Empirical Studies on Market Microstructure Models

Feng Liu, Ruiwen Zhang

Abstract
Market microstructure studies the trading process and the costs of providing transaction services on the short-run security prices. The cost of a trade depends on the asymmetric information possessed by the participants in the trade. Investors are involved in the market for securities as well as market for related information. The generalized Roll model and Kyle's model are the most cited microstructure models. We studies high-frequency intra-day transactional data and perform empirical studies on microstructure models.

Key Words: Asymmetric Information, Strategic trade model, Market Microstructure

1. Introduction

Microstructure theory focuses on how specific trading mechanisms affect the price formation process. In a trading market, financial assets are not transformed but transferred from one investor to another. The field of market microstructure studies the costs for trading securities and the impact of trading costs on the short-run behavior of security prices. Costs are reflected in bid-ask spreads. The literature on perfect market often assumes that markets operate without costs and frictions whereas market microstructure research is to analyze the impact of trading costs and various friction factors. The investors are generally involved in the market for securities and related information. The market for securities deals with the determinants of security prices such as earnings, revenues etc. The market for information deals with the supply and demand of information. It incorporates the incentives of security analysis and related information transfer. The asymmetric information is closely related to transaction services since the cost of a trade depends on the information possessed by the participants in the trade.

In this paper, we perform empirical studies on market microstructure models. When we look at the security price dynamics with respect to market microstructure, our focus has shifted from monthly or daily to minute or tick level with more features such as bid /ask price, bid information, trade price, volume etc. The additional features of price and trading dynamics reflect complexity of microstructure data.

Roll suggests a model of high frequency trade prices which incorporate trading dynamics. This model is fundamental to market microstructure models such that it illustrates the distinction between price movement due to fundamental security value and those attribute to the market organization and trading mechanism. Roll model or generalized Roll model articulates an important aspect of the bid-ask effect on trading price.

Kyle's strategic trading model is one of the celebrated microstructure models. There are large number of literature interpreting or extending Kyle's model. Holden and Subrahmanyam (1992) study the competition among multiple insiders each endowed with perfect private information. Foster and Viswanathan (1996) consider the competition with heterogeneous private signals. Huddart, Steven, Hughes and Levine (2001) study the insider's announcement of his trading volume right after submission. Cochrane (2005), Vayanos and Wang (2009) have surveyed on liquidity and asset pricing.

This paper examines Roll and Kyle's model on market microstructure data. We provide empirical studies and found Kyle's model is more suitable for market microstructure analysis. The rest of the paper is organized as follows: Section 2 presents Roll model and its analysis. Section 3 discuss Kyle's model. We present our empirical studies in Section 4. Finally, section 5 makes concluding comments.

2. Generalized Roll's Model

Roll (1984) suggests a model of high frequency trade prices which incorporate market dynamics. This model is fundamental to many market microstructure models such that it illustrates the distinction between price movement due to fundamental security value and those attribute to market organization and trading mechanism. The former arises from the earning capability and future cash flows of the underlying security, whereas the later are transient due to market behavior. The model provides meaningful economic interpretation, and in some cases, explains the market movement well.

For $t = 1, 2, ...,$

$$p_t = m_t + c\, q_t, \tag{1}$$
$$m_t = m_{t-1} + u_t, \tag{2}$$

which consists of an observation equation (1) and a state evolution equation (2), where m_t denote the martingale efficient price at tth trade, p_t is the trade price. The q_t are direction indictors, which take on the value 1 (buy) or -1 (sell) with equal probability, the shocks $u_1, u_2, ...$ are iid $N(0, \sigma^2)$ random variables, the parameters $c > 0$ and $\sigma > 0$ represent the effective cost and the volatility respectively. The two sequences $\{q_t\}$ and $\{u_t\}$ are assumed to be independent. Note that only $\{p_t\}$ are observed, while $\{m_t\}$ and $\{q_t\}$ are treated as latent variables.

The model implies

$$\Delta p_t = c\Delta q_t + u_t, \tag{3}$$

from which it follows that $c = [-cov(\Delta p_t, \Delta p_{t-1})]^{1/2}$, if $cov(\Delta p_t, \Delta p_{t-1}) < 0$, and $c = 0$, otherwise. The first-order autocovariance is non-zero. Δp_t exhibits volatility and *negative* serial correlation as the result of effective cost. The intuition is: If m_t is fixed so that prices take on only two values, the bid and the ask, and if the current price is the ask, then the price change between the current price and the previous price must be either 0 or $-2c$, and the price change between the next price and the current price must be either 0 or $2c$. The moment estimate is feasible, however, only if the first-order sample autocovariance of the price change is negative.

If the dealers compete to the point where the costs are just covered, the bid and the ask are $m_t - c$ and $m_t + c$, with the spread 2c, a constant. We collect the data of 200,000 trades for MSFT (Microsoft) on Jan, 2 to Jan, 5 2002 from TAQ, the first-order autocovariance is $\hat{\gamma}_1 = -0.00522$. This implies $c = \$0.035$, and bid-ask spread of $2c = \$0.070$; while the estimates from TAQ database shows the bid-ask spread is $\$0.0625$.

3. Kyle's Strategic Model

Kyle proves the existence and uniqueness of a linear equilibrium solution in which the parameters are derived via a set of recursive formulas. One asset that pays off, $v \sim N(p_0, \Sigma_0)$, Σ_0 is value uncertainty. One informed trader and the uninformed investors are placing orders. Trading by the uninformed traders Δu_n is exogenous and normally distributed $N(0, \sigma_u^2 \Delta t_n)$. The informed trader knows the distribution of the uninformed order flow (but not its value) and takes account of his order flow on the market clearing price. The informed trader observes v and submitted order flow Δx_n. The competitive risk-neutral market-maker determines the auction price to reflect the information contained in the aggregated order flow $\Delta y_n = \Delta x_n + \Delta u_n$.

Definition 1 *A sequential auction equilibrium is defined as a pair (X, P) such that the following conditions hold:*

(*C1*) (*profit maximization*) *For $n = 1, ..., N$ and all $X' = (\Delta x'_1, ..., \Delta x'_N)$ with $\Delta x'_i = \Delta x_i$, $i = 1, ..., n-1$, we have*

$$E[\pi_n(X, P)|\mathcal{F}^I_{n-1}] \geq E[\pi_n(X', P)|\mathcal{F}^I_{n-1}]. \tag{4}$$

(*C2*) (*market efficiency*) *For $n = 1, ..., N$ we have*

$$p_n = E(v|\mathcal{F}^U_{n-1}, \Delta y_n). \tag{5}$$

Definition 2 *A sequential auction equilibrium (X, P) is called a linear equilibrium if the component functions of X and P are linear, and a recursive linear equilibrium if there exist parameters $\lambda_1, ..., \lambda_N$ such that*

$$p_n = p_{n-1} + \lambda_n \Delta y_n, \quad n = 1, ..., N. \tag{6}$$

Theorem 1 *There exists a unique linear equilibrium (X, P), represented as a recursive linear equilibrium, characterized by (for $n = 1, ..., N$)*

$$\begin{aligned} \Delta x_n &= \beta_n \, (v - p_{n-1}) \, \Delta t_n, && (7) \\ p_n &= p_{n-1} + \lambda_n \Delta y_n, && (8) \\ \Sigma_n &= Var(v|\mathcal{F}^U_n), && (9) \\ E[\pi_n|\mathcal{F}^I_{n-1}] &= \alpha_{n-1} \, (v - p_{n-1})^2 + \delta_{n-1}; && (10) \end{aligned}$$

Given Σ_0 and σ^2_u, the parameters $\beta_n, \lambda_n, \Sigma_n, \alpha_n, \delta_n$ are the unique solutions to equations

$$\begin{aligned} \alpha_{n-1} &= [4\lambda_n(1 - \alpha_n\lambda_n)]^{-1}, && (11) \\ \delta_{n-1} &= \delta_n + \alpha_n \, \lambda^2_n \, \sigma^2_u \, \Delta t_n, && (12) \\ \beta_n \, \Delta t_n &= (1 - 2\alpha_n\lambda_n) \, [2\lambda_n(1 - \alpha_n\lambda_n)]^{-1}, && (13) \\ \lambda_n &= \beta_n \, \Sigma_n \, \sigma^{-2}_u, && (14) \\ \Sigma_n &= (1 - \beta_n\lambda_n \, \Delta t_n) \, \Sigma_{n-1}, && (15) \end{aligned}$$

subject to $\alpha_N = \delta_N = 0$ and the second order condition

$$\lambda_n(1 - \alpha_n\lambda_n) > 0. \tag{16}$$

4. Empirical Studies

"Trade volume" is the total number of trade orders or trade size at specific time frame, e.g. 10 minutes. Actual volume per trade from real markets vary in quantities. The basic sequential trade model assume one trade quantity in each trade, while Kyle's strategic model does not have such restriction. The trade volume is an important market dynamics. We obtain 10 random firms from CRSP database, data range from Jan 1988 to Dec 2004, and plot the cross-sectional daily stock return over daily trade volume in Fig. 1. The summary statistics is shown in table 1.

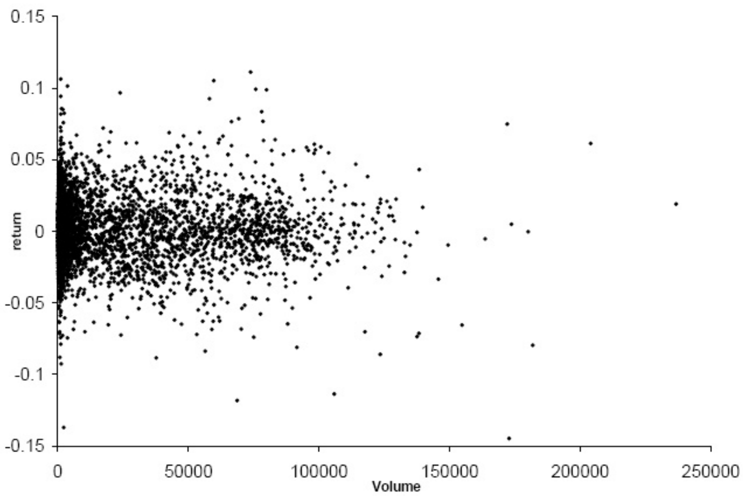

Figure 1: Relationship between trade volume and return

Table 1: Summary statistics of daily returns vs. volume

Variables	Sample period	Observations	Mean	SD
Return	01/1988 - 12/2004	32890	0.00126	0.022964
Volume	01/1988 - 12/2004	32890	22173.54	30640.38

Variables	Max	Min	Skewness	Kurtosis
Return	0.195652	-0.15598	0.06162	3.839
Volume	236675	81	1.63798	2.60493

From Fig. 1, we observe trade volume are quite symmetric across zero return and high volume does not tend to be associated with high return. In Kyle's strategic trading model, the author conjectures a relationship between stock price change and its order flow. In Pasquariello and Vega (2009) empirical study, they address cross-trading effect with daily aggregated order imbalance. We take similar approach with modified setting.

We use intraday, transaction-level data from trade and quotes (TAQ) database during regular market hours (9:30am to 4:00pm). Corresponding daily price data comes from CRSP. We obtain MSFT (Microsoft) transaction level data on January 2001. We define the trades variable as $\{q_t\}$, +1 (buy) or -1 (sell), $t = 1, 2...$ for each transaction. Then we get the signed order flow by multiple trades and order size, denote as $\hat{\Delta y_t}$, where $\Delta y_t = \Delta x_t + \Delta u_t$.

We denote "order imbalance" as the total number of signed order flows at given time period, e.g. 10 minutes. Table 2 shows MSFT first-order autocorrelation of trades from certain days of January 2001.

The results shows the sequence of the order types are more likely to pair with each other, buy after buy, sell after sell. This may imply the asymmetric information are processed by difference traders since the informed traders tend to trade in the direction of his knowledge.

In Kyle's framework, "market depth" is an important characteristics of market dynamics. It refers to the size of an order flow innovation required to change the price in a given amount. The market depth is denoted as λ_n^{-1}, with $p_n = p_{n-1} + \lambda_n \Delta Y_n$ for $n = 1, ..., N$. It deals with order imbalance with respect to the price increment. We do empirical studies on intraday transactional data.

Table 2: Sample first-order autocorrelation of trades

Date	Autocorrelation	N	P-Value
2	0.3306	40237	<0.001
3	0.3049	58859	<0.001
4	0.3331	48718	<0.001
5	0.3341	42902	<0.001
8	0.3559	41755	<0.001
9	0.3543	55388	<0.001
10	0.3837	48945	<0.001
11	0.3232	41093	<0.001
12	0.3545	36273	<0.001
13	0.3653	33158	<0.001

We present the aggregated intraday transaction level order imbalance across trade price increments. Figure 2 illustrates MSFT aggregated order imbalance vs. price changes at each trading date using microstructure data. The correlation between the two series is

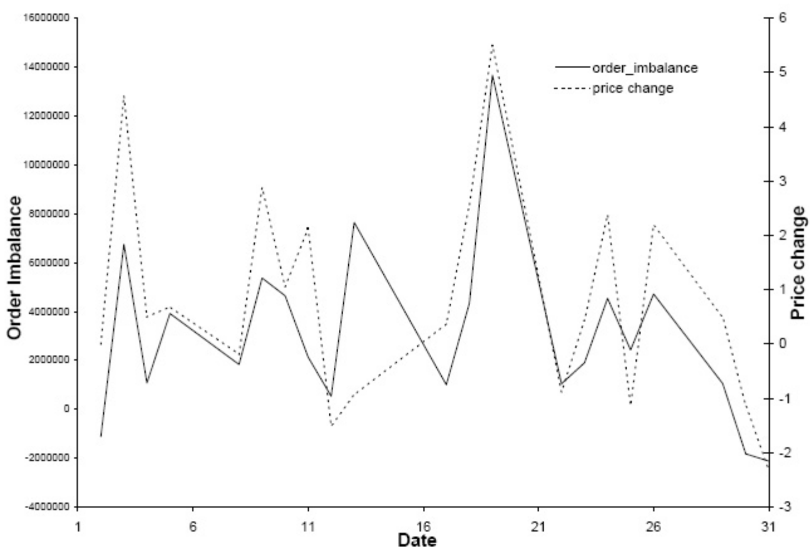

Figure 2: Aggregated intraday order imbalance vs. price change

0.76. The results show strong explanatory power of order imbalance in the price change movement.

We conjecture the market depth (or λ_n) is constant in Kyle's model. We use regression to do our analysis. The t-statistics for λ coefficient is 5.09, with p-value less than 10^{-5} which rejects zero coefficient null hypothesis.

In Kyle's model, the informed trader wants to trade aggressively on her private information, i.e., buy a large quantity if her information is positive. But the market maker knows that if he sells into a large net buy customer, he is more likely to be on the wrong side of the trade. He protects himself by setting a price that is increasing in the order flows. This acts as a brake on the informed trader's desires. If there is an imbalance between buy and sell orders, the market maker makes up the difference in the Kyle's model. The results assert that it is the order imbalance that drives the price movement.

5. Conclusion

This paper performs empirical studies on microstructure models. Our analysis shows Roll model are not adequate and Kyle's model fits market data better. It is the market order imbalance that drives the price movement.

REFERENCES

Cochrane, J. (2005). Liquidity, Trading and Asset Prices. *NBER Reporter*, 1-12.

Foster, F., and S. Viswanathan (1996). “Strategic Trading When Agents Forecast the Forecast of Others”. *The Journal of Finance*, 51, 1437-1478.

Holden, C., and A.Subrahmanyam (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.

Pasquariello, P. and Vega, C. (2009). Strategic cross-trading in the U.S. stock market. Working paper.

Vayanos, D. and Wang, J. (2009). Theories of Liquidity. *Foundations and Trends in Finance*, forthcoming.