



UNIVERSITY  
OF WOLLONGONG  
AUSTRALIA

University of Wollongong  
Research Online

---

Faculty of Commerce - Papers (Archive)

Faculty of Business

---

2009

# The probability of informed trading based on VAR model

Min Xu

*Beihang University*, xumin\_0828@sina.com

Shancun Liu

*Beihang University*, liushancun@buaa.edu.cn

---

## Publication Details

Xu, M. & Liu, S. (2009). The probability of informed trading based on VAR model. *International Conference on Business Intelligence and Financial Engineering, 2009. BIFE '09* (pp. 750-753). Piscataway, USA: IEEE.

Research Online is the open access institutional repository for the University of Wollongong. For further information contact the UOW Library:  
research-pubs@uow.edu.au

---

# The probability of informed trading based on VAR model

## **Abstract**

The paper researches the representative variable of the probability of informed trading, selecting CCER high-frequency trading data of Shanghai stock exchange from 2003.7.1 to 2003.12.31, adopting VAR model. Different from previous studies, the paper firstly accounts for the dynamic relationship between trade and price. Then, the content of information in trading volume, duration and trading direction are considered in our model. Finally, it gets the probability of informed trading and analyzes this variable. The results show: the probability of informed trading is about 0.172713; the more asymmetric information is, the larger spread is; the probability of informed trading is the well-known U-shape; it is the biggest before the announcement.

## **Keywords**

probability of informed trading, informed trader, spread, intraday pattern, announcement

## **Disciplines**

Business | Social and Behavioral Sciences

## **Publication Details**

Xu, M. & Liu, S. (2009). The probability of informed trading based on VAR model. International Conference on Business Intelligence and Financial Engineering, 2009. BIFE '09 (pp. 750-753). Piscataway, USA: IEEE.

## The Probability of Informed Trading based on VAR Model

Min Xu

School of Economics and Management,  
Beihang University,  
Beijing 100083, China  
School of Accounting & Finance,  
Wollongong University, NSW 2500, Australia  
xumin\_0828@sina.com

Shancun Liu

School of Economics and Management,  
Beihang University,  
Beijing 100083, China  
liushancun@buaa.edu.cn

**Abstract**—The paper researches the representative variable of the probability of informed trading, selecting CCER high-frequency trading data of Shanghai Stock Exchange from 2003.7.1 to 2003.12.31, adopting VAR model. Different from previous studies, the paper firstly accounts for the dynamic relationship between trade and price. Then, the content of information in trading volume, duration and trading direction are considered in our model. Finally, it gets the probability of informed trading and analyzes this variable. The results show: the probability of informed trading is about 0.172713; the more asymmetric information is, the larger spread is; the probability of informed trading is the well-known U-shape; it is the biggest before the announcement.

**Keywords**—*probability of informed trading; informed trader; spread; intraday pattern; announcement*

### I. INTRODUCTION

Based on asymmetric information theory, researches yield two important empirical predictions: (1) asymmetric information and bid-ask spread have positive relationship, and (2) asymmetric information has positive impact on the price effect of trade-related information. For the researches on the first predications, the popular approach is to find out the representative variable for asymmetric information, such as spread which has limits, because of the dispersal. Other researches on the second aspect focus on VAR model from Hasbrouck (1991) and its

related models.

Hasbrouck (1991) found out the change of price depends on trading size and the direction. Glosten and Milgrom (1985), Diamond and Verrecchia (1987) explored the importance of time. In Glosten and Milgrom (1985) research, they focused on the effect of direction of trading on asymmetric information. In summary, volume, duration and direction all have effects on price and contain essential information about stock value. The information inflow into the market through informed traders.

How many potential informed traders are in the market and how much real information they have? These questions refer to the probability of informed trading. Easley, Kiefer, O' Hara and Paperman (1996) (EKOP model) proposed PIN as a direct measurement for probability of informed trading, however their researches only took trading volume into consideration.

We investigate the price permanent effect on stock return, adopting Hasbrouck (1991) VAR model, considering the lagged endogenetic variables, putting duration and direction in the model. Based on these, the return was divided into two parts and derives the portion aroused by informed trading. Then, the paper acquires the representative variable of probability of informed trading and analyzes it.

Section II introduces the model and methodology. The selection of data is in Section III. The estimation of model is in Section IV.

Section V is the conclusion.

## II. MODEL AND METHODOLOGY

### A. Hasbrouck Model

At time  $t$ , the trading  $x_t$  happens and the market maker quotes bid and ask price  $q_t^b$  and  $q_t^a$  later.  $r_t = (q_t^b + q_t^a)/2 - (q_{t-1}^b + q_{t-1}^a)/2$  indicates the revision of quoted price. Hasbrouck assumes that the public information arrives at the market after  $x_t$  and before the revision of quoted price. Hasbrouck proposed the following VAR model:

$$\begin{aligned} r_t &= \sum_{i=1}^5 a_i r_{t-i} + \sum_{i=0}^5 b_i x_{t-i}^0 + v_{1,t} \\ x_t^0 &= \sum_{i=1}^5 c_i r_{t-i} + \sum_{i=1}^5 d_i x_{t-i}^0 + v_{2,t} \end{aligned} \quad (1)$$

$v_{1,t}$  indicates the public information and  $v_{2,t}$  indicates the unexpected trading volume which includes the trading from private information and noise owing to liquidity.  $x_t^0$  indexes trading direction.

### B. Return-volume Model

$$\begin{aligned} R_t &= \sum_{i=1}^5 a_i R_{t-i} + \lambda^Z D_t V_t + \sum_{i=0}^5 (c_i \\ &+ d_i \ln(T_{t-i}) + e_i x_{t-i}^0 \ln(T_{t-i})) V_{t-i} + v_{1,t} \\ V_t &= \sum_{i=1}^5 b_i R_{t-i} + \lambda^V D_{t-1} V_{t-1} + \sum_{i=1}^5 (f_i \\ &+ g_i \ln(T_{t-i}) + k_i x_{t-i}^0 \ln(T_{t-i})) V_{t-i} + v_{2,t} \end{aligned} \quad (2)$$

$R_t$  and  $V_t$  index the standard return and volume, and  $T_t$  indicates the duration (+1 second).  $D_t$  is a dummy variable.

### C. Return Decomposition

The ratio of  $\exp(\sum_{i=1}^5 a_i R_{t-i}) / \exp(R_t)$  measures the uninformed component of return, while  $1 - \exp(\sum_{i=1}^5 a_i R_{t-i}) / \exp(R_t) = \text{inf}_t$  is the informed component of return.

### D. The Probability of Informed Trading and Spread

The microstructure theories figure out

spread has positive relation with private information. The regression is as following:

$$s_t = \delta_0 + \delta_1 \text{inf}_t + \delta_2 V_t + \delta_3 s_{t-1} + e_t \quad (3)$$

Chordia et al. (2002) have shown that order imbalance is also an important indicator of the dealer's inventory pressure. The regressive equation is as following:

$$s_t = \delta_0 + \delta_1 \text{inf}_t + \delta_2 V_t + \delta_3 s_{t-1} + \delta_4 q_t + e_t \quad (4)$$

The effective spread is expressed by  $s_t$  and  $q_t$  indexes the order imbalance.

### E. The Intraday Pattern of Probability of Informed Trading

We can explain the U pattern of spread and volume using the asymmetric information commendably. The high spread in the opening and the closing of the market show high quantity of informed traders. In the theory of Dufour and Engle (2000), the duration is smaller and the trading is more active when there are more informed traders. Meanwhile, Hasbrouck (1991) found the more informed trading is, the more volumes are. No matter from volume and trading frequency, the volumes increase in the opening and closing of the trading day. We suppose the U pattern of probability of informed trading.

### F. The Announcement Impact on the Probability of Informed Trading

The announcement of company public the information and has impact on asymmetric information. For every announcement, we select six days as the announcement days and compare their  $\text{inf}_t$ . The two days before announcement, the announcement day and the day after announcement day, the second and third days after announcement day are defined as before announcement ( $\alpha_b$ ), during announcement ( $\alpha_d$ ) and after announcement ( $\alpha_a$ ) separately.

## III. DATA AND VARIABLES

This study adopts CCER high-frequency trading database of Beijing Sinofin Information Service, and the high-frequency trading data

between Jul. 1, 2003 and Dec. 31, 2003 of SSE 50 Index stocks as sample. We select stock 600602, because of its mid-capitalization. Other stocks' results are similar and we can provide them if needed.

Basing on Engle (2000), the standard return and volume are:  $R_t = \ln(10000 \times r_t / T_t + 1)$  and  $V_t = \ln(v_t / T_t + 1)$ . In the equation  $R_t = \frac{(q_t^b + q_t^a) / 2}{(q_{t-1}^b + q_{t-1}^a) / 2}$ ,  $v_t$  is the volume.

The trading hours of Shanghai Stock Exchange are from every Monday to Friday, 9:30am-11:30am and 13:00pm-15:00pm.  $D_t$  equals 1 when the trading happens between 9:30 and 10:00, otherwise,  $D_t$  equals 0. We use Lee and Ready (1991) method to judge the direction of trading, buying or selling. The paper truncates the lagged items at 5, basing on Akaike Information Criterion and Schwarz Criterion.

In the regression of spread,  $s_t = [(\text{price} - \text{midpoint}) / \text{midpoint}] * 200$ . The midpoint is the average of best bid and best ask price.  $q_t = \sum_{i=1}^m v_{t-i}$ ,  $m$  is the numbers of trading before  $t$  at the same direction and the same price.  $q_t$  is the cumulated bid-ask pressure ( $q_t = 0$ , in the opening of the day or if there is a change of trading direction and trading price).

#### IV. RESULTS AND DISCUSSIONS

##### A. The Representative Variable of the Probability of Informed Trading

We adopt weighted least squares method to estimate the model in (2), because the model exists heteroskedasticity testing by white heteroskedasticity test.  $a_1, a_2, a_3, a_4$  and  $a_5$  are 0.173436, 0.212511, 0.190345, 0.219645 and 0.184952 separately. The probability of informed trading is 0.172713.

##### B. The Regressive Results between Probability of Informed Trading and Spread

We regress the effective spread, probability of informed trading, one lagged spread and volume, and the result is in table 1 (Regression with spread 1). After that, we include the order

imbalance into the regression and the result is in Regression with spread 2.

TABLE I. THE PROBABILITY OF INFORMED TRADING AND SPREAD

Regression with spread 1		Regression with spread 2	
Coef.	Mean	Coef.	Mean
$\delta_0$	0.139655 (10.47506)	$\delta_0$	0.146757 (11.02404)
$\delta_1$	0.0322 (-2.6601)	$\delta_1$	0.0463 (2.952268)
$\delta_2$	-0.02841 (-11.951)	$\delta_2$	-0.03016 (-12.7036)
$\delta_3$	0.000938 (0.360136)	$\delta_3$	-0.02172 (-7.80072)
		$\delta_4$	3.32E-06 (22.80637)

The coefficient of volume is negative and that of informed trading is positive, which is the same as expectation. It shows the increase of volume reduces spread, while the increase of informative trading enlarges spread because the others are afraid of trading with informed traders. The lagged spread has negative impact on spread and order imbalance has insignificant effect on spread.

##### C. The Intraday Pattern

From figure 1, the probability of informed trading has U pattern.

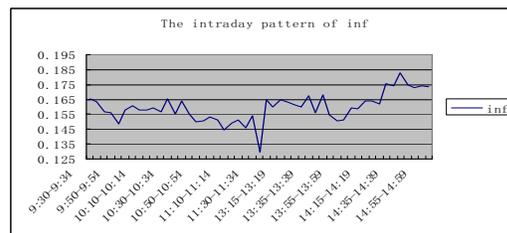


Figure 1. The intraday pattern of the probability of informed trading

In the first hour of opening, the extent of asymmetric information is high, about 15.8%. The institution and informed traders trade quickly and effectively to utilize the cumulated private information. The information becomes public information as their trading. In the middle of the trading day, the probability of informed

trading goes down to 12.9555%. At the last hour of trading, the probability of informed trading increases again, even 18.2958%.

#### D. The Announcement

The quarterly announcement happened on 30<sup>th</sup> October and the effects are shown in table II. The probability of informed trading is highest before the announcement. The probability reduces to 11.0966% as the private information turns to public.

TABLE II. THE PROBABILITY OF INFORMED TRADING BEFORE AND AFTER ANNOUNCEMENT

	Before	On	After	Others
<b>Mean</b>	0.130863	0.125882	0.110966	0.117003
<b>Median</b>	0.22426	0.192318	0.162074	0.185442
<b>Std</b>	0.430681	0.443976	0.375713	0.420791
$\alpha_b=1.118459, \alpha_d=1.075887, \alpha_a=0.948403$				

From the analysis, the publication of announcement has the meanings as a method of information transparenence. The publication of announcement can reduce the informed traders in the market, reduce trading costs, order the market and make the traders trading in a fair market.

#### V. CONCLUSION

The paper investigates volume impact on price, adopting CCER high-frequency trading data from Jul. 1, 2003 to Dec.31, 2003, selecting Shanghai Stock Exchange 600602 as sample, using VAR model in Hasbrouck (1991). After that, we decompose return and achieve the probability of informed trading and analyze the variable.

The empirical results show the probability of informed trading is 0.172713 in the sample. The asymmetric information has positive effect on spread and volume has opposite effect. The probability has U shape in the trading day. Before the announcement, the asymmetric information is highest which shows the transparenence of market after the announcement.

#### ACKNOWLEDGMENT

This research is supported by A

Foundation for the Author of National Excellent Doctoral Dissertation of PR China (No.200466) and National Natural Science Foundation of China (No.70671006).

#### REFERENCE

- [1] Hasbrouck J., "Measuring the Information Content of Stock Trades," *Journal of Finance*, vol. 46(1), pp. 179-207,1991.
- [2] Glosten Lawrence R, Milgrom Paul R, "Bid Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents," *Journal of Financial Economics*, vol. 14, pp. 71-100,1985.
- [3] Diamond D.W and Verrecchia, R.E. "Constraints on Short-Selling and Asset Price Adjustment to Private Information," *Journal of Financial Economics*, vol. 18, pp. 277-311,1987.
- [4] Easley, D.; Kiefer, N.M. and O'Hara, M., "Paperman, J.B. Liquidity, Information and Infrequently Traded Stocks," *Journal of Finance*, vol. 51, pp. 1405-1436, 1996.
- [5] Chordia, T.; Roll, R.and Subrahmanyam, A. "Order Imbalance, Liquidity, and Market Returns," *Journal of Financial Economics*, vol. 65, pp. 111-130, 2002.
- [6] Dufour, A. and Engle, R.F. "Time and the Price Impact of a Trade," *Journal of Finance*, vol. 55, pp. 2467-2498, 2000.
- [7] Engle, R.F. "The Economics of Ultra-High-Frequency Data," *Econometrica*, vol. 68, pp. 1-23, 2000.
- [8] Lee G and Ready M J. "Inferring Trading Direction from Intraday Data," *Journal of Finance*, vol. 46, pp. 733-746, 1991.