in the high-volume state, trading activity intensifies considerably for both stocks when price varies considerably. For example in Panel A, value of R is 1.53 for stock 600063. The corresponding R value when market is quiet with low trading volume is only 1.15. Similar pattern is also observed for stock 600019 as well as in Panel B. Though no formal inference can be made based on R statistics, they do suggest that the trading dynamics during a short duration in the high-volume state can be rather different from those in the low-volume state.

4.3 Linear log-ACD estimates

Throughout the paper, estimation of log-ACD parameters uses maximum likelihood estimation (MLE) method. We assume that the innovation process $\{z_i\}$ in Equation (3) follows an exponential distribution, and the associated likelihood function is given by

$$L(\Theta) = -\sum_{i=1}^{N}\left(\log(\varphi_i) + \frac{x_i}{\varphi_i}\right), \tag{10}$$

where N is the number of observations and Θ is the vector of parameters.

Here, we shall first consider the estimates of linear log-ACD model given by Equation (5). Consistent with Manganelli (2005) and most literatures, it can be seen from Table 3 that the volume coefficient ξ is significantly negative for all stocks except 600050. That is, large volume in SHSE is associated with high trading activity or short duration. The empirical result for the volatility coefficient γ is rather different. Except for 600900, larger price volatilities tend to be followed by longer durations. Adopting the Dufour and Engle's (2000) view on liquidity, positive volatility coefficients suggest that SHSE is dominated by liquidity traders.

< Insert Table 3: Linear log-ACD model >

4.4 Nonlinear log-ACD estimates

The above empirical results contradict with most literature, noticeably Dufour and Engle (2000) and Manganelli (2005) who find for NYSE stocks concentrated trading is associated with an increased presence of informed traders. For the China stock market, Fang and Wang (2005) also find that informed trading leads to short durations. Our proposed nonlinear log-ACD models resolve this contradiction and suggest that a linear log-ACD is likely a model misspecification.

< Insert Table 4: Nonlinear log-ACD model I >

Table 4 provides estimates of our first nonlinear log-ACD model specified by (6). The most striking difference lies in the fact that all volatility coefficients γ_H are now significantly negative except for stock 600697. So when volume is high, short duration (high trading intensity) implies a higher number of informed traders on the China stock market. When volume is low, all γ_L's are significantly positive. According to Dufour and Engle (2000) and Seppi (1997), a liquidity driven trade would normally have a lower impact on price, and trade-related information takes longer to be fully incorporated into prices. That is, when market is dominated by liquidity traders, large price change corresponds to longer duration. Our results in Table 4 thus suggest that it is liquidity traders who account for active trading when market is in a low-volume state. Finally, we remark that preference of the nonlinear model over its linear counterpart is supported by 9 out of 10 significant likelihood ratio statistics.

From Table 4, we can see that the volume coefficients for less liquid stocks are negative, which is consistent with the fact that high volume coincides with short duration. For the large stocks, 4 out of 5 ξ's are positive (3 of them significant). When trading volume is high, this may be regarded as signs supporting Foster and Viswanathan's (1990) prediction that presence of informed traders deters liquidity traders and results in lower trading volume (at short duration). This explanation does not hold, however, when the trading volume is below the average level. To allow for a different dynamics between volume and duration when trading is less active, we estimate our second nonlinear log-ACD model given by (7).

< Insert Table 5: Nonlinear log-ACD model II >

If Foster and Viswanathan were correct, we would expect to see a negative relationship between volume and duration when volume is low, but the relationship would become positive when volume is high. As can be seen from Table 5 above, this indeed turns out to be the case. First of all, when trading volume is low, all low-volume coefficients (ξ_L) are negative. When trading volume is high, 3 out of 5 large stocks have significantly positive high-volume coefficients (ξ_H). Though the other two large stocks have negative ξ_H, only one of them is significant.

Finally, we remark that that both the SHSE and Paris Bourse are purely order driven markets. It is interesting to see that the linear log-ACD specified by (5) yields similar (incorrect) inference as Cellier (2003). The fact that the sample NYSE stocks analyzed by Dufour and Engle (2000) and Manganelli (2005) are from an order driven market with specialists suggests that there could be a subtle difference in the dynamics of the two different market structures.[9] The important point here is that when an appropriate nonlinear model is used, the underlying economics are found to be the same for both NYSE and SHSE.

---

[9] Specialists at NYSE have dual broker-dealer roles. They trade as brokers for their clients while acting as dealers for their own accounts; see Harris and Panchapagesan (2005) for more details on the market structure of NYSE.

4.5 Robustness of results

Managanelli (2005) proposes a VAR framework to study duration, volume and returns simultaneously, which has the advantage of taking into account feedbacks among the variables concerned. We do not carry out such analysis here, partly because our main objective does not include impulse response function analysis, for which feedbacks should be more rigorously dealt with. More importantly, similar to Dufour and Engle (2000) who treats duration exogenously in their price and trade model, we believe our results are not affected by the issue of simultaneity and are qualitatively valid. This is supported by various analyses that have been carried out to check the robustness of the aforementioned empirical results. Due to constraint of space, we do not report all the numerical results.[10] Overall, the following analyses show that the findings of this paper are stable and robust.

<u>The intraday pattern of volatility</u>

Similar to the duration and volume, volatility has an L-shaped intraday pattern. To make sure that our findings are not spurious results due to intraday seasonality in volatility, we apply the smoothing methods given by (8) and (9) to the volatility series $|u_i|$, and re-estimate our nonlinear log-ACD models. Results obtained are qualitatively similar to the above findings.

<u>The influence of other factors on duration</u>

Present theoretical or empirical works on duration find that there are other factors that may affect the dynamics of the duration besides volume and volatility. We follow Bauwens and Giot (2000) to consider more control variables in our nonlinear log-ACD models. In particular, buy-to-sell ratio and spread are augmented

[10] Detailed results are available from the authors upon request.

to (7):

$$
\begin{aligned}
\varphi_i &= E(x_i \mid \Omega_{i-1}) \\
&= \mu + \kappa \ln(x_{i-1}) + \eta\varphi_{i-1} \\
&\quad + (\xi_L Volume_{i-1} + \gamma_L |u_{i-1}|) \cdot V_{i-1}^{L} + (\xi_H Volume_{i-1} + \gamma_H |u_{i-1}|) \cdot V_{i-1}^{H} \\
&\quad + \theta_1 BSration_{i-1} + \theta_2 Spread_{i-1}.
\end{aligned} \tag{11}
$$

In above, $BSratio_i$ is the ratio of buyer-initiated volume to the total volume cumulated from the first trade after market open to the current trade i. $BSratio_i$ can be regarded as a proxy for the stock price trend. Generally speaking, if $BSratio_i$ is larger than 0.5, it implies that the stock price is on an upward trend; otherwise, it is on downward trend. The other variable, $Spread_i$, can be regarded as a proxy for the presence of asymmetric information. While an uninformed liquidity trader may be deterred from trading by a large spread, a competing informed trader would be keen to trade as soon as possible before his private information become obsolete. So in the former case, duration will be lengthened, whereas in the latter case, duration will be shortened. Anyway, spread is an important variable that needs to be considered. Again, estimates of (11) reveal the same message as in the last section.

<u>Ljung-Box statistics</u>

Table 6 below lists Ljung-Box (LB) statistics for the original duration series (after adjustment for intraday periodicity) and its estimated residuals using nonlinear log-ACD model given by (7). 50 lags are used in calculating the LB statistics. We can see that there is a huge reduction in the LB statistics after fitting the nonlinear log-ACD(2,2) specification. Though most of the LB statistics are significant, two remarks are made here. First is that this is a common feature with long time series. Engle (2000) and Manganelli (2005) also face similar data fitting problems. Indeed, our data is extremely long: the longest time series has 232,364 observations, compared to 52,146 observations in Engle (2000) and 88,917 observations in

Manganelli (2005). Second, more importantly, our estimated auxiliary models with longer lags reveal the same conclusions.

< Insert Table 6: Ljung-Box statistics >

5. Conclusions

The empirical evidence obtained in this paper on the Shanghai Stock Exchange (SHSE) contributes to the literature on the microstructure of financial markets. The fact that both the SHSE and Paris Bourse are purely order driven and that both Cellier (2003) and our linear log-ACD analysis provides similar (incorrect) inference suggest there is a subtle difference in the learning process between a centralized purely-order-driven market and an order-driven-specialists market such as NYSE.[11] However, the economics that underlie the trading of SHSE are the same: a higher trading activity coincides with an increased presence of informed traders on the market. This observation is made possible by using a nonlinear log-ACD model that identifies the different dynamic of informed trading from that of liquidity trading. Since an informed trader will be equally keen to trade on an illiquid stock if there is private information to be exploited, it is probable that the Manganelli's (2005) findings (on the presence of informed trading) can be extended to less frequently traded stocks if nonlinearity is taken into account. We also validate the claim made by Dufour and Engle (2000) that their findings support Foster and Viswanathan's (1990) model. This is evidenced from the empirical results of our nonlinear econometric model: when volume is high, short duration (high trading intensity) coincides with lower volume, suggesting that the presence of informed traders deters discretionary liquidity traders from trading.

[11] Both the samples of TORQ and TAQ databases used by Dufour and Engle (2000) and Manganelli (2005) respectively use NYSE transaction data.

Reference

Admati, A.R., Pfleiderer, P., 1988. A theory of intraday patterns: Volume and price variability. Review of Financial Studies 1, 3-40.

Anand, A., Chakravarty, S., Martell, T., 2005. Empirical evidence on the evolution of liquidity: Choice of market versus limit orders by informed and uninformed traders. Journal of Financial Markets 8, 289 – 309.

Andersen, T., 1996. Return volatility and trading volume: An information flow interpretation of stochastic volatility. Journal of Finance 51, 169 – 204.

Andersen, T.G., Bollerslev, T., 1997. Intraday periodicity and volatility persistence in financial markets, Journal of Empirical Finance 4, 115-158.

Bauwens, L., Giot, P., 2000. The Logarithmic ACD model: An application to the Bid-Ask quote process of three NYSE stocks. Annales d'Economie et de Statistique 60, 117-147.

Brennan, M.J., Subrahmanyam, A., 1996. Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. Journal of Financial Economics 41, 441 – 464.

Cellier, A., 2003. Duration-Volume-Volatility: A VAR GARCH model, Working paper of JEREM.

Chakravarty, S., 2001. Stealth-trading: which traders' trades move stock prices? Journal of Financial Economics 34, 281 – 306.

Chan, K., Menkveld, A.J., Yang, Z., 2007. The informativeness of domestic and foreign investors' stock trades: Evidence from the perfectly segmented Chinese market. Journal of Financial Markets 10, 391 – 415.

Dacorogna, M., Gencay, R., Müller, U.A., 2000. An introduction to high-frequency finance, Academic press, San Diego and London.

Diamond, W., Verrecchia, R.E., 1987. Constraints on Short-selling and asset price adjustment to private information, Journal of Financial Economics 18, 277-311.

Dufour, A., Engle, R.F., 2000. Time and the price impact of a trade, The Journal of Finance 55, 2467-2498.

Easley, D., O'Hara, M., 1987. Price, trade size and information in securities markets, Journal of Financial Economics 19, 69-90.

Easley, D., O'Hara, M., 1992. Time and the process of security price adjustment, Journal of Finance 47, 577-606.

Easley, D., Kiefer, N.M., O'Hara, M., 1997. One day in the life of a very common stock, Review of Financial Studies 10, 805-835.

Engle, R.F., 2000. The Econometrics of Ultra-High-Frequency data, Econometrica 68, 1-22.

Engle, R.F., Russell, J.R., 1998. Autoregressive conditional duration: A new model for irregularly spaced transaction data, Econometrica 66, 1127-1162.

Fang, Z.M., Wang, C.F., 2005. Empirical study of Chinese stock market microstructure based on irregularly spaced data. Journal of Systems Engineering 20, 89-93. (In Chinese)

Feng, L., Seasholes, M.S., 2004. Correlated trading and location. Journal of Finance 59, 2117 – 2144.

Foster, F.D., Viswanathan, S., 1990. A theory of the interday variations in volume, variance, and trading costs in securities markets. Review of Financial Studies 3, 593-624.

Foster, F.D., Viswanathan, S., 1993. Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models. Journal of Finance 48, 377-399.

Ghysels, E., Gourieroux, C., Jasiak, J., 2004. Stochastic volatility duration models, Journal of Econometrics 119, 413-433.

Grossman, S.J., Miller, M.H., 1988. Liquidity and market structure. Journal of Finance 43, 617 – 637.

Harris, L.E., 1990. Liquidity, trading rules, and electronic trading systems. Monograph Series in Finance and Economics 4, Stern School of Business, New York.

Harris, L.E., Panchapagesan, V., 2005. The information content of the limit order book: evidence from NYSE specialist trading decisions. Journal of Financial Markets 8, 25 – 67.

Hasbrouck, J., 1991. Measuring the information content of stock trades. Journal of Finance 46, 179 – 207.

Holden, C.W., Subrahmanyam, A., 1992. Long-lived private information and imperfect competition. Journal of Finance 47, 247 – 270.

Kyle, A.S., 1985. Continuous auctions and insider trading. Econometrica 53, 1315 – 1335.

Lamoureux, C., Lastrapes, W., 1990. Heteroscedasticity in stock return data: Volume versus GARCH effects. Journal of Finance 45, 221 – 229.

Lee, C.M.C., Radhakrishna, B., 2000. Inferring investor behaviour: Evidence from TORQ data. Journal of Financial Markets 3, 83 – 111.

Madahavan, A., Richardson, M., Roomans, M., 1997. Why do security prices change? A transaction level analysis of NYSE stocks, Review of Financial Studies 10, 1035-1064.

Manganelli, S., 2005. Duration, volume and volatility impact on trades, Journal of Financial markets 8, 377-399.

Martens, M., 2001. Forecasting daily exchange rate volatility using intraday returns, Journal of International Money and Finance 20, 1-23.

Ng, L., Wu, F., 2007. The trading behavior of institutions and individuals in Chinese equity markets. Journal of Banking & Finance 31, 2695 – 2710.

O'Hara, M., 1995. Market microstructure theory, Cambridge, MA: Blackwell Publishers.

Seppi, D.J., 1997. Liquidity provision with limit orders and a strategic specialist. Review of Financial Studies 10, 103 – 150.

Xu, C.K., 2000. The microstructure of the Chinese stock market. China Economic Review 11, 79 – 97.

Zhang, M.Y., Russell, J.R., Tsay, R.S., 2001. A nonlinear autoregressive conditional duration model with applications to financial transaction data, Journal of Econometrics 104, 179-207.

Table 1: Sample stocks

Summary statistics for the sample stocks are provided in the table below. Price refers to the transacted price, duration is the time between transactions that result in a price change, spread is simply asks minus bids, and volume is the number of shares transacted in each interval.

Stock	Industry Code	Company name	Average Price	Average Duration	Average Spread	Average Volume	No. of Obs.
Large (liquid) stocks							
600019	C65	Baoshan Iron & Steel Co., Ltd.	6.47	34.0	0.011	59864.39	142891
600028	B03	China Petroleum & Chemical Corp.	4.72	32.6	0.011	71231.02	154545
600036	I01	China Merchants Bank Co., Ltd.	9.52	27.1	0.012	33201.39	186765
600050	G85	China United Tele. Corp. Ltd.	3.56	21.8	0.010	109833.3	232364
600900	D01	China Yangtze Power Co., Ltd.	8.84	25.9	0.011	39535.97	190278
Small (illiquid) stocks							
600063	C47	Amhui Wanwei High-Tech Mat. Ind. Co. Ltd.	5.41	142.3	0.015	5357.67	32150
600172	C61	Henan Huanghe Whirlwind Co., Ltd.	5.53	85.4	0.013	6733.03	54719
600426	C47	Shandong Hualu-Hengsheng Chem. Co. Ltd.	11.22	64.3	0.022	5760.99	72386
600697	H11	Chang Chun Eurasia Group Co., Ltd.	6.00	160.0	0.018	6582.38	28645
600877	C75	China Jialing Industry Co., Ltd. (Group)	4.18	138.1	0.016	9286.49	33663

Table 2: Duration and trading activity

Volume and spread statistics are provided for stocks 600019 and 600063. At each observation *i*, duration x_i is sorted into short-medium-long groups; price volatility $|u_i|$ is sorted into small-medium-large groups. Relevant statistics only for the short and long duration groups as well as small and large price variation groups are reported below for the high and low volume scenarios. The figures in Panel A are average number of shares per second transacted in an effective duration. Panel B tabulates the average volume statistics, whereas Panel C provides figures on the spread, defined as asks minus bids quotes. Finally, R is the ratio of large-$|u_i|$ to small-$|u_i|$ figures. R measures the relative increase in trading intensity when price becomes volatile.

		600019 (Large stock)			**600063 (Small stock)**		
		Small $\lvert u_i \rvert$	**Large** $\lvert u_i \rvert$	**R**	**Small** $\lvert u_i \rvert$	**Large** $\lvert u_i \rvert$	**R**
Panel A: Average volume per unit time							
High volume	**Short** x_i	23046	36685	1.59	885	1355	1.53
	Long x_i	3591	4252	1.18	55	52	0.95
Low volume	**Short** x_i	1477	1900	1.29	139	160	1.15
	Long x_i	365	363	0.99	10	10	1.00
Panel B: Average volume per price change							
High volume	**Short** x_i	160183	229588	1.43	10646	16140	1.52
	Long x_i	201677	260661	1.29	13007	13655	1.05
Low volume	**Short** x_i	9687	11737	1.21	1551	1788	1.15
	Long x_i	17324	20241	1.17	2291	2343	1.02
Panel C: Average spread							
High volume	**Short** x_i	0.0139	0.0178		0.0123	0.0223	
	Long x_i	0.0105	0.0115		0.0128	0.0189	
Low volume	**Short** x_i	0.0103	0.0126		0.0114	0.0227	
	Long x_i	0.0105	0.0115		0.0139	0.0197	

Table 3: Linear log-ACD model

Estimates of the linear log-ACD model given by equation (5),

$$\varphi_i = \mu + \kappa \ln(x_{i-1}) + \eta\varphi_{i-1} + \xi Volume_{i-1} + \gamma|u_{i-1}|,$$

are provided in the table below. The volume and volatility coefficients are highlighted for their relevance. The values in parentheses are p-values of the estimated coefficients.

Stock	μ	κ	η	ξ	γ	Likelihood
Large (liquid) stocks						
600019	0.013 (0.000)	0.029 (0.000)	0.974 (0.000)	**-0.094 (0.000)**	**1.349 (0.000)**	-126995.9
600028	0.012 (0.000)	0.024 (0.000)	0.978 (0.000)	**-0.163 (0.000)**	**0.934 (0.000)**	-140739.6
600036	0.015 (0.000)	0.031 (0.000)	0.969 (0.000)	**-0.281 (0.000)**	**0.639 0.018**	-175416.9
600050	0.004 (0.000)	0.010 (0.000)	0.990 (0.000)	**0.024 (0.000)**	**0.755 (0.000)**	-228673.5
600900	0.014 (0.000)	0.025 (0.000)	0.977 (0.000)	**-0.101 (0.000)**	**-2.341 (0.000)**	-179604.9
Small (illiquid) stocks						
600063	0.046 (0.000)	0.068 (0.000)	0.922 (0.000)	**-1.954 (0.000)**	**4.475 (0.000)**	-24167.7
600172	0.044 (0.000)	0.059 (0.000)	0.934 (0.000)	**-1.668 (0.000)**	**1.721 (0.000)**	-44136.7
600426	0.034 (0.000)	0.063 (0.000)	0.937 (0.000)	**-0.710 (0.000)**	**2.394 (0.000)**	-58080.9
600697	0.036 (0.000)	0.061 (0.000)	0.931 (0.000)	**-1.088 (0.000)**	**3.884 (0.000)**	-22467.9
600877	0.043 (0.000)	0.074 (0.000)	0.913 (0.000)	**-0.963 (0.000)**	**3.514 (0.000)**	-27563.2

Table 4: Nonlinear log-ACD model I

Estimates of the nonlinear log-ACD model given by equation (6),

$$\varphi_i = \mu + \kappa \ln(x_{i-1}) + \eta \varphi_{i-1} + \xi Volume_{i-1} + \gamma_L |u_{i-1}| \cdot V_{i-1}^L + \gamma_H |u_{i-1}| \cdot V_{i-1}^H,$$

are provided. V_i^L (V_i^H) is an indicator that equals to one if *Volume* is below (above) the average value. The values in parentheses are p-values of the estimated coefficients whereas the values in square brackets are likelihood ratio (LR) statistics for model specification (6) over (5). Under the null hypothesis, the LR statistics are distributed as Chi-squared with 2 degree of freedom with 5.99 (9.21) as 5% (1%) critical value.

Stock	μ	κ	η	ξ*100	γ_L	γ_H	Likelihood
Large (liquid) stocks							
600019	0.009 (0.000)	0.028 (0.000)	0.974 (0.000)	**0.0841** **(0.000)**	**5.075** **(0.000)**	**-3.051** **(0.000)**	-126916.8 [158.2]
600028	0.008 (0.000)	0.025 (0.000)	0.977 (0.000)	**0.0433** **(0.001)**	**4.630** **(0.000)**	**-3.178** **(0.000)**	-140651.1 [177.0]
600036	0.013 (0.000)	0.031 (0.000)	0.968 (0.000)	**-0.193** **(0.000)**	**3.796** **(0.000)**	**-1.824** **(0.000)**	-175371.8 [90.2]
600050	0.003 (0.000)	0.011 (0.000)	0.989 (0.000)	**0.044** **(0.000)**	**1.383** **(0.000)**	**-0.756** **(0.000)**	-228632.1 [82.8]
600900	0.012 (0.000)	0.025 (0.000)	0.977 (0.000)	**0.015** **(0.233)**	**1.194** **(0.005)**	**-6.732** **(0.000)**	-179548.2 [113.4]
Small (illiquid) stocks							
600063	0.039 (0.000)	0.069 (0.000)	0.921 (0.000)	**-1.342** **(0.000)**	**8.285** **(0.000)**	**-1.696** **(0.057)**	-24144.3 [46.8]
600172	0.039 (0.000)	0.060 (0.000)	0.933 (0.000)	**-1.180** **(0.000)**	**6.175** **(0.000)**	**-6.069** **(0.000)**	-44095.7 [82.0]
600426	0.031 (0.000)	0.064 (0.000)	0.936 (0.000)	**-0.500** **(0.000)**	**5.790** **(0.000)**	**-1.525** **(0.005)**	-58057.0 [47.8]
600697	0.035 (0.000)	0.062 (0.000)	0.930 (0.000)	**-0.984** **(0.000)**	**4.684** **(0.000)**	**2.898** **(0.000)**	-22466.3 [3.2]
600877	0.039 (0.000)	0.077 (0.000)	0.908 (0.000)	**-0.561** **(0.000)**	**7.744** **(0.000)**	**-4.124** **(0.000)**	-27509.4 [107.6]

Table 5: Nonlinear log-ACD model II

Estimates of the nonlinear log-ACD model given by equation (7),

$$\varphi_i = \mu + \kappa \ln(x_{i-1}) + \eta\varphi_{i-1} + (\xi_L Volume_{i-1} + \gamma_L |u_{i-1}|)V_{i-1}^L + (\xi_H Volume_{i-1} + \gamma_H |u_{i-1}|)V_{i-1}^H,$$

are provided. V_i^L (V_i^H) is an indicator that equals to one if *Volume* is below (above) the average value. The values in parentheses are p-values of the estimated coefficients whereas the values in square brackets are likelihood ratio (LR) statistics for model specification (7) over (6). Under the null hypothesis, the LR statistics are distributed as Chi-squared with 2 degree of freedom with 5.99 (9.21) as 5% (1%) critical value.

Stock	μ	κ	η	ξ_L*100	ξ_H*100	γ_L	γ_H	Likelihood
Large (liquid) stocks								
600019	0.015 (0.000)	0.029 (0.000)	0.971 (0.000)	**-3.053** **(0.000)**	**0.082** **(0.000)**	**5.321** **(0.000)**	**-3.964** **(0.000)**	-126810.2 [213.3]
600028	0.012 (0.000)	0.025 (0.000)	0.974 (0.000)	**-2.453** **(0.001)**	**0.069** **(0.000)**	**4.849** **(0.001)**	**-3.903** **(0.000)**	-140560.4 [181.4]
600036	0.019 (0.000)	0.032 (0.000)	0.966 (0.000)	**-3.396** **(0.000)**	**-0.234** **(0.000)**	**4.677** **(0.000)**	**-2.328** **(0.000)**	-175262.0 [309.8]
600050	0.004 (0.000)	0.011 (0.000)	0.988 (0.000)	**-0.623** **(0.000)**	**0.045** **(0.000)**	**1.410** **(0.000)**	**-0.635** **(0.000)**	-228597.0 [153.0]
600900	0.018 (0.000)	0.026 (0.000)	0.974 (0.000)	**-3.618** **(0.000)**	**-0.005** **(0.589)**	**1.784** **(0.489)**	**-8.089** **(0.000)**	-179328.0 [553.8]
Small (illiquid) stocks								
600063	0.048 (0.000)	0.069 (0.000)	0.921 (0.000)	**-4.970** **(0.000)**	**-1.329** **(0.000)**	**9.194** **(0.000)**	**-3.133** **(0.057)**	-24134.5 [19.6]
600172	0.050 (0.000)	0.060 (0.000)	0.932 (0.000)	**-5.697** **(0.000)**	**-1.205** **(0.000)**	**7.940** **(0.000)**	**-8.846** **(0.000)**	-44069.8 [51.8]
600426	0.039 (0.000)	0.065 (0.000)	0.935 (0.000)	**-3.890** **(0.000)**	**-0.583** **(0.000)**	**6.584** **(0.000)**	**-2.511** **(0.005)**	-58036.6 [40.8]
600697	0.047 (0.000)	0.061 (0.000)	0.930 (0.000)	**-5.585** **(0.000)**	**-1.092** **(0.000)**	**5.450** **(0.000)**	**1.784** **(0.000)**	-22450.3 [32.0]
600877	0.055 (0.000)	0.079 (0.000)	0.905 (0.000)	**-7.132** **(0.000)**	**-0.641** **(0.000)**	**9.043** **(0.000)**	**-5.960** **(0.000)**	-27486.1 [46.6]

Table 6: Ljung-Box statistics

The table below lists Ljung-Box statistics for the original duration series (after adjusted for intraday periodicity) and its estimated residuals using nonlinear log-ACD model given by (7) with various ACD auxiliary specifications. 50 lags are used in calculating the Ljung-Box statistics. The corresponding critical values at 5% and 1% significance levels are 67.5 and 76.2 respectively.

	Large (liquid) stocks				
	600019	600028	600036	600050	600900
Original duration	70568	74044	97774	30857	79403
ACD(1,1) residuals	334	2396	1392	1205	2175
ACD(2,2) residuals	74	622	370	344	335
ACD(3,3) residuals	72	586	448	313	332

	Small (illiquid) stocks				
	600063	600172	600426	600697	600877
Original duration	62773	102426	173287	51010	44999
ACD(1,1) residuals	179	290	715	230	207
ACD(2,2) residuals	181	282	760	230	207
ACD(3,3) residuals	151	257	653	170	179