UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI
FAKULTET
DEPARTMAN ZA MATEMATIKU I
INFORMATIKU

Nataša Teodorović

Mathematical models of liquidity and price impact function

master thesis

Novi Sad, 2008.

Introduction

The successful trading of securities on the stock market concedes

- the ability to trade an unlimited amount of securities with immediacy and

- with a minimal influence on future prices.

These two requirements have been broadly discussed in the recent academic and empirical literature through the concept of liquidity and price impact of the trade. Liquidity measures and impact of trade on the future price formation are discussed based on daily, weekly, monthly or annual data. The rapid development of electronic trading in the recent years significantly improved the process of transferring securities from one market participant to another. Electronic trading enables a large number of participants to interact on the market, improves the speed of trade realization, and also requires a fast reactions of participants on any new information. In such new environment, knowing current liquidity state in the market, anticipating the level of liquidity that will be left after realized transaction and estimation of degree of the influence of realized transactions on immediate and future price movements is crucial for every successful trading strategy. As a consequence, the liquidity and price impact have to be analyzed on high frequency (trade-by-trade) financial data. Analysis of different aspects of liquidity measures based on trade-by-trade data is given in Section 2 of the present thesis.

The price impact of the trade is discussed on several different levels in the literature. In many empirical applications, price impacts are measured over short horizon as a difference of the midquote immediately after the transaction and midquote just before transaction. Such approach is appropriate for an immediate level of price impact of the trade. Due to this level, if the trade is buyer-initiated, the price tends to go up, and when it is seller-initiated, the price tends to go down. However, this approach does not discuss the possible reasons of price impact and its implications. The literature concerning the price impact of the trade often make distinction between temporary and permanent price impact. Short-horizon temporary price impact has been attributed to market frictions, while the source of the permanent price impact is the presence of investors with superior information - private information. The realized price impact is then a combination of the permanent and temporary impacts. We will shortly discuss several approaches to measuring price impact.

Farmer et al. [25] discussed the price formation dependent of trade ordering. They recognize a mechanical impact and informational impact, which together make the total impact of the trade ordering. The placement of trading orders may depend of complicated factors but, once when the sequence of trading orders is given, price formation is purely mechanical. The mechanical impact of trading order can be defined as the change in the future prices that occurs even if no trading orders are changed in any way. It is realized in the absence of any information. The informational impact is the part which is left after removing the mechanical impact from the total impact. Their empirical findings based on effective market orders [1] from on-book market of the London Stock Exchange during the period of three years (2000-2002) implies that the average mechanical impact decays to zero monotonically in transaction time, as a power function with an exponent of about 1.7. In contrast, the informational impact is a concave function of transaction time and approaches a constant value.

Lillo et al. [53] discussed how much the price changes in average in response to an order to buy or sell of a given size. Their investigation was based on Trade And Quotes (TAQ) database of the 1000 stocks with the largest market capitalization traded in the New York Stock Exchange in the period 1995-1998. They investigated the average price shift as a function of the transaction size measured in dollars, doing separately for buys and sells. Their findings indicate that an average price impact is a power function of transaction size with an exponent less than one, and that it depends on the market capitalization of the stock. The behavior was roughly the same for both buy and sell orders.

Vasiliki et al. [65] empirically investigated how stock prices respond to the changes in demand. They consider the change in demand through two variables and examine them separately. The first one is the number imbalance, a difference in the number of buyer-initiated and seller-initiated trades in a certain time interval. The second one is the volume imbalance,

[1] An effective market order is defined as any order or component of order that generates transaction

a difference in the number of shares traded in the buyer-initiated and seller-initiated trades in a certain time interval. They found that the price impact function, defined as the conditional expectation function of price change, given a number of imbalance (volume of imbalance) displays odd function forming shape "S" through the first and the third quadrant of the coordinate system that seems universal for all stocks.

An extensive amount of theoretical literature is devoted to the price changes caused by the presence of informed and uninformed traders in the market, i.e. by information asymmetry. These models examine market makerŠs behavior when some of the traders are better informed than the others. The main result of such research is that the price changes due to market friction are temporary, while the presence of informed traders causes permanent price changes.

Kyle [50] gave a fundamental model of the dynamics of the market with information asymmetry. He assumes that a monopolistic insider, a market maker and a noise traders interact. The market maker observes the aggregate net order flow of insider and noise traders. It can be positive - net buy, and negative - net sell order flow. This order flow provides a signal about the liquidation value of the asset to market maker. Based on this signal, market maker revises her/his beliefs and sets price such that it equals the expected liquidation value given the observed order flow. The resulting equilibrium price change is an increasing linear function of net order flow, whose slope represents a measure of the market depth. The smaller slope indicates the deeper market. It determines how much the market maker adjusted the price in response to the net order flow. In the Kyle's model price changes are completely information induced.

Kyle's model has been extended by Back [5], who allows for a more general distribution of the private signal and formally derives an equilibrium pricing rule. In [50] orders are batched together and cleared at the predetermined points in time. Glosten et al. [30] suggested the sequential trading model which assumes that the orders arrive sequentially according to some stochastic fashion. In this model and in Glosten [29] the order arrivals are independent over time. Madhavan et al. [55] generalized Glosten [29] by allowing autocorrelation in order flow.

Hasbrouck's approach [39] to future price formation is also based on the asymmetric information theory. In a market with asymmetrically informed participants, the market environment is measured by bid-ask spread and the trade is described by its direction (positive if the trade is a buyer-initiated and negative if the trade is a seller-initiated) and by its size. The concept of an asymmetrically informed market concedes that market makers possess only public information and that they interact with the other market participants who can have some superior information - private information. The informed and uninformed traders are undistinguishable to the market makers. The main idea of the model is that the trade conveys information and that market makers posted bid and ask prices after the realized transaction and with respect to that information. The model is represented by the vector autoregression system of the return and the trade equation based on both price and order flow history. Considering such a system, the effect of public and private information on the price formation is considered and the transitory and permanent price impact is identified. Hasbrouck's empirical findings indicate that the effect of permanent price impact is not instantaneous and that it takes several transactions before it is fully realized. Using the impulse response technique Hasbrouck [40] constructed the permanent price impact as a cummulative response of return to a shock in the innovation of trade equation, where private information must arise if such exists. Also, by variance decomposition technique Hasbrouck [40] calculated the contribution of private information to the future price formation.

Hausman et al. [44] introduced one of the first empirical models that consider the effect of the time span or duration between the trades on price movement - the ordered probit model. They investigate the conditional distribution of price changes given a set of explanatory variables which includes the sequence of past prices and irregularly spaced order arrivals. Their cross-sectional analysis illustrates how the sequence of trades affects the dynamics of price changes.

A more sophisticated statistical approache to analyzing the transaction arrival times is given by Engle et al. [22]. An Autoregressive Conditional Duration (ACD) model is applied to explicitly specify the dynamics of the time duration between order arrivals. Engle et al. [21], [23] explore a statistical model suitable for analyzing transactions data, generalizing the VAR model in Hasbrouck [39] to incorporate the role of time between trades in stock price movements and trade processes.

This master thesis analyzes different liquidity aspects and the effect of information asymmetry on the future price formation according to Hasbrouck's approach [39], [40] on eighteen stocks traded on the London Stock Exchange from the FTSE 100 index. The thesis is organized as follows.

In Chapter 1 is given an survey of definitions and theorems needed for understanding the concept of Hasbrouck's vector autoregression model.

In Chapter 2, a large number of different liquidity measures based on trade-by-trade data is presented.

Chapter 3 deals with the basic univariate and multivariate time series analysis needed for understanding Hasbrouck's *VAR* approach to modelling permanent price impact. Of primary interest is the multivariate form of autoregression and its moving-average representations needed for the impulse response and variance decomposition analysis.

An economic framework of the microstructure and asymmetric information theory as a motivation for the Hasbrouck's model of price impact is delt with in Chapter 4. Hasbrouck's *VAR* model is discussed in details.

Chapter 5 presents the results of an empirical analysis of the different liquidity measures based on trade-by-trade data, as well as the results of the analysis of price changes due to the asymmetric distribution of information using Hasbrouck's approach. The specifications of the obtained results in comparison to Hasbrouck's results [39], [40] are also discussed.

◊ ◊ ◊

I would like to thank to everyone who supported me and helped during my work on this thesis.

My thanks to members of committee for evaluation master thesis, professors Nataša Krejić, Zorana Lužanin, Marko Nedeljkov and Branko Urošević.

I am very indebted to professor Stevan Pilipović, who I spent with three inspiring years in dealing with functional analysis. During the work on this thesis, professor Stevan Pilipović gave me a plenty of useful suggestions.

Finally, my thanks to my mentor professor Nataša Krejić, who opened the world of applied mathematics to me. Her constant ongoing support, helpful hints, unlimited enthusiasm and motivation made this thesis possible. With their unlimited knowledge and professional abilities, Professor Nataša Kreić and professor Stevan Pilipović had the most important influence in forming my mathematical profile.

◊ ◊ ◊

This master thesis was completed during the work on research project *Modelling and Forecasting Stock Prices Behavior - Analysis of High Frequency Financial Data*, University of Novi Sad, Serbia Department of Mathematics and Informatics, Faculty of Science, Novi Sad in cooperation with Dresden Kleinworth Securities LTD, London.

Novi Sad, April 10, 2008 Nataša Teodorović

Contents

List of Figures

List of Symbols

x	sign trade
x^0	trade indicator variable
p^a	ask price
p^b	bid price
v^b	bid volume
v^a	ask volume
m	midquote
r	return
V	trading volume
TO	turnover
D	volume depth
$Dlog$	logarithmic volume depth
$D\$$	money depth
Dur	duration
S	absolute spread
$logS$	logarithm of absolute spread
S^{rel}	relative spread
S^{prop}	proportional spread
$LogSlog^{rel}$	logarithmic relative spread of logarithmic quotes
S^{eff}	effective spread
LP	liquidity premium
$S^{relleff}$	relative effective spread
$S^{propeff}$	proportional effective spread
QS	quote slope
$LogQS$	logarithmic quote slope
$LogQSadj$	adjusted logarithmic quote slope
CL	composite liquidity
$LR1$	liquidity ratio 1
$LR2$	liquidity ratio 2

M	Martin index
FR	flow ratio
OR	order ratio
MI^{V^*}	market impact
MI^{a,V^*}	market impact for the ask side
MI^{B,V^*}	market impact for the bid side
PI^a	price impact for the ask side
PI^b	price impact for the bid side
$\mathbb{R}$	set of real numbers
$\mathbb{R}^n$	set of n-dimensional real vectors
$\mathbb{Z}$	set of integers
$\mathcal{A}$	σ-algebra
$\mathcal{B}$	Borel's σ-algebra
$(\Omega, \mathcal{A})$	measurable space
P	probability function
$(\Omega, \mathcal{A}, P)$	probability space
$\mathcal{N}(\mu, \sigma^2)$	normal distribution
$\mathcal{L}^p$	space of p-integrable functions
χ^2	chi-square distribution
Z	Student's t-distribution
F	Fisher distribution
E	mathematical expectation
E^*	linear projection, linear expectation
Var	variance
1_A	characteristic function of set A
γ	moment of random variable
α	absolute moment of random variable
c	central moment of random variable
C_X	autocovariance function of random variable X
C_{XY}	cross-covariance function of random variables X and Y
ρ_X	autocorrelation function of random variable X
ρ_{XY}	cross-correlation function of random variables X and Y
g_X	autocovariance generating function of random variables X
g_{XY}	cross-covariance generating function of random variables X and Y
π	projection
R_s	Spearman rank coefficient

L	lag operator
ε	white noise
Σ	error covariance matrix
AR	autoregression
MA	moving-average
VAR	vector autoregression
VMA	vector moving-average
I_n	$n \times n$ identity matrix
$R^2_{jk,s}$	variance decomposition coefficients of $X_{j,t}$ at horizon s.
ν_1	innovation in return equation
ν_2	innovation in trade equation
e	efficient price
err	forecast error
$\mathcal{V}$	security value
$\mathcal{F}$	public information set
ω	innovation in efficient price
σ^2_ω	variance of the innovation in efficient price
σ^2_1	variance of return innovation
Λ	variance of trade innovation
$\sigma_{\omega,x}$	absolute measure of trade informativeness
GLS	generalized least squares
WLS	weighted least sqiuares
$FGLS$	feasible generalized least squares
R^2_x	coefficient of multiple determination of trade equation
R^2_r	coefficient of multiple determination of return equation
R^2_ω	contribution of trade innovation to the total volatility of the efficient price
PI^{has}	total permanent price impact obtained from Hasbrouck's VAR

Chapter 1

Survey of definitions and theorems

In this chapter we will briefly recall the definitions and theorems from the theory of probability and statistics, necessary for the understanding the work developed in this thesis. We assume that integrable theory, theory of vector spaces and Hilbert spaces are already known.

1.1 Probability space

Suppose that Ω is a nonempty set of events and $\mathcal{A}$ is a family of subsets of Ω with the following properties.

1. $\Omega \in \mathcal{A}$
2. $A \in \mathcal{A} \Rightarrow \Omega \setminus A \in \mathcal{A}$
3. If $A_1, A_2, ...$ are sets from $\mathcal{A}$, then $A_1 \cup A_2 \cup ... \in \mathcal{A}$.

The family $\mathcal{A}$ with the properties 1-3 is called *σ-algebra* or *σ-field* of events on Ω. The pair $(\Omega, \mathcal{A})$ is a *measurable space.*

A minimal σ-field which contains $\{(-\infty, x) : x \in \mathbb{R}\}$ is a *Borel's* σ-field of the subsets from $\{(-\infty, x) : x \in \mathbb{R}\}$ in the notation $\mathcal{B} = \mathcal{B}(\mathbb{R})$. The pair $(\mathbb{R}, \mathcal{B})$ is a *Borel's* measurable space. Let $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$ and

$$(-\infty, x) = (-\infty, x_1) \times (-\infty, x_2) \times ... \times (-\infty, x_n)$$

be an interval in $\mathbb{R}^n$. A minimal σ-field which contains $\{(-\infty, x), x \in \mathbb{R}^n\}$ is a Borel's σ-field in $\mathbb{R}^n$ in the notation $\mathcal{B}^n = \mathcal{B}(\mathbb{R}^n)$. The pair $(\mathbb{R}^n, \mathcal{B}^n)$ is n-dimensional Borel's measurable space.

Let P be a function defined on the σ-field $\mathcal{A}$ such that

1. $P(A) \geq 0$ for every $A \in \mathcal{A}$
2. If $\{A_n\}$ is a sequence of the sets on $\mathcal{A}$ which are disjoint in pairs, then $P(A_1 \cup A_2 \cup ...) = P(A_1) + P(A_2) + ...$
3. $P(\Omega) = 1$.

Than P is called the *probability function* and $(\Omega, \mathcal{A}, P)$ is called the *probability space.*

Remark 1.1 *Using the terminology from the measure theory, probability function is a measure on the σ-algebra $\mathcal{A}$. A set $E \in \mathcal{A}$ such that $P(E) = 0$ is called a zero set. If some property holds always except on a set $W \in \mathcal{A}$ such that $P(\Omega \setminus W) = 0$, then we say that this property holds almost surely.*

Suppose that $(\Omega, \mathcal{A}, P)$ is a probability space and that $A, B \in \mathcal{A}$, $P(B) > 0$. Let us define

$$\hat{\Omega} := B,$$

$$\hat{\mathcal{A}} := \{C \cap B : C \in \mathcal{A}\},$$

$$\hat{P}_B(D) = \frac{P(D)}{P(B)}, \quad D \in \hat{\mathcal{A}}. \tag{1.1}$$

Then $\hat{\mathcal{A}}$ is a σ-algebra, $\hat{P}_B$ is probability function, and space $(B, \hat{\mathcal{A}}, \hat{P}_B)$ is a new probability space. *Conditional probability* of the event $A \in \mathcal{A}$ with respect to $B \in \mathcal{A}$ is given by

$$P(A|B) = \hat{P}_B(A \cap B) = \frac{P(A \cap B)}{P(B)}.$$

1.2 Random variables

Let $(\Omega, \mathcal{A}, P)$ be a probability space and $(\mathbb{R}, \mathcal{B})$ be a Borel's measurable space. The mapping X of a measurable space $(\Omega, \mathcal{A})$ into the measurable space $(\mathbb{R}, \mathcal{B})$ with the property

$$X^{-1}(B) \in \mathcal{A}, \ for\ every\ B \in \mathcal{B},$$

is called one-dimensional or real *random variable*. The n-dimensional vector $X = (X_1, X_2, ..., X_n)$ whose components are real random variables defined on the same probability space $(\Omega, \mathcal{A}, P)$ is n-dimensional random variable. More formally, n-dimensional random variable is a mapping X of a measurable space $(\Omega, \mathcal{A})$ into the measurable space $(\mathbb{R}^n, \mathcal{B}^n)$ with the property

$$X^{-1}(B^n) \in \mathcal{A}, \; for \; every \; B^n \in \mathcal{B}^n.$$

Random variables defined on a measurable space $(\Omega, \mathcal{A})$ are called $\mathcal{A}$-measurable functions. Every random variable X generates a σ-algebra on Ω. It is defined by

$$\mathcal{A}(X) = \{X^{-1}(B^n) : B^n \in \mathcal{B}^n\}.$$

It is the smallest σ-algebra which is a subalgebra of $\mathcal{A}$ such that X is measurable with respect to it. We say that $\mathcal{A}(X)$ contains all relevant information about random variable X.

Let X be an n-dimensional random variable $X = (X_1, X_2, ..., X_n)$. The function

$$F_X(x_1, x_2, ..., x_n) = P(\{\omega \in \Omega : X_1(\omega) < x_1, X_2(\omega) < x_2, ..., X_n(\omega) < x_n\})$$

is called the *joint probability distribution* of one-dimensional random variables $X_1, X_2, ..., X_n$. The random variables $X_1, X_2, ..., X_n$ are *independent* if

$$F_X(x_1, x_2, ..., x_n) = F_{X_1}(x_1) F_{X_2}(x_2) \cdot ... \cdot F_{X_n}(x_n).$$

A random variable X is *discrete* if $X(\Omega)$ is a finite or countable infinite set. It is defined by its *probability distribution*. Random variable X is *continuous* if there is a non-negative integrable function

$$f_X : \mathbb{R}^n \longrightarrow (0, \infty)$$

such that for every $x = (x_1, x_2, ..., x_n) \in \mathbb{R}^n$,

$$F_X(x) = \int_{\infty}^{x_1} \int_{\infty}^{x_2} ... \int_{\infty}^{x_n} f_X(t_1, t_2, ..., t_n) dt_1 dt_2 ... dt_n.$$

Such function is called *joint probability density function* of one-dimensional random variables $X_1, X_2, ..., X_n$.

1.2.1 Expectation and conditional expectation

Let $(\Omega, \mathcal{A}, P)$ be a probability space and $X = X(\omega)$ be a random variable defined on that space. The Lebesgue integral of X with respect to P, given by

$$E(X) = \int_\Omega X dP \equiv \int_\Omega X(\omega) dP(\omega)$$

is called *mathematical expectation* of the random variable X. The *conditional expectation* of random variable X with respect to $B \in \Omega$ is defined by

$$E(X|B) = \int_\Omega X d\hat{P}_B = \frac{1}{P(B)} \int_B X dP,$$

where $\hat{P}_B$ is given by (1). It is a mean value of the random variable $X(\omega)$, $\omega \in B$. Let us define a *characteristic function* for a given set $A \in \Omega$ as

$$1_A = \begin{cases} 1, & \omega \in A \\ 0, & \omega \in \Omega \setminus A. \end{cases}$$

Then the conditional expectation of the random variable X with respect to $B \in \Omega$ can be defined as

$$E(X|B) = \frac{E(1_B X)}{P(B)}.$$

Suppose that $X : \Omega \to \mathbb{R}^n$ is a random variable defined on the probability space $(\Omega, \mathcal{A}, P)$ such that $E(|X|) < \infty$, and suppose that $\mathcal{G} \subset \mathcal{A}$ is a subalgera of $\mathcal{A}$. The conditional expectation of X with respect to $\mathcal{G}$ in the notation $E(X|\mathcal{G})$, is an almost surely unique random variable Z with the properties

1. Z is $\mathcal{G}$-measurable,
2. $E(Z1_A) = E(X1_A)$ for every $A \in \mathcal{G}$.

Suppose that X and Y are random variables defined on the same probability space $(\Omega, \mathcal{A}, P)$, $E(|Y|) < \infty$, and suppose that $\mathcal{A}(Y)$ is a σ-algebra generated by Y. Conditional expectation of the random variable Y with respect to the random variable X is the $\mathcal{A}(X)$-measurable function defined by

$$E(Y|X) = E(Y|\mathcal{A}(X)).$$

1.2.2 Moments of random variable

The *moment of order* r of the random variable X is the mathematical expectation of X^r given by

$$\gamma_r = E(X^r) = \int_{\Omega} X^r dP.$$

The first moment of the random variable X is the mathematical expectation of the random variable X. The mathematical expectation of $|X|^r$ given by

$$\alpha_r = E(|X|^r) = \int_{\Omega} |X|^r dP$$

is called the *absolute moment of an order* r of the random variable X. The *central moment of an order* r of the random variable X is the mathematical expectation of $(X - E(X))^r$, i.e.

$$c_r = E((X - E(X))^r).$$

The second central moment of random variable X is called *variance* or *dispersion* of the random variable X in the notation $Var(X)$ or $D(X)$.

Random variables with finite absolute moments of order p define the space $\mathcal{L}^p \equiv \mathcal{L}^p(\Omega, \mathcal{A}, P)$, i.e.

$$X \in \mathcal{L}^p \;\; if \; and \; only \; if \;\; E(|X|^p) = \int_{\Omega} |X|^p dP < \infty.$$

Random variables $X \in \mathcal{L}^p$ are called *p-integrable* random variables.

Let $X \in \mathcal{L}^2$ be a real-valued random variable defined on a probability space $(\Omega, \mathcal{A}, P)$ and let $\mathcal{G}$ be a subalgebra of $\mathcal{A}$. The *conditional variance* of X with respect to $\mathcal{G}$, denoted by

$$Var(X|\mathcal{G})$$

is a $\mathcal{G}$-measurable random variable

$$E((X - E(X|\mathcal{G}))^2|\mathcal{G}).$$

The conditional variance of a random variable X with respect to the random variable Y, defined on $(\Omega, \mathcal{A}, P)$ and denoted by

$$Var(X|Y)$$

is the $\mathcal{A}(Y)$-measurable random variable

$$E((X - E(X|\mathcal{A}(Y))^2|\mathcal{A}(Y)).$$

1.2.3 Some probability distributions

Normal distribution. A real random variable X is normally distributed with the mean μ and variance σ^2 in the notation $X \sim \mathcal{N}(\mu, \sigma^2)$ if its probability density function is given by

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu)^2}{2\sigma^2}}.$$

Normally distributed random variables are also called the *Gaussian* random variables. A normal random variable with zero mean and unit variance in the notation $\mathcal{N}(0,1)$ has a *standard normal distribution.*

χ^2-distribution. If $X_i \sim \mathcal{N}(0,1)$ $i = 1, 2, ..., n$ are independent variables then $Z = \sum\limits_{i=1}^{n} X_i^2$ is χ^2-distributed with n degrees of freedom (d.f.) in the notation $Z \sim \chi^2_n$.

Student's t-distribution. If $X \sim \mathcal{N}(0,1)$ and $Y \sim \chi^2_n$ and X and Y are independent, then

$$Z = \frac{X}{\sqrt{\frac{Y}{n}}}$$

has a Student's $t-$distribution with d.f. n in the notation $Z \sim t_n$. As $n \to \infty$, Student's t-distribution approaches the standard normal distribution $\mathcal{N}(0,1)$.

Fisher's F-distribution. If $Y_1 \sim \chi^2_{n_1}$ and $Y_2 \sim \chi^2_{n_2}$ are independent, then

$$Z = \frac{Y_1/n_1}{Y_2/n_2}$$

has the F-distribution in the notation $Z \sim F_{n_1,n_2}$. The subscript n_1 refers to degree of freedom of the numerator and the subscript n_2 refers to degree of freedom of the denominator. As $n_2 \to \infty$, F-distribution approaches to the $\chi^2_{n_1}$ distribution.

1.3 Stochastic processes - basic theory

1.3.1 Properties

Let $(\Omega, \mathcal{A}, P)$ be a probability space and T be a set of indexes t. The family of random variables

$$X_t = \{X_t(\omega) : t \in T\}$$

defined on the same probability space $(\Omega, \mathcal{A}, P)$ is a *stochastic process.* Stochastic process is a function of two variables: $\omega \in \Omega$ and $t \in T$. For shortening notation it is usually written as X_t. For fixed $t \in T$, X_t represents a random variable. For fixed $\omega \in \Omega$ it represents a real function defined on T called the *trajectory.* In the text to follow we will assume that T is a time interval and t represents time point from that interval.

Let X_t be a stochastic process. The *autocovariance function* of the process X_t at two different time points $t, s \in T$ is

$$C_X(t,s) = E[(X_t - E(X_t))(X_s - E(X_s))] = E(X_t X_s) - E(X_t)E(X_s).$$

A normalized version of the autocovariance function is *autocorrelation function* (ACF), given by

$$\rho_X(t,s) = \frac{C_X(t,s)}{\sqrt{Var(X_t)Var(X_s)}}.$$

Both functions are real valued. It is obvious that for $t = s$, the covariance of process X_t is a variance of the process X_t, i.e. $C_X(t,t) = Var(X_t) \geq 0$. Hence, $\rho(t,t) = 1$. It is easy to show from the Schwartz inequality

$$(\int |f(x)g(x)|dx)^2 \leq \int (f(x))^2 dx \int (g(x))^2 dx,$$

that

$$(C_X(t,s))^2 \leq C_X(t,t)C_X(s,s) = Var(X_t)Var(X_s),$$

which implies $|\rho_X(t,s)| \leq 1$. A process X_t is *covariance stationary* or *weak stationary* if:

1. It has a time-constant mean, i.e. $E(X_t) = \mu$, for all $t \in T$.
2. It has a time-constant variances, i.e. $E((X_t - \mu)^2) = \sigma^2 < \infty$, for all $t \in T$.

3. The covariance between two observations at two different time points depends only on the distance between them, but not on the time points themselves, i.e. $C_X(t,s) = C(t-s)$ for every $t, s \in T,\ t \neq s$.

A stronger type of stationarity is *strict stationarity*. We say that the process X_t is strict stationary if the probability distribution of every finite vector of observations $(X_{t_1}, X_{t_2}, ..., X_{t_s})$ and its shifted vector of observations $(X_{t_1+h}, X_{t_2+h}, \dots, X_{t_s+h})$ are always the same for an arbitrary h. Notice that strict stationarity does not require existence of first and second moments. If a strict stationary process has a mean, it is a constant. If a strict stationary process has finite moments of second order then its covariance function does not depend on time points, but only on distance between them. Therefore, the strict stationary process is weak stationary if it has a mean and if all its second moments are finite. A weak stationary process is strict stationary only if all probability distributions are Gaussian. The concept of strict stationarity requires that the joint distributions of the variables X_t are known. This may be quite complicated in practice. For a practical purpose, the weak stationarity concept is usually sufficient. In the present paper, stationarity would always mean a weak stationarity, unless specified otherwise.

A weak stationary process can be described as a process that looks the same for any time span of a certain fixed length. Because of finite variance we say that the weak stationary process has a *mean revision* property, which means that it can never drift too far from its mean. The speed of mean revision is determined by an autocovariance function: mean revision is fast when autocovariances are small, and slow when autocovariances are large.

Since for weak stationary processes the covariance does not depend on time, but only on the distance between time points, one can write

$$C_X(t,s) = C_X(t-s, s-s) = C_X(t-s) = C_X(h) = C_X(t, t-h)$$

were $h = t - s$. Subsequently, the autocorrelation function is

$$\rho_X(h) = \frac{C_X(h)}{C(0)}$$

and it has maximum for $h = 0$. Also, the autocovariance and the autocorrelation function of a stationary process X_t are symmetric, i.e.

$$C_X(h) = C_X(-h),$$

$$\rho_X(h) = \rho_X(-h).$$

For two different stochastic processes X_t and Y_t we define a *cross-covariance* function by

$$C_{XY}(t,s) = E[(X_t - E(X_t))(Y_s - E(Y_s))] = E(X_tY_s) - E(X_t)E(Y_s)$$

at two different time points t and s. The *cross-correlation* function (CCF) of two different processes X_t and Y_t at two different time points t and s is

$$\rho_{XY}(t,s) = \frac{C_{XY}(t,s)}{\sqrt{Var(X_t)Var(X_s)}}.$$

For two strictly stationary processes X_t and Y_t the *strict joint stationarity* means that their joint distributions do not depend on the origin from which the indexes of these two processes are taken. A *weak joint stationarity* of two different weak stationary processes X_t and Y_t means that the cross-covariance function $C_{XY}(t,s)$ at two different time points t and s from an observable time interval depends only on the distance between the time points, but not on time points themselves, i.e.

$$C_{XY}(t,s) = C(t-s).$$

By analogy with stationary processes, for two different jointly stationary processes X_t and Y_t the following equalities are satisfied.

$$C_{XY}(h) = C_{XY}(t, t-h),$$

$$\rho_{XY}(h) = \frac{C_{XY}(h)}{C_X(0)C_Y(0)},$$

$$C_{XY}(h) = C_{YX}(-h),$$

$$\rho_{XY}(h) = \rho_{YX}(-h),$$

$$|\rho(h)| \leq 1,$$

but, $\rho_{XY}(0) \neq 1$. Since the autocorrelation represents a correlation of the process with itself over successive time intervals it is also called the *serial correlation*, while crosscorrelation represents the correlation between two processes over successive time intervals and it is also called the *joint correlation.*

Autocovariance generating function of the specific process X_t is a function constructed by taking the h-th autocovariance, $h \in \mathbb{Z}$ and multiplying it by the h-th power of some complex number z and summing it over all possible values of h. This function generates all autocovariances of the specific process X_t, and it is given by

$$g_X(z) = \sum_{h=-\infty}^{\infty} C_X(h) z^h. \tag{1.2}$$

Following a similar concept, for two specific processes X_t and Y_t the function which generates all their cross-covariances is a *cross-covariance generating function* given by

$$g_{XY}(z) = \sum_{h=-\infty}^{\infty} C_{XY}(h) z^h. \tag{1.3}$$

The autocovariance (autocorrelation) function has a graphical representation. Graphical representation of an autocorrelation function is called *correlogram*. Different stationary processes have different correlogram shapes, whereas non-stationary processes do not have well defined ACF. The correlograms of a stationary time series can be used for model selection.

Some processes are not stationary but they become stationary when they are observed for a long time spans. These processes are called *stable* or *asymptotically stationary*.

1.3.2 Linear projection

Let $\mathcal{L}^2$ denote a Hilbert space of the squared integrable random variables X with the inner product

$$\langle X, Y \rangle = E(XY)$$

and with the norm

$$\|X\|^2 = E(X^2).$$

Let $\mathcal{S} \subset \mathcal{L}^2$. The random variables $X, Y \in \mathcal{S}$ are equivalent if

$$E((X - Y)^2) = 0.$$

Let $X \in \mathcal{L}^2$ be a random variable. A random variable $\pi_{\mathcal{S}}$ is called the *projection* of X onto $\mathcal{S}$ if

$$E((X - \pi_{\mathcal{S}}(X))^2) = \min_{S \in \mathcal{S}} E((X - S)^2).$$

Notice that there is no apriori guarantee that $\pi_{\mathcal{S}}(X)$ exists or that it is unique. A sufficient condition for the existence is that $\mathcal{S}$ is a closed subspace. For the given random variables $Y_1, Y_2, ..., Y_k$ let

$$\mathcal{M}(Y_1, Y_2, ..., Y_k) = \{h(Y_1, Y_2, ..., Y_k) : h : R^k \longrightarrow R \textit{ is measurable}\} \cap \mathcal{L}^2$$

be a closed linear subspace of $\mathcal{L}^2$ consisting of measurable functions of $Y_1, Y_2, ..., Y_k$ with $E(h^2(Y_1, Y_2, ..., Y_k)) < \infty$. The mean-square projection of X onto $\mathcal{M}(Y_1, Y_2, ..., Y_k)$ is defined as

$$E(X|Y_1, Y_2, ..., Y_k) \equiv \pi_{\mathcal{M}(Y_1, Y_2, ..., Y_k)}(X).$$

The statistical interpretation of the mean-square projection is as follows. We consider the random variable X belonging to $\mathcal{L}^2$ and we want to have the best approximation of this variable by the family $\{Y_t : t = 1, 2, ..., k\}$. This means that we are going to look after the element of $\mathcal{M}(Y_1, Y_2, ..., Y_k)$ which minimizes the expression $\|X - Y\|^2$. Therefore, $E(X|Y_1, Y_2, ..., Y_k)$ is the best predictor of X based on $Y_1, Y_2, ..., Y_k$ in the mean-square error sense. In other words, there exists a measurable function

$$f : \mathbb{R}^k \longrightarrow \mathbb{R}$$

with

$$E(X|Y_1, Y_2, ..., Y_k) = f(Y_1, Y_2, ..., Y_k)$$

and

$$E(X - f(Y_1, Y_2, ..., Y_k)) = \min_h E(X - h(Y_1, Y_2, ..., Y_k))$$

where h's are measurable functions $h : \mathbb{R}^k \longrightarrow \mathbb{R}$. If we insist that the function f is linear, than $E(X|Y_1, Y_2, ..., Y_k)$ represents the best linear predictor of X based on $Y_1, Y_2, ..., Y_k$ and it is called *linear expectation* or *linear projection.*

1.4 Multiple linear regression

1.4.1 Definition

A linear multiple regression expresses a dependent random variable y as a linear function of independent variables $x_1, x_2, ..., x_k$ and error term ε. Dependent variable y is also called *endogenous variable* or *regressand.* Independent variables $x_1, x_2, ..., x_k$, which are possible random are called *exogenous variables* or *regressors.* The error term is the *innovation*, *shock* or *unexpected shift* in exogenous variables. Considering all observations from a sample of size n it can be written as

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + ... + \beta_k x_{ik} + \varepsilon_i$$

where $i = 1, 2, ..., n$ and $\beta_1, \beta_2, ..., \beta_k$ are *regression coefficients.* In the matrix notation it can be written as

$$\begin{bmatrix} y_1 \\ y_2 \\ . \\ . \\ . \\ y_n \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & ... & x_{1k} \\ x_{21} & x_{22} & ... & x_{2k} \\ . & & & \\ . & & & \\ . & & & \\ x_{n1} & x_{n2} & ... & x_{nk} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \\ . \\ . \\ . \\ \beta_k \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ . \\ . \\ . \\ \varepsilon_n \end{bmatrix},$$

or in shorter notation

$$y = X\beta + \varepsilon. \tag{1.4}$$

1.4.2 Estimation

First we will define some properties of estimators. Suppose that $\hat{\theta}$ is an estimator of the parameter θ. We say that $\hat{\theta}$ is

1. *unbiased*, if $E(\hat{\theta}) = \theta$. If $E(\hat{\theta}) \neq \theta$, the estimator is *biased* and the difference $E(\hat{\theta}) - \theta$ is called a *bias*.

2. *efficient*, if it is unbiased and it has a minimum variance in the class of all unbiased estimators.

A linear multiple regression coefficient can be estimated by minimizing the error sum of squares i.e.

$$\min_{\beta}((y - X\beta)^T(y - X\beta)) = \min_{\beta} \sum_{i=1}^{n} (y_i - x_i^T\beta)^2$$

where $x_i, \ i = 1, 2, ..., n$ are row vectors of matrix X in the equation (1.4). Such method is called the *least squares method.* The least squares estimator for β, denoted by $\hat{\beta}$ is

$$\hat{\beta} = (X^TX)^{-1}X^Ty.$$

Following assumptions are required to examine the distribution of $\hat{\beta}$ and the optimality of the least square procedure.

1. $y = X\beta + \varepsilon$

2. $E(\varepsilon_i|X) = 0, \ \ i = 1, 2, ..., n$

3. $Var(\varepsilon_i|X) = \sigma^2, i = 1, 2, ..., n$
4. $E(\varepsilon_i\varepsilon_j|X) = 0, \ \ i, j = 1, 2, , ...n, \ \ i \neq j$
5. $rank(X^TX) = rank(X) = k$
6. $\varepsilon|X \sim N(0, \Sigma)$

The first assumption is necessary if we want that the least square method is a reasonable method for estimation. The second assumption is called the *strict exogeneity* which provides that the least squares estimator $\hat{\beta}$ of β is unbiased, i.e. $E(\hat{\beta}|X) = E(\beta)$. The third assumption is called *homoskedasticity*. The forth assumption provides that the errors are uncorrelated. The fifth assumption provides that $(X^TX)^{-1}$ exists, and therefore $\hat{\beta}$ can be estimated. Assumptions 2 and 6 provide that

$$\Sigma = \sigma^2 I_n$$

where I_n is an $n \times n$ identity matrix. Assumptions 1-6 provide that

$$Var(\hat{\beta}|X) = \sigma^2(X^TX)^{-1}.$$

Under the same assumptions the least squares estimator $\hat{\beta}$ of β has the lowest variance among all other linear unbiased estimators, and therefore it is efficient. Under assumptions 1-6 $\hat{\beta}$ is normally distributed, i.e.

$$\hat{\beta} \sim \mathcal{N}(\beta, \sigma^2(X^TX)^{-1}).$$

For dynamics models[2], i.e. when regressors x_i are lagged y_t, the strict exogeneity is not satisfied. Therefore, this assumption has to be relaxed by the assumption of *predetermined regressors* given by

$$E(x_i\varepsilon_i) = 0, \ i = 1, 2, ..., n$$

which means that the error term is uncorrelated (orthogonal) to the contemporaneous regressors.

1.4.3 Hypothesis testing

We will consider the null hypothesis given by

$$H_0: \ B\beta - b = 0$$

[2]See section 3

where B is an $m \times k$ matrix, k is the number of regressors and m is the number of restrictions given by the hypothesis.

T- test. It is used to test the significance of the coefficients $\beta_i,\ i = 1, 2, ..., k$. Under assumption 1-6 to test $H_0 : \beta_i = 0$ against $H_1 : \beta_i \neq 0,\ i = 1, 2, ..., k$, t-test is given by

$$\frac{B\hat{\beta} - r}{\sqrt{s^2 B(X^T X)^{-1} B^T}} \sim t_{n-k},$$

where b is a $1 \times k$ vector with all components being zeros expect of the i-th component which is one. For $n \to \infty$, Student's t-distribution converges to the standard normal distribution, $\mathcal{N}(0, 1)$. The t-statistics for the specific β_i is given by

$$\frac{\hat{\beta}_i}{\sqrt{s^2 (X^T X)^{-1}_{[ii]}}}. \tag{1.5}$$

The term $(X^T X)^{-1}_{[ii]}$ denotes the i-th diagonal element of $(X^T X)^{-1}$ and s^2 is the sample error variance estimator given by

$$s^2 = \frac{\hat{\varepsilon}^T \hat{\varepsilon}}{n - k}$$

where $\hat{\varepsilon}$ is the estimated error term.

The Wald test. It tests whether an independent variable has a statistically significant relationship with a dependent variable. Under Assumptions 1-6 the Wald test is given by

$$\frac{(B\beta - b)^T [B^T (X^T X)^{-1} B]^{-1} (B\beta - b)}{s^2} \sim F_{m,n-k}. \tag{1.6}$$

If $n \to \infty$, $F_{m,n-k}$ converges to the χ^2_m distribution.

1.4.4 Large sample properties of the least squares estimator

The assumptions 1-6 are small sample assumptions of multiple linear regression. Most of them are not realistic in the applications to financial data which are generally non-normal, heteroskedastic and correlated. To avoid employing more definitions and theoretical details, large sample assumptions will not be discussed here. Assuming that large sample assumptions are satisfied, we will discus the large sample properties or the asymptotic behavior of the least squares estimators.

The asymptotic properties of the least squares estimator $\hat{\beta}$ of β discuss its behavior as $n \to \infty$. First we need the definitions of convergence in probability, convergence in distribution and consistency property of estimators.

Convergence in probability. A sequence of random variables X_n, $n \in \mathbb{N}$ converges in probability to a constant c, in the notation $X_n \xrightarrow{P} c$, if

$$\lim_{n\to\infty} P(|X_n - c| > \varepsilon) = 0$$

for any $\varepsilon > 0$.

Convergence in distribution. Let $X_1, X_2, ...$ be a sequence of random variables with the corresponding distribution functions $F_{X_1}, F_{X_2}, ...$. A sequence of random variables $X_1, X_2, ...$ converges in distribution to a random variable X with the distribution function F_X, in the notation

$$X_n \xrightarrow{d} X$$

if the sequence of distribution functions $F_{X_1}(x), F_{X_2}(x), ...$ converge to $F_X(x)$ in every point x at which F_X is continuous.

Consistency of estimator. Suppose that $\hat{\theta}$ is an estimator of θ based on a sample of size n. Then an estimator $\hat{\theta}$ is sad to be *consistent* if and only if

$$\hat{\theta} \xrightarrow{P} \theta.$$

A sufficient condition for $\hat{\theta}$ to be consistent is that its bias and variance should both tend to zero as a sample size increases.

The asymptotic properties of the least squares estimator $\hat{\beta}$ are

1. consistency, i.e. $\hat{\beta} \xrightarrow{P} \beta$.

   This property is provided by assumption of predetermined regressors.

2. asymptotic normality, i.e.

   $$\hat{\beta} \xrightarrow{d} \mathcal{N}(\hat{\beta}, \frac{\sigma^2}{n} Q^{-1})$$

   where

   $$Q = E(x_i^T x_i)$$

   and

   $$\frac{X^T X}{n} \xrightarrow{P} Q$$

which is in fact one of the large sample assumptions. We refer to [35] for more details.

1.4.5 Heteroskedasticity

We say that in the multiple regression model (1.4) the error term is *hetroskedastic* if

$$E(\varepsilon_i|X) = 0,$$
$$Var(\varepsilon_i|X) = \sigma_i^2 = \sigma^2 w_i, \;\; w_i > 0, \;\; i = 1, 2, ..., n.$$

If we assume that the error terms are not autocorrelated, i.e.

$$E(\varepsilon_i \varepsilon_j|X) = 0, \;\; i \neq j, \; i, j = 1, 2, ..., n$$

then the error covariance matrix can be written as

$$\Sigma = Var(\varepsilon|X) = E(\varepsilon\varepsilon^T|X) = \sigma^2 W \tag{1.7}$$

where $W = diag(w_1, w_2, ..., w_n)$. Under heteroskedasticity the least squares estimator $\hat{\beta}$ of β remains unbiased, consistent and asymptotically normally distributed but it is not efficient any more. The covariance matrix of $\hat{\beta}$ is given by

$$Var(\hat{\beta}) = (X^T X)^{-1} X^T \sigma^2 W X (X^T X)^{-1} \neq \sigma^2 (X^T X)^{-1}.$$

Standard tests such as the Wald or t-test are not valid. Under very general conditions the matrix

$$S = \frac{X^T \sigma^2 W X}{n} = \frac{1}{n} \sum_i \sigma_i^2 x_i x_i^T$$

can be consistently estimated by the *White's heteroskedasticity consistent covariance estimator* [69]

$$\hat{S} = \frac{1}{n} \sum_i \varepsilon_i^2 x_i x_i^T. \tag{1.8}$$

Then the White's estimator of covariance matrix $V(\hat{\beta})$ is given by

$$Var(\hat{\beta}) = n(X^T X)^{-1} \hat{S} (X^T X)^{-1}. \tag{1.9}$$

Tests based on the White's estimator hold only asymptotically as the test statistics is consistent but not unbiased.

1.4.6 Generalized least squares

If data have heteroskedastic errors, efficiency of the estimator $\hat{\beta}$ of β in multiple linear regression (1.4) can be obtained by the *generalized least squares* method, *GLS*. Let us assume that error covariance matrix (1.7) is known. Multiplying both sides of equation (1.4) by $W^{-1/2}$

$$W^{-1/2}y = W^{-1/2}X\beta + W^{-1/2}\varepsilon$$

it is transformed into

$$\widetilde{y} = \widetilde{X}\beta + \widetilde{\varepsilon}, \tag{1.10}$$

where $\widetilde{y} = W^{-1/2}y$, $\widetilde{X} = W^{-1/2}X$ and $\widetilde{\varepsilon} = W^{-1/2}\varepsilon$. Such transformation provides that errors in equation (1.10) are homoskedastic, i.e.

$$Var(\widetilde{\varepsilon}|X) = \sigma^2 I_n.$$

Application of the least squares method in equation (1.10) leads to the generalized least squares estimator $\hat{\beta}_{GLS}$ of β, given by

$$\hat{\beta}_{GLS} = (X^T W^{-1} X)^{-1} X^T W^{-1} y.$$

The generalized least squares is also called the *weighted least squares* (*WLS*) as it minimizes the sum of squared residuals weighted by $1/w_i$, $i = 1, 2, ..., n$, i.e.

$$\min_\beta W^{-1}((y - X\beta)^T(y - X\beta)) = \min_\beta \sum_{i=1}^{n} \frac{1}{w_i}(y_i - x_i^T\beta)^2$$

where x_i, $i = 1, 2, ..., n$ are row vectors of matrix X in equation (1.4).

When the error covariance matrix is unknown, multiple linear regression (1.4) has to be estimated by the *feasible generalized least squares* (*FGLS*). We refer to [35] for more details.

1.4.7 Multivariate regression

The system of K standard multiple linear regressions is called a *multivariate regression*

$$y_j = X_j\beta_j + \varepsilon_j, \quad j = 1, 2, ..., K.$$

These regressions do not necessarily have the same number of regressors. Multivariate regression can be written in the vectorized version for the all system as

$$y^{\#} = X^{\#}\beta^{\#} + \varepsilon^{\#}. \tag{1.11}$$

If n is the number of observations, vectors $y^{\#}$ and $\varepsilon^{\#}$ are the $nK \times 1$ vectors obtained by stacking vectors y_j, that is by stacking vectors $\varepsilon_j,\ j = 1, 2, ..., K$. The matrix $X^{\#}$ is a block-diagonal matrix of the form

$$X^{\#} = \begin{bmatrix} X_1 & 0 & ... & 0 \\ 0 & X_2 & ... & 0 \\ . & & & \\ . & & & \\ . & & & \\ 0 & 0 & ... & X_K \end{bmatrix}.$$

If M_j is the number of regressors in equation $j = 1, 2, ..., K$, the dimension of matrix $X^{\#}$ is $nK \times M$, where $M = \sum_{i=1}^{K} M_i$.

In the multivariate regression model, crucial assumptions concern the error covariance matrix of the model (1.11) given by $E(\varepsilon^{\#}\varepsilon^{\#T})$. If the errors are unrelated across equations and all regressors are different, the regressions are unrelated. For more general assumption of the error covariance matrix we need the following definitions.

Positive definite matrix. A symmetric matrix A ($A = A^T$) is positive definite if there is a positive constant α such that for all $x \in \mathbb{R}^n$ is satisfied

$$x^T A x \geq \alpha ||x||^2, \tag{1.12}$$

where $||\cdot||$ is the vector's norm.

Positive semidefinite matrix. A symmetric matrix A is positive semidefinite if inequality (1.12) holds for $\alpha = 0$.

Kronecker product. The Kronecker product of $k \times l$ matrix A, and $m \times n$ matrix B, in the notation $A \otimes B$ is the $km \times ln$-matrix given by

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & ... & a_{1l}B \\ a_{21}B & a_{22}B & ... & a_{2l}B \\ . & & & \\ . & & & \\ . & & & \\ a_{k1}B & a_{k2}B & ... & a_{kl}B \end{bmatrix}.$$

Under assumptions

$$Var(\varepsilon_j|X_j) = \sigma_{jj} I_n,\ \ j = 1, 2, ..., K,$$

$$E(\varepsilon_j \varepsilon_k|X_j, X_K) = \sigma_{jk} I_n,\ \ j, k = 1, 2, ..., K,\ j \neq k,$$

the error covariance matrix of the model (1.11) becomes

$$E(\varepsilon^{\#}\varepsilon^{\#T}) = \begin{bmatrix} \sigma_{11}I_n & \sigma_{12}I_n & ... & \sigma_{1K}I_n \\ . & & & \\ . & & & \\ . & & & \\ \sigma_{K1}I_n & \sigma_{K2}I_n & ... & \sigma_{KK}I_n \end{bmatrix} = \Sigma \otimes I_n,$$

where Σ is positive semidefinite matrix. Therefore, the error covariance matrix $E(\varepsilon^{\#}\varepsilon^{\#T})$ is a matrix filled with diagonal blocks. Under such assumption, the errors are uncorrelated over time, though they may be correlated across the equations. This model is called the *seemingly unrelated regression* or *SUR*, [67].

The least squares estimator $\hat{\beta}^{\#}$ of $\beta^{\#}$ for the *SUR* model is given by

$$\hat{\beta}_j^{\#} = (X^{\#T}X^{\#})^{-1}X^{\#T}y^{\#}, \tag{1.13}$$

or viewed for a individual equation it is simply the least squares estimator *equation by equation*

$$\hat{\beta}_j = (X_j^T X_j)^{-1}X_j^T y_j, \quad j = 1, 2, ..., K.$$

The generalized least squares estimator $\hat{\beta}^{\#}$ of $\beta^{\#}$ for the *SUR* model is given by

$$\hat{\beta}_{GLS}^{\#} = (X^{\#T}(\Sigma^{-1} \otimes I_n)X^{\#})^{-1}X^{\#T}(\Sigma^{-1} \otimes I_n)y^{\#}. \tag{1.14}$$

The least squares estimator is consistent in the *SUR* model if the assumption of strict exogeneity or the assumption of predetermined regressors is satisfied. To analyze the efficiency of the least squares estimator we will consider two special cases for the error covariance matrix in the *SUR* model given in [49].

Regressors are the same in all equations. Then,

$$X^{\#} = I_n \otimes X.$$

It is easy to see by classical matrix operations that putting this expression of $X^{\#}$ in the generalized least squares estimator (1.14) it becomes the least squares estimator (1.13).

The error covariance matrix is diagonal. Since $X^{\#}$ is the block-diagonal, the expression (1.14) immediately decomposes into K separate estimates for $\beta_j,\ j = 1, 2, ..., K$ of the form

$$(\sigma_{jj}^2 X_j^T X_j)^{-1}\sigma_{jj}^2 X_j^T y_j = (X_j^T X_j)^{-1}X_j^T y_j,$$

which is again the least squares estimator of $\beta^{\#}$.

Since the generalized least squares estimator is efficient, these two cases implies

1. The least squares estimator $\hat{\beta}^{\#}$ of $\beta^{\#}$ in the *SUR* model with identical regressors is equal to the generalized least squares estimator, and therefore it is efficient.

2. The least squares estimator $\hat{\beta}^{\#}$ of $\beta^{\#}$ in the *SUR* model with contemporaneously orthogonal errors is equal to the generalized least squares estimator, and therefore it is efficient.

In the economical literature this efficiency property of the least squares estimator in a multivariate regression with identical regressors or with diagonal error covariance matrix is sometimes called the Kruskal's Theorem. We refer to [48] for detail analysis of assumptions which make the least squares and the generalized least squares estimators equal.

1.5 Trimmed mean

Let $x_1, x_2, ..., x_n$ be a set of real-valued observations ordered as $x_1 \leq x_2 \leq ... \leq x_n$. The k-th trimmed mean $\overline{x}_k$ is defined as

$$\overline{x}_k = \frac{1}{n-2k} \sum_{i=k+1}^{n-k} x_i.$$

By ordering the original observations, and taking away the first k smallest observations and the first k largest obserations, the trimmed mean takes the arithmetic average of the resulting data. Trimmed mean dramatically reduces sample standard deviation. The idea of a trimmed mean is to eliminate outliers, or extreme observations that do not seem to have any logical explanations in calculating the overall mean of a population.

1.6 Spearman rank correlation test

The Spearman rank correlation [46] is a technique used to test the direction and strength of the relationship between two variables. The data sets of realization of observed variables has to be ranked from 1 to n according to their values, where n is the sample size. The null hypothesis of this test states that there is no relationship between the two sets of data. Let $d_i,\ i = 1, 2, ..., n$ be the differences of corresponding ranks of two observable sets of size n. The Spearman rank correlation is given by

$$R_s = 1 - \frac{6 \cdot \sum\limits_{i=1}^{n} d_i^2}{n(n^2 - 1)}.$$

The zero R_s indicates that null hypothesis is accepted, otherwise it is rejected. The sign of R_s indicates the sign of correlation. If $|R_s| \leq 0.5$, the correlation is weak. If $|R_s| \geq 0.5$, the correlation is strong.

Chapter 2

Liquidity measures

The universe we are interested in is a stock exchange market with a mechanism of processing transactions between different market participants. The main participants in a traditional market are market makers and market agents. A market maker is a firm or a person that stands ready to buy and sell a particular stock at a publicly quoted prices. There are two types of stock prices in the stock market. They are the *bid price* - a price at which a trader is willing to buy a number of stocks, and the *ask price* - a price at which a trader is willing to sell a number of stocks. For every bid price a *bid volume* is specified - the number of shares that can be bought at that price. Also, for every ask price an *ask volume* is specified - the number of shares that can be sold at that price. Agents interact with market makers through a sell or buy transaction proposal which is called *order*. There are two categories of orders: *market orders* and *limit orders*. If the purpose is to trade securities as soon as possible one would put the market order - a buy or sell order of a certain number of stocks at the current standing (bid or ask) price. Therefore, this kind of order specifies a number of stocks that has to be executed in that transaction, but does not specify an execution price. On the other hand, if the purpose is to trade at a specific price or better, one would put the limit order - a limit bid or ask stock price at which the transaction has to be executed. The buy limit order specifies the maximum price at which a trader is willing to buy, and the sell limit order specifies the minimum price at which a trader is willing to sell. Hence, for the limit orders there is no certainty about whether it will be executed and when, but if the order was executed it would be executed at the requested price or better. The stock market is full of participants who want to buy or sell stocks at different prices. Therefore, one who is willing to buy will try to buy from one who offers the lowest price. On the other hand, one who is willing to sell will try to sell to one who is ready to pay the highest price. The highest bid and the lowest ask prices are the best bid and best ask price for which trade can be executed in any specific time. The difference between the best ask and bid price is called a *spread*. Market makers seek to buy shares at the lowest price and sell at the highest price. The bid-ask spread is therefore intended to compensate market makers for the risk they take in dealing with a security.

A huge development of information technology enables a detail recording of stock market activities in the electronic order book. Such book contains information such as price, volume and time of realized transaction, several levels of bid and ask prices with their related volumes posted after realized transaction and time when they posted. Also, an order book contains information about orders, such as the type (buy or sell) and number of orders.

One who intends to trade securities on the stock market wants to buy or sell a certain quantity of securities at the acceptable price and in desired time. Subsequently, the questions about market ability to provide transferring the ownership of a security from one market participant to another with the lowest degree of difficulties naturally appear. This ability is called *liquidity*. Liquid market is therefore a market where market participants are able to buy or sell an unlimited amount of securities with immediacy, at the price close to the last traded price. Alternatively, illiquid market is the one where market participants will be able to trade securities only at prices different from the last traded price. In an illiquid market, the large transactions cause shifting of the price from the observed price and they may be executed only with a long time lag. Liquidity movement has a straight impact on transaction price. The worse is liquidity in the market, the more different the price will be from its "fair" or "true" value. A perfectly liquid market is the one where market participants would get the same price of the security irrespective of the transaction time, transaction quantity and of the transaction type (buy or sell).

Since in reality it is impossible to achieve perfectly liquid market one is interested in the liquidity degree of the market or individual stock. In that sense liquidity may be defined as a degree of realization of big transactions in specified time interval and with minimal influence on the future prices. Therefore, it is necessary to measure such liquidity degree in some way.

Liquidity problems are complex and require analysis of different aspects of trading activity. That is the reason why it is impossible to have one definition of liquidity or one general measure of liquidity. In the recent literature, liquidity measures consider one aspect of liquidity - *one-dimensional liquidity measures*, or combine a several liquidity aspects - *multi-dimensional liquidity measures*. The following four aspects of liquidity are distinguished in the literature.

Trading time. The ability to execute a transaction immediately at the prevailing quote. It is measured by the waiting time or duration between subsequent transaction or by its inverse - the number of trades per time unit.

Tightness. The ability to buy and sell an asset at about the same price at the same time. Tightness shows the cost associated with transacting, or the cost of immediacy. The natural measure of tightness are different versions of the spread.

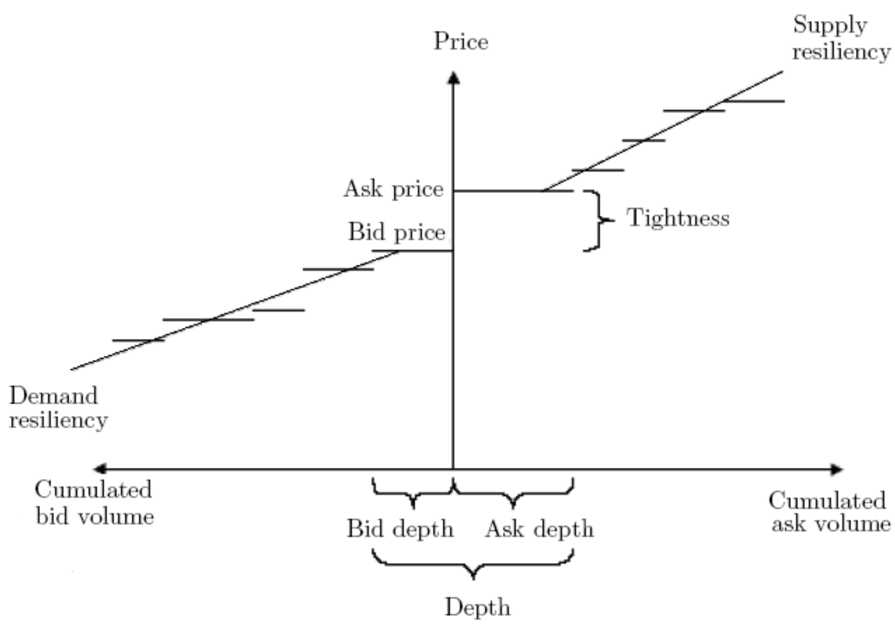

Figure 2.1: Four liquidity aspects in a static image of the limit order book.

Depth. The ability to buy or sell a certain amount of an asset at a particular bid-ask prices. It is measured by the volume of shares available for immediate trade on both sides of the market, by order ratio, trading volume or the flow ratio.

Resiliency. The ability to buy and sell a certain amount of an asset with a small influence on quoted prices. It indicates the speed at which prices revert to the previous level if there is a disturbance in the process of price formation. This aspect of liquidity can be described by intraday returns, or the liquidity ratio.

These four aspects of liquidity, represented in a static image of the limit order book, are depicted in Fig. 2.1. Every dimension of liquidity is subject to change at every point in time.

In the subsequent section we will introduce some of one-dimensional and multi-dimensional liquidity measures based on intraday transaction data and their properties.

2.1 One-dimensional liquidity measures

One-dimensional liquidity measures capture the following liquidity dimensions: volume of trade, the time between consecutive trades and spread.

2.1.1 Volume-related liquidity measures

We will represent several measures that are based on some volume dimension of trade and their reversal measures which calculate the time needed to realize that volume dimension.

The *trading volume* [52] in a certain time interval I is

$$V_I = \sum_{i=1}^{N_I} v_i$$

where N_I denotes the number of trades in the time interval I and v_i is the number of shares of trade i. A higher trading volume measure indicates a higher liquidity.

The *volume imbalance* or the *net traded volume* [66] measures the difference between the buyer initiated volume and the seller initiated volume during a certain period of time. The volume imbalance gives an information about direction of price changes. Positive volume imbalance indicates excess demand over supply and the price will increase. Negative volume imbalance indicates excess supply over demand and the price will decrease. Also, positive/negative volume imbalance indicates higher liquidity on the ask/bid side than on the bid/ask side.

The *turnover* [13] is the total money value of transactions over the time interval I. It is given with

$$TO_I = \sum_{i=1}^{N_I} p_i v_i$$

where p_i is the price of trade i and N_I denotes the number of trades in the time interval I. The advantage of this measure is that it makes different stocks comparable.

The *volume duration* [33] is a reverse measure of the trading volume measure. It is time needed to trade a certain number of shares. Similarly, the *volume imbalance duration* and the *turnover duration* represent the time needed to trade a certain volume imbalance, that is a certain turnover. A higher duration indicates a lower liquidity.

The measures described are defined only if the transaction is realized and in the time of realization. The following measures exist at every point of time even if no transaction take place.

The *volume depth* [10] represents the sum of the best bid and best ask volumes at the time point t, i.e.

$$D_t = v_t^a + v_t^b.$$

This measure captures the total available volume to trade at the best bid and ask prices. A higher depth indicates a higher liquidity. If the depth is divided by two it is called the *average volume depth* [61], [32], [14]. Since the best bid and ask volumes do not necessarily move in common, it is interesting to investigate them separately [47]. Another version of volume depth measure is the *logarithmic volume depth* [12], defined as a sum of the logarithms of the best bid and ask volume, given by

$$Dlog_t = ln(v_t^a) + ln(v_t^b) = ln(v_t^a v_t^b).$$

This measure improves distributional properties of the volume depth.

The *money depth* [66] is defined as

$$D\$_t = \frac{v_t^a p_t^a + v_t^b p_t^b}{2}.$$

Similarly to the turnover measure, the money depth makes different stocks comparable.

Larger bid or ask orders may exceed the depth on the bid or ask size. Such orders cannot be executed at the best bid or ask price and therefore they have to "walk the book". This situations are considered in the following measures.

2.1.2 Time-related liquidity measures

Time-related liquidity measures are concerned with trading intensity.

The *number of transaction per time unit* N_I [66], is the number of trades in the observable time interval I. The reverse of this measure is the waiting time between two subsequent transactions, or *duration*. The waiting time or duration between a trade executed at the time point t_i and the next trade executed at the time point t_{i+1} is simply

$$Dur_i = t_{i+1} - t_i.$$

The duration conveys the same information as the number of trades. Also, it may be interesting to calculate the *number of orders per time unit* and its reversal measure - the *duration between subsequent orders*. The higher the number of transactions/orders per time unit the higher is the liquidity. A higher duration between two subsequent transaction/orders indicates a lower liquidity.

2.1.3 Spread-related liquidity measures

Spread related liquidity measures represent transaction cost measure of liquidity. The smaller are the spread related measures the more liquid is the market. The basic one is the *absolute spread*, defined as a difference between the lowest ask and the highest bid price

$$S_t = p_t^a - p_t^b.$$

It is always positive and its lower limit is the minimum of the price change - the *tick size*. Absolute spread contains two components of transaction cost. The first one compensates market maker for inventory costs and order processing cost. Inventory costs represent compensation to market makers to bearing the risk of balancing inventory level due to market microstructure and stock market fluctuations. The order processing costs are the fees charged by market makers for matching buy and sell orders. This component is unrelated to underlying value of securities and therefore it is transient. The other one is information asymmetry cost that arises because market maker may trade with unidentified informed traders. The market makers compensate for their losses to informed traders by widening the spread when dealing with uninformed traders. Consequently, a higher proportion of informed traders in the market will cause an increase of the bid-ask spread.[3]

The *logarithm of absolute spread*, [36] is defined as

$$logS_t = ln(p_t^a - p_t^b).$$

The distribution of this version of spread is closer to the normal distribution than the distribution of the absolute spread.

Let p_t denote the price at which trade is executed at the time point t. Then the notations p_{t-1}^a and p_{t-1}^b stand for the pre-trade bid and ask prices, i.e. bid and ask prices prevailing before that trade. The following spread measures are defined in terms of the given notation.

The *relative spread* [26] is calculated as

$$S_t^{(rel)} = \frac{p_{t-1}^a - p_{t-1}^b}{p_t}.$$

Since p_t may be ask or bid price, this measure takes the market movements in consideration. The ask price will move market upward, and bid price will move market downward.

To make spreads of different stocks comparable the *proportional spread* is defined as

$$S_t^{(prop)} = \frac{p_t^a - p_t^b}{m_t} \tag{2.1}$$

where m_t is a *midquote* calculated as

$$m_t = \frac{p_t^a + p_t^b}{2}.$$

The *relative spread of logarithmic quotes* is calculated by analogy to the logarithmic return of the asset

$$Slog_t^{(rel)} = ln(p_t^a) - ln(p_t^b) = ln(\frac{p_t^a}{p_t^b}).$$

This measure represents a good proxy for the proportional spread [56]. Again, the *logarithmic relative spread of logarithmic quotes* defined as

$$LogSlog_t^{(rel)} = ln(Slog_t^{(rel)}) = ln(ln(\frac{p_t^a}{p_t^b})).$$

[3] See section 4.

improves the distributional properties.

The *effective spread* measures the actual cost of trading for investors. It is calculated as an absolute value of the difference between the transaction price at time t and pre-trade midquote m_{t-1}, that is

$$S_t^{(eff)} = |p_t - m_{t-1}|.$$

Since the trades sometimes occur at prices that are better than the posted quotes, the effective spread measure captures this "improved pricing." The effective spread smaller than half the absolute spread reflects trading inside of the bid-ask spread.

Using effective spread the *liquidity premium* [8] is defined as

$$LP_t = x_t^0(p_t - m_{t-1})$$

where x_t^0 is the *trade indicator variable* which takes value 1 if the trade is a buyer initiated and value -1 if trade is a seller initiated. It can be determined by the Lee and Ready rule [51]

$$x_t^0 = \begin{cases} 1, & p_t > m_{t-1} \quad (\textit{buyer initiated}) \\ 0, & p_t = m_{t-1} \quad (\textit{undeterminated}) \\ -1, & p_t < m_{t-1} \quad (\textit{seller initiated}). \end{cases} \tag{2.2}$$

The liquidity premium is positive if the buyer pays more, or if the seller pays less than the midquote.

By an analogy to the proportional spread and relative spread, the *relative effective spread* and the *proportional effective spread* are defined by

$$S_t^{(releff)} = \frac{|p_t - m_{t-1}|}{p_t},$$

$$S_t^{(propeff)} = \frac{|p_t - m_{t-1}|}{m_{t-1}}.$$

These measures are comparable for different stocks.

2.2 Multi-dimensional liquidity measures

Multi-dimensional liquidity measures combine properties of the different one-dimensional liquidity measures.

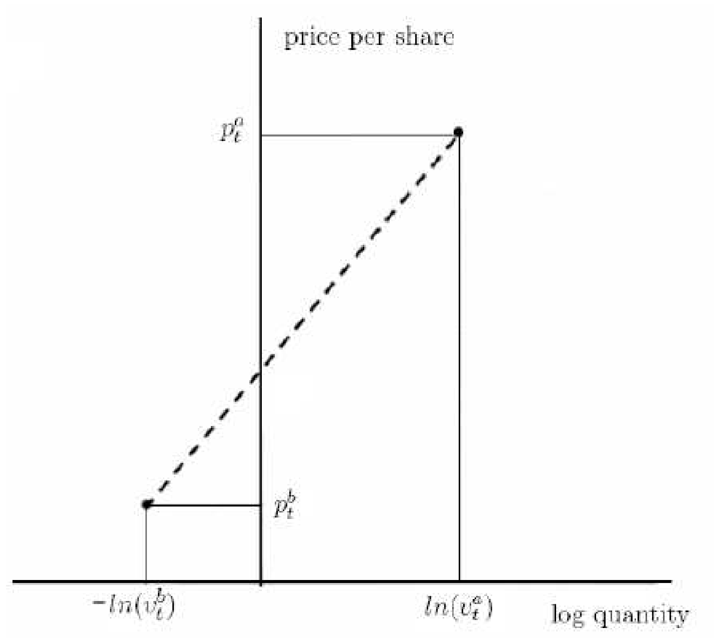

Figure 2.2: The quote slope.

The *quote slope* measure [42] is the spread divided by the logarithmic volume depth

$$QS_t = \frac{S_t}{Dlog_t} = \frac{p_t^a - p_t^b}{ln(v_t^a) + ln(v_t^b)} = \frac{p_t^a - p_t^b}{ln(v_t^a v_t^b)}.$$

As depicted in Fig. 2.2, the quote slope is the slope of the line connecting the bid and ask price/quantity pairs. The larger bid and ask volumes or the closer are bid and ask quotes to each other, the flatter is the slope of the quote and market becomes more liquid.

The *logarithmic quote slope* [42] has the relative spread of logarithmic quotes in the numerator instead of the absolute spread

$$LogQS_t = \frac{Slog_t^{(rel)}}{Dlog_t} = \frac{ln(\frac{p_t^a}{p_t^b})}{ln(v_t^a v_t^b)}.$$

Since the ask price is always higher than the bid price, the quote slope and the logarithmic quote slope are always positive.

Correction of the logarithmic quote slope for a market moving in one direction is *adjusted logarithmic quote slope* [62], defined as

$$LogQSadj_t = \frac{ln(\frac{p_t^a}{p_t^b})}{ln(v_t^a v_t^b)} + \frac{|ln(\frac{v_t^b}{v_t^a})|}{ln(v_t^a v_t^b)} \cdot ln(\frac{p_t^a}{p_t^b})$$

$$= LogQS_t \cdot (1 + |ln(\frac{v_t^b}{v_t^a})|).$$

The correction factor $|ln(\frac{v_t^b}{v_t^a})|$ is zero if volumes on the bid and ask side are equal. If either bid or ask volume are higher than the other, the correction factor is larger than one. Therefore, if this measure is increasing, then the liquidity is decreasing.

The another version of the quote slope liquidity measure is the *composite liquidity* [14]. It is the proportional spread divided by the money depth

$$CL_t = \frac{S_t^{(prop)}}{D\$_t} = \frac{2(p_t^a - p_t^b)}{m_t(v_t^a p_t^a + v_t^b p_t^b)}.$$

A higher composite liquidity measure indicates a lower liquidity.

The liquidity ratios combine some of volume-based liquidity measures and return. The *liquidity ratio* 1 compares the turnover to the absolute price change during the time interval I,

$$LR1_I = \frac{TO_I}{|r_I|} = \frac{\sum_{i=1}^{N_I} p_i v_i}{|r_I|}.$$

If $r_I = 0$ the liquidity ratio 1 is set to be zero. The higher is volume the more price movement can be absorbed. Therefore, the higher liquidity ratio 1 indicated the higher liquidity in the market.

The reverse of the defined liquidity ratio 1 is the *return per turnover* [2], given by

$$\frac{1}{LR1_I},$$

or the *Martin index* [7], given by

$$M_I = \sum_{i=2}^{N} \frac{(p_i - p_{i-1})^2}{TO_I}.$$

The *liquidity ratio* 2 [4] is defined as an average price change of transaction

$$LR2_I = \frac{\sum_{i=1}^{N_I} |r_I|}{N_I}.$$

The higher is the liquidity ratio 2, the lower is the liquidity.

The *flow ratio* [60] is a combination of the turnover and average duration in the analyzed time interval I,

$$FR_I = \frac{\sum_{i=1}^{N_I} p_i v_i}{\frac{1}{N_I - 1} \sum_{i=2}^{N_I} Dur_{i-1}}.$$

The flow ratio measures whether trading takes place in a few but large transaction or in lots of small trades. A higher flow ratio indicates a higher liquidity.

The *order ratio* [60] measures the size of market imbalance relative to the turnover in the transaction time t

$$OR_t = \frac{|v_t^b - v_t^a|}{p_t v_t}$$

A higher market imbalance causes a higher order ratio and therefore a lower liquidity. If the turnover in a certain transaction time is equal to zero the order ratio is set to be zero.

The information about increase or decrease of the absolute spread is not sufficient to determine whether the liquidity increases or not. The *market impact*, defined as the spread for a given volume V^*, is much better indicator of liquidity movement. It is calculated as

$$MI_t^{V^*} = p_t^{a,V^*} - p_t^{b,V^*}.$$

The market impact may be calculated separately for the two sides of the market. The market impact for the ask side and the bid side separately is given by the following two formulas:

$$MI_t^{a,V^*} = p_t^{a,V^*} - m_t,$$

$$MI_t^{b,V^*} = m_t - p_t^{b,V^*}.$$

The grater is the increase of spread per additional volume, the lower is the liquidity.

Let transaction of the size v be executed at K different prices with v_i shares traded at the price p_i where $i = 1, 2, ..., K$, that is

$$\sum_{i=1}^{K} v_i = v.$$

The *price impact* [15] is the execution cost dependent on the prevailing demand and supply schedules in the market. The price impact of the trade is then defined in terms of the appropriately signed percentage difference between the weighted-average execution price and the pre-trade midquote for the bid and ask side separately. For the ask side of the order book (buy order) the price impact is

$$PI^a(v) = ln(\frac{\sum_{i=1}^{K} p_i v_i}{v m_{t-1}}). \tag{2.3}$$

For the bid side (sell order) the price impact is

$$PI^b(v) = -ln(\frac{\sum_{i=1}^{K} p_i v_i}{v m_{t-1}}) \tag{2.4}$$

where m_{t-1} is, as before, the pre-trade midquote. A negative sign in front of the logarithm in the equation (2.4) is set because the expression in the brackets is in the interval $(0, 1)$. A higher price impact indicates a lower liquidity.

The *depth for price impact* measures the number of shares to be traded before the price move a certain amount of k ticks away from the midquote. This measure can be calculated for the bid and ask side of the market separately. The grater depth for the price impact means that the market can absorb a greater volume without significant movement in price, meaning more liquidity for the security. Obviously, the price impact is an inverse measure of the depth for the price impact.

Chapter 3

Time series

In this chapter we introduce a basic theory of univariate and multivariate time series of interest in the present thesis - the autoregression and moving-average time series.

3.1 Univariate time series

A *time series* is a stochastic process $\{X_t(\omega),\ t \in T\}$ defined on the probability space $(\Omega, \mathcal{A}, P)$, where T is a set of discrete points in time. In other words, it is a sequence of numbers that represents observations of some system's features that are usually taken at equidistant time points. In fact, the requirement of equidistant time points is just for simplification sake. In real problems, there is often a need to observe some feature in non-equidistant time points as, for example, observing price changes transaction by transaction. Even when we observe daily price changes, we have to notice that stock prices are not observed on weekends and holidays. In the present section we will consider a sequence of real-valued random variables indexed by $\mathbb{Z}$. The notation X_t means that the variable X has a realization at the time point t and that it comes before X_{t+1}, the realization of variable X at the time point $t+1$. As first we will define two important stochastic processes - the white noise and the random walk.

Definition 3.1 *The sequence $\{\varepsilon_t\}$, consisting of independent (uncorrelated) random variables with equal mean of zeros and equal finite variances σ^2 is the white noise, i.e.*

$$E(\varepsilon_t \varepsilon_s) = 0$$

$$E(\varepsilon_t) = 0$$
$$E(\varepsilon_t^2) = \sigma^2 < \infty$$

for every $t, s \in \mathbb{Z}$ *and* $t \neq s$. *If* $\{\varepsilon_t\}$ *are not only independent, but also identically distributed with zero means and variance* σ^2 *(i.i.d.*$(0, \sigma^2)$ *property) the white noise is called the strict white noise.*

The white noise is an extremely irregular and unpredictable process. In statistics, it is known as the *purely random process.*

Definition 3.2 *The random walk process is a process whose first differences are white noises, i.e.*

$$X_t - X_{t-1} = \varepsilon_t.$$

Apart from the white noise process, the random walk must be defined as being initiated at $t = 0$ *or* $t = 1$. *If its first differences are strict white noises, the random walk is called the strict random walk.*

In the present thesis, the white noise and random walk would always mean the strict white noise and strict random walk.

We have seen a simple relationship between these two processes but it has to be noticed a strong difference between them. While the white noise is an extremely irregular and unpredictable process, the random walk is characterized by slow changes and high predictability.

Now we can represent two time series models of interest in the present thesis - autoregressive and moving average models with their most important properties.

3.1.1 Autoregressive and moving-average time series

Definition 3.3 *An autoregressive process of order p,* $AR(p)$ *is defined as*

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + ... + \phi_p X_{t-p} + \varepsilon_t$$

where ε_t *is the white noise and* ϕ_i, $i = 1, 2, ...p$ *are fixed real numbers.*

A moving-average process of interest here is defined as follows.

Definition 3.4 *A moving-average process of order q,* $MA(q)$ *is defined as a weighted sum of the subsequent observations of white noise*

$$X_t = \varepsilon_t + \theta_1 \varepsilon_{t-1} + ... + \theta_q \varepsilon_{t-q}.$$

Moving-average processes are *finite independent*, that is X_t and X_{t-j} are uncorrelated for all $j > q$. For any time series X_t and $n \in \mathbb{Z}$, the *lag operator* L is defined as

$$L^n X_t = X_{t-n}.$$

Obviously,

$$L^0 X_t = X_t,$$

$$L^{-n} X_t = X_{t+n}.$$

Using the lag operator L, the $AR(p)$ and $MA(q)$ processes can be rewritten as

$$AR(p) : \varepsilon_t = (1 - \phi_1 L - \phi_2 L^2 - ... - \phi_p L^p) X_t, \tag{3.1}$$

$$MA(q) : X_t = (1 + \theta_1 L + \theta_2 L^2 + ... + \theta_q L^q) \varepsilon_t. \tag{3.2}$$

Setting z instead of L in the expressions in brackets on the right-hand sides of equations (3.1) and (3.2) we define the *lag polynomial* for an $AR(p)$ process

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - ... - \phi_p z^p$$

and the lag polynomial for the $MA(q)$ process

$$\Theta(z) = 1 + \theta_1 z + \theta_2 z^2 + ... + \theta_q z^q.$$

Therefore, both processes have shorter representations given by

$$AR(p) : \varepsilon_t = \Phi(L) X_t,$$

$$MA(q) : X_t = \Theta(L) \varepsilon_t.$$

The advantage of using lag polynomials is not only in shortening notation. The first advantage is checking the stability of autoregression and moving-average processes. If we consider the lag polynomial of the $AR(p)$ process

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - ... - \phi_p z^p,$$

according to The Fundamental Theorem of Algebra, the characteristic equation $\Phi(z) = 0$ has p solutions or roots in $\mathbb{C}$ if roots are counted by their multiplicity. If the modulus of all p roots are larger than one, i.e. if all roots lie outside of the unit circle $AR(p)$ process is stable, otherwise it is not. For example, the lag polynomial of the $AR(1)$ model is

$$\Phi(z) = 1 - \phi z.$$

Its root is $1/\phi$, so the condition that the root of the $AR(1)$ process lies outside of the unit circle is equivalent to the condition that $|\phi| < 1$. The same holds for a moving-average. It is stable if all roots of its characteristic equation $\theta(z) = 0$ lie outside of the unit circle.

The lag polynomials have all algebraic properties like ordinary polynomials, which enables establishing of a relationship between autoregressive and moving-average processes. If $AR(p)$ process, written as

$$\Phi(L)X_t = \varepsilon_t$$

is stable, its inverted lag polynomial converges as a power series (polynomial of infinite order). Then one can write

$$X_t = \Phi^{-1}(L)\varepsilon_t.$$

Since for an $AR(1)$ process, the stability implies that $|\phi| < 1$, using the geometric series one may obtain

$$\Phi^{-1}(z) = 1 + \phi z + \phi^2 z + \dots .$$

Hence

$$X_t = \sum_{i=0}^{\infty} \phi^i L^i \varepsilon_t = \sum_{i=0}^{\infty} \phi^i \varepsilon_{t-i},$$

which is the moving-average representation of the infinite order $MA(\infty)$ of an $AR(1)$ process. More generally, every stable finite order $AR(p)$ process, defined with

$$\Phi(L)X_t = \varepsilon_t,$$

has $MA(\infty)$ representation

$$X_t = \Phi(L)^{-1}\varepsilon_t.$$

Similarly, under the same condition a moving average process $MA(q)$ of finite order, given by

$$X_t = \Theta(L)\varepsilon_t$$

has an autoregression representation of infinite order given by

$$\varepsilon_t = \Theta^{-1}(L)X_t.$$

If an autoregressive (moving-average) process of finite order has a moving average (autoregression) representation of infinite order, we say that it has an *invertible* property. A finite-order stationary autoregressive process has roots larger than one in the modulus, therefore it is stable, i.e. invertible. Because of the finite variance, every moving-average process of finite order is stationary whatever its coefficient might be. A moving average process of infinite order

$$X_t = \sum_{i=0}^{\infty} \theta_i \varepsilon_{t-i}$$

is stationary only if its coefficients form a convergent series, that is if

$$\sum_{i=0}^{\infty} \theta_i^2 < \infty.$$

Apart from autoregressive processes, stationary moving average processes may not be stable, and therefore they may be non-invertible. For example, a finite moving average process

$$Y_t = \varepsilon_t - \varepsilon_{t-1} = (1 - L)\varepsilon_t$$

is stationary, but it is not stable (invertible) since its characteristic root is one. Invertible representations have the most satisfactory properties in forecasting since they provide that the future is generated by the past. If the process is non-invertible, one would need that the past is generated by the future which is not a reasonable strategy. The following theorem provides that every weakly stationary process may be decomposed into two mutually uncorrelated parts, one purely deterministic and the other purely non-deterministic, with satisfactory forecast properties.

Theorem 3.1 *(Wold's Theorem)[68] Any weakly stationary stochastic process X_t can be uniquely represented in the form*

$$X_t = X_t^d + \sum_{i=0}^{\infty} \theta_j \varepsilon_{t-1}$$

where

(i) $\theta_0 = 1$ *and* $\sum_{i=0}^{\infty} \theta_i^2 < \infty$

(ii) ε_t *are white noises*

(iii) $E(\varepsilon_t X_s^d) = 0$ *for every* $t, s > 0$*;* ε_t *is the error in forecasting* X_t*, i.e.*

$$\varepsilon_t = X_t - E^*(X_t | X_{t-1}, X_{t-2}, ...)$$

where E^* *stands for a linear projection.*

(iv) X_t^d *is a deterministic process and it can be predicted from a linear function of the lagged* X_t*.*

The Wold's theorem says that any weakly stationary process (even non-linear) has a linear representation: a deterministic part and a purely non-deterministic part represented with stationary infinite moving-average process. Hence, this theorem provides that some non-invertible processes have a moving average representation with satisfactory forecast properties, which is very important for dynamic analysis.

The main models of interest in the present section are autoregressive models. There are two important questions: lag order determination and model estimation. Lag order can be determined according to *partial autocorrelation function* ($PACF$), which is the highest order autoregression coefficient that is significantly different from zero according to its t-value. Except for this criterion, there are also *information criterions* like AIC-Akaike information criterion, AIC_c-corrected Akaike information criterion and BIC-Bayesian information criterion. We refer to [49] for more details about these criteria. Estimating of $AR(p)$ model is simple - it can be estimated by the least squares method. In fact, autoregressive models are the only time series models that can be estimate by the least squares method.

3.2 Multivariate time series

In the following subsection we will introduce the generalization of autoregressive processes - *vector autoregressive* process, *VAR*. All variables and error term in this process are vectors. Regarded as a system of individual equations, a vector autoregressive process represents the system of linear multiple regressions - multivariate regression.

3.2.1 Vector autoregressions

The vector autoregression *VAR* (p) model is a multivariate form of an $AR(p)$ process. At first it was proposed by Granger [34], then used by Sims [64] and Doan et al. [17]. It is based on the following ideas. The distinction between endogenous and exogenous variables is artificial, so all variables should be regarded as endogenous; restriction of coefficients imposed by traditional econometric models should be avoided. A model should rise possible restriction by itself at the time of estimation and analysis. A vector autoregresssion model of order p, $VAR(p)$ is defined by equation

$$X_t = \Phi_1 X_{t-1} + ... + \Phi_p X_{t-p} + \varepsilon_t, \tag{3.3}$$

where $X_t = (X_{1,t}, X_{2,t}, ..., X_{n,t})^T$ is a vector of n variables, $\varepsilon_t = (\varepsilon_{1,t}, \varepsilon_{2,t}, ..., \varepsilon_{n,t})^T$ is a vector of unobservable white noise errors and $\Phi_j,\ j = 1, ..., p$ are coefficients matrices of dimension $n \times n$. For example, a bivariate $VAR(p)$ is given by

$$X_{1,t} = \Phi_1^{11} X_{1,t-1} + \Phi_1^{12} X_{2,t-1} + ... + \Phi_p^{11} X_{1,t-p} + \Phi_p^{12} X_{2,t-p} + \varepsilon_{1,t}$$

$$X_{2,t} = \Phi_1^{21} X_{1,t-1} + \Phi_1^{22} X_{2,t-1} + ... + \Phi_p^{21} X_{1,t-p} + \Phi_p^{22} X_{2,t-p} + \varepsilon_{2,t}$$

or in the matrix form

$$\begin{bmatrix} X_{1,t} \\ X_{2,t} \end{bmatrix} = \begin{bmatrix} \Phi_1^{11} & \Phi_1^{12} \\ \Phi_1^{21} & \Phi_1^{22} \end{bmatrix} \begin{bmatrix} X_{1,t-1} \\ X_{2,t-1} \end{bmatrix} + ...$$

$$... + \begin{bmatrix} \Phi_p^{11} & \Phi_p^{12} \\ \Phi_p^{21} & \Phi_p^{22} \end{bmatrix} \begin{bmatrix} X_{1,t-p} \\ X_{2,t-p} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{bmatrix}.$$

Components $\varepsilon_{1,t}, \varepsilon_{2,t}, ..., \varepsilon_{n,t}$ are uncorrelated over time, but they are not necessarily contemporaneously orthogonal, hence the *VAR* is a variant of the *SUR* model. An error covariance matrix is given by

$$\Sigma = E(\varepsilon_t \varepsilon_t^T)$$

where $E(\varepsilon_t \varepsilon_{t-j}^T) = 0$ for $j \neq 0$. Since in a *VAR* model the variables are modelled as depending on its own lags only it is a *dynamic* system, apart from *simultaneous* system where variables are modelled being dependent on other variables at the same time point t. The *VAR* is weakly stationary if its both mean vector and covariance matrix are independent of time t. The covariance matrix for $VAR(p)$ is defined by

$$C_X(h) = C_X(t, t-h) = \begin{bmatrix} C_{X_1}(h) & C_{X_1X_2}(h) & ... & C_{X_1X_n}(h) \\ C_{X_2X_1}(h) & C_{X_2}(h) & ... & C_{X_2X_n}(h) \\ . & & & \\ . & & & \\ . & & & \\ C_{X_nX_1}(h) & C_{X_nX_2}(h) & ... & C_{X_n}(h) \end{bmatrix}.$$

Equation (3.3) can be rewritten as

$$\Phi(L) X_t = \varepsilon_t$$

where L is the lag operator and

$$\Phi(L) = I_n - \sum_{i=1}^{p} \Phi_i L^i$$

is a matrix lag polynomial. Checking the *VAR* stability is based on the same idea as in the univariate case, but it is slightly more difficult. In the univariate case, stability is checked by analyzing roots of the characteristic equation. In multivariate case the lag polynomial is the matrix polynomial

$$\Phi(z) = 1 - \sum_{i=1}^{p} \Phi_i z^i$$

that consists of n^2 separate polynomials in its rows and columns. Considering the determinant of $\Phi(z)$, which is an univariate polynomial, it can be shown that *VAR* is stable if all roots of $det\Phi(z) = 0$ lie outside of the unit circle. Sufficient condition for stability of *VAR* model is that all variables in the system are stationary [54].

As in the univariate case, there are also lag ordering and model estimation problems. Since the *VAR* is a special variant of *SUR* multivariate regressions with predetermined regressors, the least squares estimator is consistent. Since it is a multivariate regression with identical regressors, by the Kruskal's Theorem [48] it can be efficiently estimated by the least squares method or equation-by-equation. In order to keep efficiency, the lag order should be determined blockwise - fixing a common factor p for all matrix elements. As in univariate case, the methods for lag ordering are visual, i.e. using the *ACF* or *PACF* functions, or using a multivariate version of some information criteria mentioned in subsection 3.1.

3.2.2 Relationship between *VAR* variables

Traditional reporting of the estimated parameters or standard test statistics is not useful for the *VAR* model. The relationship between *VAR* variables cannot be analyzed by interpretation of estimated coefficients because the *VAR* estimated coefficients are matrices. The (j, k) entry of the matrix Φ_l cannot be viewed as the marginal reaction of the variable X_j to a unit change in X_k after l times periods. There are several concepts of interpreting relationship between *VAR* variables such as Granger's casuality, impulse response analysis [64] and variance decomposition [37], [28], [54]. While the Granger's casuality concept gives information about relationship between *VAR* variables based on predictability, impulse response and variance decomposition give information about the dynamic relationship between variables.

3.2.3 Granger's casuality

The Granger's casuality concept was introduced by Clive Granger [34]. The idea is simple and can be formulated as that cause cannot come after effect. The definition of casuality given by Granger is focussed on predictability - an event X causes the event Y, if Y can be better predicted by observing X than without observing X. Denoting the set $\{x_j,\ \ tj \leq j \leq t2\}$ by x_{t1}^{t2} , the formal definition of Granger's casuality is as follows.

Definition 3.5 *[34] Suppose that a universe consists of the vector of time series variables $(X, Y, Z)^T$, where Z may have an arbitrary dimension. Then, X is said to Granger cause Y if and only if*

$$E(Y_{t+1}|X_1^t, Y_1^t, Z_1^t)$$

is a better forecast than

$$E(Y_{t+1}|Y_1^t, Z_1^t),$$

where the better forecast is determined by the prediction error variance

$$E((Y_{t+1} - E(Y_{t+1}|...))^2).$$

The one-step forecast of the $VAR(p)$ model is defined by

$$X_{t+1|t} = \hat{\Phi}_1 X_t + \hat{\Phi}_2 X_{t-1} + ... + \hat{\Phi}_p X_{t-p+1}.$$

Clearly, the-two step forecast would be

$$X_{t+2|t} = \hat{\Phi}_1 X_{t+1|t} + \hat{\Phi}_2 X_t + ... + \hat{\Phi}_p X_{t-p+2},$$

etc.

In linear VAR models, the Granger's casuality can be easily checked by testing its coefficients. We will say that a VAR variable Y causes the VAR variable X if and only if at least one matrix coefficient of Y is not zero, or as Granger formulated it in the following theorem:

Theorem 3.2 *[34] Suppose $(X, Y, Z)^T$ has a VAR representation*

$$X_t = \sum_{j=1}^{p_{11}} \Phi_j^{11} X_{t-j} + \sum_{j=1}^{p_{12}} \Phi_j^{12} Y_{t-j} + \sum_{j=1}^{p_{13}} \Phi_j^{13} Z_{t-j} + \varepsilon_{1,t}$$

$$Y_t = \sum_{j=1}^{p_{21}} \Phi_j^{21} X_{t-j} + \sum_{j=1}^{p_{22}} \Phi_j^{22} Y_{t-j} + \sum_{j=1}^{p_{23}} \Phi_j^{23} Z_{t-j} + \varepsilon_{2,t}$$

$$Z_t = \sum_{j=1}^{p_{31}} \Phi_j^{31} X_{t-j} + \sum_{j=1}^{p_{32}} \Phi_j^{32} Y_{t-j} + \sum_{j=1}^{p_{33}} \Phi_j^{33} Z_{t-j} + \varepsilon_{3,t}.$$

Then,

$$X \textit{ causes } Y \Leftrightarrow \Phi_j^{21} \neq 0 \textit{ for at least one } j \in \{1, 2, ..., p_{21}\}.$$

$$Y \textit{ causes } X \Leftrightarrow \Phi_j^{12} \neq 0 \textit{ for at least one } j \in \{1, 2, ..., p_{12}\}.$$

A complete picture about the interaction between system's variables cannot be embraced only by Granger's casuality concept. It is necessary to examine a dynamic interaction between *VAR* variables.

3.2.4 Impulse response analysis

In order to analyze a dynamic interaction between *VAR* variables we will consider a $VAR(p)$ model with the lag polynomial $\Phi(L)$ written as

$$\Phi(L)X_t = \varepsilon_t, \tag{3.4}$$

$$\Phi(L) = I_n - \sum_{i=1}^{p} \Phi_i L^i. \tag{3.5}$$

If *VAR* is stable, that is if all its variables are stationary, it has a *VMA* (∞) representation

$$X_t = \Phi^{-1}(L)\varepsilon_t = \Psi(L)\varepsilon_t,$$

where

$$\Psi(L) = I_n + \sum_{i=1}^{\infty} \Psi_i L^i.$$

Hence the $VMA(\infty)$ representation of *VAR* model defined with (3.4) and (3.5) is

$$X_t = \sum_{i=0}^{\infty} \Psi_i \varepsilon_{t-i}, \quad \Psi_0 = I_n. \tag{3.6}$$

The vector moving average coefficient matrices can be obtained from the identity

$$\Phi(L)\Psi(L) = (I_n - \sum_{i=1}^{p} \Phi_i L^i)(I_n + \sum_{i=1}^{\infty} \Psi_i L^i) = I_n,$$

which leads to

$$\sum_{i=1}^{\infty}\Psi_i L^i - \sum_{i=1}^{p}\Phi_i L^i - (\sum_{i=1}^{p}\Phi_i L^i)(\sum_{i=1}^{\infty}\Psi_i L^i) = 0.$$

Thus, to find the coefficient matrices Ψ_i one has to set all resulting coefficient matrices of each power L equal to zero. For L^1 we have

$$\Psi_1 - \Phi_1 = 0 \Rightarrow \Psi_1 = \Phi.$$

The obtained equality can be rewritten as

$$\Psi_1 = \sum_{i=1}^{1}\Psi_{1-i}\Phi_i.$$

For L^2 we have

$$\Psi_2 - \Psi_1\Phi_1 - \Phi_2 = 0 \Rightarrow \Psi_2 = \Psi_1\Phi_1 + \Phi_2 = \sum_{i=1}^{2}\Psi_{2-i}\Phi_i.$$

Continuing with the same procedure, we obtain all coefficient matrices Ψ_j for $j \geq 1$. Hence

$$\Psi_0 = I_n,$$

$$\Psi_j = \sum_{i=1}^{j}\Psi_{j-1}\Phi_i$$

where $\Phi_j = 0$ when $j > p$. A moving-average representation of the $VAR(p)$ model gives a straightforward form to analyze the dynamic relations among the $VAR(p)$ variables. To see that we will consider $VMA(\infty)$ representation of $VAR(p)$ given in (3.6) for s steps ahead

$$X_{t+s} = \Psi(L)\varepsilon_{t+s} = \sum_{i=0}^{\infty}\Psi_i\varepsilon_{t+s-i}.$$

The effect of a unit change in ε_t on X_{t+s} is

$$\frac{\partial X_{t+s}}{\partial \varepsilon_t} = \Psi_s.$$

The ε_t's represent shocks or innovation in the system. Therefore the Ψ_s, $s = 1, 2, ...$ matrices represent the model's response to a unit shock at time point t in each of the variables which will be realized s periods ahead. They are called the *dynamic multipliers*.

The response of X_j to a unit innovation in the k-th variable that has occurred s periods ago is given by

$$\frac{\partial X_{j,t+s}}{\partial \varepsilon_{k,t}} = \Psi_{jk,s}$$

that is with the (j,k) entry of Ψ_s. Viewed as a function of $s = 0, 1, 2, \dots$, (j,k) entries of matrix Ψ_s represent the *impulse response functions* of a variable X_j to a unit innovation in the k-th variable. The *VAR* with n variables has n^2 such functions. Generally, an impulse response function traces the effect of a one-time shock to one of the innovations on current and future values of the endogenous variables. A plot of these values as a function of s is the graphical representation of the impulse response function. If

$$\lim_{s\to\infty} \frac{\partial X_{j,t+s}}{\partial \varepsilon_{k,t}} = 0$$

is satisfied, then the shock to $X_{j,t}$ has no permanent impact on the level of $X_{k,t}$. Also, one may be interested in the accumulated effect over several or more periods of a shock in variable. This effect can be determined by summing up the *VMA* coefficient matrices. More precisely, the (j,k) entry of

$$\Upsilon_n = \sum_{i=0}^{n} \Psi_i$$

is the accumulated response variable X_j over n periods to a unit shock in the variable X_k. The total accumulated effect for all future periods is given by *long-run-effect* or *total multipliers*, obtained by summing up all *VMA* coefficients

$$\Upsilon_\infty = \sum_{i=0}^{\infty} \Psi_i.$$

If the *VAR* variable $X_{k,t}$ does not Granger cause the *VAR* variable $X_{j,t}$, then the variable $X_{j,t}$ does not react to a shock in $X_{k,t}$, i.e. the impulse response variable $X_{j,t}$ to the shock in $X_{k,t}$ is zero. To see that, suppose that $X_{2,t}$ does not Granger cause $X_{1,t}$. Then, all coefficient matrices Φ_i, $i = 1, 2, \dots, n$ have zero on $(1,2)$ place. Therefore, the coefficient matrices Ψ_i, $i = 1, 2, \dots, n$ in the moving-average representation will also have zero on $(1,2)$ place.

The problem with the described impulse response functions is that components of ε_t are usually contemporaneously correlated. That is the reason why it is difficult to recognize an effect of a single change in one component of ε_t keeping all other components unchanged, since they usually come at the same time. That problem can be overcamed by orthogonalizing the errors. If we find a matrix Q such that $QQ^T = \Sigma$ and define $\varepsilon_t^* = Q^{-1}\varepsilon_t$, then

$$E(\varepsilon_t^* \varepsilon_{t^*}^T) = E(Q^{-1}\varepsilon_t\varepsilon_t^T(Q^{-1})^T) = I_n.$$

Since the error covariance matrix Σ is real, symmetric and if it is positive definite, it can be decomposed as

$$\Sigma = ADA^T$$

where A is a lower triangular matrix with ones on the main diagonal and D is a positive diagonal matrix. Setting $u_t = A^{-1}\varepsilon_t$, we obtain

$$E(u_t u_t^T) = A^{-1}\Sigma(A^{-1})^T = A^{-1}ADA^T(A^{-1})^T = D.$$

We say that components of the vector u_t are *mutually uncorrelated* and u_t is called *orthogonal innovation*. The covariance matrix Σ can be decomposed into the product of a left triangular matrix and its transpose

$$\Sigma = ADA^T = AD^{1/2}D^{1/2}A^T = PP^T.$$

It is the Cholesky decomposition of the covariance matrix Σ. Setting $\varepsilon_t^* = P^{-1}\varepsilon_t$ we obtain

$$\varepsilon_t^* = P^{-1}\varepsilon_t = D^{-1/2}A^{-1}\varepsilon_t = D^{-1/2}u_t.$$

Therefore $E(\varepsilon_t^*\varepsilon_{t^*}^T) = I_n$. Now, the orthogonalized *VMA* representation is

$$X_t = \sum_{i=0}^{\infty} \Psi_i\varepsilon_i = \sum_{i=0}^{\infty} \Theta_i\varepsilon_i^*$$

where $\varepsilon^* = P^{-1}\varepsilon$ and $\Theta_i = \Psi_i P$. Hence, the orthogonalized dynamic multipliers are

$$\frac{\partial X_{t+s}}{\partial \varepsilon_t^*} = \Theta_s = \Psi_s P$$

and orthogonalized impulse response functions are

$$\frac{\partial X_{j,t+s}}{\partial \varepsilon_{k,t}^*} = \Theta_{jk,s} = \sum_{l=1}^{n} \Psi_{jl,s}P_{lk}.$$

The triangularity of the matrix P implies triangularity of the matrices Θ_i, $i = 1, 2, ...$. That is the reason why the shocks $\varepsilon_{k,t}, ..., \varepsilon_{n,t}$ do not have contemporaneous affect on $X_{1,t}, ..., X_{k-1,t}$, but have on $X_{k,t}, ..., X_{n,t}$. Subsequently, the ordering of variables is important and different ordering produces different impulse response functions. In order to find the ordering for which the resulting interpretations will be consistent to, practitioners usually proceed as follows. The first variable should be selected such that it is the only one with potential immediate impact on all other components of X_t. The second variable may have immediate impact on the last $n-2$ components, but not of the first one, and so on. The other suggestion is the ordering of variables according to Granger-casual ordering. A practitioners also can order variables in their own way, if it is reasonable for her/him, usually due to her/his theoretical knowledge about variables.

3.2.5 Variance decomposition

Consider now again stable $VAR(p)$ with n components and its orthogonal VMA representation

$$X_t = \sum_{i=0}^{\infty} \Theta_i \varepsilon^*_{t-i}.$$

Then, the s step-ahead forecast for X_t is given by

$$E_t(X_{t+s}) = \sum_{i=s}^{\infty} \Theta_i \varepsilon^*_{t+s-i}.$$

where E_t denotes the expectations formulated at a time t, based on the estimated VAR model. Hence, the s step-ahead forecast error is given by

$$err_{t+s} = X_{t+s} - E_t(X_{t+s}) = \sum_{i=0}^{s-1} \Theta_i \varepsilon^*_{t+s-i}.$$

The j-th component of the forecast error is given by

$$err_{j,t+s} = \sum_{i=0}^{s-1} \sum_{k=1}^{n} \Theta_{jk,i} \varepsilon^*_{k,t+s-i} = \sum_{k=1}^{n} \sum_{i=0}^{s-1} \Theta_{jk,i} \varepsilon^*_{k,t+s-i}.$$

If the shocks are both serially and contemporaneously uncorrelated, the error variance is

$$Var(err_{j,t+s}) = \sum_{k=1}^{n} \sum_{i=0}^{s-1} Var(\Theta_{jk,i} \varepsilon^*_{k,t+s-i})$$

$$= \sum_{k=1}^{n} \sum_{i=0}^{s-1} \Theta_{jk,i}^2 Var(\varepsilon_{k,t+s-i}^*).$$

The sum

$$\sum_{i=0}^{s-1} \Theta_{jk,i}^2 Var(\varepsilon_{k,t+s-i}^*)$$

corresponds to the error variance generated by innovations to one specific X_k. The sum

$$\sum_{k=1}^{n} \sum_{i=0}^{s-1} \Theta_{jk,i}^2 Var(\varepsilon_{k,t+s-i}^*)$$

corresponds to the error variance generated by the sum of all innovation responses. Subsequently, the ratio

$$R_{jk,s}^2 = \frac{\sum\limits_{i=0}^{s-1} \Theta_{jk,i}^2 Var(\varepsilon_{k,t+s-i}^*)}{\sum\limits_{k=1}^{n} \sum\limits_{i=0}^{s-1} \Theta_{jk,s}^2 Var(\varepsilon_{k,t+s-i}^*)}$$

is a relative measure of how important innovations of the k-th variable are in the explaining the variation in the variable j at an s step-ahead forecast. The described procedure is called *variance decomposition*, and $R_{jk,s}^2$ are called the *variance decomposition coefficients* of $X_{j,t}$ at the horizon s.

The impulse response functions traces the effect of a shock to one endogenous variable on the other variables in the *VAR*. Variance decomposition separates the variation in an endogenous variable into the components shock to the *VAR*. In other words, the impulse response functions give the answer to the question how the system's endogenous variables respond dynamically to the exogenous shock. The variance decomposition gives the answer to the question which shocks are the primary causes of variability in the exogenous variables.

Chapter 4

Hasbrouck's *VAR* model

4.1 Asymmetrically informed market

The market microstructure concept has been defined in different ways, by focusing on different aspects of it. Following the definition given by O'Hara [58], "market microstructure is a study of the process and outcomes of exchanging assets under a specific set of rules. Microstructure theory focuses on how specific trading mechanisms affect the price formation process." For the purpose of analysis in this thesis, we will focus on the information aspect of market microstructure, and on the impact of the information on the price formation.

The theory of efficient market assumes that a market is anonymous and all the participants in the market are equally informed about traded instrument. Therefore, no participant can make economic profit by trading such information, and information contained in trades is immediately reflected in stock prices. However these assumptions would hardly hold in practice. In reality, all information are not available to all participants at the same time, hence some market participants have a definite advantage over the others. Moreover, although information is public, there is still difference in speed of processing them by different participants, which produces lag effect between the news announcement and trade realization.

Traders may be classified into informed traders - traders with superior information and uninformed traders or liquidity traders - traders with only public information. Informed traders may possess information about true value of the security, fundamentals, quantities or about who is informed. They tend to trade the specific stock in which they have private information.

Liquidity traders trade to smooth consumption or to adjust the risk return profiles on their portfolios. They buy stocks if they have excess of cash or have become more risk tolerant, and they sell stocks if they need cash or have become less risk tolerant.

Information influences the bid-ask spread in the market and hence liquidity. Before the information is publicly available the spread tends to be wider, producing a lot of volatility in the market. The informed traders knowing that the spread will narrow down once when the information become public, tend to take the liquidity from the market by executing trades at available price. The uninformed traders or traders who do not have access to the information that can affect the market value of the trading instrument, tend to adopt different trading strategies compared to the informed traders. Informed traders will make a benefit at an expense of the uninformed traders. Consequently, the uninformed traders seek to identify the counterpart, while informed traders seek to hide their identity.

The presence of informed and uninformed traders causes an asymmetric distribution of information among market participants which is the basis of asymmetric information market theory. Bagheot [6] was the first to consider market with heterogeneously informed traders. Then this problem is analyzed by Copeland et al. [16] and formulated and developed by Kyle [50], Glosten et al. [30], Easely et al. [18] Admati et al. [1], Foster et al. [27].

The market makers possessing only public information expose bid and ask quotes to the trading participants and faces informed and uninformed traders, but they cannot distinguish between them. Because market makers have an access only to the public information, they compensate for the loss that appears from trading with informed traders by fixing a spread.

Even if the difference between informed and uninformed traders is undistinguishable to market makers, informed traders can be recognized by observing trading activity. Informed traders transact only when they have superior information and tend to trade quickly larger quantities in order to benefit from their information before it becomes public. Under the presence of information, trades are clustered in time, so that the duration between trades are shorter and the prices adjusted very quickly in the calendar time. Also, in the presence of private information the bid-ask spread tends to be wider. Consequently, trades convey information and that is the driving force idea in the asymmetric information theory.

Hasbrouck [39] in his analysis of market microstructure found empirical evidence that in a market with asymmetrically informed traders, trades convey information which causes a persistent impact on the security price. According to Hasbrouck the magnitude of the price effect for a given trade size is generally held to be a positive function of the proportion of potentially informed traders in population, the probability that a private information signal has in fact been observed, and the precision of the private information. Also, information asymmetry is positively correlated to price impact of the trade and spread. However, some market imperfections like price discreteness, clearing fees, inventory control, order fragmentation, vitiates this correlation.

4.2 The model

In this subsection we will introduce Hasbrouck's model of the asymmetric information effect on the future prices. According to this model, trades are motivated by private information and/or liquidity needs. Price impact of the trade can appear on both transitory and permanent levels. The transitory price impact of the trade is attributed to the trades effects that drive current transaction prices away from the efficient price, i.e. from the price updated based on all available public information. Those effects are non-information based microstructure effects caused by market imperfections as price discreteness, market makers costs such as costs associated with inventory control and order processing, demand and supply effects, etc. The permanent response of security prices to trading activity is caused by information asymmetry between the public and private information, i.e. by the agent's belief about private information content of the trade. Hasbrouck modelled the quote revision and trade dynamics as a non-standard bivariate vector autoregressive system to examine the public and private information components in the price changes and the speed of price adjustment to these effects.

Hasbrouck considered the following trading mechanism: the transaction characterized by the signed volume x_t is realized at the time t and at price p_t. The signed volume is the size of transaction multiplied by trade indicator variable defined by (2.2). After the transaction realization and announcement of trade x_t the market makers post bid and ask quotes denoted by p_t^b and p_t^a. If these quotes are posted in the absence of trade x_t, they are posted due to non-trade public information and then x_t is set to be zero. In this notation, the transaction realized at the time t is realized at the bid or ask price prevailing before that transaction, denoted by p_{t-1}^b and p_{t-1}^a. The initial assumption in Hasbrouck's model is that the quotes are set symmetrically about the expected value of the security conditional on all public information expressed as

$$E((p_t^b + p_t^a)/2 - \mathcal{V}_\tau | \mathcal{F}_t) = (p_t^b + p_t^a)/2 - E(\mathcal{V}_\tau | \mathcal{F}_t) = 0, \qquad (4.1)$$

where $\mathcal{V}_\tau$ is the security value at some convenient terminal time τ in the distant future and $\mathcal{F}_t$ is the public information set at the time t. That makes reasonable to consider the midquote

$$m_t = \frac{p_t^b + p_t^a}{2}$$

as an unbiased proxy of the efficient price. Subsequent revision of the midquote given by

$$r_t = \frac{p_t^b + p_t^a}{2} - \frac{p_{t-1}^b + p_{t-1}^a}{2}$$

contains the information inferred from the trade x_t that occurs at a time t. Hasbrouck assumed that the public information arrives after the t-th trade, but before associate quote revision. Subsequently, the quote revision will reflect the public information as well as the private information.

Initially, Hasbrouck assumed that the relationship between the quote revision r_t and signed trade x_t is contemporaneously linear

$$r_t = bx_t + \nu_{1,t}.$$

The price impact of the trade is represented by the coefficient b while $\nu_{1,t}$ is the disturbance which reflects the public information. Many market imperfections mentioned above require some corrections of the proposed model. Price discreteness, inventory control effect, lagged adjustment to information, price smoothing, all of them cause serial dependence in quote revision. On the other hand, the order fragmentation causes serial dependence in trades. Therefore, neither quote revision nor signed trade can be viewed as exogenous variables. Hasbrouck suggested joint analysis of quote revision and trade by the bivariate *VAR* system

$$r_t = a_1 r_{t-1} + a_2 r_{t-2} + ... + b_0 x_t + b_1 x_{t-1} + b_2 x_{t-2} + ... + \nu_{1,t}, \qquad (4.2)$$

$$x_t = c_1 r_{t-1} + c_2 r_{t-2} + ... + d_1 x_{t-1} + d_2 x_{t-2} + ... + \nu_{2,t}. \qquad (4.3)$$

Theoretically, this model can be of infinite order, but for practical purposes it is truncated at some lag. Under serial correlation condition of the quote revision the initial assumption of quote symmetry from equation (4.1) has to be relaxed. The weaker version of asymmetry is given by

$$E((p_s^b + p_s^a)/2 - \mathcal{V}_\tau | \mathcal{F}_t) \to 0 \tag{4.4}$$

when $s \to \tau$. Therefore, the deviations of the midquote from the efficient price are transient. Intuitively, this relation expresses expectation that during some time quotes on average will return to their fair value.

Coefficient b_0 in the quote revision equation represents immediate impact of contemporaneous trade x_t. Coefficients $b_i, 1, 2, ...$ in the same equation capture transitory trade effects on prices and the innovation term $\nu_{1,t}$ represents the effect of non-trade public information. The innovation in the trade equation $\nu_{2,t}$ captures an unexpected transaction activity where the private information resides if such exists. It is modelled as a component of the trade relative to an expectation formed from a linear projection based on the history of previous transactions and quote revisions which is entirely known. It is not purely deterministic because the presence of liquidity traders will cause the noise component of $\nu_{2,t}$ that is uncorrelated with private information. Since the predictable portion of the trade conveys no new information, Hasbruck formally defined the informational impact of the trade as the ultimate impact on the stock price resulting from an unexpected component of the trade, i.e. the persistent price impact of the trade innovation. This impact probably would not be instantaneous, but rather occurs over a long period of time. The model assumes predetermined regressors, i.e. that the innovations $\nu_{1,t}$ and $\nu_{2,t}$ are uncorrelated with regressors. Also, it is assumed that they have zero mean

$$E(\nu_{1,t}) = E(\nu_{2,t}) = 0$$

and that they are jointly and serially uncorrelated

$$E(\nu_{1,t}\nu_{1,s}) = E(\nu_{2,t}\nu_{2,s}) = E(\nu_{1,t}\nu_{2,s}) = 0, \quad for\ every\ \ t \neq s.$$

The described *VAR* model is not entirely standard since it assumes that a market maker has information about all lagged quote revisions and lagged trades, as well as an information about contemporaneous trade available at time t. That means that the quote revision r_t contains all publicly available information at the time t, and that market makers act primarily on this information set. This model permits Granger's casuality running from trade to quote revision both contemporaneously and with lags. The model also permits Granger's casuality running from the lagged quote revision to trades, but it does not permit contemporaneous casuality running from quote revisions to trades. The presence of contemporaneous trade x_t, and assumption of predetermined regressors imply that errors are contemporaneously orthogonal, i.e. $E(\nu_{1,t}\nu_{2,t}) = 0$, which does not hold for the standard *VAR* model in general.

Under assumptions of predetermined regressors and contemporaneous orthogonality of errors $\nu_{1,t}$ and $\nu_{2,t}$, the least squares estimation of the described *VAR* model is consistent and efficient. The innovation in the quote revision equation represents a transitory effect of public information. Hasbrouck [43] and Stoll [63] argued that many market imperfections are of a transient character. On the other hand, information inferred from a trade due to asymmetric information is permanently impounded in the stock prices. The main point of Hasbrouck's analysis is that the innovation component of quote revision equation is effected by the public information, causing transitory price impact.

4.3 Cumulative impulse response - price impact

The original *VAR* system given by (4.2) and (4.3) may be rewritten in the matrix form

$$\begin{bmatrix} 1 & -b_0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} r_t \\ x_t \end{bmatrix} = \begin{bmatrix} a_1 & b_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} r_{t-1} \\ x_{t-1} \end{bmatrix} + \\ + \begin{bmatrix} a_2 & b_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} r_{t-2} \\ x_{t-2} \end{bmatrix} + ... + \begin{bmatrix} \nu_{1,t} \\ \nu_{2,t} \end{bmatrix}.$$

After multiplying both sides of the last equation by

$$\begin{bmatrix} 1 & -b_0 \\ 0 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} 1 & b_0 \\ 0 & 1 \end{bmatrix}$$

it becomes

$$\begin{bmatrix} r_t \\ x_t \end{bmatrix} = \begin{bmatrix} a_1 + b_0 c_1 & b_1 + b_0 d_1 \\ c_1 & d_1 \end{bmatrix} \begin{bmatrix} r_{t-1} \\ x_{t-1} \end{bmatrix} + \\ + \begin{bmatrix} a_2 + b_0 c_2 & b_2 + b_0 d_2 \\ c_2 & d_2 \end{bmatrix} \begin{bmatrix} r_{t-2} \\ x_{t-2} \end{bmatrix} + ... + \begin{bmatrix} \nu_{1,t} + b_0 \nu_{2,t} \\ \nu_{2,t} \end{bmatrix}.$$

Notice that after the last transformation the trade innovation $\nu_{2,t}$ remains the same, but the quote revision innovation is modified, and it contains a trade innovation part. Denoting

$$\nu_{1,t} + b_0\nu_{2,t} = \nu'_{1,t}$$

and assuming that these two time series are weakly stationary, by the Wold's theorem we have that the VAR is invertible and it has the vector-moving average representation of a infinite order given by

$$\begin{bmatrix} r_t \\ x_t \end{bmatrix} = \begin{bmatrix} \nu'_{1,t} \\ \nu_{2,t} \end{bmatrix} + \begin{bmatrix} a'_1 & b'_1 \\ c'_1 & d'_1 \end{bmatrix} \begin{bmatrix} \nu'_{1,t-1} \\ \nu_{2,t-1} \end{bmatrix} + \begin{bmatrix} a'_2 & b'_2 \\ c'_2 & d'_2 \end{bmatrix} \begin{bmatrix} \nu'_{1,t-2} \\ \nu_{2,t-2} \end{bmatrix} + ... \quad (4.5)$$

The coefficients in the $VMA(\infty)$ representation form the impulse response functions described in Section 3. Writing the last matrix equation as a system of equations, and after some grouping, we obtain

$$r_t = \nu_{1,t} + a'_1\nu_{1,t-1} + a'_2\nu_{1,t-2} + ... + b_0\nu_{2,t} + (a'_1 b_0 + b'_1)\nu_{2,t-1} + (a'_2 b_0 + b'_2)\nu_{2,t-2} + ...$$

$$x_t = c'_1\nu_{1,t-1} + c'_2\nu_{1,t-2} + ... + \nu_{2,t} + (c'_1 b_0 + d'_1)\nu_{2,t-1} + (c_2 b_0 + d'_2)\nu_{2,t-2} + ...$$

Finally, the system becomes

$$r_t = \nu_{1,t} + a^*_1\nu_{1,t-1} + a^*_2\nu_{1,t-2} + ... + b^*_0\nu_{2,t} + b^*_1\nu_{2,t-1} + b^*_2\nu_{2,t-2} + ... \quad (4.6)$$

$$x_t = c^*_1\nu_{1,t-1} + c^*_2\nu_{1,t-2} + ... + \nu_{2,t} + d^*_1\nu_{2,t-1} + d^*_2\nu_{2,t-2} + ... \quad (4.7)$$

where

$$a^*_i = a'_i, \quad c^*_i = c'_i, \quad b^*_0 = b_0,$$
$$d^*_i = c'_i b_0 + d'_i$$
$$b^*_i = a'_i b_0 + b'_i, \qquad i = 1, 2, ...$$

The coefficients b^*_i, $i = 1, 2, ...$ give the effect of a unit trade innovation on the midquote revision at a i period horizon. Since the model operates with the data indexed in tick time, the price impact is measured in units of transactions. The sum

$$\sum_{i=0}^{l} b^*_i$$

represents the impact of an unexpected trade on quote revisions after l transactions. The long-run impact of trade on quote revisions is given by

$$\sum_{i=0}^{\infty} b^*_i.$$

Hasbrouck's empirical findings indicate that the price impact takes many periods before it is fully realized. The cumulative sum of impulse responses creates the price impact function. Hasbrouck empirically found that price impact function is concave and it has positive horizontal asymptote. The asymptote of price impact function represents the total price impact of the trade.

4.4 Variance decomposition - trade informativeness

Hasbrouck assumed that the midquote may be divided in two unobservable components

$$m_t = e_t + s_t. \tag{4.8}$$

The term e_t is efficient price, that is the expected value of the asset conditional on all currently available public information modelled as a random walk process

$$e_t = e_{t-1} + \omega_t.$$

The efficient price is the permanent component of the midquote. The innovation ω_t reflects updates to the public information set. Being a white noise process, it has the following properties:

$$E(\omega_t) = 0,$$

$$E(\omega_t^2) = \sigma_\omega^2,$$

$$E(\omega_t \omega_\tau) = 0, \quad \tau \neq t.$$

The second component s_t is a zero mean stochastic process jointly covariance stationary with ω_t. It is a transitory component of the midquote. It represents the disturbance term that incorporates inventory control, price discreteness and other market imperfections that drive the midquote away from the efficient price. As we mentioned above, these imperfections are of transient character, which is compatible with the implication of weakly stationarity assumption

$$E_t(s_{t+k}) \to E(s_{t+k}) = 0, \quad k \to \infty$$

where E_t is the expectation formulated at a time t based on equation (4.8). The market's signal private information is defined as

$$x_t - E(x_t | \mathcal{F}_{t-1})$$

where $\mathcal{F}_{t-1}$ is the public information set available by the time t. The impact of the trade innovation on the efficient price innovation is

$$E(\omega_t|x_t - E(x_t|\mathcal{F}_{t-1})).$$

Therefore, the absolute measure of trade informativeness is

$$Var(E(\omega_t|x_t - E(x_t|\mathcal{F}_{t-1}))),$$

and its relative measure to the total public information is

$$Var(E(\omega_t|x_t - E(x_t|\mathcal{F}_{t-1})))/Var(\omega_t).$$

The problem with these measures is that the random walk decomposition is unobservable. It can be captured from Hasbrouck's VAR given by equations (4.2) and (4.3), that is by its moving average representation given by (4.6) and (4.7). We will denote the variance of quote revision innovation $\nu_{1,t}$ and trade innovation $\nu_{2,t}$ by

$$Var(\nu_{1,t}) = \sigma_1^2,$$
$$Var(\nu_{2,t}) = \Lambda.$$

If the public information set is defined by the trade and quote history $\mathcal{F}_t = x_t, r_t, x_{t-1}, r_{t-1}, ...$ then

$$Var(E(\omega_t|x_t - E(x_t|\mathcal{F}_{t-1}))) = Var(E^*(\omega_t|v_{2,t})) \equiv \sigma^2_{\omega,x},$$
$$Var(E(\omega_t|x_t - E(x_t|\mathcal{F}_{t-1})))/Var(\omega_t) = \sigma^2_{\omega,x}/\sigma^2_\omega \equiv R^2_\omega \tag{4.9}$$

where E^* stands for a linear projection. Therefore, R^2_ω tells us which part of variance in the random walk component of stock price, that is in efficient price, is attributable to the trade innovation. Theorem 4.1 gives the computational details. Before defining and proving the following theorem we will show that for every $MA(p)$ process, defined by $X_t = \theta(L)\varepsilon_t$, the autocovariance generating function is

$$g_X(z) = \sigma^2_\varepsilon \theta(z)\theta(z^{-1}) \tag{4.10}$$

where σ^2_ε is the variance of white noise process ε_t. Let us consider a double infinite moving average process given by

$$X_t = \sum_{j=-\infty}^{\infty} \theta_j \varepsilon_{t-j} = \theta(L)\varepsilon_t.$$

Since ε_t is white noise, then

$$\begin{aligned}
E(X_t) &= 0,\\
C_X(h) &= E(X_t X_{t-h})\\
&= E(\sum_{j=-\infty}^{\infty} \theta_j \varepsilon_{t-j} \sum_{k=-\infty}^{\infty} \theta_k \varepsilon_{t-h-k})\\
&= \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \theta_j \theta_{j-h}.
\end{aligned}$$

The last equality follows from

$$E(\varepsilon_{t-j}, \varepsilon_{t-k-h}) = \begin{cases} \sigma_\varepsilon^2, & j = h + k \\ 0, & otherwise \end{cases}$$

The autocovariance generating function is then

$$\begin{aligned}
g_X(z) &= \sum_{h=-\infty}^{\infty} C_X(h) z^h\\
&= \sum_{h=-\infty}^{\infty} (\sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \theta_j \theta_{j-h}) z^h.
\end{aligned}$$

If we set $j - h = k$, then

$$\begin{aligned}
g_X(z) &= \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \theta_j \theta_k z^{j-k}\\
&= \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} \theta_j \theta_k z^j z^{-k}\\
&= \sigma_\varepsilon^2 \sum_{j=-\infty}^{\infty} \theta_j z^j \sum_{k=-\infty}^{\infty} \theta_k z^{-k}\\
&= \sigma_\varepsilon^2 \theta(z) \theta(z^{-1}).
\end{aligned}$$

To see that this function generates all autocovariances for every $MA(p)$ process let us observe the $MA(1)$ process given by

$$X_t = \varepsilon_t + \theta \varepsilon_{t-1} = (1 + \theta L)\varepsilon_t.$$

Its autocovariance generating function is

$$\begin{aligned}
g_X(z) &= \sigma_\varepsilon^2 \theta(z) \theta(z^{-1})\\
&= \sigma_\varepsilon^2 (1 + \theta z)(1 + \theta z^{-1})\\
&= \sigma_\varepsilon^2 ((1 + \theta^2) z^0 + \theta z + \theta z^{-1}).
\end{aligned}$$

From the last equality there follows

$$\begin{aligned}
C_X(0) &= 1 + \theta^2\\
C_X(1) &= C_X(-1) = \theta\\
C_X(k) &= C_X(-k) = 0, \quad k = \pm 2, \pm 3, ...
\end{aligned}$$

Theorem 4.1 *[40] For the trade/quote-revision VMA, given by (4.5) and (4.6), the random walk variance, and the contribution of trades to the variance, defined by (4.8) are given by*

$$\sigma_\omega^2 = (1 + \sum_{i=1}^{\infty} a_i^*)^2 \sigma_1^2 + (\sum_{i=0}^{\infty} b_i^*)^2 \Lambda, \tag{4.11}$$

$$\sigma_{\omega,x}^2 = (\sum_{i=0}^{\infty} b_i^*)^2 \Lambda. \tag{4.12}$$

Proof. The *VMA* representation (4.6) and (4.7) implies

$$r_t = a^*(L)\nu_{1,t} + b^*(L)\nu_{2,t}$$

where a^* and b^* are the lag polynomials

$$a^*(L) = 1 + a_1^* L + a_2^* L^2 + \dots ,$$

$$b^*(L) = b_0^* L^0 + b_1^* L + b_2^* L^2 + \dots .$$

The return r_t also can be written as

$$r_t = m_t - m_{t-1} = (1-L)m_t = (1-L)e_t + (1-L)s_t = \omega_t + (1-L)s_t. \tag{4.13}$$

Now, the result (4.10) implies

$$g_r(z) = a^*(z)a^*(z^{-1})\sigma_1^2 + b^*(z)b^*(z^{-1})\Lambda. \tag{4.14}$$

From equation (4.13) the covariance of r_t will be

$$\begin{aligned} C_r(h) &= E(r_t r_{t-h}) \\ &= E[(\omega_t + (1-L)s_t)(\omega_{t-h} + (1-L)s_{t-h})] \\ &= E(\omega_t \omega_{t-h}) + E(\omega_t(1-L)s_{t-h}) + \\ &\quad + E((1-L)s_t\omega_{t-h}) + E((1-L)s_t(1-L)s_{t-h}). \end{aligned}$$

Then, the autocovariance generating function of r_t is

$$\begin{aligned} g_r(z) = &\sum_{h=-\infty}^{\infty} [E(\omega_t\omega_{t-h})]z^h + \sum_{h=-\infty}^{\infty} [E(\omega_t(1-L)s_{t-h})]z^h + \\ &+ \sum_{h=-\infty}^{\infty} [E((1-L)s_t\omega_{t-h})]z^h + \sum_{h=-\infty}^{\infty} [E((1-L)s_t(1-L)s_{t-h})]z^h. \end{aligned}$$

We will calculate every term on the right-hand side of the last equation.

(1) From $E(\omega_t\omega_{t-h}) = C_\omega(h)$ there follows

$$\sum_{h=-\infty}^{\infty} E(\omega_t\omega_{t-h})z^h = g_\omega(z). \tag{4.15}$$

(2) From

$$E(\omega_t(1-L)s_{t-h}) = E(\omega_t s_{t-h}) - E(\omega_t s_{t-h-1}) = C_{\omega s}(h) - C_{\omega s}(h+1),$$

implies

$$\begin{aligned}
&\sum_{h=-\infty}^{\infty} C_{\omega s}(h)z^h - \sum_{h=-\infty}^{\infty} C_{\omega s}(h+1)z^h = \\
&\qquad = \sum_{h=-\infty}^{\infty} C_{\omega s}(h)z^h - \sum_{h=-\infty}^{\infty} C_{\omega s}(h)z^{h-1} \\
&\qquad = (1-z^{-1})g_{\omega s}(h).
\end{aligned}$$

(3) Using

$$\begin{aligned}
E((1-L)s_t\omega_{t-h}) &= E(s_t\omega_{t-h}) - E(s_{t-1}\omega_{t-h}) \\
&= C_{s\omega}(h) - C_{s\omega}(h-1),
\end{aligned}$$

we get

$$\begin{aligned}
&\sum_{h=-\infty}^{\infty} C_{s\omega}(h)z^h - \sum_{h=-\infty}^{\infty} C_{s\omega}(h-1)z^h \\
&\qquad = \sum_{h=-\infty}^{\infty} C_{s\omega}(h)z^h - \sum_{h=-\infty}^{\infty} C_{s\omega}(h)z^{h+1} = (1-z)g_{s\omega}(h).
\end{aligned}$$

(4) Taking into account

$$\begin{aligned}
&E((1-L)s_t(1-L)s_{t-h}) = \\
&\qquad = E(s_t s_{t-h}) - E(s_t s_{t-h-1}) - E(s_{t-1}s_{t-h}) + E(s_{t-1}s_{t-h-1}) \\
&\qquad = C_s(h) - C_s(h+1) - C_s(h-1) + C_s(h),
\end{aligned}$$

one can see that

$$\begin{aligned}
&\sum_{h=-\infty}^{\infty} C_s(h)z^h - \sum_{h=-\infty}^{\infty} C_s(h+1)z^h - \sum_{h=-\infty}^{\infty} C_s(h-1)z^h + \sum_{h=-\infty}^{\infty} C_s(h)z^h \\
&= \sum_{h=-\infty}^{\infty} C_s(h)z^h - \sum_{h=-\infty}^{\infty} C_s(h)z^{h-1} - \sum_{h=-\infty}^{\infty} C_s(h)z^{h+1} + \sum_{h=-\infty}^{\infty} C_s(h)z^h \\
&= (1-z)(1-z^{-1})g_s(\omega).
\end{aligned}$$

Finally, we get

$$g_r(z) = g_w(z) + (1-z^{-1})g_{ws}(z)(1-z)g_{sw} + (1-z)(1-z^{-1})g_s(z).$$

Putting $z = 1$ into the last equation we obtain

$$g_r(1) = g_w(1). \tag{4.16}$$

Equality (4.15) implies

$$g_w(1) = \sigma_w^2. \tag{4.17}$$

Finally, from (4.14), (4.16), (4.17) there follows

$$g_r(1) = \sigma_w^2 = [a^*(1)]^2\sigma_1^2 + [b^*(1)]^2\Lambda = (1 + \sum_{i=1}^{\infty} a_i^*)^2\sigma_1^2 + (\sum_{i=0}^{\infty} b_i^*)^2\Lambda.$$

The final result has to be understood as follows. Public information events are incorporated into quote revision via the innovation $\nu_{1,t}$. The permanent effect on midquotes of a unit quote revision innovation is given as the sum of one (the contemporaneous impact) and $\sum\limits_{i=1}^{\infty} a_i^*$. Hence, the variation in efficient price implied by public information is given by the first term of the equation (4.11). A permanent effect on midquotes of the unexpected unite trade is $\sum\limits_{i=0}^{\infty} b_i^*$. Therefore, the variation in efficient price implied by private information is the second term of the equation (4.11). The variation in efficient price caused by both public and private information is then a sum of variations in the efficient price caused by public and private information separately.

Remark 4.1 *The described VAR system may be generalized taking a vector of trade attributes instead of the signed trade variable x_t. In that case, b_i and c_i are coefficient matrices, and Λ is the variance-covariance matrix of the vector $\nu_{2,t}$, and the equations from Theorem 4.1 will be*

$$\sigma_\omega^2 = (1 + \sum_{i=1}^{\infty} a_i^*)^2\sigma_1^2 + (\sum_{i=0}^{\infty} b_i^*)\Lambda(\sum_{i=0}^{\infty} b_{i^*}^T)$$

$$\sigma_{\omega,x}^2 = (\sum_{i=0}^{\infty} b_i^*)\Lambda(\sum_{i=0}^{\infty} b_{i^*}^T).$$

Chapter 5

Empirical results

5.1 The properties of London Stock Exchange

The two types of markets can be distinguished. A *quote-driven market* and an *order-driven market*. The first one is traditional market, where market makers have an obligation to quote continuously two-way prices at which they are prepared to buy and sell a security. In that way they fill gaps arising from imperfect synchronization between the arrivals of buyers and sellers. They are the counterpart in all transactions at the quoted prices: the bid price, at which they are willing to buy securities and the ask price, at which they willing to sell. They are the only providers of liquidity in the quote-driven market.

The development of electronic trading technology in the recent years has led to a rapid spread of so-called order-driven trading. In an order-driven market there are no designated market makers. Any trader can choose to execute trade via a limit or a market order. They input buy and sell orders for a security into a central computer system where they are automatically executed whenever they can be matched in terms of price and amount. The London Stock Exchange is predominantly an order-driven market.

In the present chapter we have used high frequency trading data of eighteen stocks listed on the London Stock Exchange from the FTSE 100 index. The FTSE 100 is the share index of the 100 largest publicly quoted UK companies. Taken together, these shares are worth around 80% of the UK stock market. The trading process of FTSE stocks is provided by the stock electronic order driven system, called the *Stock Exchange Trading System* or shortly SETS. The SETS is an order matching system based on the concept of priority trading, where orders are ranked in priority of price, then in time within the price. The order book is conveyed publicly in real time. As a result, market benefits from pre-trade transparency, which means that participants have an access to the whole order book and the post-trade transparency, which means that participants can immediately observe the last trades recorded by the system. On the other hand, order and trades are mostly anonymous. Regarding liquidity problem and process of price setting, the order-driven market is significantly different from the quote-driven market. Since limit orders allow a trader to set a limit price at which the order might be filled, but there is a risk the order will not be executed, the liquidity and establishing the bid-ask spreads in an order-driven market rely only on limit orders.

5.2 Data and cleaning

Over a period of 62 days, from March 1, 2006 to May 31, 2006 the trading attributes of interest for the liquidity analysis are viewed trade-by-trade for a group of eighteen stocks listed on FTSE 100 index. The 62 day trading sample is long enough to allow reasonably precise estimations [19], [20]. We used data organized in the order book containing following columns.

1. date
2. symbol
3. local time
4. five levels of bid prices
5. five levels of ask prices
6. bid volumes related to the given five levels of bid prices
7. ask volumes related to the five levels of ask prices
8. the number of bid orders related to the five levels of bid prices
9. the number of ask orders related to the five levels of ask prices
10. is a trade
11. transaction price

12. transaction size

13. trade time

The trading attributes of interest for our analysis are time, price and volume of executed trade, pre-trade bid and ask prices with their related volumes. In Table 1 is given a part of the order book which includes only eight columns of interest.

bid1	ask1	bsize1	asize1	is a trade	price	size	trade time
129.5	129.75	27211	836899	1	129.5	122789	8:02:19
129.5	129.75	0	836899	1	129.5	27211	8:02:19
129.25	129.75	750000	836899	0	129.5	27211	8:02:19
129.25	129.75	950000	836899	0	129.5	27211	8:02:23
129.25	129.75	950000	836899	0	129.5	27211	8:02:23
129.25	129.75	950000	836899	0	129.5	27211	8:02:23
129.25	129.75	950000	686899	0	129.5	27211	8:02:23
129.25	129.75	850000	686899	1	129.25	100000	8:02:23
129.25	129.75	650000	686899	0	129.25	100000	8:03:06
129.5	129.75	200000	686899	0	129.25	100000	8:03:06
129.5	129.75	100000	686899	1	129.5	100000	8:03:06
129.5	129.75	100000	686899	0	129.5	100000	8:03:07
129.5	129.75	86410	686899	1	129.5	13590	8:03:07
129.5	129.75	86410	686899	0	129.5	13590	8:03:15
129.5	129.75	186410	686899	0	129.5	13590	8:03:15

Table 1: The part of the order book for Vodafone at day March 14, 2006.

The column *is a trade* takes value 1, if trade was executed, and 0 if it was not. For example, the trade which was executed at time 8:02:23, was executed at pre-trade bid price 129.25 and therefore it was seller initiated. The size of executed trade is 100000, which decreased pre-trade bid size from 950000 to 850000. Using Matlab software we proceeded the following cleaning.

All data that occur outside the normal trading hours, i.e. before 8:00 a.m. and after 4:30 p.m. were deleted from the sample. Then we matched the executed trades with their related pre-trade bid and ask prices and bid and ask sizes. After that, we excluded all rows in the order book for which *is a trade* was zero. We eliminated all anomalous data caused by human and system errors, such as negative spreads, zero bid prices and spreads that are larger than 10% of actual stock prices. To eliminate any observation which does not reflect the market activity we applied the cleaning procedure for eliminating outliers taken from [11]. Let $\{p_i\}^N$ be an ordered tick-by-tick series. The procedure for removing outliers is as follows. If

$$|p_i - \overline{p}_i(k)| < 3s_i(k) + \gamma, \tag{5.1}$$

p_i is kept, otherwise p_i is removed. Here $\overline{p}_i(k)$ and $s_i(k)$ denote respective α-trimmed sample mean and sample standard deviation of a neighborhood of k observations around i and γ is a granularity parameter. The neighborhood of observations is always chosen so that an observation considered is compared with observations belonging to the same trading day. If the observation is at the very beginning or end of a trading day, then first (and respectively last) k ticks are used, whereas for an observation in the middle of a day approximately $k/2$ ticks before and $k/2$ ticks after, are considered. The percentage of trimming α should be based on the frequency of outliers - the higher the frequency, the higher is α to be chosen. The choice of k depends on frequency of trading, and it should be chosen in such a way that it does not include "too distant" prices. Therefore, for very frequent stocks the value for k should be significantly bigger than the value for some other less frequently traded stock.

The procedure is heuristic and it depends heavily on the proper choice of α, γ and k. In order to decrease the level of heuristic dependence we apply the same procedure iteratively, applying described algorithm iteratively - two or three times. The data cleaning procedure described above is applied for all eighteen shares considered. Since our price data are discrete with minimal allowed change or *tick*, the granulation parameter for each stock equal to its tick size is the only reasonable choice by our empirical experience. On the other hand, filtering procedure is sensitive to k. We tried $k = 20, 40, 60$; the best results are almost always obtained with $k = 40$. For Vodafone, the number of outliers is always less than 1% with γ equal to half of tick size and $k = 40$, which is consistent with the results in [11]. Also, we concluded that two filter iterations were sufficient in the vast majority of cases.

Since there might be several transactions reported at the same time that were executed at different price levels, we applied some form of aggregation which is consistent with liquidity analysis and Hasbrouck's *VAR* model. The trades that occur at the same time, with same price and in the same direction are treated as one trade. The volume of such trade is then simply a sum of the volumes corresponding to individual trades. After such aggregations procedure the number of observation for each stock decreased in average more than two times. The information about the number of transactions after the aggregation procedure and average midquote for each stock are given in Table 2. Midquote is a much better price variable than the actual transaction price because it limits the bid-ask bounce problem. The bid ask bounce occurs when there is no news in some observable period, but the buy and sell orders come successively, which makes an impression that the prices changed more than they actually did.

symbol	company name	number of transactions	average midquote
ABF	Associated British Foods	27384	800.68
AZN	Astra Zenaca	88232	2878.9
BARC	Barclays	75709	655.28
CPI	Capita Group	33198	462.61
GSK	Glaxosmithkline	86972	1515.4
HBOS	Hbos	71450	961.46
HSBA	Hsbc Hldgs-Uk	87085	963.23
IAP	Icap	22854	493.3
KAZ	Kazakhmys	29439	1107.6
LLOY	Lloyds Tsb	66833	533.46
PRU	Prudential	63535	642.6
RB	Reckit Bencksr	54738	2008
RIO	Rio Tinto	125383	2943.3
SHP	Shile	41806	862.57
SLOU	Slough Estates	23189	625.65
VOD	Vodafone	85138	124.63
WPP	Wpp Group	40899	677.89
XTA	Xstata	85921	1994.2

Table 2. The basic information of stocks after the aggregations procedure.

5.3 Variables and estimation

Although a large number of liquidity measures is presented in Chapter 2, there is no need to consider all of them. Many of the presented measures are just modifications of some other measures and they carry the same information. We chose the following ten measures which can be calculated trade-by-trade to examine different liquidity aspects of eighteen chosen stocks.

1. Volume per trade
2. Turnover per trade
3. Volume depth available immediately before trade
4. Duration between two successive trades
5. Flow ratio between two successive trades
6. Absolute spread
7. Proportional spread
8. Quote Slope
9. Price impact for the ask side per trade
10. Price impact for the bid side per trade

symbol	tick size	tick size at the base points
ABF	0.5	6.24
AZN	1	3.47
BARC	0.5	7.63
CPI	0.25	5.40
GSK	1	6.60
HBOS	0.5	5.20
HSBA	0.5	5.19
IAP	0.25	5.07
KAZ	0.5	4.51
LLOY	0.25	4.69
PRU	0.5	7.78
RB	1	4.98
RIO	1	3.40
SHP	0.5	5.80
SLOU	0.5	7.99
VOD	0.25	20.06
WPP	0.5	7.38
XTA	1	5.01

Table 3. The tick size for each stock in absolute values and in the base points values.

To compare proportional spreads of different stocks we modify the formula (2.1) by

$$S_t^{prop_{tick}} = \frac{S_t^{prop}}{tick}$$

where *tick* is a tick size for the corresponding stock with respect to average midquote. In such way the proportional spread of eighteen analyzed stocks is expressed in the tick size. The tick size for each stock calculated for the observation period of 62 days in absolute values and relatively to average midquote multiplied by 10^4, i.e. in the base points values are given in Table 3.

Price impact for the ask and the bid side per trade is given by

$$PI_t^a = ln(\frac{p_t}{m_{t-1}}) \tag{5.2}$$

$$PI_t^b = -ln(\frac{p_t}{m_{t-1}}). \tag{5.3}$$

Since the formulas (5.2) and (5.3) are modifications of the formulas (2.3) and (2.4) for the case when the volume is excluded, they represent the immediate impact of the last transaction, where m_{t-1} is the pre-trade midquote. To see the real effect of price impact for the ask and the bid side, and to make them comparable for different stocks, we scaled them by average proportional spread, that is we modified the formulas (5.2) and (5.3) into

$$PI_t^a = \frac{1}{\bar{S}^{prop}} \cdot ln(\frac{p_t}{m_{t-1}}),$$

$$PI_t^b = -\frac{1}{\bar{S}^{prop}} \cdot ln(\frac{p_t}{m_{t-1}}).$$

The average proportional spread $\bar{S}^{prop}$ is calculated over the period of 62 observation days.

The flow ratio between two successive trades is given by

$$FR_t = \frac{TO_{t-1}}{Dur_t}$$

where Dur_t is time between a transaction at the time point t and a transaction before that, and TO_{t-1} is turnover of the first transaction realized before the time point t. It measures the ability of stock to absorb large volumes between two successive trades. If duration between two successive trades is zero, the flow ratio is set to be zero.

The variables of interest for Hasbrouck's VAR model are the sign trade x_t and quote revision r_t, as it described in Chapter 4. Under assumptions of Hasbrouck's VAR model given in Chapter 4, it can be estimated consistently and efficiently by the least squares method. Hasbrouck's empirical findings show that due to the least squares estimation coefficients of sign trade variables are positive, but of highly variable magnitude in the quote revision equation. That is the reason why Hasbrouck suggested a better behaved model resulting from replacing the sign trade variable x_t with the trade indicator variable x_t^0, defined by (2.2). In the text to follow the "trade variable" will denote the trade indicator variable x_t^0. As a quote revision variable he used the change in the natural logarithm of the midquote that follows the current trade at time t,

$$r_t = ln(m_{t+1}) - ln(m_t)$$

or midquote return of current trade at time t. For simplicity sake we say "return" instead of "quote return" and we will call the equation (4.2) the return equation. For the same reasons as for price impact for the ask and the bid side, we expressed the return variable as a part of average proportional spread by

$$r_t = \frac{ln(m_{t+1}) - ln(m_t)}{\bar{S}^{prop}}. \tag{5.4}$$

All described variables was calculated and analyzed using the Matlab software. In computing autoregressions the sign trade and quote revision prior to the first observation of the day are assumed to be zero. The proportion in percents of buyer initiated, seller initiated and undeterminate trades for each of eighteen analyzed stocks are given in Table 4.

Following Hasbrouck [39], [40], and Engle et al. [21], we assumed that the model given by equations (4.2) and (4.3) can be truncated at five lags i.e.

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}, \tag{5.5}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t}. \tag{5.6}$$

We will make some remarks about the sign trade variable. It is a dummy variable and it is quite unusual to have such variable in vector autoregression. Dummy variable presents no econometric difficulties when it is an explanatory variable which is case for the return equation, but in the case of the trade equation the linear specification is potentially inappropriate. The least squares estimation yields to an inefficient estimation of the trade coefficients and standard errors are biased. To avoid this problem, we will correct the standard errors by using the White's heteroskedasticity consistent covariance estimator (1.8) to correct the Wald and t-statistics.

symbol	buyer initiated	seller initiated	undeterminate
ABF	45.52	47.90	6.58
AZN	45.09	48.75	6.17
BARC	44.81	46.20	8.99
CPI	45.92	47.50	6.58
GSK	44.40	48.98	6.62
HBOS	46.40	50.00	3.60
HSBA	45.88	48.61	5.51
IAP	44.81	51.37	3.82
KAZ	48.15	49.61	2.25
LLOY	46.19	45.59	8.22
PRU	44.86	47.10	8.03
RB	45.21	47.76	7.04
RIO	47.00	47.26	5.73
SHP	46.03	46.97	6.99
SLOU	48.01	46.86	5.13
VOD	42.69	54.89	2.42
WPP	45.45	45.62	8.93
XTA	47.17	48.43	4.39

Table 4. The proportion of the buyer initiated, seller initiated and undeterminate trades for each of eighteen analyzed stocks.

5.4 Results

5.4.1 Liquidity measures

Summary statistics for ten liquidity measures are given in Appendix 1. The proportional spread is given in tick size. The price impact for the ask and the bid side are scaled by average proportional spread. Since for each stock the liquidity measures have very high standard deviations, liquidity measures for all stocks show high variability. All measures for all eighteen stocks are positively skew. All measures for all eighteen stocks have kurtosis larger than three, hence the distribution of calculated measures are elongated. No one of the measures is normally distributed. For each liquidity measure we ranked stocks from the most liquid to the most unliquid in Table 5. The last column in Table 5 represents the average ranks for each stock. Therefore, the stock with the smallest average rank is the most liquid stock. The list of stocks ranking from the most liquid to the most unliquid according to ten calculated measures is: VOD, HSBA, BARC, LLOY, GSK, HBOS, PRU, WPP, AZN, CPI, RIO, SHP, RB, ABF, XTA, SLOU, IAP and KAZ. Vodafone is the most liquid stock according to almost all of ten liquidity measures. An exception is the duration which indicates trading intensity.

symbol	V	TO	D	$D\$$	Dur	FR	S	QS	S^{prop}	PI^a	PI^b	average
ABF	14	16	12	14	16	16	10	10	8	9	13	12.5
AZN	12	3	14	7	2	2	14	14	10	11	12	9.2
BARC	4	6	4	5	7	6	5	5	4	3	2	4.6
CPI	8	15	9	15	14	14	2	2	13	12	11	10.5
GSK	9	4	6	3	4	5	13	13	3	4	4	6.2
HBOS	7	8	8	8	8	8	7	7	6	7	8	7.5
HSBA	3	2	2	2	3	3	4	4	2	2	5	2.9
IAP	10	17	11	18	17	17	9	9	17	17	17	14.5
KAZ	15	13	18	17	15	15	17	17	18	18	18	16.5
LLOY	2	7	3	4	9	9	3	3	15	5	3	5.7
PRU	5	10	5	6	10	10	8	8	7	8	7	7.6
RB	18	12	15	11	11	12	15	15	9	10	9	12.5
RIO	17	5	17	10	1	4	16	16	14	15	15	11.8
SHP	11	14	13	13	12	13	12	12	12	13	10	12.3
SLOU	13	18	10	16	18	18	11	11	11	14	14	14.0
VOD	1	1	1	1	6	1	1	1	1	1	1	1.5
WPP	6	11	7	9	13	11	6	6	5	6	6	7.8
XTA	16	9	16	12	5	7	18	18	16	16	16	13.5

Table 5. The stocks ranked from the most liquid to the most unliquid for each liquidity measure with average rank.

The stocks which are traded more frequently than Vodafone are RIO, AZN, HSBA, GSK and XTA by order. On the other hand the average trade volume and average turnover for Vodafone is several times bigger than the average trade volume and average turnover for these five stocks. This ability of Vodafone to absorb large trades in a short time interval can be seen by the flow ratio.

5.4.2 Hasbrouck's *VAR*

For each day from the sample of 62 observation days, we calculated the return vector given by (5.4) and the trade vector given by Lee and Ready rule (2.2). The return and trade prior to the first observation of each day is set to be zero. Adding the return vector of day $i+1$ at the and of the return vector of day i, $i = 1, ..., 61$, we obtained the $n \times 1$ return vector r, where n is the sample size. In the same way, the $n \times 1$ trade vector x^0 is obtained from the trade vectors of each day. For estimating Hasbrouck's $VAR(5)$ given by (5.5) and (5.6) we made vectors r_t and x_t^0 by cutting the first five entries of return vector r and trade vector x^0. The lagged vectors r_{t-k} and x_{t-k}^0, $k = 1, 2, ..., 5$ are obtained by cutting the first $5-k$, and the last k entries of return and trade vector by order. Therefore, to obtain the least squares coefficients of the return equation we regressed vector r_t on the matrix

$$[r_{t-1}, r_{t-2}, r_{t-3}, r_{t-4}, r_{t-5}, x_t^0, x_{t-1}^0, x_{t-2}^0, x_{t-3}^0, x_{t-4}^0, x_{t-5}^0]. \qquad (5.7)$$

To obtain the least squares coefficients of the trade equation, we regressed vector x_t on the matrix

$$[r_{t-1}, r_{t-2}, r_{t-3}, r_{t-4}, r_{t-5}, x_{t-1}^0, x_{t-2}^0, x_{t-3}^0, x_{t-4}^0, x_{t-5}^0]. \qquad (5.8)$$

Matlab function

$$[b, bint, r, rint, stats] = regress(y, X),$$

gives the least squares coefficients of the equation (1.4). This function returns b, the least squares estimation of β with its 95% confidence interval *bint*; the vector of residuals r with its 95% confidence interval *rint*; a 1×3 vector *stats* which contains the coefficient of multiple determination R^2 along with F-statistics and P-value for the regression. The estimated least squares coefficients for the return and trade equation together with their corresponding t-statistics for all eighteen stocks calculated by Matlab function *regress* are given in Appendix 2. The t-statistics given by (1.5) in the trade equation are corrected by using the White's heteroskedastcity consistent covariance estimator given by (1.7). The stars above t-statistics denote significance at the 5% level. For each equation, the coefficient of multiple determination R^2 is given. For the return equation, the variance of innovation term σ_1^2 is calculated as well as the variance of innovation term of trade equation Λ. We proceed with the Wald test (1.6) for the following hypothesis.

1. Coefficients of trade variables in the return equation are jointly equal to zero i.e.
$$H_0: \; b_i = 0, \;\; i = 0, 1, ..., 5.$$
To calculate the Wald statistics (1.6) for this test, a matrix B has to consist of two blocks
$$B = [Z \; I]$$
where Z is the 6×5 zero matrix, and I is the the 6×6 identity matrix. Vector b has to be 6×1 zero vector, and matrix X is matrix (5.7).

2. Sum of the coefficients of the trade variables in the return equation is equal to zero, i.e.
$$H_0: \; \sum_{i=0}^{5} b_i = 0.$$
For this test the return equation (5.5) is transformed into
$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + \sum_{i=1}^{5} b_i x_t^0 + \\ + b_1(x_{t-1}^0 - x_t^0) + ... + b_5(x_{t-5}^0 - x_t^0) + \nu_{1,t}.$$
Then, after calculating new variables $x_{t-k}^0 - x_t^0, \;\; k = 1, 2, ..., 5$ we regressed vector r_t on the matrix
$$[r_{t-1}, \; ... \;, r_{t-5}, (x_{t-1}^0 - x_t^0), \; ... \;, (x_{t-5}^0 - x_t^0)], \tag{5.9}$$
and tested hypothesis
$$H_0: \; \xi = 0, \;\; \xi = \sum_{i=1}^{5} b_i.$$
For such test, B is the 1×11 vector with all entries zero, except the sixth which is one. The matrix X is matrix (5.9), and $b = 0$.

3. Coefficients of return variables in the trade equation are jointly equal to zero i.e.
$$H_0: \; c_i = 0, \;\; i = 1, ..., 5.$$
Because of heteroskedasticity problem caused by the presence of dummy variable in trade equation, the Wald statistics (1.6) has to be corrected

by the White's consistent covariance estimator (1.8). Writing the Wald satistics (1.6) as

$$(B\beta - b)^T[B^T s^2(X^T X)^{-1}B]^{-1}(B\beta - b) \sim \chi^2_n,$$

it can be seen from the subsection (1.4.2) that

$$s^2(X^T X)^{-1} = Var(\hat{\beta})$$

where s is the sample error variance estimator given by

$$s^2 = \frac{\hat{\varepsilon}^T\hat{\varepsilon}}{n},$$

and $\hat{\varepsilon}$ is the estimated vector of residuals. It is clear now, that $Var(\hat{\beta}) = s^2(X^T X)^{-1}$ in the Wald test has to be replaced by the White's estimator of covariance matrix $V(\hat{\beta})$ given by (1.9). After such correction the Wald test becomes

$$\frac{(B\beta - b)^T[B^T(X^T X)^{-1}\hat{S}(X^T X)^{-1}B]^{-1}(B\beta - b)}{n},$$

where $\hat{S}$ is the White heteroskedasticity consistent covariance estimator given by

$$\hat{S} = \frac{1}{n}\sum_i \varepsilon_i^2 x_i x_i^T.$$

To calculate the corrected Wald statistics for this test matrix B has to consist of two blocks

$$B = [I\ Z]$$

where I is the 5×5 identity matrix and Z is the 5×5 zero matrix. Vector b has to be the 5×1 zero vector. Matrix X is given by (5.8).

The algorithm for calculating the $VMA(q)$ of Hasbrouck's $VAR(p)$ described in Chapter 4 is as follows.

Step 1. Using obtained the least squares coefficients of equations (5.5) and (5.4) forme matrices

$$A_i = \begin{bmatrix} a_i + b_0c_i & b_i + b_0d_i \\ c_i & d_i \end{bmatrix},\ i = 1, 2, ..., p$$

Step 2. Construct the $2q \times 2$ matrix given by

$$\varphi = [A_1\ A_2\ A_3\ A_4\ A_5\ Z]^T$$

where Z is the $2(q-p) \times 2$ zero matrix.
Step 3. Construct the $2 \times 2q$ matrix

$$\Psi = [I\ Z]$$

where I is the 2×2 identity matrix, and Z is the $2(q-1) \times 2$ zero matrix.
Step 3. For $i = 1, ..., q$

1. Calculate
$$\Psi' = \Psi_i \cdot \varphi$$

2. Construct a matrix Ψ'' by putting matrix Ψ' at begin of the matrix Ψ and by cutting the 2×2 zero matrix at the end of Ψ.

3. $\Psi' = \Psi''$

After the described algorithm the resulting $2 \times 2q$ matrix is

$$\begin{bmatrix} a'_q & b'_q & ... & a'_2 & b'_2 & a'_1 & b'_1 \\ c'_q & d'_q & ... & c'_2 & d'_2 & c'_1 & d'_1 \end{bmatrix},$$

where coefficients $a_i, b_i, c_i, d_i,\ i = 1, 2, ..., q$ are impulse responses given in (4.6). The effect of a unit trade innovation on the return at a i period horizon is calculated by

$$b_i^* = a'_i b_0 + b'_i,\ i = 1, 2, ..., q$$

as it is described in Section 4.

Following the described algorithm we calculated the the moving-average representation of *VAR* (4.6), (4.7) truncated at 40 lags. The price impact function is calculated as a cumulative impulse b_i^* responses of return to the innovation in the trade equation. The graphs of price impact functions are given for each stock. The total price impact is calculated as well as the number of transaction needed for its realization. At the end, the variance decomposition coefficient R^2_ω given by (4.9) is calculated.

The most important coefficients are coefficients of trade variables in the return equation and in the trade equation. The coefficient b_0 in the return equation for each of eighteen stocks measures an average rise of return relative to the proportional spread immediately after the buy order. The coefficients $b_i, \quad i = 1, 2, ..., 5$ in the return equation for all eighteen stocks tend to be positive, meaning that the buys tend to increase and sells tend to decrease the return. According to the Wald test the hypothesis that the coefficients $b_i, i = 1, 2, ..., 5$ are jointly equal to zero is rejected. The sum of them is positive, and according to the Wald test, significantly different from zero at the 1% level, indicating that the order flow has a positive influence on the return.

The coefficients of lagged trade variables in the return equation for each of eighteen stocks are positive, reflecting a positive autocorrelation in trades, which indicates that a purchase tends to follow a purchase, and a sell tends to follow a sell. They are also significantly different from zero, even at the 1% level.

symbol	liquidity	permanent price impact	R^2_ω
ABF	12	8	12
AZN	9	9	5
BARC	3	15	3
CPI	10	5	8
GSK	4	14	7
HBOS	5	7	2
HSBA	2	17	10
IAP	17	1	17
KAZ	18	2	16
LLOY	7	16	9
PRU	6	11	14
RB	14	6	4
RIO	11	4	6
SHP	13	10	13
SLOU	16	13	18
VOD	1	18	15
WPP	8	12	11
XTA	15	3	1

Table 6. The stocks ranked from the highest to the lowest liquidity, total permanent price impact calculated by Hasbrouck's *VAR*, and variance decomposition coefficient R^2_ω.

The negative coefficients of lagged return variables in the trade equation indicate negative autocorrelation in the returns. This negative autocorrelation is predominant for stocks with the symbols ABF, CPI, KAZ, SHP, SLOU, and VOD. For other stocks, this behavior is weaker.

The Wald test of the hypothesis that the coefficients of return variables in the trade equation are jointly zero is rejected at the 1%, level indicating Granger's casuality running from returns to trades.

In Table 6, eighteen observed stocks are ranked from the highest to the lowest liquidity, total permanent price impact calculated by Hasbrouck's VAR, and variance decomposition coefficient R^2_ω.

To test the null hypothesis

H_0: There is no relationship between the different liquidity measures, permanent price impact PI^{has} and variance decomposition coefficient R^2_ω,

against the alternative hypothesis

H_1: There is a relationship between the different liquidity measures, permanent price impact PI^{has} and variance decomposition coefficient R^2_ω,

we use the Sperman rank correlation test, whose results with related P-values in the brackets are given in Table 7. To have correlation between sets of data, the null hypothesis has to be rejected. For this test all measures were ranked as in Table 5, from the highest to the lowest liquidity. That means that for example, stocks are ranked from the lowest to the highest duration, and from the highest to the lowest trade volume. Therefore, stocks are ranked from the lowest to the highest Hasbrouck's price impact. We ranked stocks from the highest to the lowest variance decomposition coefficient. Significant correlations (at the 1% or 5% level) between different liquidity measures are mostly under 50%. Correlations significant at the 10% level are between 40% and 50%. As it was expected, those correlations positive, implying that more liquid stocks according to one liquidity measure, are also more liquid according to some other liquidity measure. From all calculated correlations we were the most interested in correlations between the Hasbruck's price impact PI^{has} and all liquidity measures, and between the variance decomposition coefficient R^2_ω and all liquidity measures. Correlations between the Hasbrouck's price impact and almost all considered liquidity measures are positive and significant. Exceptions are the duration and the flow ratio which are positively, but not significantly correlated with the Hasbrouck's price impact. Such results imply that more liquid stocks according to some liquidity measure have lower Hasbrouck's price impact, as it was expected.

	V	TO	D	$D\$$	Dur	FR
V		0.4262 (0.0789)	0.9443 (0.0001)	0.6677 (0.0059)	0.1496 (0.5372)	0.3560 (0.1421)
TO	0.4262 (0.0789)		0.4159 (0.0864)	0.8885 (0.0002)	0.9195 (0.0001)	0.9814 (0.0001)
D	0.9443 (0.0001)	0.4159 (0.0864)		0.7193 (0.0030)	0.1496 (0.5372)	0.3437 (0.1565)
$D\$$	0.6677 (0.0059)	0.8885 (0.0002)	0.7193 (0.0030)		0.7337 (0.0025)	0.8431 (0.0005)
Dur	0.1496 (0.5372)	0.9195 (0.0001)	0.1496 (0.5372)	0.7337 (0.0025)		0.9525 (0.0001)
FR	0.3560 (0.1421)	0.9814 (0.0001)	0.3437 (0.1565)	0.8431 (0.0005)	0.9525 (0.0001)	
S	0.8885 (0.0002)	0.1723 (0.4773)	0.8658 (0.0004)	0.4345 (0.0732)	-0.0857 (0.7240)	0.1249 (0.6067)
QS	0.8885 (0.0002)	0.1723 (0.4773)	0.8658 (0.0004)	0.4345 (0.0732)	-0.0857 (0.7240)	0.1249 (0.6067)
S^{prop}	0.5562 (0.0218)	0.5335 (0.0278)	0.6945 (0.0042)	0.7358 (0.0024)	0.3498 (0.1492)	0.5129 (0.0345)
PI^a	0.7750 (0.0014)	0.6285 (0.0096)	0.8679 (0.0003)	0.8782 (0.0003)	0.3911 (0.1068)	0.5645 (0.0199)
PI^b	0.7750 (0.0014)	0.6285 (0.0096)	0.8679 (0.0003)	0.8782 (0.0003)	0.3911 0.1068)	0.5645 (0.0199)
PI^{has}	0.7399 (0.0023)	0.5170 (0.0330)	0.8369 (0.0006)	0.7833 (0.0012)	0.2817 (0.2454)	0.4489 (0.0642)
R^2_w	-0.1393 (0.5657)	0.4407 (0.0692)	-0.1187 (0.6246)	0.3209 (0.1857)	0.6244 (0.0100)	0.5088 (0.0359)

	S	QS	S^{prop}	PI^a	PI^b	PI^{has}	R^2_w
V	0.8885 (0.0002)	0.8885 (0.0002)	0.5562 (0.0218)	0.7750 (0.0014)	0.7750 (0.0014)	0.7399 (0.0023)	-0.1393 (0.5657)
TO	0.1723 (0.4773)	0.1723 (0.4773)	0.5335 (0.0278)	0.6285 (0.0096)	0.6285 (0.0096)	0.5170 (0.0330)	0.4407 (0.0692)
D	0.8658 (0.0004)	0.8658 (0.0004)	0.6945 (0.0042)	0.8679 (0.0003)	0.8679 (0.0003)	0.8369 (0.0006)	-0.1187 (0.6246)
$D\$$	0.4345 (0.0732)	0.4345 (0.0732)	0.7358 (0.0024)	0.8782 (0.0003)	0.8782 (0.0003)	0.7833 (0.0012)	0.3209 (0.1857)
Dur	-0.0857 (0.7240)	-0.0857 (0.7240)	0.3498 (0.1492)	0.3911 (0.1068)	0.3911 (0.1068)	0.2817 (0.2454)	0.6244 (0.0100)
FR	0.1249 (0.6067)	0.1249 (0.6067)	0.5129 (0.0345)	0.5645 (0.0199)	0.5645 (0.0199)	0.4489 (0.0642)	0.5088 (0.0359)
S		1.0000 (0.0000)	0.4819 (0.0469)	0.6821 (0.0049)	0.6821 (0.0049)	0.6078 (0.0122)	-0.1765 (0.4669)
QS	1.0000 (0.0000)		0.4819 (0.0469)	0.6821 (0.0049)	0.6821 (0.0049)	0.6078 (0.0122)	-0.1765 (0.4669)
S^{prop}	0.4819 (0.0469)	0.4819 (0.0469)		0.8762 (0.0003)	0.8762 (0.0003)	0.7399 (0.0023)	0.1249 (0.6067)
PI^a	0.6821 (0.0049)	0.6821 (0.0049)	0.8762 (0.0003)		1.0000 (0.0000)	0.8638 (0.0004)	0.1806 (0.4565)
PI^b	0.6821 (0.0049)	0.6821 (0.0049)	0.8762 (0.0003)	1.0000 (0.0000)		0.8638 (0.0004	0.1806 (0.4565)
PI^{has}	0.6078 (0.0122)	0.6078 (0.0122)	0.7399 (0.0023)	0.8638 (0.0004)	0.8638 (0.0004)		-0.0980 (0.6860)
R^2_w	-0.1765 (0.4669)	-0.1765 (0.4669)	0.1249 (0.6067)	0.1806 (0.4565)	0.1806 (0.4565)	-0.0980 (0.6860)	

Table 7. The Sperman rank correlation test for different liquidity measures, the Hasbruck's price impact and variance decomposition coefficient R^2_ω.

On the other hand, correlations between the variance decomposition coefficient R^2_ω and different liquidity measures are mostly insignificant. It is interesting that there is no significant correlation between the Hasbruck's price impact and the variance decomposition coefficient. The duration and the flow ratio are only two measures that significantly correlated to R^2_ω. These correlations are positive, implying that intensively traded stocks show larger contribution of unexpected trade in variation of efficient price. Also, the turnover is positively correlated to the variance decomposition coefficient at the 10% level.

Hasbrouck formally suggested that the variance decomposition coefficient R^2_ω indicates the proportion of volatility in the efficient price caused by the presence of informed traders represented in the unexpected component of the trade. Following such suggestion we would expect that more liquid stocks have a lower variance decomposition coefficient R^2_ω. However, our results are not consistent with such suggestion. For example, the most unliquid stocks according to Table 5, IAP, KAZ, and SLOU have the lowest coefficient R^2_ω. They also take 17, 15 and 18 rank by order according to the duration and the flow ratio. Excluding these three stocks, Vodafone has the smallest R^2_ω. According to the Spearman rank correlation test it seems that for intensively traded stocks, the coefficient R^2_ω is highly overestimated.

Our findings indicate that the variance decomposition coefficient strongly depends on the predictability of the return equation and trade equation. According to the coefficients of multiple determination R^2_r and R^2_x, the return equation shows a higher predictability than the trade equation for each stock except for Vodafone. Such results are reasonable, since in the trade equation the trade variable x^0_t as a dummy variable takes only three values, $-1, 0, 1$. It can be seen from Appendix 2 that the stocks with higher predictability of the return equation compared to the predictability of the trade equation have a higher R^2_ω. The predictability of the return equation strongly depends on the volatility in return. The proportions of positive returns, zero returns, and negative returns for each of eighteen analyzed stocks are given in Table 8. Vodafone has the extremely large proportion of zero returns of all eighteen stocks. This can explain the low predictability of its return equation compared to the predictability of the trade equation. The Pearson correlation coefficient between R^2_r/R^2_x and proportion of zero returns is -0.7211 with a P-value of 0.0007, which indicates the significance at the 1% level.

symbol	positive return %	zero return %	negative return %
ABF	25.34	49.99	24.67
AZN	28.95	41.84	29.21
BARC	21.31	56.87	21.82
CPI	27.51	44.73	27.77
GSK	19.44	61.08	19.48
HBOS	25.16	49.49	25.35
HSBA	16.33	67.45	16.22
IAP	30.72	38.02	31.26
KAZ	33.39	32.73	33.87
LLOY	20.48	58.83	20.69
PRU	25.35	48.56	26.09
RB	27.77	44.14	28.09
RIO	32.75	34.43	32.81
SHP	27.70	43.58	28.72
SLOU	24.90	49.80	25.30
VOD	6.81	86.25	6.94
WPP	23.24	53.59	23.17
XTA	32.91	33.87	33.22

Table 8. The proportions of positive returns, zero returns, and negative returns for each stock.

All presented results show that the variance decomposition coefficient R^2_ω cannot be interpreted as a proportion of volatility in the efficient price caused by the presence of informed traders. Also, the predictability of the return equation for all eighteen stocks shows strong dependence of the contemporaneous trade x^0_t. The coefficients of multiple determination for the return equation R^2_r for all eighteen stocks when the contemporaneous trade x^0_t is included and when it is excluded are given in Table 9. Such results imply that the last trade carries the most important information in prediction of the future price.

The results obtained by applying Hasbrouck's [39], [40] model of total price impact and contribution of private information to the volatility in efficient price reveals some differences from Hasbrouck's results [39], [40]. The most important differences are those concerning the variance decomposition coefficient R^2_ω. It is expected that the stocks which have higher total permanent price impact have a lager the variance decomposition coefficient R^2_ω.

symbol	with %	without %
ABF	17.90	3.90
AZN	27.28	4.41
BARC	24.64	5.30
CPI	20.76	4.18
GSK	22.89	5.33
HBOS	26.50	4.51
HSBA	20.92	5.89
IAP	15.92	2.96
KAZ	17.16	9.39
LLOY	22.02	5.22
PRU	21.71	5.59
RB	24.72	3.73
RIO	26.48	4.51
SHP	18.56	4.09
SLOU	15.19	4.58
VOD	11.69	6.41
WPP	22.98	4.27
XTA	25.66	4.94

Table 9. The coefficients of multiple determination for the return equation R_r^2 each stock with and without contemporaneous trade x_t^0.

Figure 5.1 shows the variance decomposition coefficient versus total permanent price impact across 18 analyzed stocks. Some kind of stock clustering can be noticed. The most isolated stocks are the three the most unliquid stocks according to Table 5, IAP, KAZ and SLOU, and they have the lowest R_ω^2. The rest, the more liquid ones have higher R_ω^2. The Vodafone is quite separated from this group showing its own behavior. It seems that Hasbrouck's findings are more appropriate for less liquid than for more liquid stocks.

Some explanations can be found in the characteristics of an order driven market, i.e. electronic trading. Hasbrouck's model assumes quite simply trading mechanism between market participants. It assumes the presence of market makers who post bid and ask quotes after the realized transaction at time t, and according to the information contained in the recent order flow. With such mechanism, every trade that model cannot predict is taken as a unexpected trading activity caused by the presence of traders with private information. In an order driven market interaction between market participants is much more complicated. First, there are no classical market makers.

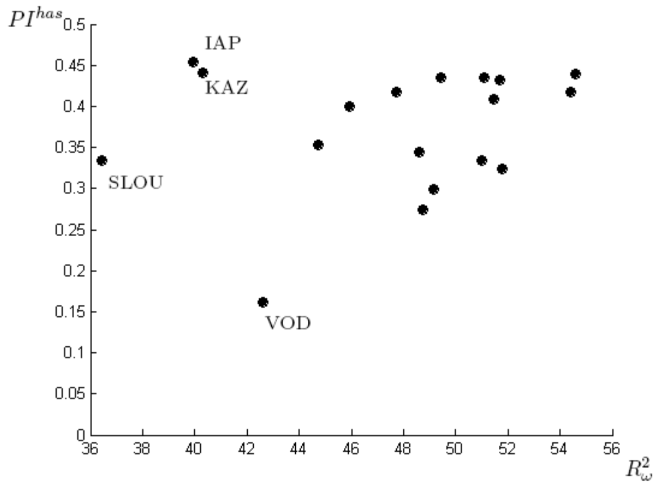

Figure 5.1: Variance decomposition $R^2_\omega(\%)$ coefficient vs. total permanent price impact with respect to the average proportional spread PI^{has}.

The transactions are realized by matching price and amount of different orders putted in the central computer system by different participants. The large number of transaction can be realized in a short time interval, or even more, in the same time. Also, electronic system enables recording almost every trading activity, which significantly enlarges the information set available to market participants. The possibility to record every price movement produces a higher predictability for the return equation. It is also expected that price of very liquid stocks cannot dramatically vary in a small time interval. That explains a higher predictability of the return equation for the more liquid stocks than for the less liquid. The exception is Vodafone, because of large proportion of zero returns. An electronic trading providing very frequent trading, requires fast reaction on any trading activity. It is possible that traders make fast decisions according to the recent information and for the short time horizon, which can explain the extreme decrease of the predictability of the return equation after excluding contemporaneous trade from the return equation.

In the last ten years trading becomes much more sophisticated because of development of so called algorithmic trading. Algorithmic trading usually refers to employment of computer algorithms for breaking up large order into sequence of smaller orders and engagement of automated trading strategies for their execution with respect to a numerous user-defined parameters such as time horizon, liquidity constraints, depth of market, volatility, etc. Moreover, the big trend in the recent years is creating unique trading algorithms. They have to be created, tested and applied in the market as quickly as possible, because market changes fast. In 2006 at the London Stock Exchange, over 40% of all orders were entered by algorithmic traders, with 60% predicted for 2007.

The larger variance decomposition coefficient of the intensive traded socks can be explained by the algorithmic trading. Such trading can be viewed as a superior trading which will behave as an unexpected trade, i.e. as a trade caused by the superior information in the Hasbrouck's VAR model. Clearly, the proportion of algorithmic trading has to be larger for intensively traded stocks.

The total price impact is defined as a response of the return to the trade innovation. Such innovation is viewed as a trade initiated by the private information. As it was expected, our results show that for more liquid stocks this impact is lower. However, our results show that the bivariate system (4.2) i (4.3) cannot isolate trades which are truly initiated by private information. Significant at the 5% level Spearman rank correlations of the variance decomposition coefficient R_ω^2 to the duration and the flow ratio require an extra research. The same is with the 10% level significant correlation between the variance decomposition coefficient R_ω^2 and the turnover.

References

[1] Admati, R., Pfeiderer, P., A theory of intraday patterns: Volume and price variability, Review of Financial Studies 1 (1988), pp. 3-40.

[2] Amihud, Y., Illiquidity and stock returns, Cross-section and time-series efects, Journal of Financial Markets 5 (2002), pp. 31-56.

[3] Amihud, Y., Mendelson H., Pedersen L. H., Liquidity and Asset Prices, Foundations and Trends in Finance, Vol. 1, No 4 (2005), pp. 269-364.

[4] Brunner, A., Messkonzepte zur Liquidität auf WertpapiermÄarkten, Beiträge zur Theorie der Finanzmärkte, Universität Frankfurt (1996).

[5] Back, K., Insider Trading in Continuous Time, Review of Financial Studies 5 (1992), pp. 387-409.

[6] Bagheot, W., The only game in town, Financial Analysts Journal, 38 (1971), pp. 31-53.

[7] Baker, H. K., Trading location and liquidity: An analysis of U.S. dealer and agency markets for comon stocks, Financial Markets, Institutions Instruments 5(4) (1996), pp. 1-51.

[8] Battalio, R., Greene, J., Jennings, R., Order flow distribution, bid-ask spreads, and liquidity costs: Merril Lynch's decision to cease routinely routing orders to regional and liquidity costs, Journal of Financial Intermediation 7 (1998), pp. 338-358.

[9] Brüggemann, R., Krolzig, H., Lütkepohl, H., Comparison of Model Reduction Methods for VAR Processes, (2003).

[10] Brockman, P., Chung, D. Y., An empirical investigation of trading on asymmetric information and heterogeneous prior beliefs', Journal of Empirical Finance 7 (2000), pp. 417-454.

[11] Brownlees, C. T., Giampiero, Gallo, M., Financial Econometric Analysis at Ultra-High Frequency: Data Handling Concerns (2006), Technical Report.

[12] Butler, A. W., Grullon, G., Weston, J. P., Stock market liquidity and the cost of raising capital (2002), Working Paper.

[13] Chan, K., Chung, Y. P., Fong, W. M., The informational role of stock and option volume, The Review of Financial Studies 15(4) (2002), pp. 1049-1075.

[14] Chordia, T., Roll, R., Subrahmanyam, A., Market liquidity and trading activity, The Journal of Finance 56(2) (2001), pp. 501-530.

[15] Coppejans, M., Domowitz, I., An empirical analysis of trades, orders, and cancellations in a limit order market, Economics Department Working Paper Duke University (2002).

[16] Copeland, T. E., Galai, D., Informational efects on the bid ask spread, Journal of Finance 38 (1983), pp. 1457-1469.

[17] Doan, T., Litterman, R., Sims, C., Forecasting and Conditional Projection Using Realistic Prior Distributions, Econometric Review, 3 (1984), pp. 1-100.

[18] Easely, D., O'Hara, M., Price, trade size, and information in securities markets, Journal of Financial Economics 19 (1987), pp. 69-90.

[19] Easley, D., Kiefer, N. M., O'Hara, M., One day in the life of a very common stock, Working paper, Cornell University, (1993).

[20] Easley, D.,Kiefer, N. M., O'Hara, M., Paperman,J. B., Liquidity, information, and infrequently traded stocks, Journal of Finance, 51 (1996), pp. 1405-1436.

[21] Engle, R. F., Dufour, A., Time and the Price Impact of a Trade, Journal of Finance, 55 (2000), pp. 2467-2498.

[22] Engle, R. F., Russell, J., Autoregressive Conditional Duration: A new model for irregularly spaced transaction data, Econometrica, 66(5) (1998), pp. 1127-1163.

[23] Engle, R. F., Russell, J., Analysis of High Frequency Financial Data, Econometrica, 66 (2005), pp. 1127-1162.

[24] Engle, R. F., The Econometrics of Ultra High Frequency Data, Econometrica 68 (2000), pp. 1-22.

[25] Farmer, J. D., Zamani, N., Mechanical vs. informational components of price impact, The Europian Phisical Journal B, (2006).

[26] Fleming, M. J., Remolona, E. M., Price formation and liquidity in the U.S. Treasury market: The response to public information, The Journal of Finance 54(5) (1999), pp. 1901-1915.

[27] Foster, D., Viswanathan, S., A Theory of the Intraday Variations in Volume, Variance, and Trading Costs in Securities Markets, Review of Financial Studies, 3 (1990), pp. 593-624.

[28] Franses, P. H., Time Series Models for Business and Economic Forecasting, Cambridge University Press, 1998.

[29] Glosten, L. R., Components of the bid-ask spread and the statistical properties of transaction prices, Journal of Finance 42(5) (1987), pp. 1293-1307.

[30] Glosten, L. R., Milgrom, P., Bid, ask and transaction prices in a specialist market with heterogeneously informed traders. Journal of Financial Economics (1985), pp. 71-100.

[31] Gomber, P., Schweickert, U., Der Market Impact: Liquidittsmass im elektronis- chen Wertpapierhandel', Die Bank 7, (2002) 485-489,

[32] Goldstein, M. A., Kavajecz, K. A., Eighths, sixteenths and market depth, Changes in tick size and liquidity provision on the NYSE, Journal of Financial 56 (2000), pp. 125-149.

[33] Gouriéroux, C., Jasiak, J., Le Fol, G., Intra-day market activity, Journal of Financial Markets 2 (1999), pp. 193-226.

[34] Granger, C.W.J., Investigating Causal Relations by Econometric Models and Cross-Spectral MethodsŠ Econometrica 37 (1969), pp. 424-438.

[35] Greene, W. H., Econometric Analysis, 5th edition, Prentice Hall, 2003.

[36] Hamao, Y., Hasbrouck, J., Securities trading in the absence of dealers: Trades and quotes on the Tokyo Stock Exchange, The Review of Financial Studies 8(3) (1995), pp. 849-878.

[37] Hamilton, J. D., Time Series Analysis, Princeton University Press, (1994).

[38] Hasbrouck J., Inferring Trading Costs from Daily data: US equities for 1962 to 2001, Working Paper, NYU Stern, (2005).

[39] Hasbrouck, J., Measuring the Information Content of Stock Trades, Journal of Finance, vol. 46, No. 1. (1991), pp. 179-207.

[40] Hasbrouck J., The Summary Informativeness of Stock Trades: An Econometric Analysis, The Review of Financial Studies, Vol. 4, No. 3 (1991), pp. 571-595.

[41] Hasbrouck, J., Modelling market microstructure time series, in Handbook of Statistics, vol. 14, Statistical Methods in Finance, eds. G.S. Maddala, C.R. Rao, Elsevier, North-Holland, Amsterdam, (1996).

[42] Hasbrouck, J., Seppi, D. J., Common factors in prices, order flows and liquidity, Journal of Financial Economics 59 (2001), pp. 383-411.

[43] Hasbrouck, J., Trades, quotes, inventories, and information, Journal of Financial Economics 22 (1988), pp. 229-252.

[44] Hausman, J., Lo, A. W., MacKinlay, A. C., An Ordered Probit Analysis of Transaction Stock Prices, Journal of Financial Economics 31 (1992), pp. 319-379.

[45] Irvine, P., Benston, G., Kandel, E., Liquidity beyond the inside spread: Measuring and using information in the limit order book, Working Paper (2000).

[46] Kanji, G. K., 100 Statistical Tests, Sage, London, (1999).

[47] Kavajecz, K. A., A specialist's quoted depth and the limit order book, The Journal of Finance 54(2) (1999), pp. 747-771.

[48] Kruskal, W., When are Gauss-Markov and Least Squares Estimators Identical?, Annals of Mathematical Statistics, 39 (1968), pp. 70-75.

[49] Kunst M. R., Econometrics II-A lecture course for the Institute for Advanced Studies, (2007).

[50] Kyle, A. S., Continuous auctions and insider trading. Econometrica, 53 (1985), pp. 1315-1335.

[51] Lee, C., Ready, M., Inferring trade direction from intraday data, Journal of Finance 46 (1991), pp. 733-746.

[52] Lee, C., Swaminathan, B., Price momentum and trading volume, The Journal of Finance 55(5) (2000), pp. 2017-2069.

[53] Lillo, F., Farmer, J. D., Mantegna, R. N., Master curve for price-impact function Nature 421 (2003), pp. 129-130.

[54] Lütkepohl, H., New Introduction to Multiple Time Series Analysis, Springer-Verlag, 1993.

[55] Madhavan, A., Richardson, M., Roomans, M., Why do security prices change? A transaction-level analysis of NYSE stocks, Review of Financial Studies 10(4) (1997), pp. 1035-1064.

[56] Müller, U., Dacorogna, M., Olsen, R., Pictet, O., Schwarz, M., Morgenegg, C., Statistical study of foreign exchange rates, empirical evidence of a price change scaling law, and intraday analysis, Journal of Banking and Finance 14, (1990) pp. 1189-1208.

[57] Newey, W. K., McFadden, D., Large Sample Estimation and Hypothesis Testing, in Handbook of Econometrics, ed. by R. Engle, and D. McFadden, vol. IV, chap. 36 (1994), pp. 2111-2245. Elsevier Science.

[58] O'Hara, M., Market Microstructure Theory, Cambridge, MA, Blackwell Publishers, (1995).

[59] Øksendal, B., Stochastic Differential Equations - An Introduction with Applications, Springer Verlag,(2000).

[60] Ranaldo, A., Intraday Trading Activity on Financial Markets, The Swiss Evidence, PhD thesis, University of Fribourg, (2000).

[61] Sarin, A., Shastri, K. A., Shastri, K., Ownership structure and stock market liqudity, Working Paper, Santa Clara University, (1996).

[62] Schoch, J., Über das Pricing von Liquidität auf Aktienmärkten, Master's thesis, University of St. Gallen, (2001).

[63] Stoll, H. R., Inferring the Components of the Bid-Ask Spread: Theory and Empirical Tests, Journal of Finance 44 (1989), pp. 115-134.

[64] Sims, C. A., Macroeconomics and Reality, Econometrica 48 (1980), pp. 1-48.

[65] Vasiliki P., Parameswaran G., Xavier G., Eugene S., Quantifying stock-price response to demand fluctuations, Phisical Review E 66 (2002).

[66] von Wyss, R., Measuring and Predicting Liquidity in the Stock Market, Dissertation Nr. 2899 of the University of St. Gallen, (2004).

[67] Zellner, A., An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias, Journal of the American Statistical Association 57 (1962), pp. 348-368.

[68] Wold, H. O., A Study in the Analysis of Stationary Time Series, Stockholm, Almqvist and Wiksell, 1938.

[69] White, H., A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity, Econometrica, 48 (1980), pp. 817-838.

[70] Wright, J. H., Bollerslev, T., High frequency data, frequency domain inference and volatility forecasting, International Finance Discussion Papers 64, (1999).

Appendix 1

Table A.1.1
The summary statistics of ten liquidity measures for ABF

ABF							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	14801	6960	$9.30{\cdot}10^5$	1	26831	8.2	142.76
TO	$1.17{\cdot}10^7$	$5.54{\cdot}10^6$	$7.01{\cdot}10^8$	800.5	$2.11{\cdot}10^7$	7.9375	130.64
D	14538	11961	208670	104	11074	2.6637	17.981
$D\$$	$57.61{\cdot}10^5$	$47.99{\cdot}10^5$	$79.45{\cdot}10^6$	45423	$42.82{\cdot}10^5$	2.5683	17.185
Dur	0.0163	0.0058	0.4358	0	0.027	3.6892	24.439
FR	$8.50{\cdot}10^9$	$6.96{\cdot}10^8$	$1.56{\cdot}10^{12}$	0	$2.99{\cdot}10^{10}$	12.54	361.51
S	0.7816	0.5	44	0.5	0.6577	19.081	910.35
QS	0.0473	0.0322	2.8523	0.0227	0.0425	21.309	1007.9
S^{prop}	1.5681	1.0617	78.829	0.9062	1.2739	15.955	678.5
PI^a	0.5425	0.4999	24.175	0.0117	0.4558	21.143	871.79
PI^b	0.5675	0.5001	60.308	0.0249	0.7507	48.094	3412.4

Table A.1.2
The summary statistics of ten liquidity measures for AZN

AZN							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	16209	6477	$4.14{\cdot}10^6$	1	37769	25.882	2064.2
TO	$4.67{\cdot}10^7$	$1.86{\cdot}10^7$	$1.09{\cdot}10^{10}$	2791	$1.08{\cdot}10^8$	23.238	1654.8
D	12675	9529	352610	29	12016	3.4798	29.09
$D\$$	$18.29{\cdot}10^6$	$13.69\ {\cdot}10^6$	$50.41{\cdot}10^7$	42180	$17.43{\cdot}10^6$	3.4696	28.571
Dur	0.0054	0.0019	0.1497	0	0.0089	3.7665	24.843
FR	$4.60{\cdot}10^{10}$	$6.59{\cdot}10^9$	$4.92{\cdot}10^{12}$	0	$1.48{\cdot}10^{11}$	11.026	201.15
S	1.6544	1	53	1	1.0776	4.1679	87.283
QS	0.1034	0.0698	3.2074	0.0442	0.0685	4.0244	74.22
S^{prop}	1.659	1.0287	55.09	0.9386	1.0863	4.259	94.474
PI^a	0.5515	0.5	23.597	0.0277	0.4601	16.714	619.25
PI^b	0.5636	0.5	20.434	0.0385	0.46	13.756	411.2

Table A.1.3
The summary statistics of ten liquidity measures for BARC

BARC							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	44970	19998	$3.25{\cdot}10^6$	1	87771	9.0223	174.73
TO	$2.93{\cdot}10^7$	$1.30{\cdot}10^7$	$2.26{\cdot}10^9$	597.5	$5.72{\cdot}10^7$	9.0989	179.22
D	72360	61657	752020	235	50039	1.8929	10.547
$D\$$	$23.74{\cdot}10^6$	$20.21{\cdot}10^6$	$23{\cdot}10^7$	76970	$16.42\ {\cdot}10^6$	1.8143	9.6202
Dur	0.0063	0.0028	0.1719	0	0.0097	3.7215	25.909
FR	$2.62{\cdot}10^{10}$	$3.74{\cdot}10^9$	$7.17{\cdot}10^{12}$	0	$9.07{\cdot}10^{10}$	20.166	988.15
S	0.62582	0.5	4	0.5	0.26065	2.4132	11.209
QS	0.0316	0.0259	0.2407	0.0203	0.0131	2.6177	14.681
S^{prop}	1.2566	1.0032	7.67	0.9351	0.5349	2.4446	11.4750
PI^a	0.5144	0.4999	21.685	0.0716	0.3140	23.3260	1082
PI^b	0.5106	0.5001	27.213	0.0713	0.3033	32.051	2198.3

Table A.1.4
The summary statistics of ten liquidity measures for CPI

CPI							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	25994	11522	$2.56{\cdot}10^6$	1	52436	12.534	379.17
TO	$1.20{\cdot}10^7$	$5.32{\cdot}10^6$	$1.15{\cdot}10^9$	465.5	$2.42{\cdot}10^7$	12.249	358.05
D	23659	18126	338050	146	20276	2.9221	23.432
$D\$$	$54.75{\cdot}10^5$	$41.97{\cdot}10^5$	$80.29{\cdot}10^6$	33682	$46.95{\cdot}10^5$	2.9821	25.037
Dur	0.0137	0.0044	0.5561	0	0.0243	4.4613	39.923
FR	$9.13{\cdot}10^9$	$8.02{\cdot}10^8$	$1.26{\cdot}10^{12}$	0	$3.47{\cdot}10^{10}$	13.115	282.64
S	0.45823	0.25	13	0.25	0.38901	5.6076	89.485
QS	0.0265	0.0161	0.7494	0.0110	0.0228	5.8447	90.133
S^{prop}	1.8365	1.0279	51.8	0.9412	1.5586	5.5582	89.231
PI^a	0.5601	0.4999	16.004	0.0314	0.4504	7.7028	149.54
PI^b	0.5598	0.5001	26.803	0.0206	0.4982	18.821	781.64

Table A.1.5
The summary statistics of ten liquidity measures for GSK

GSK							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	24610	10728	$3.30{\cdot}10^6$	1	49964	13.555	478.87
TO	$3.73{\cdot}10^7$	$1.62{\cdot}10^7$	$5.08{\cdot}10^9$	1471	$7.57{\cdot}10^7$	13.738	495.68
D	46470	38567	520500	103	34625	2.6127	17.717
$D\$$	$35.24{\cdot}10^6$	$29.20{\cdot}10^6$	$40.85\ {\cdot}10^6$	77862	$26.34\ {\cdot}10^6$	2.6385	18.109
Dur	0.0055	0.0025	0.1089	0	0.008	3.3206	19.87
FR	$3.44{\cdot}10^{10}$	$5.20{\cdot}10^9$	$7.13{\cdot}10^{12}$	0	$1.11{\cdot}10^{11}$	13.666	411.83
S	1.2321	1	11	1	0.5285	3.131	20.026
QS	0.0655	0.0537	0.6275	0.0402	0.0289	3.3989	25.114
S^{prop}	1.2328	1.0118	10.868	0.9532	0.5307	3.115	19.707
PI^a	0.5179	0.4999	17.092	0.0627	0.3068	18.047	641.17
PI^b	0.5183	0.5001	17.404	0.0623	0.2657	18.854	819.25

Table A.1.6
The summary statistics of ten liquidity measures for HBOS

HBOS							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	28913	13399	$2.57{\cdot}10^6$	1	55840	10.646	262.07
TO	$2.77{\cdot}10^7$	$1.29{\cdot}10^7$	$2.63{\cdot}10^9$	937.5	$5.35{\cdot}10^7$	10.796	269.85
D	35543	29208	521310	20	26853	2.1437	13.825
$D\$$	$17.10{\cdot}10^6$	$14.01{\cdot}10^6$	$24.32\ {\cdot}10^7$	9420.8	$12.95{\cdot}10^6$	2.1342	13.427
Dur	0.0066	0.0028	0.1606	0	0.0103	3.4467	20.687
FR	$2.37{\cdot}10^{10}$	$3.60{\cdot}10^9$	$3.73{\cdot}10^{12}$	0	$7.62{\cdot}10^{10}$	12.825	311.34
S	0.7016	0.5	10	0.5	0.3752	2.8851	20.509
QS	0.0385	0.0287	0.5331	0.0214	0.0210	3.127	23.542
S^{prop}	1.4055	1.0183	18.844	0.9365	0.7518	2.8238	18.869
PI^a	0.5292	0.4999	24.06	0.0557	0.3797	21.89	974.7
PI^b	0.5387	0.5001	36.154	0.0499	0.5025	35.989	2015.5

Table A.1.7
The summary statistics of ten liquidity measures for HSBA

HSBA							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	49367	18386	$6.32{\cdot}10^6$	1	114470	12.648	344.38
TO	$4.75{\cdot}10^7$	$1.77{\cdot}10^7$	$6.13{\cdot}10^9$	957	$1.11{\cdot}10^8$	12.794	352.11
D	128200	108090	$17.46{\cdot}10^5$	223	92905	2.2165	14.574
$D\$$	$61.92{\cdot}10^6$	$52{\cdot}10^6$	$85.80{\cdot}10^7$	102860	$45.25{\cdot}10^6$	2.2395	14.854
Dur	0.0055	0.0025	0.1603	0	0.0079	3.4392	22.449
FR	$4.23{\cdot}10^{10}$	$5.65{\cdot}10^9$	$8.29{\cdot}10^{12}$	0	$1.49{\cdot}10^{11}$	14.6	401.98
S	0.5715	0.5	6.5	0.5	0.203	4.0814	38.75
QS	0.0275	0.0241	0.3763	0.0188	0.0102	4.6948	56.379
S^{prop}	1.1445	1.0007	12.7120	0.9665	0.4086	4.0389	37.389
PI^a	0.5138	0.4999	30.732	0.1	0.3467	49.396	3651.5
PI^b	0.5183	0.5	40.075	0.0624	0.4402	52.94	3924.5

Table A.1.8
The summary statistics of ten liquidity measures for IAP

IAP							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	20619	7984.5	$3.29{\cdot}10^6$	1	55762	17.83	686.46
TO	$1.01{\cdot}10^7$	$3.91{\cdot}10^6$	$1.50{\cdot}10^9$	474.25	$2.73{\cdot}10^7$	16.463	565.1
D	14563	11421	973720	125	17774	29.275	1455.6
$D\$$	$35.89{\cdot}10^5$	$27.97{\cdot}10^5$	$24.59{\cdot}10^7$	28995	$44.48{\cdot}10^5$	30.032	1510.2
Dur	0.0188	0.0061	0.6714	0	0.0341	4.5605	38.23
FR	$6.47{\cdot}10^9$	$3.80{\cdot}10^8$	$1.44{\cdot}10^{12}$	0	$3.17{\cdot}10^{10}$	17.999	534.39
S	0.77291	0.5	30.25	0.25	0.8222	7.1893	154.9
QS	0.0472	0.0306	1.6353	0.0109	0.0498	6.1567	112.7
S^{prop}	3.1079	2.061	119.06	0.8577	3.3098	7.0399	145.85
PI^a	0.6675	0.4999	58.878	0.0168	1.2352	23.832	867.05
PI^b	0.6552	0.5001	30.765	0.0154	0.9238	13.5	312.9

Table A.1.9
The summary statistics of ten liquidity measures for KAZ

KAZ							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	14508	5764	$2.76{\cdot}10^6$	1	40389	23.274	1067.7
TO	$1.60{\cdot}10^7$	$6.36{\cdot}10^6$	$3.21{\cdot}10^9$	1044	$4.50{\cdot}10^7$	24.7	1201.9
D	7886	6355	108550	97	6396.3	3.1084	22.526
$D\$$	$43.56{\cdot}10^5$	$35.15\ {\cdot}10^5$	$62.86{\cdot}10^6$	41108	$35.43{\cdot}10^5$	3.0903	22.702
Dur	0.0153	0.0044	1.5778	0	0.031	9.4172	279.44
FR	$8.89{\cdot}10^9$	$7.61{\cdot}10^8$	$3.45{\cdot}10^{12}$	0	$4.21{\cdot}10^{10}$	35.293	2196.6
S	2.1605	1.5	309.5	0.5	3.0297	38.346	3625.9
QS	0.1412	0.0954	22.238	0.0243	0.2074	43.828	4403.1
S^{prop}	4.3141	2.9349	122.05	0.7722	4.9444	6.1194	82.579
PI^a	0.7197	0.4999	93.592	0.0043	1.6115	28.022	1202
PI^b	0.7015	0.5001	38.054	0.0095	1.1309	15.179	400.65

Table A.1.10
The summary statistics of ten liquidity measures for LLOY

LLOY							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	52411	22270	$6.96{\cdot}10^6$	1	116860	14.52	476.45
TO	$2.78{\cdot}10^7$	$1.19{\cdot}10^7$	$3.75{\cdot}10^9$	541.5	$6.19{\cdot}10^7$	14.545	482.42
D	102570	77960	$19.38{\cdot}10^5$	270	91543	2.9379	20.595
$D\$$	$27.30{\cdot}10^6$	$20.74{\cdot}10^6$	$50.96{\cdot}10^7$	68983	$24.32{\cdot}10^6$	2.9127	20.162
Dur	0.0071	0.0031	0.1922	0	0.0107	3.4099	21.934
FR	$2.14{\cdot}10^{10}$	$2.92{\cdot}10^9$	$4.32{\cdot}10^{12}$	0	$7.74{\cdot}10^{10}$	16.444	505.8
S	0.5454	0.5	3.5	0.25	0.2551	1.9329	9.3727
QS	0.0268	0.0241	0.1766	0.0102	0.0124	2.1183	11.494
S^{prop}	2.1893	2.0049	13.839	0.9572	1.0347	1.8845	8.9693
PI^a	0.519	0.4999	15.559	0.0626	0.3234	17.013	568.33
PI^b	0.5159	0.5001	22.485	0.0555	0.3328	27.355	1414.4

Table A.1.11
The summary statistics of ten liquidity measures for PRU

PRU							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	36252	16363	$3.25\cdot10^6$	1	69341	9.077	187.45
TO	$2.34\cdot10^7$	$1.05\cdot10^7$	$1.89\cdot10^9$	592.5	$4.49\cdot10^7$	8.5219	153.08
D	67383	38024	$60\cdot10^7$	158	$33.66\cdot10^5$	178.19	31756
$D\$$	$21.48\cdot10^6$	$12.09\cdot10^6$	$18.23\cdot10^{10}$	53573	$10.23\cdot10^8$	178.18	31754
Dur	0.0074	0.0028	0.2217	0	0.0121	3.8373	26.852
FR	$2.09\cdot10^{10}$	$2.57\cdot10^9$	$4.37\cdot10^{12}$	0	$7.16\cdot10^{10}$	15.062	504.12
S	0.75444	0.5	34	0.5	0.5655	14.828	581.82
QS	0.0401	0.0284	2.0558	0.0175	0.0311	16.717	713.38
S^{prop}	1.5139	1.0707	64.032	0.8426	1.1007	13.306	498.3
PI^a	0.539	0.4999	37.533	0.018	0.504	30.347	1607.7
PI^b	0.5308	0.5001	24.706	0.0314	0.401	22.465	1031.5

Table A.1.12
The summary statistics of ten liquidity measures for RB

RB							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	9686.8	4569	$1.01\cdot10^6$	1	17977	11.18	329.32
TO	$1.95\cdot10^7$	$9.18\cdot10^6$	$2.07\cdot10^9$	1925	$3.62\cdot10^7$	11.325	341.24
D	9104	7445	79360	38	6664.9	1.9751	10.039
$D\$$	$91.49\cdot10^5$	$74.74\cdot10^5$	$79.95\cdot10^6$	39008	$67.21\cdot10^5$	1.9934	10.196
Dur	0.0086	0.0031	0.2686	0	0.0146	4.0318	29.328
FR	$1.67\cdot10^{10}$	$2.08\cdot10^9$	$2.56\cdot10^{12}$	0	$5.33\cdot10^{10}$	12.817	342.08
S	1.6549	1	42	1	1.1332	5.5351	104.28
QS	0.1059	0.0713	2.9563	0.0482	0.0751	6.254	128.61
S^{prop}	1.6563	1.0237	41.261	0.94119	1.1334	5.3731	98.165
PI^a	0.5432	0.4999	22.131	0.0294	0.4641	17.049	577.76
PI^b	0.5476	0.5001	40.55	0.0311	0.4562	31.602	2389.3

Table A.1.13
The summary statistics of ten liquidity measures for RIO

RIO							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	11704	5214	$1.58{\cdot}10^{6}$	1	25437	14.972	483.02
TO	$3.45{\cdot}10^{7}$	$1.54{\cdot}10^{7}$	$4.75{\cdot}10^{9}$	2955	$7.50{\cdot}10^{7}$	15.016	487.19
D	8522.4	6489	$20.\ {\cdot}10^{6}$	33	80141	247.62	61769
$D\$$	$12.55{\cdot}10^{6}$	$95.24{\cdot}10^{5}$	$32.16{\cdot}10^{9}$	49997	$12.88{\cdot}10^{7}$	248.09	61926
Dur	0.0038	0.0014	0.1431	0	0.0067	4.3738	33.444
FR	$3.78{\cdot}10^{10}$	$6.63{\cdot}10^{9}$	$9.22{\cdot}10^{12}$	0	$1.22{\cdot}10^{11}$	17.9	727.1
S	2.0864	2	60	1	1.5922	3.7533	47.696
QS	0.1363	0.1137	3.7016	0.0473	0.1061	3.7731	43.161
S^{prop}	2.0959	1.844	57.482	0.8685	1.6193	3.8583	49.219
PI^{a}	0.5908	0.5	112.74	0.0179	1.0568	65.231	5730.4
PI^{b}	0.5891	0.5	176.19	0.0138	0.9449	116.99	20582

Table A.1.14
The summary statistics of ten liquidity measures for SHP

SHP							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	17374	7634	$1.45{\cdot}10^{6}$	1	33504	9.7114	218.12
TO	$1.50{\cdot}10^{7}$	$6.58{\cdot}10^{6}$	$1.33{\cdot}10^{9}$	815.5	$2.92{\cdot}10^{7}$	10.156	243.05
D	14527	11531	156950	125	11775	2.6716	15.587
$D\$$	$62.85{\cdot}10^{5}$	$49.43{\cdot}10^{5}$	$70.90{\cdot}10^{6}$	55566	$51.50{\cdot}10^{5}$	2.6792	15.554
Dur	0.011	0.0036	0.2714	0	0.0191	3.7688	24.244
FR	$1.20{\cdot}10^{10}$	$1.24{\cdot}10^{9}$	$2.04{\cdot}10^{12}$	0	$4.17{\cdot}10^{10}$	12.527	302.34
S	0.86047	0.5	57.5	0.5	0.78415	22.836	1367.7
QS	0.0521	0.0336	4.2083	0.0227	0.0508	28.781	1977.4
S^{prop}	1.7268	1.062	127.06	0.8992	1.6053	26.649	1812.5
PI^{a}	0.5622	0.4999	44.042	0.0293	0.6494	35.828	2032.3
PI^{b}	0.5588	0.5001	32.209	0.0121	0.5504	22.214	932.15

Table A.1.15
The summary statistics of ten liquidity measures for SLOU

SLOU							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	14982	7004	$1.21\cdot10^6$	1	28598	10.179	231.47
TO	$9.39\cdot10^6$	$4.39\cdot10^6$	$7.90\cdot10^8$	614.5	$1.81\cdot10^7$	10.716	258.21
D	15469	13259	128190	93	10472	1.9969	11.743
$D\$$	$48.38\cdot10^5$	$41.49\cdot10^5$	$41.60\cdot10^6$	27666	$32.83\cdot10^5$	2.0203	12.017
Dur	0.019	0.0069	0.5639	0	0.0313	3.8003	26.831
FR	$6.07\cdot10^9$	$4.35\cdot10^8$	$1.57\cdot10^{12}$	0	$2.63\cdot10^{10}$	19.529	753.54
S	0.85512	0.5	42	0.5	0.89155	16.716	529.62
QS	0.051	0.0326	2.3891	0.0233	0.0546	16.773	496.82
S^{prop}	1.7135	1.0434	77.645	0.8918	1.7571	15.344	442.88
PI^a	0.578	0.4999	46.371	0.0137	0.7858	34.665	1791.4
PI^b	0.5756	0.5001	43.73	0.0141	0.778	33.331	1556.4

Table A.1.16
The summary statistics of ten liquidity measures for VOD

VOD							
	mean	med	max	min	st.dev	sqewness	kurtoses
V	448390	128300	$7.90\cdot10^7$	1	$1.22\cdot10^6$	14.955	481.04
TO	$5.58\cdot10^7$	$1.60\cdot10^7$	$9.95\cdot10^9$	112.5	$1.53\cdot10^8$	15.171	498.98
D	$39.01\cdot10^5$	$33.89\cdot10^5$	$30.02\cdot10^6$	11952	$23.61\cdot10^5$	2.1147	11.685
$D\$$	$24.21\cdot10^7$	$21.22\cdot10^7$	$17.28\cdot10^8$	706290	$14.30\cdot10^7$	1.9325	10.191
Dur	0.0056	0.0025	0.1278	0	0.0079	3.2357	19.502
FR	$5.60\cdot10^{10}$	$5.16\cdot10^9$	$2.09\cdot10^{13}$	0	$2.47\cdot10^{11}$	27.879	1563.8
S	0.25961	0.25	1.5	0.25	0.0511	6.3137	60.421
QS	0.0093	0.0089	0.0718	0.0077	0.0019	6.5241	81.627
S^{prop}	1.0397	0.9940	6.2969	0.9155	0.2092	6.1979	62.429
PI^a	0.5027	0.4998	21.99	0.0993	0.2265	56.684	4312
PI^b	0.5017	0.5003	7.4499	0.1247	0.1079	24.908	1199.9

Table A.1.17
The summary statistics of ten liquidity measures for WPP

	WPP						
	mean	med	max	min	st.dev	sqewness	kurtoses
V	31490	13600	$1.80{\cdot}10^6$	1	61476	7.909	117.66
TO	$2.13{\cdot}10^7$	$9.21{\cdot}10^6$	$1.22{\cdot}10^9$	675	$4.17{\cdot}10^7$	7.9527	119.35
D	41067	33009	578480	18	32546	2.7199	18.134
$D\$$	$13.94{\cdot}10^6$	$11.19{\cdot}10^6$	$19.71{\cdot}10^7$	5932.3	$11.09{\cdot}10^6$	2.7149	18.046
Dur	0.0114	0.0044	0.2861	0	0.018	3.4896	22.329
FR	$1.74{\cdot}10^{10}$	$1.58{\cdot}10^9$	$4.40{\cdot}10^{12}$	0	$6.69{\cdot}10^{10}$	16.902	637.38
S	0.6987	0.5	21.5	0.5	0.4055	7.3697	221.38
QS	0.0377	0.0282	1.4912	0.0205	0.0228	11.319	497.89
S^{prop}	1.3985	1.0048	43.891	0.9566	0.8163	7.4834	231.27
PI^a	0.5241	0.4999	25.76	0.0555	0.4108	28.955	1460.6
PI^b	0.521	0.5001	16.538	0.0523	0.3419	18.864	708.25

Table A.1.18
The summary statistics of ten liquidity measures for XTA

	XTA						
	mean	med	max	min	st.dev	sqewness	kurtoses
V	12868	5752.5	$3.98{\cdot}10^6$	1	34242	39.825	3386.3
TO	$2.56{\cdot}10^7$	$1.13{\cdot}10^7$	$6.76{\cdot}10^9$	1632	$6.70{\cdot}10^7$	34.139	2458
D	8955.8	7067	196860	14	7616.3	3.8527	44.179
$D\$$	$88.91{\cdot}10^5$	$69.92{\cdot}10^5$	$20.33{\cdot}10^7$	13714	$76.49{\cdot}10^5$	4.0116	47.904
Dur	0.0055	0.0019	0.1956	0	0.0103	4.7774	39.009
FR	$2.43{\cdot}10^{10}$	$3.59{\cdot}10^9$	$7.74{\cdot}10^{12}$	0	$8.94{\cdot}10^{10}$	25.659	1384.7
S	2.4139	2	73	1	1.9192	3.5268	44.191
QS	0.1561	0.1223	3.8089	0.0451	0.1266	3.4362	32.865
S^{prop}	2.4304	1.9655	79.168	0.7903	1.9316	3.8115	55.604
PI^a	0.6078	0.4999	52.149	0.0174	0.792	26.144	1212.8
PI^b	0.6007	0.5001	23.887	0.0207	0.5656	8.7799	212.98

Appendix 2

Table A.2.1
Estimation of the Hasbrouck's VAR model for ABF

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for ABF based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

ABF					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0789	-12.84*	c_1	-0.3281	-8.79*
a_2	-0.0077	-1.26	c_2	-0.0202	-0.92
a_3	-0.0109	-1.76	c_3	-0.0187	-1.35
a_4	0.0209	3.39*	c_4	-0.0209	-2.03*
a_5	-0.0421	-6.85*	c_5	-0.0119	-0.93
b_0	0.2423	68.32*	d_1	0.2618	23.67*
b_1	0.0339	8.59*	d_2	0.1008	12.15*
b_2	0.0124	3.13*	d_3	0.0739	10.31*
b_3	0.0009	0.22	d_4	0.0442	6.47*
b_4	-0.0172	-4.34*	d_5	0.0386	5.55*
b_5	-0.0028	-0.72			
$\sum_{i=0}^{5} b_i = 0.2695$					
$R_r^2 = 17.90\%$			$R_x^2 = 10.57\%$		
$\sigma_1^2 = 0.2876$			$\Lambda = 0.8354$		
$R_\omega^2 = 47.72\%$					

Table A.2.1′
Hypothesis tests for estimated Hasbrouck's VAR model for ABF

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	5670.3	1808.8	83.8

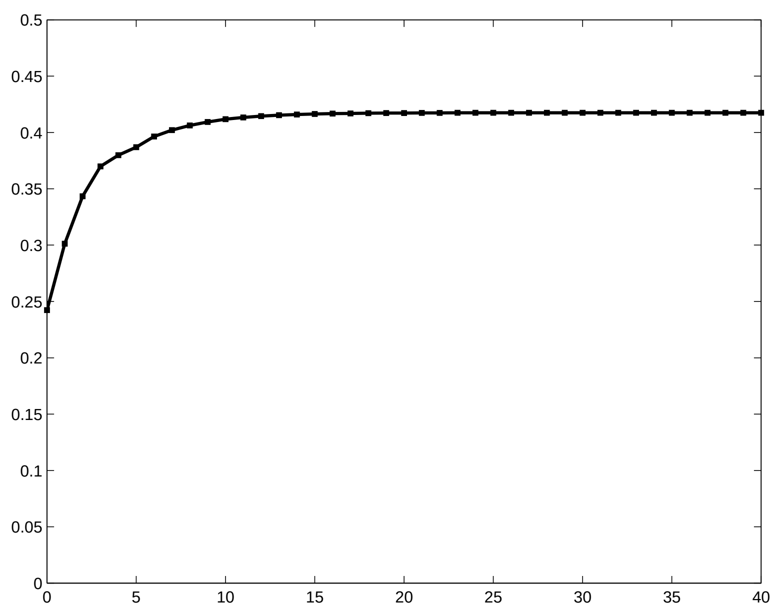

Figure 5.2: Price impact function for ABF.
The total impact is 0.4175 of average proportional spread. It is fully realized after 27 transactions.

Table A.2.2
Estimation of the Hasbrouck's VAR model for AZN

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for AZN based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

AZN					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0807	-23.22*	c_1	-0.4847	-55.23*
a_2	-0.008	-2.31*	c_2	-0.0379	-5.12*
a_3	0.0168	4.83*	c_3	-0.004	-0.55
a_4	0.0228	6.55*	c_4	-0.0027	-0.4
a_5	0.0122	3.51*	c_5	-0.0081	-1.21
b_0	0.2961	166.53*	d_1	0.2609	59.95*
b_1	0.0272	13.01*	d_2	0.0754	18.72*
b_2	0.0082	3.93*	d_3	0.0415	10.25*
b_3	-0.0048	-2.30*	d_4	0.0275	6.91*
b_4	-0.0097	-4.61*	d_5	0.0268	6.86*
b_5	-0.0082	-3.97*			
$\sum_{i=0}^{5} b_i = 0.3089$					
$R_r^2 = 27.28\%$			$R_x^2 = 8.64\%$		
$\sigma_1^2 = 0.2391$			$\Lambda = 0.8572$		
$R_\omega^2 = 51.48\%$					

Table A.2.2′
Hypothesis tests for estimated Hasbrouck's VAR model for AZN

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	31065	6911.2	3171

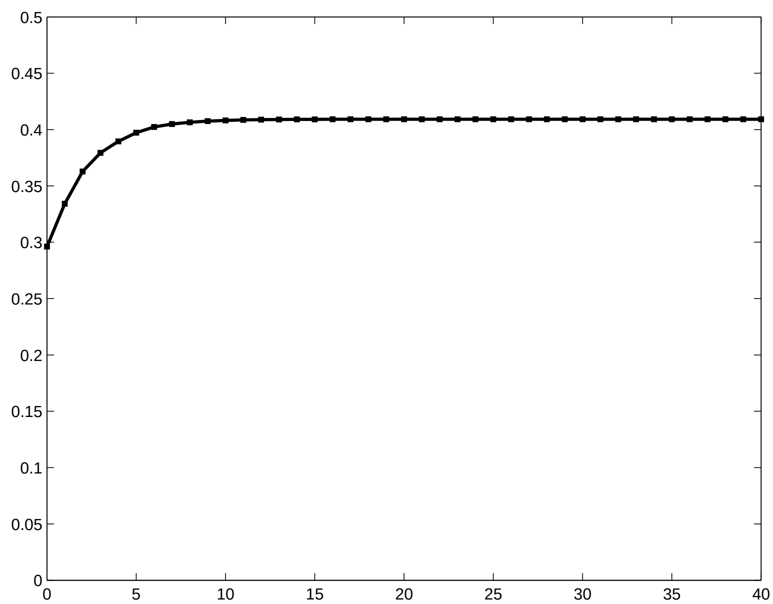

Figure 5.3: Price impact function for AZN
The total impact is 0.4092 of average proportional spread. It is fully realized after 18 transactions.

Table A.2.3
Estimation of the Hasbrouck's VAR model for BARC

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for BARC based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

BARC					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0915	-24.21*	c_1	-0.6537	-75.09*
a_2	-0.0074	-1.94	c_2	-0.04	-4.55*
a_3	0.0076	2.00*	c_3	-0.0236	-2.67*
a_4	0.0143	3.78*	c_4	-0.0222	-2.54*
a_5	0.0136	3.61*	c_5	-0.0083	-0.96
b_0	0.2222	139.37*	d_1	0.2914	69.26*
b_1	0.0305	16.55*	d_2	0.0667	15.71*
b_2	0.0074	4.02*	d_3	0.0494	11.56*
b_3	-0.0033	-1.76	d_4	0.038	8.95*
b_4	-0.0062	-3.35*	d_5	0.0265	6.40*
b_5	-0.0107	-5.90*			
$\sum_{i=0}^{5} b_i = 0.2401$					
$R_r^2 = 24.64\%$			$R_x^2 = 10.84\%$		
$\sigma_1^2 = 0.1562$			$\Lambda = 0.8115$		
$R_\omega^2 = 51.79\%$					

Table A.2.3′
Hypothesis tests for estimated Hasbrouck's VAR model for BARC

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	22984	6139.4	5660.8

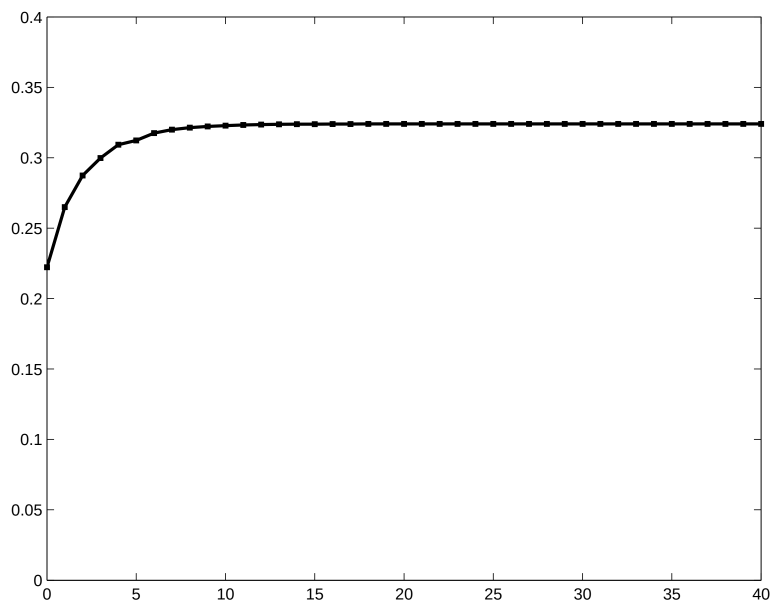

Figure 5.4: Price impact function for BARC
The total impact is 0.3240 of average proportional spread. It is fully realized after 18 transactions.

Table A.2.4
Estimation of the Hasbrouck's VAR model for CPI

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for CPI based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

CPI					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0864	-15.40*	c_1	-0.3585	-28.95*
a_2	0.0028	0.5	c_2	-0.0052	-0.49
a_3	0.0202	3.58*	c_3	-0.0223	-2.26*
a_4	-0.0203	-3.61*	c_4	-0.0236	-2.39*
a_5	-0.0014	-0.24	c_5	-0.0237	-2.42*
b_0	0.2695	83.32*	d_1	0.2611	40.92*
b_1	0.0407	11.13*	d_2	0.08	12.64*
b_2	0.0089	2.42*	d_3	0.0637	10.17*
b_3	-0.0092	-2.52*	d_4	0.0437	6.94*
b_4	-0.0019	-0.53	d_5	0.0283	4.60*
b_5	-0.0032	-0.89			
$\sum_{i=0}^{5} b_i = 0.3047$					
$R_r^2 = 20.76\%$			$R_x^2 = 9.17\%$		
$\sigma_1^2 = 0.2946$			$\Lambda = 0.8485$		
$R_\omega^2 = 49.42\%$					

Table A.2.4′
Hypothesis tests for estimated Hasbrouck's VAR model for CPI

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	8243.5	2395.3	860.5

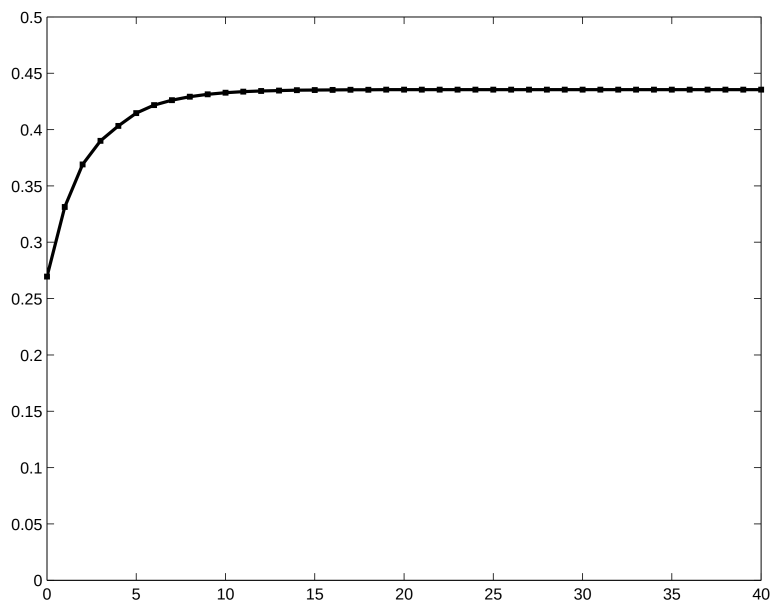

Figure 5.5: Price impact function for CPI
The total impact is 0.4355 of average proportional spread. It is fully realized after 23 transactions.

Table A.2.5
Estimation of the Hasbrouck's VAR model for GSK

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for GSK based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

GSK					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0972	-27.70*	c_1	-0.6173	-74.42*
a_2	0.0059	1.67	c_2	-0.0347	-4.18*
a_3	0.024	6.82*	c_3	0.0058	0.69
a_4	0.0124	3.53*	c_4	0.0111	1.34
a_5	0.0113	3.22*	c_5	0.0198	2.40*
b_0	0.2051	140.70*	d_1	0.286	74.03*
b_1	0.0338	20.26*	d_2	0.0689	17.56*
b_2	0.0079	4.70*	d_3	0.0552	14.04*
b_3	-0.0028	-1.68	d_4	0.0331	8.46*
b_4	-0.0045	-2.67*	d_5	0.0255	6.65*
b_5	-0.0065	-3.96*			
$\sum_{i=0}^{5} b_i = 0.2330$					
$R_r^2 = 22.89\%$			$R_x^2 = 10.49\%$		
$\sigma_1^2 = 0.1544$			$\Lambda = 0.8357$		
$R_\omega^2 = 50.99\%$					

Table A.2.5′
Hypothesis tests for estimated Hasbrouck's VAR model for GSK

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	23947	7276.6	5545.9

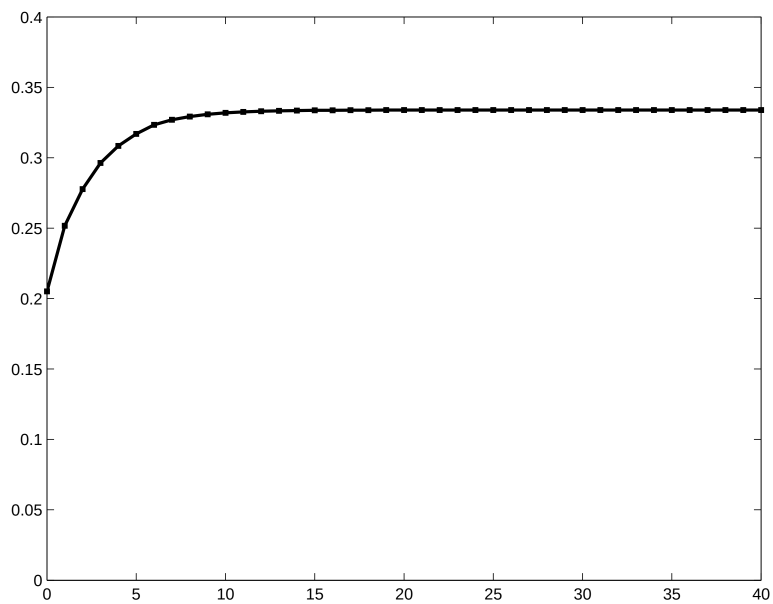

Figure 5.6: Price impact function for GSK
The total impact is 0.3339 of average proportional spread. It is fully realized after 21 transactions.

Table A.2.6
Estimation of the Hasbrouck's VAR model for HBOS

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for HBOS based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

HBOS					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0666	-17.32*	c_1	-0.4957	-57.14*
a_2	0.0163	4.23*	c_2	0.0138	1.62
a_3	0.024	6.23*	c_3	0.0074	0.9
a_4	0.0132	3.43*	c_4	0.0042	0.51
a_5	0.0057	1.5	c_5	-0.0043	-0.53
b_0	0.2631	146.20*	d_1	0.2992	67.63*
b_1	0.0287	13.54*	d_2	0.0645	14.34*
b_2	0.0054	2.54*	d_3	0.0448	9.97*
b_3	-0.0035	-1.67	d_4	0.0312	7.01*
b_4	-0.0068	-3.20*	d_5	0.0276	6.35*
b_5	-0.007	-3.37*			
$\sum_{i=0}^{5} b_i = 0.2799$					
$R_r^2 = 26.50\%$			$R_x^2 = 9.65\%$		
$\sigma_1^2 = 0.1934$			$\Lambda = 0.8358$		
$R_\omega^2 = 54.40\%$					

Table A.2.6′
Hypothesis tests for estimated Hasbrouck's VAR model for HBOS

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	24744	5828.9	3273.4

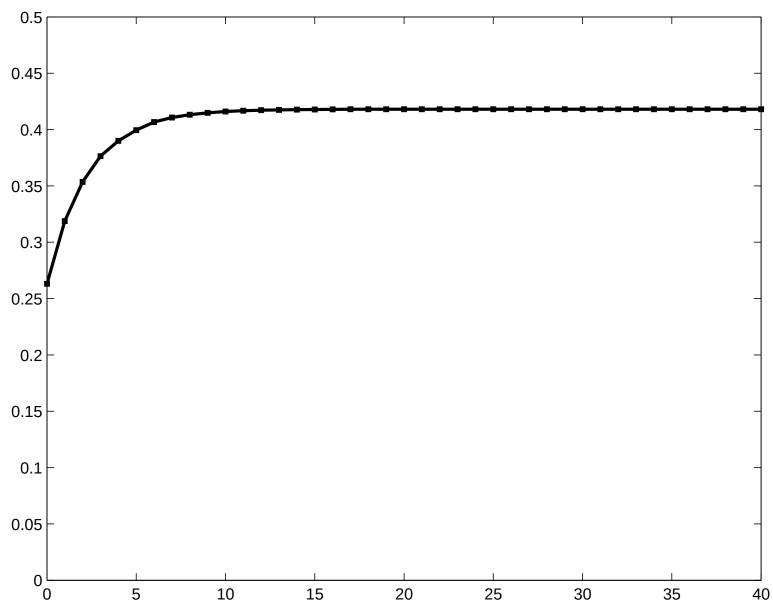

Figure 5.7: Price impact function for HBOS
The total impact is 0.4181 of average proportional spread. It is fully realized after 20 transactions.

Table A.2.7
Estimation of the Hasbrouck's VAR model for HSBA

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for HSBA based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

HSBA					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.1208	-34.20*	c_1	-0.6958	-70.95*
a_2	-0.0217	-6.09*	c_2	-0.1113	-12.49*
a_3	0.0053	1.48	c_3	-0.0543	-6.27*
a_4	0.004	1.13	c_4	-0.0217	-2.57*
a_5	0.0028	0.78	c_5	-0.0135	-1.57
b_0	0.1838	128.61*	d_1	0.2852	72.57*
b_1	0.0291	18.09*	d_2	0.0797	20.41*
b_2	0.0102	6.31*	d_3	0.0688	17.71*
b_3	-0.0051	-3.18*	d_4	0.0353	9.10*
b_4	-0.0052	-3.19*	d_5	0.0378	10.03*
b_5	-0.01	-6.32*			
$\sum_{i=0}^{5} b_i = 0.2028$					
$R_r^2 = 20.92\%$			$R_x^2 = 12.23\%$		
$\sigma_1^2 = 0.14745$			$\Lambda = 0.8293$		
$R_\omega^2 = 48.75\%$					

Table A.2.7′
Hypothesis tests for estimated Hasbrouck's VAR model for HSBA

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0,$ $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0,$ $i = 1, 2, ..., 5$
Wald test	20286	7338.8	5037.9

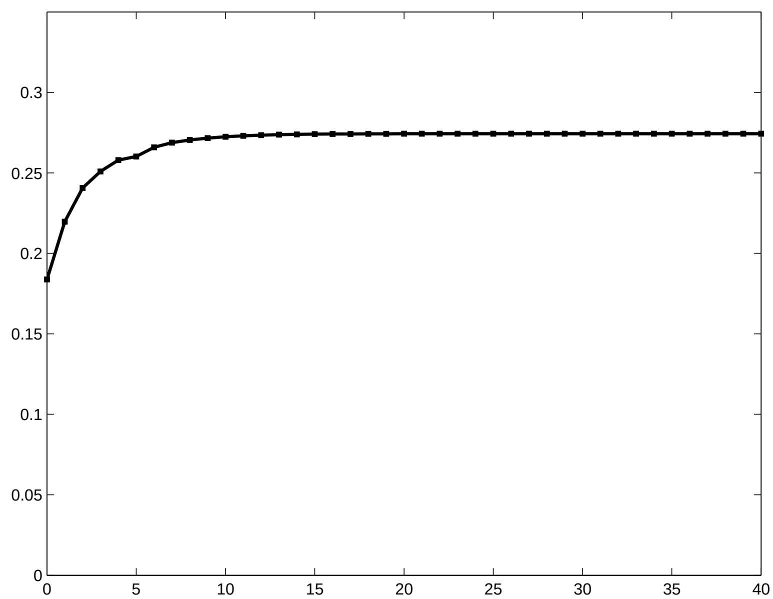

Figure 5.8: Price impact function for HSBA
The total impact is 0.2743 of average proportional spread. It is fully realized after 20 transactions.

Table A.2.8
Estimation of the Hasbrouck's VAR model for IAP

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for IAP based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

IAP					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0603	-8.91*	c_1	-0.3031	-19.99*
a_2	0.0013	0.19	c_2	-0.0348	-3.44*
a_3	0.0037	0.54	c_3	-0.0262	-2.51*
a_4	0.0149	2.20*	c_4	-0.0051	-0.47
a_5	-0.0025	-0.37	c_5	-0.0068	-0.72
b_0	0.2748	59.34*	d_1	0.238	31.03*
b_1	0.0371	7.28*	d_2	0.0837	11.28*
b_2	0.0106	2.08*	d_3	0.0661	8.91*
b_3	0.0058	1.14	d_4	0.0345	4.61*
b_4	-0.0154	-3.02*	d_5	0.0427	5.97*
b_5	-0.0074	-6.32*			
$\sum_{i=0}^{5} b_i = 0.3056$					
$R_r^2 = 15.92\%$			$R_x^2 = 9.03\%$		
$\sigma_1^2 = 0.4286$			$\Lambda = 0.8744$		
$R_\omega^2 = 39.95\%$					

Table A.2.8′
Hypothesis tests for estimated Hasbrouck's VAR model for IAP

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	4143	1153.1	405.3

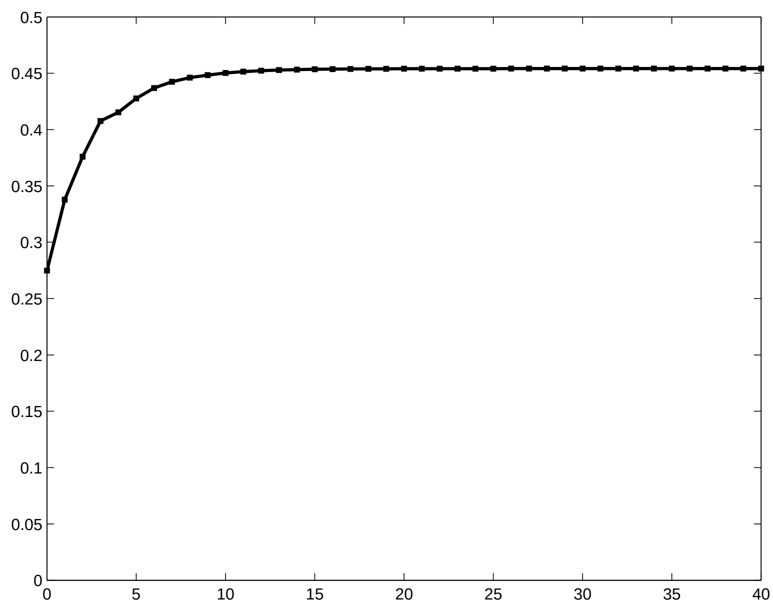

Figure 5.9: Price impact function for IAP
The total impact is 0.4541 of average proportional spread. It is fully realized after 23 transactions.

Table A.2.9
Estimation of the Hasbrouck's VAR model for KAZ

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for KAZ based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

KAZ					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.2677	-45.20*	c_1	-0.2031	-5.45*
a_2	-0.0996	-16.24*	c_2	-0.0596	-2.32*
a_3	-0.0468	-7.60*	c_3	-0.0322	-2.93*
a_4	-0.0142	-2.32*	c_4	-0.0196	-2.47*
a_5	0.011	1.86	c_5	-0.0084	-1.07
b_0	0.2713	52.53*	d_1	0.1992	14.68*
b_1	0.0877	15.95*	d_2	0.106	11.30*
b_2	0.0315	5.69*	d_3	0.07	10.43*
b_3	0.0063	1.13	d_4	0.0465	7.10*
b_4	0.0131	2.37*	d_5	0.0368	5.78*
b_5	-0.0157	-2.87*			
$\sum_{i=0}^{5} b_i = 0.3942$					
$R_r^2 = 17.16\%$			$R_x^2 = 8.00\%$		
$\sigma_1^2 = 0.7062$			$\Lambda = 0.8994$		
$R_\omega^2 = 40.30\%$					

Table A.2.9′
Hypothesis tests for estimated Hasbrouck's VAR model for KAZ

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0,$ $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0,$ $i = 1, 2, ..., 5$
Wald test	3856.5	2366	43.9

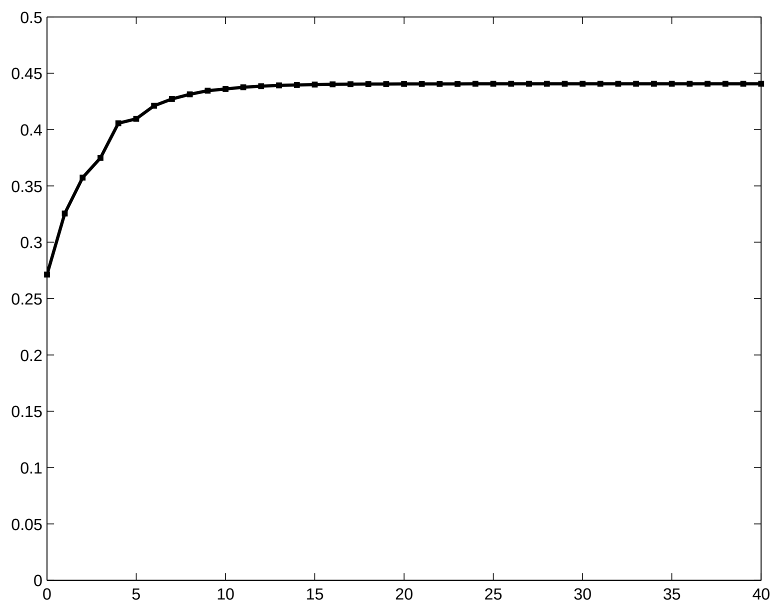

Figure 5.10: Price impact function for KAZ
The total impact is 0.4406 of average proportional spread. It is fully realized after 22 transactions.

Table A.2.10
Estimation of the Hasbrouck's VAR model for LLOY

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for LLOY based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

LLOY					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0967	-24.05*	c_1	-0.6585	-72.17*
a_2	-0.0063	-1.56	c_2	-0.0414	-4.40*
a_3	-0.0021	-0.52	c_3	-0.0409	-4.34*
a_4	0.0095	2.35*	c_4	-0.0241	-2.56*
a_5	0.0039	0.98	c_5	-0.0013	-0.14
b_0	0.1999	119.99*	d_1	0.2814	64.23*
b_1	0.0334	17.65*	d_2	0.0538	12.12*
b_2	0.0068	3.56*	d_3	0.0514	11.53*
b_3	0.0018	0.96	d_4	0.0432	9.73*
b_4	-0.007	-3.68*	d_5	0.0233	5.37*
b_5	-0.0033	-1.78			
$\sum_{i=0}^{5} b_i = 0.2316$					
$R_r^2 = 22.02\%$			$R_x^2 = 10.55\%$		
$\sigma_1^2 = 0.1523$			$\Lambda = 0.8210$		
$R_\omega^2 = 49.17\%$					

Table A.2.10′
Hypothesis tests for estimated Hasbrouck's VAR model for LLOY

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	17364	5392.1	5243.9

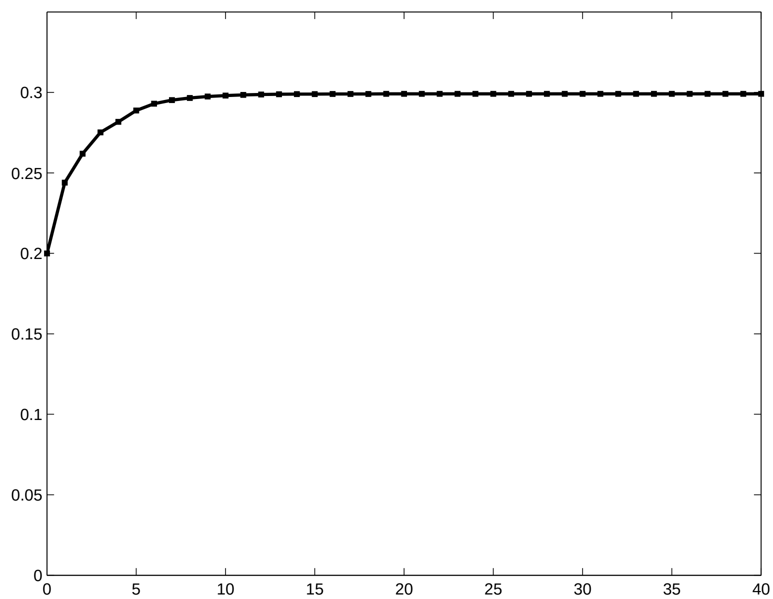

Figure 5.11: Price impact function for LLOY
The total impact is 0.2991 of average proportional spread. It is fully realized after 19 transactions.

Table A.2.11
Estimation of the Hasbrouck's VAR model for PRU

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for PRU based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

PRU					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.132	-32.39*	c_1	-0.4363	-10.93*
a_2	-0.0508	-12.37*	c_2	-0.0372	-2.72*
a_3	0.0086	2.09*	c_3	-0.0256	-1.93
a_4	0.0488	11.89*	c_4	0.0095	0.66
a_5	-0.0302	-7.42*	c_5	0.0225	1.8
b_0	0.2465	114.41*	d_1	0.2531	21.63*
b_1	0.033	13.62*	d_2	0.0661	13.30*
b_2	0.0176	7.22*	d_3	0.0535	9.49*
b_3	0.0032	1.3	d_4	0.0334	5.76*
b_4	-0.023	-9.47*	d_5	0.0241	4.83*
b_5	0.0012	0.52			
$\sum_{i=0}^{5} b_i = 0.2784$					
$R_r^2 = 21.71\%$			$R_x^2 = 8.75\%$		
$\sigma_1^2 = 0.2474$			$\Lambda = 0.8392$		
$R_\omega^2 = 44.75\%$					

Table A.2.11′
Hypothesis tests for estimated Hasbrouck's VAR model for PRU

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0,$ $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0,$ $i = 1, 2, ..., 5$
Wald test	15388	5105.3	446.7

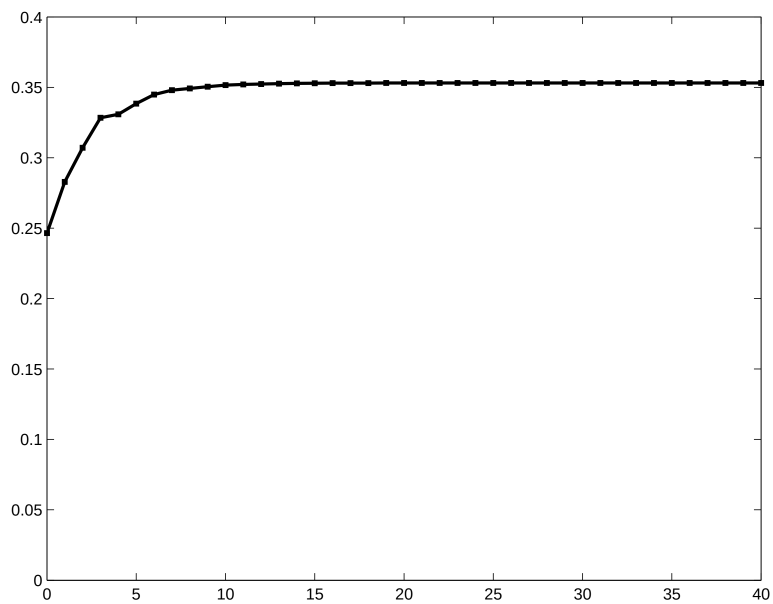

Figure 5.12: Price impact function for PRU
The total impact is 0.3531 of average proportional spread. It is fully realized after 20 transactions.

Table A.2.12
Estimation of the Hasbrouck's VAR model for RB

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for RB based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

RB					
return equation			trade equation		
		$t - value$			$t - value$
a_1	-0.0492	-11.27*	c_1	-0.4174	-37.00*
a_2	0.0258	5.89*	c_2	0.005	0.54
a_3	0.0097	2.22*	c_3	0.0113	1.27
a_4	0.0045	1.03	c_4	-0.0045	-0.46
a_5	0.0066	1.5	c_5	-0.0139	-1.56
b_0	0.2705	123.53*	d_1	0.2589	49.46*
b_1	0.0331	13.02*	d_2	0.0835	16.56*
b_2	0.0066	2.57*	d_3	0.0388	7.70*
b_3	-0.0008	-0.32	d_4	0.0344	6.67*
b_4	-0.0073	-2.85*	d_5	0.0294	5.95*
b_5	-0.0078	-3.13*			
$\sum_{i=0}^{5} b_i = 0.2942$					
$R_r^2 = 24.72\%$			$R_x^2 = 8.13\%$		
$\sigma_1^2 = 0.2241$			$\Lambda = 0.8540$		
$R_\omega^2 = 51.71\%$					

Table A.2.12′
Hypothesis tests for estimated Hasbrouck's VAR model for RB

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0,$ $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0,$ $i = 1, 2, ..., 5$
Wald test	17482	4114.3	1422.1

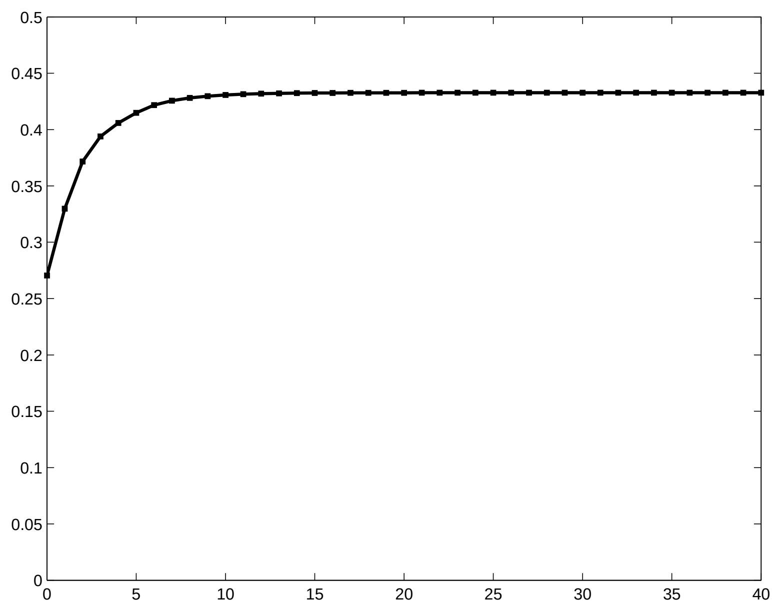

Figure 5.13: Price impact function for RB
The total impact is 0.4327 of average proportional spread. It is fully realized after 20 transactions.

Table A.2.13
Estimation of the Hasbrouck's VAR model for RIO

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for RIO based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

RIO					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.095	-32.80*	c_1	-0.4303	-64.09*
a_2	-0.0045	-1.55	c_2	-0.0185	-3.13*
a_3	0.0184	6.32*	c_3	0.0211	3.60*
a_4	0.0233	8.01*	c_4	0.0226	3.93*
a_5	0.0148	5.11*	c_5	0.0189	3.38*
b_0	0.2921	193.56*	d_1	0.2187	62.49*
b_1	0.0343	19.60*	d_2	0.0706	21.11*
b_2	0.0128	7.31*	d_3	0.0385	11.52*
b_3	0.0047	2.67*	d_4	0.0315	9.44*
b_4	-0.003	-1.7	d_5	0.0216	6.60*
b_5	-0.0029	-1.68			
$\sum_{i=0}^{5} b_i = 0.3381$					
$R_r^2 = 26.48\%$			$R_x^2 = 6.86\%$		
$\sigma_1^2 = 0.2507$			$\Lambda = 0.8780$		
$R_\omega^2 = 51.09\%$					

Table A.2.13′
Hypothesis tests for estimated Hasbrouck's VAR model for RIO

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$i = 0$, $c_i = 1, 2, ..., 5$
Wald test	41816	10600	4152.7

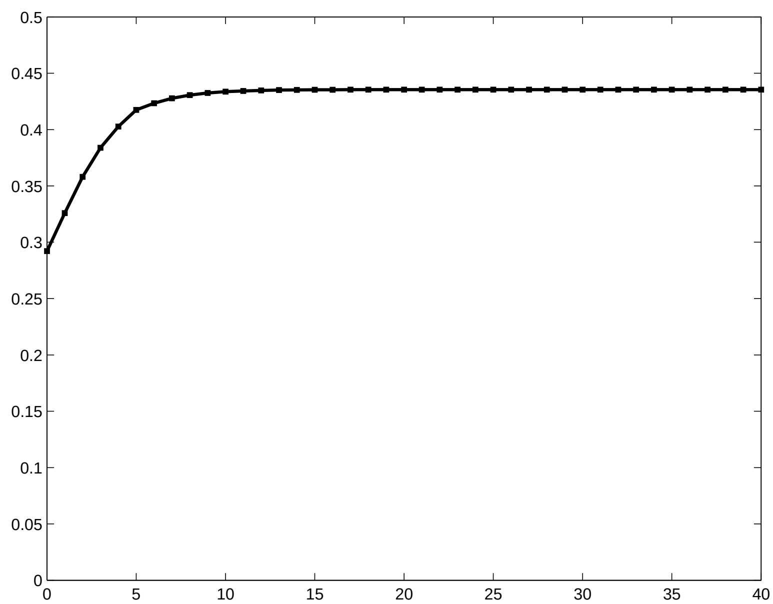

Figure 5.14: Price impact function for RIO
The total impact is 0.4355 of average proportional spread. It is fully realized after 20 transactions.

Table A.2.14
Estimation of the Hasbrouck's VAR model for SHP

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for SHP based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

SHP					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.1137	-22.89*	c_1	-0.2823	-8.00*
a_2	-0.0743	-14.87*	c_2	-0.0364	-3.39*
a_3	-0.0048	-0.967	c_3	-0.0363	-1.59
a_4	-0.0197	-3.949*	c_4	-0.0194	-2.17*
a_5	-0.0005	-0.1	c_5	-0.0332	-4.14*
b_0	0.2649	86.18*	d_1	0.2004	19.22*
b_1	0.0345	10.17*	d_2	0.0846	13.70*
b_2	0.0244	7.18*	d_3	0.0617	7.92*
b_3	0.0094	2.77*	d_4	0.0365	6.55*
b_4	0.0057	1.67	d_5	0.037	6.80*
b_5	-0.0033	-0.97			
$\sum_{i=0}^{5} b_i = 0.3357$					
$R_r^2 = 18.56\%$			$R_x^2 = 6.43\%$		
$\sigma_1^2 = 0.3437$			$\Lambda = 0.8703$		
$R_\omega^2 = 45.91\%$					

Table A.2.14′
Hypothesis tests for estimated Hasbrouck's VAR model for SHP

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$i = 0$, $c_i = 1, 2, ..., 5$
Wald test	8682.8	3170.9	99.8

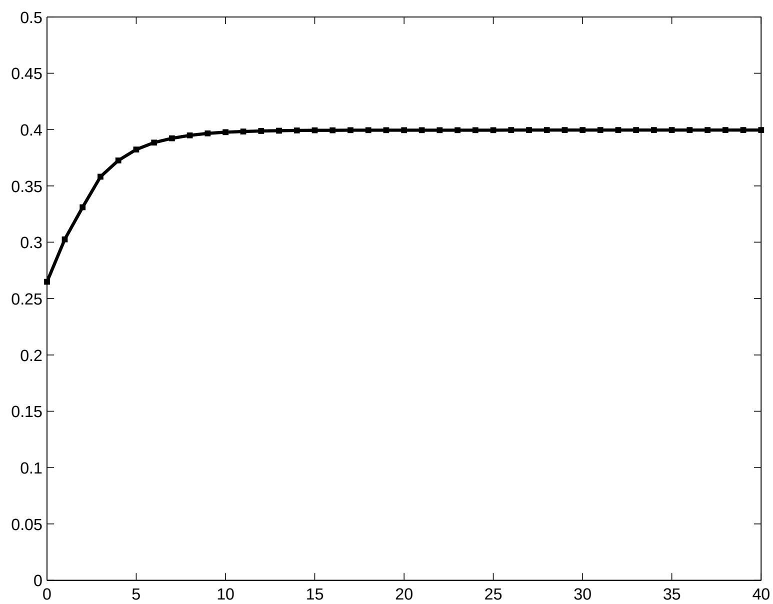

Figure 5.15: Price impact function for SHP
The total impact is 0.3995 of average proportional spread. It is fully realized after 23 transactions.

Table A.2.15
Estimation of the Hasbrouck's VAR model for SLOU

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for SLOU based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

SLOU					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.1294	-19.36*	c_1	-0.2831	-12.05*
a_2	-0.0919	-13.63*	c_2	-0.045	-3.47*
a_3	-0.0472	-6.98*	c_3	-0.037	-2.33*
a_4	-0.0181	-2.69*	c_4	-0.0336	-3.08*
a_5	-0.0291	-4.36*	c_5	-0.0076	-0.71
b_0	0.2376	53.85*	d_1	0.2086	24.28*
b_1	0.0386	8.09*	d_2	0.0662	8.76*
b_2	0.0214	4.48*	d_3	0.0535	6.92*
b_3	0.0162	3.38*	d_4	0.0295	4.09*
b_4	-0.0015	-0.31	d_5	0.0246	3.46*
b_5	0.0018	0.38			
$\sum_{i=0}^{5} b_i = 0.3141$					
$R_r^2 = 15.19\%$			$R_x^2 = 6.65\%$		
$\sigma_1^2 = 0.3995$			$\Lambda = 0.8857$		
$R_\omega^2 = 36.44\%$					

Table A.2.15′
Hypothesis tests for estimated Hasbrouck's VAR model for SLOU

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0,$ $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$i = 0,$ $c_i = 1, 2, ..., 5$
Wald test	3483.8	1555.3	169.1

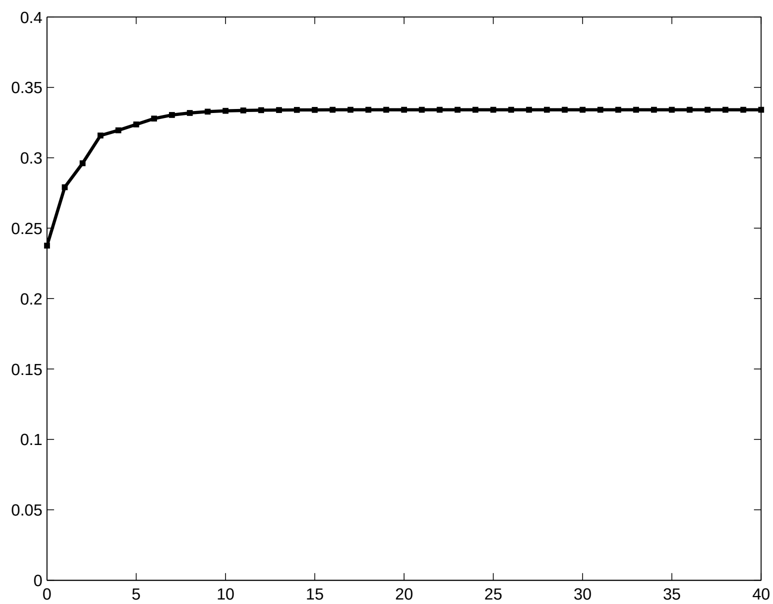

Figure 5.16: Price impact function for SLOU
The total impact is 0.3341 of average proportional spread. It is fully realized after 20 transactions.

Table A.2.16
Estimation of the Hasbrouck's VAR model for VOD

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for VOD based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

VOD					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.1301	-36.26*	c_1	-0.9751	-65.47*
a_2	-0.0598	-16.48*	c_2	-0.3274	-26.85*
a_3	-0.0368	-10.12*	c_3	-0.1958	-16.80*
a_4	-0.0034	-0.94	c_4	-0.1168	-9.52*
a_5	-0.0048	-1.35	c_5	-0.064	-4.71*
b_0	0.0775	71.33*	d_1	0.2816	74.40*
b_1	0.0313	27.03*	d_2	0.0863	22.54*
b_2	0.0122	10.46*	d_3	0.0772	20.29*
b_3	0.0071	6.09*	d_4	0.0532	14.01*
b_4	-0.001	-0.84	d_5	0.0574	15.39*
b_5	-0.0057	-5.01			
$\sum_{i=0}^{5} b_i = 0.1214$					
$R_r^2 = 11.69\%$			$R_x^2 = 16.22\%$		
$\sigma_1^2 = 0.0821$			$\Lambda = 0.8155$		
$R_\omega^2 = 42.60\%$					

Table A.2.16′
Hypothesis tests for estimated Hasbrouck's VAR model for VOD

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	89148	10432	4578.9

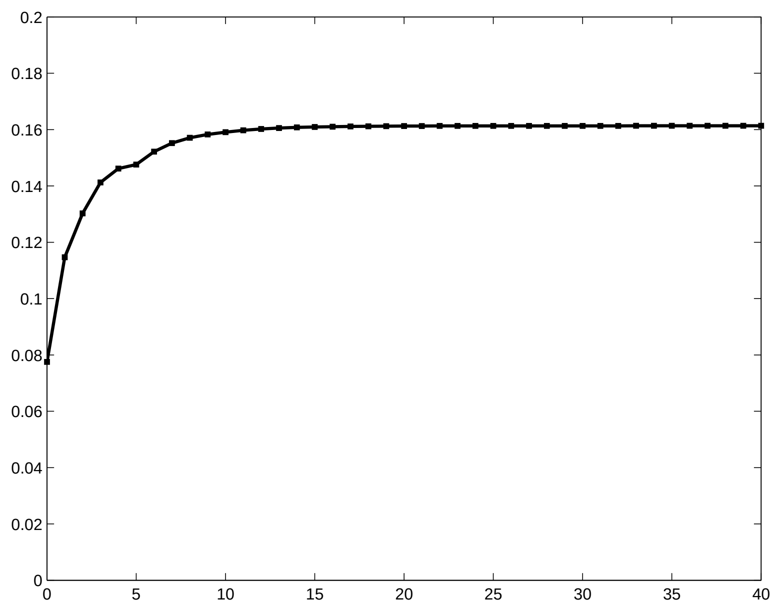

Figure 5.17: Price impact function for VOD
The total impact is 0.1613 of average proportional spread. It is fully realized after 21 transactions.

Table A.2.17
Estimation of the Hasbrouck's VAR model for WPP

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for WPP based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

WPP					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.0698	-13.68*	c_1	-0.563	-33.22*
a_2	0.0009	0.17	c_2	-0.0267	-2.01*
a_3	0.0214	4.18*	c_3	0.0102	0.79
a_4	0.0153	2.99*	c_4	0.0177	1.54
a_5	0.0021	0.4	c_5	-0.0055	-0.49
b_0	0.223	99.64*	d_1	0.2835	43.10*
b_1	0.0245	9.51*	d_2	0.0799	13.65*
b_2	0.0106	4.09*	d_3	0.0446	7.56*
b_3	-0.0076	-2.94*	d_4	0.0298	5.19*
b_4	-0.0039	-1.5	d_5	0.0232	4.15*
b_5	-0.0094	-3.75*			
$\sum_{i=0}^{5} b_i = 0.2371$					
$R_r^2 = 22.98\%$			$R_x^2 = 10.08\%$		
$\sigma_1^2 = 0.1677$			$\Lambda = 0.8189$		
$R_\omega^2 = 48.60\%$					

Table A.2.17′
Hypothesis tests for estimated Hasbrouck's VAR model for WPP

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	11576	2971.3	1238.7

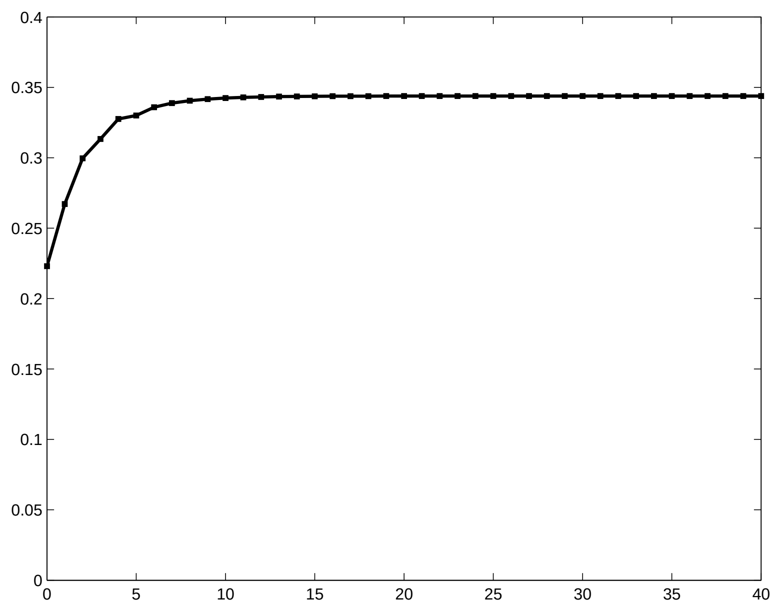

Figure 5.18: Price impact function for WPP
The total impact is 0.3438 of average proportional spread. It is fully realized after 19 transactions.

Table A.2.18
Estimation of the Hasbrouck's VAR model for XTA

The coefficient estimations and t-statistics for bivariate vector autoregression system of return and trade equation for XTA based on trade-by-trade data from March 1, 2006 to May 31, 2006. The t-statistics of coefficients of trade equation are computed using White's heteroskedasticity consistent covariance estimator. The equations are

$$r_t = a_1 r_{t-1} + ... + a_5 r_{t-5} + b_0 x_t^0 + b_1 x_{t-1}^0 + ... + b_5 x_{t-5}^0 + \nu_{1,t}$$

$$x_t = c_1 r_{t-1} + ... + c_5 r_{t-5} + d_1 x_{t-1}^0 + ... + d_5 x_{t-5}^0 + \nu_{2,t},$$

r_t is the midquote return of current trade at time t scaled by average proportional spread; x_t^0 is a trade indicator variable (+1 for buy order, -1 for sell order and 0 for undeterminate); R_r^2 and R_x^2 are the coefficients of multiple determination for return and trade equation by order; σ_1^2 and Λ are the variances of innovation terms for return and trade equation by order; R_ω^2 is the variance decomposition coefficient.

XTA					
return equation			trade equation		
		$t-value$			$t-value$
a_1	-0.1015	-28.96*	c_1	-0.4282	-56.08*
a_2	-0.0075	-2.14*	c_2	-0.0326	-4.78*
a_3	0.0132	3.76*	c_3	0.003	0.44
a_4	0.0014	0.39	c_4	-0.0052	-0.79
a_5	0.0086	2.46*	c_5	-0.0051	-0.77
b_0	0.2888	154.72*	d_1	0.2521	61.18*
b_1	0.0395	18.28*	d_2	0.0796	19.78*
b_2	0.0136	6.26*	d_3	0.0457	11.39*
b_3	0.0039	1.79	d_4	0.0301	7.54*
b_4	-0.0025	-1.16	d_5	0.0262	6.62*
b_5	-0.0036	-1.7			
$\sum_{i=0}^{5} b_i = 0.3397$					
$R_r^2 = 25.66\%$			$R_x^2 = 8.06\%$		
$\sigma_1^2 = 0.2631$			$\Lambda = 0.8790$		
$R_\omega^2 = 54.58\%$					

Table A.2.18′
Hypothesis tests for estimated Hasbrouck's VAR model for XTA

The Wald test of three hypothesis of estimated coefficients. The Wald statistics of trade equation coefficients are computed using White's heteroskedasticity consistent covariance estimator.

H_0 :	$b_i = 0$, $i = 0, 1, ..., 5$	$\sum_{i=0}^{5} b_i = 0$	$c_i = 0$, $i = 1, 2, ..., 5$
Wald test	27776	7597.9	3190.6

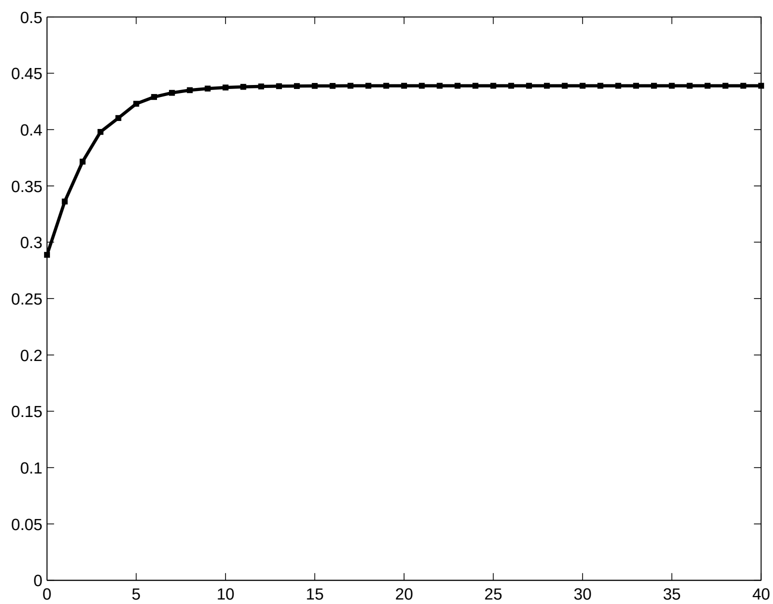

Figure 5.19: Price impact function for XTA
The total impact is 0.4389 of average proportional spread. It is fully realized after 19 transactions.

Biography

I was born on 27 November, 1970 in Šabac. I finished elementary school in Štitar and secondary school in Šabac with the highest degrees. In 1997 I graduated at the University of Novi Sad, Department of Mathematics and Informatics, as the Major Bachelor of Sciences, (Graduate Mathematician), average degree 8.58. I enrolled the master studies - Mathematical Analysis in 2002. I have been attending the master studies Applied Mathematics - Economath since 2005. I passed all exams on Economath with average degree 10.00 until the January, 2007. During the period 1997-2005 I worked as a teacher of mathematics in the Secondary School of Economics "Svetozar Miletić" in Novi Sad. Since 2003 I have been working as a teaching assistant, scientific field of Mathematics and Statistics, Braca Karic University, Faculty of Entrepreneurial Management in Novi Sad.

I participated at the conference in *Generalized functions,* Novi Sad, 2004. I passed the *DAAD intensive course on Multivariate Splines, Wavelet Analysis and Conservation Laws*, 2004, Novi Sad and the course in *Mathematics and Statistics in Management and Industry*, 2005, Novi Sad. Also, I participate at the summer school *Numerical Optimisation and its Application, DAAD Project*, 2006, Novi Sad. I am an author of two papers which were represented at PRIM conference in Kragujevac, 2006.

Novi Sad, April 10, 2008 Nataša Teodorović

UNIVERZITET U NOVOM SADU
PRIRODNO-MATEMATIČKI FAKULTET
KLJUČNA DOKUMENTACIJSKA INFORMACIJA

Redni broj:

RBR

Identifikacioni broj:

IBR

Tip dokumentacije: Monografska dokumentacija

TD

Tip zapisa: Tekstualni štampani materijal

TZ

Vrsta rada: Magistarska teza

VR

Autor: Nataša Teodorović

AU

Mentor: prof. dr Nataša Krejić

MN

Naslov rada: Matematički modeli likvidnosti i funkcije impakta cene

NR

Jezik publikacije: Srpski (latinica)

JP

Jezik izvoda: Srpski/Engleski

JI

Zemlja publikovanja: Srbija

ZP

Uže geografsko područje: Vojvodina

UGP

Godina: 2008.

GO

Izdavač: Autorski reprint

IZ

Mesto i adresa: Novi Sad, Prirodno-matematički fakultet, Trg Dositeja Obradovića 4

MA

Fizički opis rada:(4/80/37/21/27/0/0)
(broj poglavlja/strana/lit. citata/tabela/slika/grafika/priloga)

FO

Naučna oblast: Matematika

NO

Naučna disciplina: Primenjena matematika u ekonomiji

ND

Ključne reči: mere likvidnosti, funkcija impakta cene, vektor autoregresivni procesi, procesi pokretnih sredina, impuls odziva, dekompozicija varijanse, asimetrično informisano tržište, javne informacije, privatne informacije, Hasbrouckov VAR sistem jednačina prinosa i trgovanja.

PO

UDK:

Čuva se: Biblioteka Departmana za matematiku i informatiku, Novi Sad

ČU

Važna napomena:

VN

Izvod: Magistarska teza je posvećena analizi različitih mera likvidnosti i funkcije impakta cene. U periodu od 1.3.2006 do 31.5.2006, atributi trgovanja bitni za analizu likvidnosti i funkcije impakta cene su posmatrani na podacima trgovanje-po-trgovanje za grupu od 18 UK akcija indeksa FTSE 100. Polazeći od teorije asimetrično informisanog tržišta, funkcija impakta cene je dobijena iz Hasbrouckovog vektor autoregresivnog sistema jednačina trgovanja i prinosa, kao kumulanta impuls odziva inovacije trgovanja na prinos. Ovakav impakt cene je trajno ugradjen u buduće cene i u potpunosti se realizuje kroz nekoliko narednih transakcija. Učešće inovacije trgovanja na varijabilnost efikasne cene je mereno tehnikom dekompozicije varijanse. Iznete su specifičnosti dobijenih rezultata koji ukazuju na karakteristike londonske berze.

IZ

Datum prihvatanja teme od NN veća:

DP

Datum odbrane:

DO

Članovi komisije:

KO

Predsednik:

Član:

Član:

Član:

UNIVERSITY OF NOVI SAD
FACULTY OF SCIENCES AND MATHEMATICS
KEY WORDS DOCUMENTATION

Accession number:

ANO

Identification number:

ino

Document type: Monograph type

DT

Type of record: Printed text

TR

Contents Code: Master thesis

CC

Author: Nataša Teodorović

AU

Mentor: PhD, Full Professor Nataša Krejić

MN

Title: Mathematical models of liquidity and price impact function

XI

Language of text: English

LT

Language of abstract: Serbian/English

LA

Country of publication: Serbia

CP

Locality of publication: Vojvodina

LP

Publication year: 2008.

PY

Publisher: Author's reprint

PU

Publ. place: Novi Sad, Faculty of Science and Mathematics, Dositeja Obradovica 4

PP

Physical description:(4/80/37/21/27/0/0)
(chapters/pages/literature/tables/pictures/graphics/appendices)

PD

Scientific field: Mathematics

SF

Scientific discipline: Applied Mathematics in Economy

SD

Key words: liquidity measures, price impact function, vector autoregression, moving average, impuls response, variance decomposition, asymmetrically informed market, public information, private information, Hasbrouck's VAR vector autoregression system of return and trade equations.

UC:

Holding data: The library of the Department of Mathematics and Computer Science, Novi Sad

HD

Note:

N

Abstract: This master thesis is devoted to the analysis of different liquidity measures and price impact function. Over a period of 62 days, from March 1, 2006 to May 31, 2006 the trading attributes of interest for the liquidity analysis are viewed trade-by-trade for a group of eighteen UK stocks listed on FTSE 100 index. Based on asymmetrically informed market theory, the price impact function is obtained from Hasbrouck's vector autoregressive system of return and trade equation, as a cumulative impulse response of trade innovation to the return. This price impact is permanently impound in the future prices and it takes several transactions until it is fully realized. The contribution of trade innovation in variation of an efficient price is measured by decomposition variance technique. The specifications of obtained results are understood through the characteristics of London Stock Exchange.

AB

Accepted by the Scientific Board: 12.07.2006.

Defended:

Thesis defend board:

President:

Member:

Member:

Member: