“Review: algorithmic trading”

AUTHORS
Jan Fraenkle
Svetlozar T. Rachev

ARTICLE INFO

JOURNAL
"Investment Management and Financial Innovations"

FOUNDER
LLC “Consulting Publishing Company “Business Perspectives”

© The author(s) 2019. This publication is an open access article.
Jan Fraenkle (Germany), Svetlozar T. Rachev (Germany)

Review: algorithmic trading

Abstract

Automated trading systems play an increasingly important role in equity markets. The challenge of trading large order volumes and baskets is often met today with automated trading algorithms. An investment decision generally leads to orders with specifications such as security, size, and urgency but also specifications (e.g., expected profit), and constraints (e.g., dollar neutrality) may be given. The trader has to choose a trading strategy ensuring best execution under given marginal constraints. This review provides an overview of how such trading algorithms work. The ideas behind some standard strategies are presented, as well as approaches to enhance them. For developing automated trading strategies for stock markets, a deep understanding in market microstructure is necessary, so we review this topic as well. We have a look on the issue how market quality is affected by market designs of trading platforms and fragmentation of the market. Trading costs are the main attribute of market quality. Trading strategies are implemented to optimize trading costs and execution risk by taking market microstructure aspects into account. Empirical analyses of trading volume and order book characteristics help to adjust the trading strategies.

Keywords: market microstructure, trading costs, algorithmic trading.

JEL Classification: G10, G12.

Introduction

Algorithmic trading is one possible connection between a market participant and the market. Algorithmic trading systems are generally used to ensure a smooth interaction between these two parties in the sense that orders of the trader are tried to be executed with minimal market influence. Imagine a trading decision done by portfolio management which should be executed at favorable prices. The larger the order proportional to the provided liquidity, the more challenging the execution without large trading costs including price impact, broker commissions and exchange fees. The most interesting part of trading costs from a researcher’s point of view is the implicit trading costs associated with the complexity of the topic. Each order in the market has an influence on the security price also when it is not executed. The magnitude of the influence mainly depends on the order specifications relative to the market, i.e. a large order in a less liquid market leads to high market impact as well as fast order execution.

Trading algorithms generally try to reduce the market impact of a trade by splitting large orders into several smaller slices or execute the order at moments where favorable prices can be realized. These sub-orders are sent in general to markets over a period of time and to one or several execution venues.

Potential market places where orders can be routed are exchanges, electronic communication networks (ECNs), dark pools or internal matching of broker houses. The stock market landscape has changed dramatically in recent years with several new electronic trading platforms introduced that compete with the established exchanges. The price impact results from the presence and interaction of the order with the market. In this paper we describe the interaction as friction, trading costs, and observable changes in the order book.

The paper proceeds as follows. Section 1 provides an introduction to market microstructure. In section 2 the idea of trading strategies based on special benchmarks is presented. An introduction into empirical analysis of some important microstructure variables such as trading volume and order book content is covered in Section 3. Section 4 provides an overview of the aspects of algorithmic trading.

1. Market microstructure

Several definitions of market microstructure have been suggested in the literature. Two of the more notable ones are provided by O’Hara (1995) and Stoll (2001). O’Hara defines market microstructure as “the study of the process and outcomes of exchanging assets under explicit trading rules”. Stoll defines market microstructure as the study of trading costs and the impact costs resulting in the short-run behavior of security prices. As we will show, both definitions are very similar in their meaning. Moreover, we will explain why trading costs are a basic element in market microstructure. This section introduces market microstructure theory and gives a short overview of the issue and literature.

A general overview and introduction in market microstructure theory is given by O’Hara (1995).
Besides an introduction to price determination, inventory models of market makers are presented as well as theory behind bid-ask spreads. The author identifies the present influence of trading strategies of market microstructure and the information of trades in price process. An examination of market design and market performance, including liquidity and market transparency is provided in our review. Harris (2002) provides a more practical view on market microstructure, explaining the background for some key elements of market microstructure as well as the investment objectives and activities of different market participants. Harris also presents a review of trading platforms and the role they play. Cohen et al. (1986) provide a detailed cross-sectional comparison of worldwide equity markets. Stoll (2001) focuses on trading costs, describing market designs and the forces leading to centralization of trading in a single market versus the forces leading to multiple markets. Madhavan (2000) provides a review of theoretical, empirical, and experimental literature on market microstructure with a focus on informational issues.

1.1. Nature of market. One of the principal functions of financial markets is bringing together the parties interested in trading in a security. Trading platforms are the most efficient way to bring counterparts together. Such trading platforms can be accomplished via the physical presence of brokers and traders trading on the floor of an exchange. But it can also be realized as an electronic platform where the physical location is unimportant and market participants are just connected electronically. A hybrid market wherein there is both a trading floor and an electronic platform is a third alternative. The best example of a hybrid market is the New York Stock Exchange (NYSE).

The trading process itself is similar for all financial markets. All market participants express their trading interest with an order which is sent to the market. An order contains the information regarding which security to trade, the direction (buy or sell), the quantity of shares, and a limit price expressing the worst price the party is willing to accept. When the limit price is not identified as part of the order, this results in a so-called market order in which the party to a trade is willing to accept all prices.

The task of financial markets is to match compatible orders and execute them. Most of the markets define their trading rules to enable high liquidity and fast execution with low price volatility. A very basic idea for the trading process is the Walrasian auctioneer. Each market agent provides a demand-price function to the auctioneer who first aggregates these orders and then computes a price where demand and supply are equal (i.e., the market-clearing price).

Walrasian auctions are discrete auctions; that is, trading takes places only at specified times during the trading day. Modern exchanges provide continuous trading, and therefore market participants have the opportunity to trade at any time during the trading day. But for each trading interest a counterparty has to be found, willing to trade the same position in the contrarian direction. In the limiting case of iterating Walrasian auctions with infinite frequency, continuous trading would be realized but the probability of executing a trader’s order would be equal to zero. The probability of two orders reaching the same auction declines with the increase in the frequency of auctions if the trader’s order is valid for exactly one auction. So there is a need for orders which are valid for more than one auction. Such orders do not satisfy investor’s needs to be executed immediately, but their existence enables immediate execution of other orders. So besides market participants preferring immediate execution, market participants providing liquidity are needed. Traditionally, market participants providing continuous liquidity are market makers. Their profit arises from the existence of the implicit premium that the party to the trade who seeks liquidity is willing to pay. The premium increases with the volume of the trade and reflects the expected risk the market maker incurs.

1.2. Continuous trading and open limit order book. Most stock markets provide continuous trading. Some markets have additional discrete call auctions at specified times when uncertainty is large such as at the open, close, and reopen after a trading halt caused by large price movements. The economic justification is that call auctions are especially helpful in uncertain times during the trading day because of the information aggregation argument (see Madhavan, 2000).

Open limit order books are the core of most continuous trading systems. A limit order book contains limit orders of market participants, including the information about the limit price, quantity of shares, and trading direction (buy or sell). The content of open order books is provided to market participants in contrast to closed order books where no information about the status of the market is published, it is realized in so-called “dark pools”. The most relevant measure of order books is the bid-ask spread. It is the difference between the lowest provided sell price (ask) and the highest buy price (bid), where the ask is always higher than the bid. The bid-ask spread is a good measure for the liquidity of a security, i.e. in actively traded securities the spread is smaller than in inactive markets. Implicit trading costs arise in continuous trading through the existence of the spread. Liquidity takers have to cross the spread for
trading which is the premium for liquidity provision. This premium is justified by the risks and costs the liquidity provider faces, such as inventory risk and order handling costs. On the other hand, competition between liquidity providers forces the market in the direction to lower spreads. Some theoretical studies concerning liquidity provision are provided by Biais et al. (1995), Biais et al. (1999), Harris and Hasbrouck (1996), and Foucault (1998).

Trading takes place when an order arrives the order book matching at least one position, i.e. the limit price of the incoming buy (sell) order is equal or higher (lower) than the ask (bid) of the current order book. Otherwise the order is inserted into the book and provides hence the best bid or ask. The execution price of a trade is always the limit price of the order book position which is involved. This leads to jumps in security prices from bid to ask prices depending on the direction the initiator trades (see Garman (1976) and Madhavan et al. (1997) for models describing time series behavior of prices and quotes).

Each limit order book position is determined by limit price and provided volume. For the best bid and ask positions, the volume is quite small compared to the entire order book volume and also small compared to typical order sizes of institutional investors. Submitting a large order to continuous trading systems leads to a sharp price movement and a rebuilding of the order book afterwards, resulting in huge implicit trading costs because of a large realized bid-ask spread. So optimal trading in continuous trading systems requires adapted strategies where large orders are split into several smaller orders which are traded over a period of time. In the time between execution of the slices, the order book can reclaim in the sense that liquidity providers narrow the spread after it has widened through a trade (see Obizhaeva and Wang (2005)).

1.3. Trading costs. With its presence in a market, a buy or sell intention has an influence on the future price process. Perold (1988) introduces the implementation shortfall which is the performance difference between the paper portfolio and the realized one. Implementation of investment strategies leads to friction losses. This difference in performance is dominated by three blocks of costs. One consists of fees and commissions for brokers and exchanges, the second part is market impact costs, and the third is opportunity costs.

Market impact costs arise from the information acquisition and liquidity taking of orders. It is mainly a function of the aggressivity of the trade, liquidity of the security and the amount of ordered shares. Market impact increases when trading large volumes in a short time span. On the other hand, opportunity costs arise when less volume than originally wanted is traded or a longer period of time is needed because of the loss of profit and volatility risks. An investor has to find the trade-off leading to optimal costs (see Kissell (2006) and Wagner and Edwards (1993) for further introduction in different kinds of trading costs).

Market impact is a quite interesting part because of its complexity reflecting the interaction between one market participant and the rest of the market. Opportunity costs are investor specific and after understanding market impact, one can try to find trade-off between the two. Market impact is the influence of trading activities on the market, i.e. the realized price for a security is worse than the security price before the beginning of the trading activity of the investor. A possibility of measuring market impact is to calculate the difference of realized average execution price and security price before trading activity has begun (arrival price). The reasons for market impact are, as already mentioned, information acquisition and demand for liquidity. If an informed trader is willing to buy a security expecting a higher price in the future, he is also willing to pay a higher price than the current one with the constraint that the price has to be lower than the future expected price. The investor’s information is anticipated by the market resulting in market impact. The liquidity demanding component of market impact arises from the risk and costs the trading counterpart is faced with (see section 1.2). These effects differ in the sustainability of their impact, and while the information component is a permanent effect, the liquidity component is just a temporary effect. Further description of market impact and the differentiation of temporary and permanent impact can be found in Kissell (2006), Kissell and Malamut (2005), Madhavan (2000) and Almgren and Chriss (1999).

1.4. Market design. In this passage, how a market should be designed to provide an attractive environment for traders, i.e. how trading rules should be defined resulting in good market quality is discussed. It is obvious that the design of the market determines the market microstructure. The microstructure influences investing strategies, patterns of trade, liquidity, and volatility. Therefore, exchanges have to find their setup to attract traders. There are several studies in literature describing the impact of market design on the market characteristics. Levecq and Weber (2002) and Stoll (2001) give a general overview of different possibilities how a market can be organized. Levecq and Weber (2002), Levecq and Weber (1995) and Barclay et al. (2001) have focused on information technology and electronic systems in financial markets.
To evaluate the quality of trading at a special exchange, metrics for market quality should first be defined. Madhavan (2000) mentions spreads, liquidity, and volatility. Others like Boehmer (2005) add availability and execution speed to the list of quality measures. Availability expresses the reliability of the exchange. Execution speed is the time an investor needs to get a trading decision executed (within the trading hours) depending on the size of the order. Also the reaction time is an important quality measure for some special traders who are interested in ultra-high frequency trading, as it is described by Byrne (2007).

Market structure choices are elementary for exchanges to offer a market conforming the investor’s needs in a competitive environment.

1.4.1. Market architecture. Market architecture refers to the set of rules governing the trading process (Madhavan, 2000). These rules cover the market type including degree of continuity, choice between order-driven and quote-driven markets as well as the degree of automation. Most stock markets have continuous systems combined with discrete auctions at special times when uncertainty is high. Most stock markets are organized as a mixture of order- and quote-driven markets. That means that every market participant can provide prices via limit order, can trade with other traders directly, and additionally there are market makers providing quotes that take the opposite side of traders. Another aspect in market architecture is price discovery, e.g., is the price independently discovered or is the price from another market used. Another very important aspect is the transparency. Most stock markets provide pre-trade and post-trade information such as quotes and related volumes, as well as times and sales. This information can be used by an investor as a basis for trading decisions and trading optimization. Some special markets, called dark pools, do not provide any market information except trading confirmations for directly involved trading parties. It is assumed that trading has less price impact when the order information is not published because other market participants cannot react on the presence of an invisible order.

More aspects and their detailed information concerning market architecture can be found in Madhavan (2000). Levecq and Weber (2002) focus on aspects of the market architecture of electronic trading systems.

Electronic trading systems have their origin in the 1960s and 1970s with NASDAQ and Instinet. They have experienced strong growth until today and dominate stock trading today. Two parallel evolutions occurred concerning electronic markets; there are the traditional markets such as NYSE which use electronic trading systems to support their traditional trading system. Automation helps to improve efficiency because it lowers trading costs and satisfies more the investor’s needs. It is necessary in an increasingly competitive environment. With the spread of electronic networks in the finance industry, a new type of market has arisen called the ECN (electronic communication network). These trading platforms concentrate only on electronic trading mainly in liquid securities such as stocks and currencies. They provide very fast trading systems with low fees. For institutional investors it is easy, and inexpensive to connect and market data are real-time and often available for free. ECNs are established for years in the US and cover a significant fraction of NASDAQ trading volume. In Europe ECNs are quite new but very successful with a fast growth in trading volumes. Some important examples are Chi-X, BATS, and Turquoise. They have similar trading tariffs working in the way that you have to pay a fee for aggressive execution and get money for passive execution. With this trading tariff concept ECNs attract liquidity potentially from all market participants and therefore, do not need explicit market makers.

This market design is different from the design of traditional markets in raising fees, execution speed, and liquidity contribution. They also provide some other order types like pegged-limit order for example, where the limit complies with the price at the primary market to satisfy their role of secondary market. ECNs are in line with the traditional market in the price discovery process, i.e. continuous trading and open limit order book.

1.5. Fragmentation of market. Today we are faced with a widespread fragmentation of the stock market. Besides exchanges as primary markets, there are many ATSSs (alternative trading systems) playing an important role in the stock market. These are organizations, persons or systems that provide a market place for bringing together purchasers and sellers. Examples for ATSSs are ECNs, broker/dealer internal crossing and dark pools. An investor or trader can decide where to send the order with the expectation of being executed well. Investors have different requirements to exchanges and the various ATSSs try to provide optimal execution for their clients. Therefore, they focus on optimizing special aspects of trading characteristics. Some try to provide a market with very fast and continuous execution like most of the ECNs. Others do not provide any market data (dark pools).

The fragmentation of the order flow has increased in recent years because of the strong growth of ATSSs. The main question concerning fragmentation is the
overall market quality. There is some literature describing the effects of reducing liquidity by fragmentation of market (see Mendelson (1987), Chowdhry and Nanda (1991), Grossman (1992), Madhavan (1995) and Hendershot and Mendelson (2000). Bennett and Wei (2006) chose stocks which switched from listed on the NYSE to NASDAQ and vice versa. They measure the market quality before and after the switches and find that the NYSE has better market quality than the NASDAQ for illiquid stocks. NYSE is one market where NASDAQ is a pool of different ECNs and exchanges, while NASDAQ is, in itself, a fragmented market.

On the other hand, there are several effects leading to better market quality in a fragmented market. There is more competition between trading platforms leading to lower trading fees and more innovations and thus to more efficient execution (Levecq and Weber, 1995). Barclay et al. (2001) find that increased trading on ECNs improves most measures of overall market quality. As an explanation, they find that ECNs attract a higher fraction of informed orders reducing adverse selection costs faced by the market makers. This leads to lower spreads in competitive markets. Another effect of fragmentation is the lower level of trade disclosure. An investor trading large positions can benefit from this effect (see Madhavan (1995)). In a consolidated market the effect of “front run” their own order can also be much more significant.

As described above, there are opposing influences on market quality from fragmentation. Because of the interests of market participants to be well executed, there are forces in the direction of maximal market quality. Both extreme scenarios of a complete consolidation as well as a highly fragmented market are not optimal scenarios, because of the above reasons. There are several ways of linking and consolidating fragmented stock markets, for example, by regulation. One idea for quasi-consolidation is that every trade has to occur between the nationwide best bid and ask. If a marketplace does provide a worse price, the order has to be sent to another market with a better quote. On the other hand, there is much effort on the part of market participants to do pre-trade analysis to find out how to split the order and where to send it to have the best possible execution. These systems are called “smart order routing” and are provided from most brokerage firms. In recent years also, many startups arise with the business idea of doing arbitrage by high frequent trading on different markets.

These linkages of markets are a kind of consolidation with different impact on competition (see, for example, Blume (2007)).

2. Trading strategies

2.1. Benchmarks. Evaluating a trading strategy with regard to execution quality, benchmarks are usually compared with the realized values. Reasonable measures for execution quality are the executed fraction of order volume, execution price, and the execution price uncertainty where the execution price is most important. Thus, various definitions of benchmarks and reference prices are referred in literature which are used to compare with execution price. These benchmarks can be categorized into pre-, intra-, and post-trade prices (see Kissell, 2006). The most common benchmark price is VWAP (volume weighted average price) or TWAP (time weighted average price) over the trading horizon, being intra-trade prices. The arrival price (price of the security during the arrival of the order) is an example for a pre-trade price. An example for a post-trade benchmark is the day’s closing price. There are a variety of more benchmark definitions and also a spectrum of similar but slightly different definitions for each kind of benchmark (see Madhavan (2002) for various definitions for VWAP).

Different kinds of benchmarks have diverse characteristics so investors have to take care by choosing their benchmark with regard to their trading strategy and preferences. Pre-trade benchmarks are suitable for measuring market impact because they are not influenced by the price movement induced by their own trade. Measuring execution costs as part of the implementation shortfall, introduced by Perold (1988), has to be done by pre-trade benchmarks. Intra-trade benchmarks are a good indicator (Berkowitz et al., 1988) for the quality of the trading algorithm and market impact in the case of passive trading. If a market participant plays a dominant role on the market, the VWAP is heavily influenced by the trades of the dominant trader. In limiting case of a completely dominant trader, the VWAP is equal to the average execution price, but the market impact is very high anyway. VWAP as a benchmark has the advantage of representing reality better in the sense that the benchmark is calculated over a period of time like large orders, which are split and distributed over a given time span as it is usual in algorithmic trading. VWAP benchmarks contain the market movement in price inside the period when VWAP is calculated, whereas pre-trade benchmarks do not. The residuals of intra-trade benchmarks and execution price of a sample of trades generally have a significantly smaller width than residuals of pre-trade benchmarks and execution price. Post-trade benchmarks aren’t suitable for measuring market impact. But some investors, e.g., mutual fund managers may desire execution near closing price to
coincide with valuation of the fund and may con-
sider closing price as a reasonable benchmark (Kis-
sell, 2006).

Having a maximal objective view on execution
quality, several benchmarks should be taken into
account. Only one benchmark is not able to repre-
sent execution quality as a whole.

A basic concept behind all execution benchmarks is
the fact that trading is a zero sum game. The sum of
all market impact costs of all market participants is
zero which has to be considered by any measure of
market impact costs. Otherwise the benchmark is
biased and there are unexploited arbitrage opportu-
nities (see Berkowitz et al., 1988).

2.2. Trading strategies. The cost-efficient imple-
mentation of investment decisions is quite impor-
tant for successful realization of most investment
strategies. Depending on the frequency of realloca-
tion of the portfolio, trading costs can reduce per-
formance significantly. Especially large trading
volumes cannot be executed instantly and the trade
has to be split over a period of time. Trading
strategies are used to disperse the volume over
time. Because of the strong dependence of the exe-
cution quality from order volume, order types like
market or limit orders, and many other variations it
is challenging and provides opportunities to de-
velop an optimal execution strategy. Theoretical
knowledge about market dynamics and the de-
pendence of market impact from the trading traject-
ory is the basis of strategy development (see Obizhaeva and Wang, 2005).

Based on their different characteristics, Domowitz
and Yegerman (2005) describe the spectrum of trad-
ing strategies from unstructured, opportunistic li-
quidity search to highly structured, precisely sched-
uled sequences of trading activity, generally linked
to a certain benchmark. An example of a highly
structured trading algorithm is VWAP strategy
which is specified later in this chapter.

Unstructured strategies have a disadvantage that
they may generate either large trading costs or
large execution risks and tend to extremes. Be-
cause of marginal constraints of investors for exe-
cution more sophisticated strategies are necessary
to satisfy the investor’s needs better. The goal of
an enhanced trading strategy is done by achieving
a favorable execution and taking the marginal
constraints of the investor and market into ac-
count. The idea behind most of these strategies is
to define a benchmark and design a strategy trying
to beat, or at least, reach the benchmark with
preferably less systemic risk and volatility risk
with respect to that benchmark. Coggins et al.
(2006) give some introduction in algorithmic exe-
cution strategies; Obizhaeva and Wang (2005)
provide the possibility of optimal execution tak-
ing market dynamics into account.

2.2.1. Examples of algorithmic strategies. Some
examples of common execution strategies are pre-
sent in the following:

- **Arrival price** is the price of the security at the
  moment before the first order is sent. The basic
  idea of execution strategies with this bench-
  mark, also known as implementation shortfall
  (Perold, 1988), is to concentrate trading vol-
  ume at the beginning of the trade, thus near the
  arrival price to minimize volatility risk. Mini-
  mization of volatility risk leads to fast execu-
  tion and thus to high market impact, so every
  trader has to find his optimal point on the effi-
  cient frontier of execution, introduced by Alm-
gren and Chriss (1999).

- An enhanced strategy is the adaptive arrival
  price strategy of Almgren and Lorenz (2007)
  where execution speed is updated in response
to observed price motions leading to a more
realistic formulation of the mean-variance
tradeoff.

- **TWAP** trading strategy tries to beat the time
  weighted average price. Such a strategy divides
the trading period into equal sized time slots and
distributes the order volume equally over these
slots. The order volume in each time slot is gen-
erally given via limit order to the market becom-
ing more aggressive when the end of the time
slot approaches and may end in a market order
when execution is forced.

- **VWAP** trading strategy is very popular and is
  often used in the finance industry. The underly-
ing benchmark is the volume weighted average
  price (VWAP) of the security during a speci-
fied period including all trades. For some de-
tailed information and some variation of
  VWAP definitions, see Berkowitz et al. (1988)
  and Madhavan (2002). VWAP strategies work
similarly to the TWAP strategy. The given
time horizon where the trade ought to take
place is divided in n (equal sized) time slots and
every time slot gets allocated a special fraction
of entire trading volume. How large this fraction is depends on the historical trading
volume of the special security in this period of
time. Trading volume in equities is normally u-
shaped over the trading day, i.e. in the first and
in the last trading minutes, trading volume is
extremely large and the minimum is at about
noon. Within a time slot the algorithm may
send limit order to the market and wait for being executed to favorable prices. When the end of the time slot nears, limit may become more aggressive and finally a market order may be sent if the execution is forced. VWAP strategies have the advantage of opportunistic components which may lead to favorable prices. Because of the volume profile of trading volume taken into account, market characteristics are incorporated adequately and provide an appropriate basis for improvement of the plain vanilla VWAP strategy. Instead of using a static historical mean of trading volume, more sophisticated trading volume predictions may lead to better performance by raising the opportunistic component.

- **TVOL** (target volume) strategy is more opportunistic and trades a constant fraction of the actual overall trading volume in the security. Thus it is a modification of the VWAP strategy and only takes actual and not historic volume into account. There is no benchmark the strategy tries to beat. Before trading the volume and the duration of trading respectively are not known.

Examples for opportunistic trading algorithms cannot be easily mentioned because there is no industry standard. Using these algorithms is much more challenging because, they may provide lower execution costs, but the handling of the marginal constraints of the trade is more complicated or impossible.

By trading especially with schedule-driven algorithms one issue can play a significant role, if the algorithm always acts under special and clear rules. Other market participants may be able to observe special patterns and take advantage of leading to worse execution quality.


3. Empirical analysis

3.1. Trading volume. The traded volume of a security in a given period of time is a quite important measure for the liquidity of a security. The mean of trading volume depends heavily on the volume of free float stocks and thus on the market capitalization of the company. Temporary trading volume fluctuations can be influenced by strong interest in trading the security triggered by news, change in an index composition or market movements. There are also significant intra-day and inter-day seasonalities.

For VWAP trading algorithms the trading volume in future, i.e. trading volume in the trading period is of importance and has to be forecast. Therefore, empirical studies of trading volume are necessary. Static volume pattern as well as trading volume dynamic are studied in literature and provide a basis for competitive VWAP trading algorithms.

3.1.1. Seasonality of trading volume. Seasonality in trading volume of stocks is observed in different time spreads. The most significant is the intra-day volume u-shape pattern (see Lockwood and Linn, 1990). Very high trading volume is in the morning after markets open and in the evening before closing, the minimum occurring around lunchtime. An example is given in Madhavan (2002) and with the two intra-day volume pattern averaged over the year March 2007 until July 2008 (see Figure 1 and Figure 2).

![Fig. 1. Volume profile of the Vodafone stock traded at the LSE](image)
Most exchanges in Europe provide continuous trading and additional auctions at the beginning and at the end of the trading day. The closing auction is the more important one with regard to the volume in the auction. For example, Vodafone, one of the most liquid European stocks, has an average volume of about 84 Mio GBP in the closing auction and during the mentioned period.

Seasonality with a much greater frequency is described by Fishin’ (2007) where a lower stock turnover during the summer months is observed. It is pulled together with stock price returns which are lower in the summer months too.

3.1.2. Dynamics in trading volume. A look at the average stock turnover is quite useful and important but it doesn’t tell the whole truth. For describing the large variance in trading volume, several models are provided in literature. Volume is decomposed in an analog way as it is done for price returns. Lo and Wang (2000) suggest that stock turnover is well-approximated by a two factor model. Darlles and LeFol (2003) pick up the model and extends with justification of liquidity arbitrageurs and a screening tool that allows practitioners to extract information from volume time series. In the following the volume decomposition is introduced.

Principal component analysis (PCA) on stock turnover leads to data reduction and to factors whose dynamic characteristics are of interest. To identify the factors, significant correlation with known influences has to be shown empirically. PCA is used to describe variance-covariance matrix through a few linear combinations. In current example, the variance-covariance matrix of turnover series \( x_{it} \) has to be calculated, where \( x_{it} \) denotes the number of traded shares divided by the number of float shares per asset \( i \) and time \( t \). The spectral decomposition of the \( I \times I \) variance-covariance matrix leads to \( I \) orthogonal eigenvectors and eigenvalues. The turnover series decomposition can be written as:

\[
\frac{x_{it} - \bar{x}_i}{\sigma_i} = \sum_k u_i^k C_t^k ,
\]

where \( u_i^k \) is the \( i \)-th component of the \( k \)-th eigenvector and \( C_t^k = x_{it} u_k \) with \( \text{Cov}(C_t^k, C_t^l) = \lambda_k \sigma_{kl} \), where \( \lambda_k \) is the \( k \)-th eigenvalue. It can also be written in the form:

\[
x_{it} - \bar{x}_i = \frac{1}{\lambda_k} \text{Cov}(x_{it}, C_t^l) C_i^l + \sum_{k \neq l} \text{Cov}(x_{it}, C_t^k) C_i^k .
\]

The leading term is interpreted as the market turnover while the following terms are interpreted as short-term arbitrage activity in Darlles and LeFol (2003). Then the two components of the decomposition have different dynamic behavior. The first captures all the trend in turnover whereas the second should be stationary. Different interpretations of the decomposition are possible. Lo and Wang (2000) see the second component as a hedging strategy against risk of market condition modifications.

Bialkowski (2008) provides an approach for modeling dynamics in trading volume for improving VWAP trading strategies. It is an extension of Darlles and LeFol (2003) in the sense that the results of decomposing trading volume is used to improve VWAP trading strategies. To discriminate between the seasonal and dynamic part of stock turnover opens up the possibility of forecasting stock specific dynamics independently from medial seasonality. The static seasonal part of the model is given by a historical average of the common components of intra-day volume. The second component of the
model represents specific trading volume for each equity and is realized by the use of an ARMA(1,1) with white noise or alternatively by a SETAR. The application of the model leads to significant reduction of the execution risk in VWAP orders.

3.1.3. Intra-day patterns of market variables. An interesting question with regard to trading volume may be unanswered up to now. What are the reasons for the intra-day shape and how do correlations with other variables look like?

To answer these questions, two theoretical models and their empirical tests are presented. Admati and Pfleiderer (1988) provide a theory as to why concentrated trading patterns arise endogenously. It explains intra-day volatility patterns and correlation of trading volume and volatility as well as anti-correlation of volume and market depth. The model takes the behavior of some informed traders and liquidity traders into account. The predictions of the model arise from finding a Nash equilibrium of the trading game including the trader’s behavior based on their special preferences. A similar model is provided by Brock and Kleidon (1992). This model is based on the idea of portfolio re-balancing on the assumption that an optimal portfolio is a function of the ability to trade. A volume intra-day u-shape is predicted with this model as well as a correlation of bid-ask spreads and volume.

An empirical test of these theoretical hypotheses is given by Abhyankar et al. (1997). The results can be summarized as follows. Bid-ask spreads are larger near open and close which is in line with Brock and Kleidon (1992). For heavily traded stocks, a u-shape of trading volume pattern is found. For less traded stocks the volume pattern rises from open to mid-day, falling to the lowest level at lunch time and rises until the end of the trading day. Volatility is observed higher near open and close of the trading day which is in line with Admati and Pfleiderer (1988). Another empirical result is that average volume of stock traded per transaction is quite constant through the day with some rising at the beginning and end of the trading day.

3.1.4. Trading volume and market impact. A key question in trading large orders is the dependence of market impact and size of the trade. It is obviously a function of liquidity and overall trading volume in the security. Generally, when large orders are broken into smaller slices and distributed over a period of time, then the issue generalizes to the market impact behavior dependence on order size and timeframe of trading. For large funds it is of fundamental meaning having estimates for caused market impact when doing large portfolio re-balancing. But it is also a measure for the possibility of converting an inventory of a security into cash.

Empirical studies of market impact are only possible when trading information is available. Price movement observable in public market data is the result of trading activity but for measuring the impact of single and specified trades, private information is also needed. Publicly available databases, such as the NYSE TAQ database, generally contain trade and quote information like price, time and trade size but no information about the involved market participants. So chronologically following orders of the same investor and thus the determination of market impact dependence on ordered volume per time is not possible. Empirical studies on market impact dependence on single orders can be found in literature. Breen et al. (2002) develop a measure of liquidity or price impact quantifying the change of the stock price as an answer of the net trading volume and by taking predetermined firm characteristics into account. They use the data of NYSE TAQ database for adjusting their approach. Dufour and Engle (1999) find the waiting time between consecutive transactions is a significant measure for price impact of trades as well as for autocorrelation of signed trades. Lillo et al. (2003) fit their model on data of price reaction over trading volume normalized by some liquidity measures. They show that their model describes the data for stocks with different liquidity. Rydberg and Shephard (2003) propose a decomposition of price movement. So different dynamics can be modeled by different simple models where volume is also used as an explanatory variable. Bouchaud et al. (2003) developed a model for price movement as a result of the impact of previous trades. Almgren et al. (2005) provide an analysis of market impact depending on trades initiated by an identified party. Therefore, a dataset from Citigroup US equity trading desk is used. This enables the examination of the time component when a large order is split over a period of time. They differentiate between temporary and permanent price impact on the basis of the model provided by Almgren and Chriss (1999).

Studying price movements with a close look at microstructure processes, i.e. bringing price formation together with almost all elementary actions which can occur on financial markets is a complex topic resulting in studying huge datasets. But it is simultaneously a source of large potential for improving execution strategies for traders willing to trade large positions within a security.

3.2. Order book. 3.2.1. Resilience of order book. What happens with a limit order book during and
after the execution of a market order? The following is going to give an overview of the interaction between order book and aggressive order. An aggressive order takes liquidity from the order book, that is, all market order but limit order can also be aggressive. The direct reaction of the order book of such an aggressive order is quite simple, the incoming order is matched against the waiting passive orders in the order book. This results in widening of the spread and reduction of the provided volume in the book. The more interesting effect will be the reaction of the market after the execution, how the spread will narrow and how the provided volume in the book will re-rise, called the resilience of the limit order book.

Alfonsi et al. (2007) present two approaches of modeling the resilience. An exponential recovery of the limit order book is assumed. One approach models the recovery of the limit order book inventory and the second, the narrowing of the bid-ask spread. For measuring the effect, a reference limit price has to be defined. This unaffected limit (best bid or ask) is modeled by a Brownian motion. A similar model together with an empirical test on TAQ data is provided by Dong and Kempf (2007). They do not look inside the order book and take best bid and ask for the analysis, they just use the last price and the following model:

\[ S(t) = F(t) + Y(t), \]

where

\[ F(t) = \mu + F(t-1) + \epsilon(t) \]

and

\[ \Delta Y(t) = Y(t) - Y(t-1) = -\alpha Y(t-1) + \Phi(t) \]

and

\[ \Phi(t) \sim N(0, \sigma^2_\Phi), \quad \epsilon(t) \sim N(0, \sigma^2_\epsilon). \]

\( F(t) \) represents a random walk with drift describing the underlying price process. The other term describes the price recovery approach and \( \Delta Y(t) \) is interpreted as "pricing error" which tends towards zero because of market forces. The resiliency is depicted by the mean-reversion parameter \( \alpha \).

Dong and Kempf (2007) fit their model on 1-minute NYSE TAQ data using a Kalman-filter smooth estimation procedure to estimate the resiliency measure \( \alpha \). The mean value of all the resiliency estimates is \( \alpha = 0.60 \) and is significantly different from both zero and one. This means that the pricing error is stationary. Around 60% of the pricing error is corrected on average in every 1-minute interval. Further, the determinants of the resiliency measure are determined. The price level (inverse of average price) has a negative effect on resiliency indicating that lower tick size leads to more resiliency. The number of trades is positively correlated to resiliency whereas average trading size is negatively correlated as well as volatility of stock price.

3.2.2. The open limit order book and execution probability. The functioning of a limit order book is described in section 1.2. The following focuses on the dynamics of limit orders in order books and thus the interaction between book and order flow.

Theoretical models, which are provided by Kyle (1985) or Glosten and Milgrom (1985), focus on market maker quotations. Glosten (1992) analyzes limit order markets by modeling the price impact of trades reflecting their informational content.

Biais et al. (1995) provide an empirical analysis of order book characteristics, starting with descriptive statistics of an order book. The slope of an order book of a special stock is the supply and demand curve where (time-series) average of depth is drawn over average quote. They find that the bid-ask spread is twice the difference between adjacent quotes on each side of the book. The depth increases with the distance from the best bid/ask. They find that the bid-ask spread and the relative spreads on each side of the book show a intra-day u-shape pattern. The descriptive order book measure presented secondly is price discreteness. They compute the number of ticks between bid and ask quotes as well as between adjacent quotes and find a tick size dependency. The median difference between neighboring limits is larger than one tick size.

Besides order book characteristics, order flow is analyzed in detail in Biais et al. (1995). Orders can be classified according to their direction, aggressiveness and size while trades do not have a direction because it is always a buy and a sell, but it can be buyer or seller initiated. They cluster orders in different categories, for example as "large buy" which is an aggressive order larger than the volume behind the best ask. For each of these categories the unconditional probabilities are calculated using a data sample of stocks included in CAC 40 in 1991. Because of the strong intra-day pattern (u-shape), the probabilities of different orders are also proposed to calculate depending on the time of the day. The probabilities calculated of orders and trades are conditioned by the last order or trade which can be written in a matrix form. This matrix shows an interesting diagonal effect, i.e. the probability of a given order or trade is higher after this event has
just occurred than it would be unconditionally. Furthermore, they try to connect further orders or trades with the current state of the order book by calculating order and trade probabilities conditioned by the state of the book. Besides the probabilities of occurrence of a special event, they also provide an approach to predict the time interval between order and trade events.

The analysis shown in Biais et al. (1995) provides very interesting empirical approaches to describe market microstructure in limit order books. Knowledge about probability of further events in the book can be used to calculate execution probability of one’s order which can be used to optimize execution strategies.

4. Algorithmic trading

Algorithmic trading is automated trading, i.e. a computer system is completing all work from trading decision to execution. Algorithmic trading has become possible with the existence of fully electronic infrastructure in stock trading systems from market access, exchange and market data provision. The following gives an overview of chances and challenges of algorithmic trading as well as an introduction of several components needed to set up a competitive trading algorithm.

4.1. Chances and challenges. There are several advantages in contrast from algorithmic trading to trading by human beings. Computer systems have in general a much shorter reaction time and reach a very high level of reliability. The decisions reached by a computer system rely on the underlying strategy with specified rules. This leads to reproducibility of the decisions. Thus, back-testing and improving the strategy by variation of underlying rules are allowed. Algorithmic trading ensures objectivity in trading decisions and is not exposed to subjective influences (such as panic, for example). When trading many different securities at the same time, a computer system may substitute many human traders. So the observation and trading securities of a large universe become possible for companies without dozens of traders. Altogether these effects may result in better performance of the investment strategy as well as in lower trading costs. For further information concerning algorithmic trading and artificial agents, see Boman et al. (2001), Kephart (2002), LeBaron (2000) and Gudjonsson and MacRitchie (2005).

On the other hand, it is challenging to automatize the complete process from deriving investment decisions to execution because of the need of system stability. The algorithm has to be robust against numerous possible errors in services the system is dependent on, such as market data provision, connection to market and the exchange itself. These are technical issues which can be achieved by spending some effort in the implementation. Even more complex is the development of an investment strategy, i.e. deriving trading decisions, and strategies to realize these decisions. This work is focused on the realization and thus the execution strategy by assuming given investment decisions. It is beyond this work to introduce in how to derive investment decisions. All necessary information for the input of the execution algorithm is assumed to be available. Input variables may be the security names, the number of shares, and the trading direction. But also assumed available are variables like aggressivity and constraints, such as market neutrality when trading a portfolio.

The main challenge for trading algorithms is the realization of low trading costs in preferably all market environments independent from falling or rising markets as well as high and low liquid securities. Another critical point which has to be taken into account is the transparency of the execution strategy for other market participants. If a structured execution strategy acts in repeating processes, for example, orders are sent in periodical iterations; other market participants may then observe patterns in market data and may take an advantage of the situation.

4.2. Components of automated trading system. A fully automated trading system is complex with regard to technical requirements, but the numerous different research issues which have to be considered lead to even more effort and potential for improvement. An automated stock trading algorithm has to take many aspects into account which are addressed in this work. Reaching favorable trading costs, numerous cognitions of market microstructure theory have been incorporated into such a system. Strategies mentioned in 2.2. are just simple formalizations of market attributes. They are seen as an approximation of the strategy leading to minimal execution costs, but by far do not take all microstructure aspects into account. Probably all currently existing systems do not contain much more than such an approximation.

A suggestion for an automated trading system can be constructed of three components as it is denoted, for example, in Investment Technology Group (2007) or Kissell and Malamut (2005).

A pre-trade analysis component provides a previous estimate of transaction costs of a given order. Therefore, an econometric model based on historical trading data is used. The pre-trade analysis can be
used to optimize the expected transaction costs by varying the parameters or even the trading strategy. The expected trading costs do not have to necessarily be minimized, but it can be any function representing the trader’s preferences, for example:

\[(1 - \lambda) \cdot E(C) + \lambda \cdot Var(C) \rightarrow min,\]

where \(C\) is the total execution cost of the trade, \(E(C)\) is the expected value of \(C\), and \(Var(C)\) is the variance of \(C\). \(\lambda\) is the traders risk aversion parameter (Investment Technology Group, 2007). The expected cost of trade \(E(C)\) can contain opportunity costs if the trader allows the algorithm executing not the complete position and provides the expected profit of the investment. Jian Yang (2006) provides an empirical approach of selecting algorithms satisfying the trader’s needs best. The approaches introduced in 3.2.2 can be used to optimize real-time order placement in the order book to achieve favorable prices. Ian Domowitz (2005) explains how to compare performance of algorithms and specify some algo parameter. An approach to forecast and optimize execution is also provided in the work of Richard Coggins (2006). The second component is the trading algorithm itself. It’s the part executing orders according to the underlying strategy (see 2.2). The optimal strategy and values of their parameters have to be found in pre-trade analysis, but further improvement can be reached by adjusting parameters during the trading period. Therefore Jedrzej Białkowski (2005) and Białkowski (2008) provide a model of decomposing trading volume and model the components to forecast the trading volume. This can then be taken into account by the trading algorithm if it is based on volume like VWAP. Anna Obizhaeva (2005) shows the relationship of supply and demand dynamics of a security in the market and the execution performance of a given order. They provide a model of the impact of supply/demands dynamics on execution costs. Post-trade analysis is the third component of the system. After all information of the trades is available, a performance measurement can be done and compared to the pre-trade estimation. This is very important information to improve pre-trade analysis for further trades. For an example of a post-trade analysis framework, see Investment Technology Group (2007). Robert Kissell (2005) suggests a two part post trade analysis of cost measurement and algorithm performance measurement. Trading costs are measured as the difference of realized execution price and the specified benchmark to critique the accuracy of the trading cost model. Secondly, algorithmic performance is analyzed to assess the usability of the algorithm to adhere to the optimally prescribed strategy.

**Conclusion**

Algorithmic trading has become important in recent years in the finance industry and this trend probably will continue. There are numerous advantages in contrast to human traders and many possibilities arose as automated trading became available. Actually, electronic trading platforms have been founded in recent past attracting primarily algorithmic traders because of their tariffs and extremely fast reaction. These so called ECNs are now responsible for a significant percentage of daily stock turnover.

The implementation of automated trading systems requires some technical effort but also great knowledge in market microstructure. A human trader can use his knowledge and feeling for trading and is able to react on new situations by using just the human intelligence. An automated system does not have this possibility, so the knowledge of market microstructure has to be included in the system by using the models and empirical results of microstructure research.

This review tries to give an overview over the most important microstructure aspects and respective literature. The implementation shortfall, i.e. trading costs plays a central role in trading. So this aspect is analyzed in detail, the arise of implicit trading costs in order book markets and the dependence of these costs from other observable measures.

Execution strategies can work opportunistically, i.e. the strategy tries to execute orders according to the market environment to reach minimal transaction costs. The other approach is schedule driven acting, i.e. the strategy acts after strict specifications and execute every minute a given number of stocks for example. To achieve good results a mixture has to be used in reality. Thereby the detection of moments where trading results in favorable prices as well as the prediction of market reaction of a traders order require a deep understanding in market microstructure. Dominating intra-day pattern of most microstructure measures is important. Also a look inside the order book before and after an order arrives to observe the dynamics is of great interest.

From our point of view, market fragmentation and its effects on execution of single orders as well as the market reaction of a trader’s activities are very complex and not well understood up to now. So in all of these directions future research will be interesting. With increasing computational power it will become easier to study these huge amounts of tick-data produced by the market places.
References


