THESIS

Kata Váradi

Liquidity risk on stock markets
Statistical analysis and possible applications of the Budapest Liquidity Measure

Ph.D. dissertation

Supervisors:

Berlinger, Edina Ph.D.
Associate professor, Head of Department of Finance

Lublóy, Ágnes Ph.D.
Associate professor

Budapest, 2012
Department of Finance

THESIS

Kata Váradi

Liquidity risk on stock markets

Statistical analysis and possible applications of the Budapest Liquidity Measure

Ph.D. dissertation

Supervisors:

Berlinger, Edina Ph.D.
Associate professor, Head of Department of Finance

Lublóy, Ágnes Ph.D.
Associate professor

© Váradi, Kata
# Table of contents

1. **The aim of the dissertation and the theoretical background** ........................................ 4  
   1.1. The aim of the dissertation ..................................................................................... 4  
   1.2. Theoretical background and hypotheses ................................................................. 6  
      1.2.1. Characteristics of the order book ..................................................................... 6  
      1.2.2. Basic concepts of market liquidity ................................................................. 7  
      1.2.3. Liquidity adjusted Value-at-Risk models ......................................................... 9  
      1.2.4. Estimation and analysis of the price impact function ...................................... 10  

2. **Research methodology** ............................................................................................ 12  
   2.1. Statistical analysis of the Budapest Liquidity Measure ......................................... 12  
   2.2. Liquidity adjusted Value-at-Risk model ................................................................. 13  
   2.3. Estimation and time series analysis of the virtual price impact function .......... 19  

3. **Main results** ............................................................................................................. 23  
   3.1. Statistical analysis of the Budapest Liquidity Measure ......................................... 23  
   3.2. Liquidity adjusted Value-at-Risk model ................................................................. 25  
   3.3. Estimation and time series analysis of the virtual price impact function .......... 25  

4. **References** .............................................................................................................. 27  

5. **Own publications** ..................................................................................................... 32
1. The aim of the dissertation and the theoretical background

Liquidity is the essential condition of the normal functioning of financial markets and financial system. Only the appropriately liquid financial markets are able to function effectively, i.e. to transmit the savings to the users, and to aggregate market actors’ expectations and available information. The liquidity of markets, more precisely the lack of it affects the whole financial system, and indirectly through it the whole economy, thus inhibiting their normal, operational way of functioning. The financial crisis of 2008 has pointed to the outstanding importance of the liquidity of the financial system, and at the same time it pushed this question to the limelight. The revision and/or supplementation of the standard equilibrium and no-arbitrage models which assume the existence of infinite market liquidity has become a necessity, as there is an evident need to develop new pricing models and risk management techniques.

1.1. The aim of the dissertation

In my dissertation I am especially concerned with market liquidity and trading risk from different aspects: both from theoretical and also empirical points of view. Parallel with my research we have made a series of interviews supported by the Budapest Stock Exchange, during which stock-traders and portfolio managers were asked about the practical ways they manage liquidity risk (see Szücs and Váradi, 2012). The responses I got from the interviews contributed to a large extent to the formation of and refining my research questions and hypotheses. During the interviews a view gradually emerged that dynamic portfolio optimization on illiquid markets is a remarkably complex problem, which cannot be regarded as solved either from the practical or from the theoretical aspect for the time being. Market participants (in the absence of anything else) attempt to simplify the question, e.g. some are only willing to trade on liquid markets exclusively, while others decide on the portfolio they intend to create, then they give orders to traders who are specialized in carrying out the transaction of the requested size within a given time frame in a way that they are able to minimize the price impact of the transaction. Many others yet attempt to decrease liquidity risk during the build-up and/or the liquidation of a portfolio by setting up simple rules of thumb. In my dissertation I do not undertake the task of precisely describing and solving the optimization task, either, instead I attempt to take the first steps towards it by presenting the nature of liquidity risk and the options to manage it.
On illiquid markets trading costs are significantly higher than on liquid markets, i.e. transactions can only be executed with a notably higher cost and time. Therefore it is not surprising that market participants’ basic requirement is that each stock’s liquidity should be comparable and the transaction costs quantifiable. Measuring liquidity is a complex problem in itself, it is difficult to express all of its aspects with one single indicator, and it is also hard to estimate how much cost illiquidity generates during the trade, since liquidity can be interpreted along different dimensions and thus at any given time one or another of its different attributes can come to the forefront.

During my research I put strong emphasis on a liquidity indicator which quantifies the transaction costs of trading in the hypothetic and considerably extreme case when the buyer/seller is not willing to wait at all i.e. they intend to realize the transaction immediately, without any delay. This index is the so called Budapest Liquidity Measure (BLM), which has been created according to the pattern of the liquidity indicator firstly introduced and constantly published by Frankfurt Stock Exchange, the Xetra Liquidity Measure (XLM). The database has been provided to me by the Budapest Stock Exchange.

My main goal was to help liquidity as a concept to be incorporated into the daily practice of risk management, i.e. to elaborate solutions which can be easily intruded into the daily practice, but also properly developed from a theoretical point of view. From the series of interviews it evidently turned out that a prerequisite for dynamic portfolio optimization would be to get a clear view on how the return, the volatility and market liquidity of risky assets are correlated, i.e. what are the main attributes of this aggregated stochastic process. Accordingly during my research I have focused on three main issues:

1) I have examined the cross- and horizontal sectional statistical attributes of the BLM time series;

2) I have shown how the index can be integrated to a VaR-based risk management system;

3) I deducted the correlation between the BLM and the so called price-impact function, which is one of the essential concepts of liquidity and is often analyzed in literature. With the help of this latter I have also empirically analyzed how price impact evolved on the Hungarian stock markets between 2007 and 2011. This period is especially interesting because it includes the escalation and the run-down of a major liquidity crisis.
1.2. Theoretical background and hypotheses

According to my main issues, I have divided the theoretical background of my dissertation also to three parts. But I have added a fourth chapter, which includes those research that emphasize the importance of market liquidity. In sum I have split the summary of the theoretical background into four parts, which I will introduce in the order of appearance in my dissertation.

1.2.1. Characteristics of the order book

In the first chapter I have presented the stock market operation and the main attributes of quote driven and order driven markets. Apart from this, I have also given a detailed description about the statistical characteristics of the order book which is used on order driven markets based on the results of the foregoing empirical research.

Research have recently started to be substantively concerned with analyzing the statistical attributes of the order book, namely the consequence of the order book change is that the market prices on the markets will change, thus the given orders can be regarded as the most basic part of the price-formation.

A part of scientific articles concerned with the order book attributes tend to approach their examination mostly from a theoretical point of view. These studies are, among others, the following articles: Bouchaud et al. (2002), Bak et al. (1997), Chan et al. (2001), Luckock (2001), Slanina (2001), Daniels et al. (2002), Challet és Stinchcombe (2001), Willmann et al. (2003), Maslov (2000) and the works of Maslov és Mills (2001).

The order book has been statistically analyzed from various perspectives, of which the most important ones are the following: Bouchaud and Potters (2002), Zovko and Farmer (2002), Bouchaud et al. (2008), Maslov and Mills (2001), Lillo and Farmer (2004), Mike and Farmer (2008), Gopikrishnan et al. (2000), Gabaix et al. (2003), Margitai (2009), Plerou et al. (2002), Bouchaud et al. (2004), Chordia and Subrahmanyam (2002).

These research are important from the viewpoint of my dissertation, because they highlight the importance of market liquidity. Namely the main results were that the order direction and the transaction size are a predictable, long memory processes; we can anticipate the returns from the supply-demand formation in the short term. These empirical findings cannot be coordinated with the commonly observed fact that returns and thus the price-formation still cannot finally be predicted and we can characterize the latter process with a random walk process. The solution for this contradiction is market liquidity: this ensures that
the market functions efficiently and market prices cannot be predicted. Namely, Farmer et al. (2006) state that buy/sell side imbalance move together with the liquidity imbalance of the two sides, thus a certain amount can be bought or sold with a different price impact on the buy and the sell sides. This statement was based on Bouchaud et al.’s (2004), and Lillo and Farmer’s (2004) findings, who have also concluded the same.

The other reason I think that it was significant to analyze the order book, because the BLM database is based on the order book, namely it actually condenses the pieces of information in the order book through a special transformation.

1.2.2. Basic concepts of market liquidity

In the second chapter I have presented the basic concepts concerning market liquidity, the dimensions along which it can be measured and the main indexes which can quantify market liquidity from a certain aspect. I have also given a detailed description about the build-up of BLM index and its calculation process. The next listing shows which topics I have analyzed in detail, and which are those research I have mainly built on:

- Research done so far on BSE: Kutas and Végh (2005), Michaletzky (2010).
- Analysis of the relation between liquidity and volatility based on Hungarian research: Csávás and Erhart (2005), Michaletzky (2010).

During my research I have emphasized equally the international and Hungarian research, though my research questions and my hypotheses were built mainly on those questions which have not been answered by Hungarian researchers in the last few years.

Based on the Hungarian research and on the interview series, I have compared BLM with the two most commonly used liquidity indicators, the bid-ask spread, and the turnover. I have examined how the three liquidity indicators (BLM, bid-ask spread, turnover) rank the stock from the viewpoint of liquidity and I have also analyzed that under what market condition can be misleading to use only the bid-ask spread or the turnover.

Moreover I have examined the relationship between the volatility and liquidity based on the Hungarian research done by Csávás and Erhart (2005), and Michaletzky (2010). They have found that there is a strong positive correlation between the volatility and illiquidity. I
have examined this question as well to verify Csávás and Erhart’s (2005) statement: the increase of the bid-ask spread is the consequence of the new information’s volatility increase effect, since the expected volatility is already built in the value of the bid-ask spread. Based on my results I also wanted to conclude whether the crisis of 2007/2008 was a liquidity crisis as well, or the increase of illiquidity was a “natural” increase as a consequence of the higher volatility.

The questions I have examined:

– What are the average values of the BLM on the five available transaction sizes for different stocks?
– What is the relation between the bid-ask spread, turnover and BLM?
– What is the relation between liquidity and volatility?

I think these questions are important for at least three reasons. On one hand, based on the results I will be able to show which liquidity indicator is worth using, which one gives us the most reliable information on market liquidity. On the other hand it is important, because I will like to set the basics of being able to trade with liquidity in the future – whether it is quantified by the BLM. If it is possible, then it could even be the underlying asset of derivative products, allowing to hedge liquidity risk. To achieve this it is necessary to know the relation between the liquidity and volatility. Lastly, I think these questions are important, because during the dynamic portfolio optimization process it is not sufficient to bring the decision along only the return-volatility dimension, one should consider liquidity as well, since market risk is built up from two parts, the price risk (the change of the mid price) and the liquidity risk.

I will examine the following hypotheses based on the research questions:

**H1:** BLM, and the most commonly used liquidity indicators in practice (bid-ask spread, turnover) provide different ranking from the aspect of liquidity for individual stock:

   **H1/a:** during a calm period, and
   **H1/b:** during a crisis.
   **H1/c:** in the case of a liquid and
   **H1/d:** an illiquid stock.

**H2:** There is a positive relation between volatility (standard deviation, true range) and BLM.
1.2.3. Liquidity adjusted Value-at-Risk models

In the third chapter I have presented a possible model with which the BLM index can easily be integrated to the systems based on value at risk (VaR), aiming market risk management. Here the underlying idea is that on illiquid markets a value of an asset is not equal to its last market price. Namely, the buy/sell transaction reacts to the price and shifts it into the opposite direction. In this situation it is reasonable to determine the return in a way that we take the expected opposite price impact into consideration. I have given a detailed description about the models adjusted with liquidity (liquidity adjusted value at risk - LAVaR) which can be found in the literature, and then about my own theoretical model in the second part of the chapter. I regard this latter as one of my most important new results.

In the recent years several LAVaR models have been created, in which researchers tried to quantify liquidity risk, and to integrate it into the conventional VaR framework. These models can be split into two main groups:

1) Models based on order book:
   - Models based on exogenous liquidity risk (e.g.: Bangia et al., 1998; Radnai and Vonnáč, 2009),
   - Models based on endogenous liquidity risk (e.g.: Giot and Gramming, 2005; Stange and Kaserer, 2009a),
   - Models based on transactions or volume (e.g.: Berkowitz, 2000).

2) Models based on optimal execution:
   - Models based on stochastic time horizon (e.g.: Lawrence and Robinson, 1997),
   - Models based on price impact functions (e.g.: Jarrow and Subrahmanyam, 1997, 2001).

In my dissertation I will introduce only the models of the first group, since my own LAVaR model will be built on a model that considers endogenous liquidity risk. I have chosen a model that is based on the order book, because it has the advantage not to estimate numerous parameters as it is in the case of the models based on the optimal execution (Stange and Kaserer, 2009b). From the models based on the order book I have chosen one that is based on endogenous liquidity risk, because Ernst et al. (2009) made a research where they compared the models based on order book. The result of Ernst et al. (2009) demonstrate that those models produced the best results in terms of forecasting which have taken endogenous liquidity risk into account, and where the liquidity risk has been calculated from the XLM type of transaction cost liquidity indicators. In my dissertation I have based my model on Giot and Gramming’s (2005), and also on Stange and Kaserer’s (2009a) model. The difference in
my model was that I have built my model on a Hungarian database, and I have calculated LAVaR for single stocks and for differently weighted stock portfolios as well. I have also described in detail how to calculate the liquidity risk adjusted returns, and I have also shown how liquidity risk can be diversified. According to this, I have the following hypotheses:

**H3:** Market risk can be underestimated at least by 5% even for liquid stocks at the order size of EUR 20,000 on the Budapest Stock Exchange, if we do not take the liquidity risk into consideration.¹

**H4:** In case of stock portfolios not only price risk but liquidity risk can be diversified.

### 1.2.4. Estimation and analysis of the price impact function

In the fourth chapter I have described one of the central concepts of market liquidity, the so-called price impact function, which shows the expected relative price-shift caused by a particular order (Bouchaud et al., 2008). The knowledge of the behavioral attributes of the price impact function has an accentuated significance for market actors, since with the aid of it they can predict the expected price impact concerning their orders in the future, i.e. the expected surplus cost caused by price-shift. In this chapter I have described the difference between the virtual (vPIF) and the empirical price-impact functions (ePIF), and I have also presented a method of the estimation of a price impact function with the aid of Budapest Liquidity Measure. Based on the method I have elaborated, market actors can simply and quickly estimate a virtual price impact function without knowing the whole order book. During the estimation it was an important condition, that the BLM(q) function is linear, and as a consequence the virtual price impact function became linear as well. In the literature several research have been done so far, which have analyzed the shape of the empirical and virtual price impact functions, though the number of the research done on the ePIF are higher than the one made on the virtual one. I have found only four research on the vPIF function: Challet and Stinchcombe, 2001; Maslov and Mills, 2001; Smith et al., 2008; Weber and Rosenow, 2005. My opinion is that it is the consequence that the main goal of the research was to find the reason for the big price changes, namely that the price change because the transaction size is big, or because the lack of liquidity. In sum the research have found the price impact function can be described mainly with a power-law concave function (e.g.: Gabaix et al.,

¹ I have chosen 20,000 EUR, because this was the smallest available transaction size, on which the BSE calculates the BLM.
2003, 2006; Plerou et al., 2002; Farmer és Lillo, 2004; Margitai, 2009), or with a linear function. Based on the results I have chosen to assume the virtual price impact function to be linear.

Market participant can estimate easily and fast a price impact function without knowing the data of the order book, with the method I have worked out. I have estimated a virtual price impact function from the BLM database, which noone has done before, since most of the research have estimated empirical price impact functions. I have chosen to estimate a virtual PIF, because one can make time series analysis only on this function. The empirical PIF is being estimated from the average of transaction data of a longer time period (a month, a year, or even longer), so it cannot be analyzed with a time series analyses. So I have looked at the PIF-s from another viewpoint, because my opinion is, that during trading it is worth knowing how PIF evolves in time, because the trader will build his trading strategy on that. Knowing the behavior of the PIF in time could help the market participants in the timing of the transaction. When traders decide to postpone an order, in order to minimize the price impact, then he has to have a notion what the price impact will be in the future. Time series analysis can be made only on the vPIF, since we have the sufficient amount of information in that case. According to this, after the estimation of the vPIF, I have made a time series analysis, which also noone has done before. During my research I will answer the following questions:

1. What are the basic statistics of the vPIF (average, standard deviation, minimum, maximum, skewness, curtosis and distribution)?
2. Is there a trend in the time series?
3. How does the volatility change over time?
4. Are there outlier data, and are there structural breaks?
5. Can the time series data of vPIF be described as a mean reverting process?

My hypotheses based on the research questions will be the following:

**H5:** The dynamics of the virtual price impact function can be described by the following:

H5/a: symmetry,
H5/b: trend,
H5/c: cycles,
H5/d: mean reverting,
H5/e: shock resistance.
2. Research methodology

The three research issues also differ in the respect of the applied methodology: (1) I make a traditional descriptive statistical analysis on BLM database; (2) I build up a theoretical model which can be used in the field of risk management; (3) I make a time series analysis on the time series of the estimated price impact function.

2.1. Statistical analysis of the Budapest Liquidity Measure

The statistical analysis of the BLM can be split into three main parts:

1. First I show how the BLM database looks like, what are the average values of it during 1st January 2007 and 16th July 2010. I have made a cross-sectional analysis in this part of my research.

2. As a second step I have compared the average values of the BLM, bid-ask spread and turnover on different time frames: whole time series, before/during/after the crisis. After that I have tested the rank correlation (Spearman’s rank correlation and with Kendall’s rank correlation method) on different time frames. Then I have calculated also the correlation between the liquidity indicators, and also I calculated linear regression, where I have tested the explanatory power of the bid-ask spread and the turnover on the BLM. I have also tested the relation between the changes of the bid-ask spread/turnover and the change of the BLM.

3. Thirdly I have analyzed the relation between the volatility and liquidity – which will be represented by the BLM. I made the analysis with linear regression. Volatility can be measured in different ways, so I have defined it in my essay as following:
   a. Standard deviation of logreturns;
   b. Standard deviation estimated with GARCH model;
   c. Difference between the daily maximum and minimum price in percentage;
   d. True range ($TR$).

Since I have analyzed the relation between the BLM and the volatility with a linear regression I had to have volatility data for every trading day. So I couldn’t analyze the standard deviation because of the lack of data: I would have needed to know the intraday prices, which I didn’t have. Instead I have estimated daily standard deviations with a GARCH model. In this case I had an implicit assumption, that the
returns can be described with \( t \)-distribution, since GARCH model assumes this distribution during the estimation.

I have made the estimation with the following AR(1)-GARCH(1,1) model (see it in detail Bollerslev, 1986; Tulassay, 2009):

\[
\begin{align*}
    r_t &= c + \phi r_{t-1} + \epsilon_t \\
    \epsilon_t &= \sigma_t \eta_t \\
    \sigma_t^2 &= a_0 + a_1 \epsilon_{t-1}^2 + b_1 \sigma_{t-1}^2,
\end{align*}
\]

where \( r_t/r_{t-1} \) is the daily logreturn, \( \epsilon_t \) is the value of the residual, \( \sigma_t/\sigma_{t-1} \) is the standard deviation, \( \eta_t \) is an IID(0,1) random variable, while the other variables are different parameters of the estimation.

I have analyzed another volatility indicator above the GARCH model, the true range. I have chosen this, because technical analysts usually use this indicator to measure volatility (Makara, 2004). The applied equation is the following:

\[
TR = \frac{\text{max} \{ P^H_t; P^C_t \} - \text{min} \{ P^L_t; P^C_t \} }{ P^M_t },
\]

where \( P^H_t/P^L_t \) shows the highest/lowest price on the certain period, while \( P^C_t \) is the closing price of the previous period, and the \( P^M_t \) is the average price of the examined period (Wilder, 1978).

### 2.2. Liquidity adjusted Value-at-Risk model

The starting point of a LAVaR model is a conventional value at risk (VaR) calculation frequently used in everyday risk management. The VaR measure shows us the maximum loss of the portfolio over a predefined time horizon (\( t \)) at a given significance level (\( \alpha \)). It can be expressed either in forint or as a percentage of the portfolio value (Jorion, 2007). VaR calculation can be carried out according to the following formulae, for the returns (equation 5) and the prices (equation 6):

\[
\text{VaR}_{\text{return}}^{\alpha, \Delta t} = r_t^{\alpha, \Delta t} = \mu_{t+\Delta t} + \sigma_{t+\Delta t} q_{1-\alpha},
\]
where returns are considered on continuous time horizon, thus \( r^\Delta t = \ln \left( \frac{P^{t+\Delta t}}{P^{t}} \right) \), \( \mu_{t+\Delta t} \) is the expected value of the return in \( \Delta t \) time, \( \sigma_{t+\Delta t} \) is the standard deviation of the estimation, and \( q_{1-\alpha} \) is the \( 1 - \alpha \)-th quantile of a chosen distribution.

\[
\text{VaR}^{\alpha, \Delta t} = \frac{P^t_{\text{mid}} - P^{t}_{\text{mid}} \cdot \exp \left( r^\alpha_{t, \Delta t} \right) }{P^{t}_{\text{mid}}} = 1 - \exp \left( r^\alpha_{t, \Delta t} \right),
\]

where \( P^t_{\text{mid}} \) is the mid price at time \( t \), while \( P^{t+\Delta t}_{\text{mid}} = P^{t}_{\text{mid}} \cdot \exp \left( r^\Delta t \right) \). If, for example, \( \text{VaR}^{95\%, 1\text{day}} = 5\% \), then with 95% probability the loss in one day due to the change in mid price will not be larger than 5% (Jorion, 2007).

The work of Giot and Gramming (2005), Stange and Kaserer (2009a) were the starting point of my own model, who made their models based on XLM type measures. The basic idea of a LAVaR model is to incorporate the liquidity measure into the returns, and to determine the VaR value based on these new returns, as follows:

\[
\text{LAVaR}^{\alpha, \Delta t} (q) = 1 - \exp \left( r^\alpha_{\text{actual}, t} (q) \right),
\]

where \( r^\alpha_{\text{actual}, t} (q) \) is the net return including the BLM figure, thus allowing for the implicit costs of trading at a given \( q \) sized trade. So during the estimation of the LAVaR model I had to determine the net return. In the literature I haven’t found in detail how to calculate the net return. In the following I show broadly how I have determined the net return for a single stock and for an equal volume stock (EVS) portfolio. (In my dissertation I have also determined the net return for equal value stock portfolio as well.)

For a single stock the return taking only the price risk into account at a given \( \nu \) trade size is as follows:

\[
r_{\text{hypothetic}} = \ln \left( \frac{P_{\text{mid}, t} \cdot \nu}{P_{\text{mid}, t-1} \cdot \nu} \right) = \ln \left( \frac{q_t}{q_{t-1}} \right),
\]

where

\[
14
\]
where \( r_{\text{hypothetic}} \) denotes the return we would realize if trading with the asset were possible at the mid price. Accordingly, \( P_{\text{mid}, t} \cdot v \) and \( q_t \) stand for the value we were to get for selling „v‟ quantity of stock, if they were traded at the mid price, \( P_{\text{mid}, t-1} \cdot v \) and \( q_{t-1} \) denotes the same but a period earlier.

We must take the implicit cost of trading into consideration to calculate the net or \textit{actual return}. For this, first based on Equation 9 the weighted average price should be determined:

\[
b_t(v) = \frac{\sum b_{k,t} \cdot v_{k,t}}{v},
\]

where \( b_t(v) \) is the weighted average price on the bid side of the book for a given „v‟ quantity, \( b_{k,t} \) is the price in the order book at level \( k \) at time \( t \), while \( v_{k,t} \) is the quantity available at level \( k \) of the order book at time \( t \) and „v‟ is the total quantity to be traded.

The total proceeding from selling a stock at time \( t \) is \( b_t(v) \cdot v \). This can be expressed as follows:

\[
b_t(v) \cdot v = q_t^{\text{net}} = q_t \cdot \left(1 - \frac{\text{BLM}(q_t)}{2}\right),
\]

where \( q_t^{\text{net}} \) stands for the value we get when selling the stocks, and \( q_t \) is the value we would get if we were able to trade at the mid price. This latter must be adjusted for the transaction cost stemming from illiquidity, which is represented by the BLM. During the adjustment I take only half of the BLM, since the BLM represents the implicit transaction costs of turning around a position at the same time. By doing this I implicitly assume that the bid and the ask sides are symmetric. Based on the above, the actual return is to be determined by the following formula:

\[
r_{\text{actual}} = \ln \left( \frac{b_t(v) \cdot v}{P_{\text{mid}, t-1} \cdot v} \right) = \ln \left( \frac{q_t^{\text{net}}}{q_{t-1}} \right) = \ln \left( \frac{q_t^{\text{net}}}{q_t} \cdot \frac{q_t}{q_{t-1}} \right) = \\
= \ln \left( \frac{q_t \cdot \left(1 - \frac{\text{BLM}(q_t)}{2}\right)}{q_t} \cdot \frac{P_{\text{mid}, t}}{P_{\text{mid}, t-1}} \right) = \ln \left(1 - \frac{\text{BLM}(q_t)}{2}\right) + r_{\text{hypothetic}}
\]
The actual return has been split into two components; the first showing the effect of illiquidity: 
\[ \ln \left( 1 - \frac{BLM(q_t)}{2} \right) \]; the second is the return we were to realize if trading at mid price were possible: \( r_{hypothetic} \).

The return of a portfolio consisting of “n” number of stocks is calculated similarly to the return of a single stock. This is showed by Equation 12:

\[
r_{hypothetic} = \ln \left( \frac{\sum p^i_{mid,t} \cdot \cdot v_i}{\sum p^i_{mid,t-1} \cdot \cdot v_i} \right) = \ln \left( \frac{\sum q^i_t}{\sum q^i_{t-1}} \right)
\]

During my analysis I have calculated an EVS (Equal Volume Stock) portfolio’s returns. This is a portfolio comprising of the same amount of each stock, that is \( v_i = v \).

For calculating the actual returns I need the value of the portfolio at different times:
- The value of the portfolio at time t if there perfect liquidity:
  \[
  \sum_{i=1}^{N} q^i_t \cdot v_i = \sum_{i=1}^{N} v_i \cdot p^i_{mid,t} = v \cdot \sum_{i=1}^{N} p^i_{mid,t} ;
  \]
- The proceedings from selling the portfolio at time t, considering transaction costs arising from illiquidity: \( q^i_{net} = \sum b_i(v_i) \cdot v_i = \sum q^i_t \cdot \left( 1 - \frac{BLM(q^i_t)}{2} \right) ; \)
- The value of the portfolio at the previous period’s mid price:
  \[
  q_{t-1} = \sum p^i_{mid,t-1} \cdot v_i = v \cdot \sum p^i_{mid,t-1} .
  \]

Determining the three values above is necessary, as during the calculation of the portfolio’s return I have once again split the return into two components: the first coming from illiquidity, the second arising from the change in mid price. In order to determine the return from illiquidity I need the value of the portfolio with and without transaction costs. For the return from the change in mid price I need the value of the portfolio at time t and in the previous period, supposed that there is no loss from illiquidity. I have arrived at the following actual return:
$r_{\text{actual}} = \ln \left( \frac{\sum b_i(v_i) \cdot v_i}{\sum p_{\text{mid},t-1}(v_i) \cdot v_i} \right) = \ln \left( \frac{\sum b_i}{\sum p_{\text{mid},t-1}} \right) = \ln \left( \frac{q_{\text{net}}}{q_{t-1}} \right) = \ln \left( \frac{q_i}{q_t} \right) + \ln \left( \frac{q_t}{q_{t-1}} \right) = \ln \left( \frac{1}{v} \sum p_{\text{mid},t} \left( 1 - \frac{\text{BLM}(q_i)}{2} \right) \right) + \ln \left( \frac{1}{v} \sum p_{\text{mid},t-1} \right) = \ln \left( \frac{\sum p_{\text{mid},t} \left( 1 - \frac{\text{BLM}(q_i)}{2} \right)}{\sum p_{\text{mid},t}} \right) + r_{\text{hypothetic}} \quad (13) \]

In case of portfolios it is an important question whether the liquidity risk can be diversified. In Equation 14 I have determined how we can quantify the diversification of the liquidity risk. Noone in the literature has defined this equation before.

$$\gamma(q) = \frac{\sum \text{LAVaR}^{\alpha,\Delta t}(q_i) - \text{LAVaR}^{\alpha,\Delta t} \left( \sum q_i \right)}{\sum \text{VaR}^{\alpha,\Delta t}(q_i) - \text{VaR}^{\alpha,\Delta t} \left( \sum q_i \right)} \quad (14)$$

The $\gamma(q)$ shows the additional effect of diversification, as a percentage of the price diversification impact of portfolios, if we consider illiquidity. Namely, the $\sum \text{LAVaR}^{\alpha,\Delta t}(q_i) - \text{LAVaR}^{\alpha,\Delta t} \left( \sum q_i \right)$ formula gives the difference between LAVaR values (at a given confidence level and for a predefined time period) for single stocks added together and for portfolios. The $\sum \text{VaR}^{\alpha,\Delta t}(q_i) - \text{VaR}^{\alpha,\Delta t} \left( \sum q_i \right)$ formula uses the same logic for conventional the VaR measure. As a result, Equation 14 demonstrates the diversification effect as a percentage of the price diversification impact.

The technical tools I have used to estimate LAVaR were the same for single stocks and for stock portfolios. The difference in the analysis is that for portfolios it is not sufficient to know the BLM values at five different trade sizes (20, 40, 100, 200, and 500 thousand Euros), since in this case not the value but the quantity of the stocks is fixed. Accordingly, we must have BLM figures for all “$q$”-s. This can be carried out in two simple ways: 1) to use
linear interpolation based on the available BLM data for each day, or 2) to use a linear regression.

In my modeling I have taken the second approach.\(^2\) Obviously, this is a serious simplification, but based on the available data, it is appropriate for a first approximation. Linear regression is a practical, easy to use method, and qualitative consequences can surely be drawn from the analysis. Furthermore daily BLM data can be well approximated by a straight line.

In my dissertation I have calculated the conventional and also the liquidity adjusted VaR, in order to be able to compare them. In order to account for the clustering volatility of returns and net returns, I have fitted again an AR(1)-GARCH(1,1) model, where I have used t-distribution. The sample used to estimate the model was the first 2.5 years (1\(^{\text{st}}\) January, 2007-15\(^{\text{th}}\) July 2009), while the last year (16\(^{\text{th}}\) July 2009-16\(^{\text{th}}\) July 2010) was used as a control period. I calculated the daily 95% and 99% VaR using forecasts from the GARCH model. I used a rolling window of 2.5 years to continuously re-estimate the GARCH model, i.e. I have estimated a GARCH model for the first 2.5 years and have made a forecast for the next year, and then I have repeated the procedure while rolling the sample period with one day.

Figure 1 shows one of the LAVaR estimations of my dissertation, where we can see the LAVaR and VaR forecasts for an EVS portfolio, where the four stocks in the portfolio are the OTP, MOL, MTelekom and Richter (bluechip stocks on BSE):

\^2\ I will use the same approach during the estimation of the price impact function.
2.3. Estimation and time series analysis of the virtual price impact function

The BLM(q) in itself is not a price impact function, as the BLM does not inform the trader about the new mid price realized after the transaction. Instead, the BLM measures the implicit cost of trading (in basispoints) stemming from the illiquidity of the markets. Since BLM’s calculation is based on the order book, it is possible to estimate a marginal supply-demand curve (MSDC) (Acerbi, 2010), than to estimate the virtual price impact function. The estimation of the vPIF from the BLM database is my own result. Figure 2 shows the relation between the MSDC and the BLM.

In accordance with Figure 2, the BLM can be calculated on the basis of Equation 15. In Equation 15 „q” stands for the total value of the transaction in Euros, as the BLM shows the implicit cost in the function of the value, not the volume.

\[
BLM(q) = \frac{\int_{0}^{q} MSDC_{ask}(x)dx - \int_{0}^{q} MSDC_{bid}(x)dx}{q}
\]  \hspace{1cm} (15)

If I assume that the daily BLM(q) function can be approximated by a linear regression – as I did during the LAVaR calculation –, then the BLM(q) function is as follows:

\[
BLM(q) = a \times q + b
\]  \hspace{1cm} (16)
The BLM(q) function is estimated separately for the bid and the ask side of the limit order book. In the following equations BLM\textsuperscript{b} stands for the buy (bid) side, while BLM\textsuperscript{a} for the sell (ask) side.

\begin{align*}
BLM &= 2 \cdot LP + APM\textsubscript{bid} + APM\textsubscript{ask}, \\
BLM\textsuperscript{a} &= LP + APM\textsubscript{ask}; BLM\textsuperscript{b} = LP + APM\textsubscript{bid}.
\end{align*}

(17)

(18)

The linear regressions are defined as follows:

\begin{align*}
BLM\textsuperscript{a}(q) &= a\textsubscript{ask} \cdot q + b\textsubscript{ask} ; \quad BLM\textsuperscript{b}(q) = a\textsubscript{bid} \cdot q + b\textsubscript{bid}
\end{align*}

(19)

The estimation of the MSDC by means of the BLM(q) function requires the following steps on the ask side:

\begin{align*}
BLM\textsuperscript{a}(q) &= \frac{\int\limits_0^q MSDC\textsubscript{ask}(x)dx - q \cdot P\textsubscript{mid}}{q} \rightarrow \\
BLM\textsuperscript{a}(q) \cdot q &= \int\limits_0^q MSDC\textsubscript{ask}(x)dx - q \cdot P\textsubscript{mid} \rightarrow \\
dBLM\textsuperscript{a}(q) \cdot q + BLM\textsuperscript{a}(q) &= MSDC\textsubscript{ask}(q) - P\textsubscript{mid} \rightarrow \\
a\textsubscript{ask} \cdot q + a\textsubscript{ask} \cdot q + b\textsubscript{ask} + P\textsubscript{mid} &= MSDC\textsubscript{ask}(q) \rightarrow \\
2 \cdot a\textsubscript{ask} \cdot q + b\textsubscript{ask} + P\textsubscript{mid} &= MSDC\textsubscript{ask}(q)
\end{align*}

(20)

The estimation of the MSDC by means of the BLM(q) function requires the following steps on the bid side:

\begin{align*}
BLM\textsuperscript{b}(q) &= \frac{q \cdot P\textsubscript{mid} - \int\limits_0^q MSDC\textsubscript{bid}(x)dx}{q} \rightarrow \\
BLM\textsuperscript{b}(q) \cdot q &= q \cdot P\textsubscript{mid} - \int\limits_0^q MSDC\textsubscript{bid}(x)dx \rightarrow \\
dBLM\textsuperscript{b}(q) \cdot q + BLM\textsuperscript{b}(q) &= P\textsubscript{mid} - MSDC\textsubscript{bid}(q) \rightarrow \\
P\textsubscript{mid} - (a\textsubscript{bid} \cdot q + a\textsubscript{bid} \cdot q + b\textsubscript{bid}) &= MSDC\textsubscript{bid}(q) \rightarrow \\
P\textsubscript{mid} - (2 \cdot a\textsubscript{bid} \cdot q + b\textsubscript{bid}) &= MSDC\textsubscript{bid}(q)
\end{align*}

(21)
Finally, the virtual price impact function can be expressed in the function of MSDC(q):

\[ \text{VPIF}(q) = \frac{\text{MSDC}(q)}{P_{\text{mid}}} - 1 \]  

(22)

On the basis of the vPIF the empirical price impact function cannot be estimated, as the BLM database does not provide information on the probability of the occurrence of the price impacts. The ePIF can be estimated, for example, from the TAQ (trades and quotes) database (Margitai, 2009). Estimating the ePIF from the TAQ database is a time- and calculation consuming task. In my dissertation my main goal was to provide the market participants a method that enables them to estimate the price impact function easily.

Figure 3 shows the estimated virtual price impact functions for OTP for both the bid and the ask side for a few trading days. The trading days have been chosen with the intention to show how the price impact behaves in calm period (1\textsuperscript{st} January 2007 and 2\textsuperscript{nd} June 2011) and during crisis (20\textsuperscript{th} October 2008 and 9\textsuperscript{th} January 2009). Figure 3 demonstrate that during a crisis the price impact function is more slopped that refers to the fact, that the transaction cost of trading is higher: Obviously, during crisis the markets are more illiquid, then during normal times.

![Figure 3: Virtual price impact function](image)

Source: Váradi et al., 2012.
Besides plotting the vPIF for certain trading days, it is worth plotting the time series of the vPIF values for a few order sizes. The time series are shown on Figure 4 for the time period of 1 January 2007 and 2 June 2011. Figure 4 also demonstrates that the crisis of 2008 was coupled with higher price impacts, thus, with lower market liquidity.

![Figure 4: The time series of the virtual price impact function](image)

After the estimation of the vPIF I have made a time series analysis on the database. The most important research methods I have used were the following:

- Descriptive statistics: average, standard deviation, median, minimum, maximum, curtosis, skewness.
- Trend analysis: fitting of polynomial trend, calculating moving average.
- Symmetry of bid and ask side: correlation of the two sides.
- Outlier data: analyzing boxplot figures.
- Mean reverting: using an extended Dickey-Fuller test (Dickey és Fuller, 1979).

Source: Váradi et al., 2012.
3. Main results

3.1. Statistical analysis of the Budapest Liquidity Measure

Ranking of stocks based on the liquidity indicators:
- In the case of medium liquid and illiquid stocks, bid-ask spread does not give the same ranking as BLM, however the difference is not significant.
- In the case of liquid, medium liquid and illiquid stocks, turnover does not give the same ranking as BLM, however the difference is not significant.
- In a calm period i.e. before and after crisis, ranking differs less from the ranking provided by BLM based on turnover than from the one based on bid-ask spread.
- During a crisis, the ranking based on bid-ask spread differs less from the ranking provided by BLM than from the one based on turnover.
- During the crisis the rank-correlation has decreased between BLM and the spread and between BLM and the turnover.
- In the case of the medium liquid and illiquid stocks it would be worthwhile to take also the BLM into consideration as a liquidity indicator, because in their case the ranking in the wrong order is more significant. In respect of these stocks I have also shown during my analysis that there is a chance that a particular stock is sorted into a wrong liquidity category.

Change of liquidity indicators during crisis:
- In the case of liquid stocks, the values of BLM and the bid-ask spread returned to their pre-crisis level, while in the case of turnover it could only be observed in the case of OTP and MTelekom.
- In the case of medium liquid and illiquid stocks liquidity of some stocks did not return to the pre-crisis level according to the BLM and bid-ask spread, while it did not happen to any stocks according to turnover.

Relation between liquidity indicators:
- The correlation between bid-ask spread and BLM can be regarded as strongly positive, while the correlation of BLM and turnover shows a slightly negative relation.
- The less liquid a stock is, the lower the correlation between the liquidity indicators.
- The change of bid-ask spread has a strong explanatory power about BLM change in the case of a liquid stock, whilst in the case of medium liquid stocks this explanatory power is not significant. In the case of illiquid stocks, bid-ask spread change has very
limited explanatory power, which cannot even be considered as significant before the crises.
- The turnover change cannot explain BLM change in the case of liquid and illiquid stocks, whilst it has also only a low explanatory power in the case of a medium liquid one.
- Turnover and liquidity do not co-move intraday, for instance at the beginning of the day liquidity is low in every case regardless whether the turnover is big or small.
- BLM can be important for those market participants who invest in illiquid stocks or intraday.
- Each stock’s liquidity related to one another can significantly differ in the case of different liquidity indicators.

Relation between liquidity and volatility:
- On the Budapest Stock Exchange it has been justified, that there is a positive correlation between BLM and volatility, namely that the more volatile markets are, the transaction cost caused by the lack of liquidity is higher.
- The less liquid a stock is, the lower the correlation between liquidity and volatility tends to be.
- Before and during the crisis, the correlation between the true range and liquidity was stronger than the one between standard deviation and liquidity. However, after the crisis this has reversed.
- The crisis of 2008 can be regarded as a liquidity crisis based on the liquidity estimated from volatility, i.e. the estimated BLM value is lower than the actual BLM value.
- After the crisis, the estimated BLM value is typically higher than the actual value, i.e. liquidity is higher after the crises than it had been expected. The less liquid a stock is, the typically lower the correlation is between liquidity and volatility.

In sum my most important results were during the analysis of BLM, that I have pointed out that the rule of thumbs market participants use do not always lead to a proper investment decision. I have showed that the BLM is a liquidity indicator that can measure liquidity in several dimensions, and gives a more reliable picture of the actual liquidity than if we would base our decision only on the bid-ask spread or the turnover. In case of the medium liquid and illiquid stocks it worthwhile to calculate BLM, since in the case of these stocks the ranking of the stocks differ notably. Moreover the correlation between the liquidity indicators is low, which decreases during crisis. So BLM can be important for those, who are trading
with illiquid stocks, or for those who are trading intraday. But during crisis it is worth take
into account the BLM for those as well, who are investing only in liquid stocks.

Based on the results of the analysis of the relation between the volatility and liquidity
it can be said that the crisis of 2007/2008 was a liquidity crisis as well, which means that the
cause of the increased implicit cost was not only the increased volatility. My results also
prove the statement of Csávás and Erhart (2005), that the decrease of liquidity reflects the
increase of the unexpected volatility.

In sum I reject the H1 hypothesis, namely that there would be a significant difference
in the ranking. Although the differences are low it worth taking them into account. I cannot
reject the H2 hypothesis, that there is a positive correlation between liquidity and volatility.

3.2. Liquidity adjusted Value-at-Risk model

- I determined the net return, namely how return calculation changes if we take into
  consideration the cost that occur because the lack of liquidity. I have determined both
  for the individual stocks and for the volume and value weighted portfolios.

- Taking liquidity into consideration means a significant risk increase even in the case
  of the most liquid stocks both on the level of individual stocks and portfolios.
  Therefore it is not advisable to ignore this.

- In the case of portfolios, liquidity risk can be decreased by diversification; therefore it
  is worthwhile to hold various stocks in a portfolio, because thus not only the price risk,
  but also the liquidity risk decreases.

BLM and the method presented above offer an easy and rapid way to incorporate
liquidity in capital requirement calculation. Bearing in mind the deficiencies and calculation
methodology of the BLM the results should be treated with care. Nevertheless, the presented
model is able to reproduce main empirical observation like OTP is by far the most liquid
stock at BSE, therefore I advise its integration into risk management systems. In sum I cannot
reject the H3 and H4 hypotheses.

3.3. Estimation and time series analysis of the virtual price impact function

- Estimation of a virtual price impact function from the BLM database.

- The value of the descriptive statistics i.e. the mean, the median or the standard
deviation have shown a higher value in every case on the bid side of the function than
on the ask side. I have explained the phenomenon with the herd effect, namely that the virtual price impact reflects that usually traders buy stocks separately from each other, but selling stocks is often concentrated, for example because of a panic situation.

- The time series data of the virtual price impact function do not contain trends, however quarterly cyclicity can be discovered in the data.

- During the cycles the price impact values reach their minimum level in the time of quarterly reports, while their maximum values are halfway between two quarterly reports.

- By examining outlier data I have identified 52 turbulent days. All these days fall into the period of the 2008 crisis, since they can be found between 17 October 2008 and 9 April 2009.

- I have also identified a structural break in the time series with the aid of formalized statistical tests.

- There is a significant autocorrelation in the dataset, from which I draw the conclusion that the impact of an incidental shock prevails in the market data for a longer period of time.

- When liquidity ceases on one side of the order book, then liquidity will be lower on the other side of the book as well, i.e. the correlation between the buy and sell side price impact is very high.

- The vPIF process can be described as a mean reverting process. The time series data of the virtual price impact function do not contain trends, but quarter-year cyclicity can be discovered in the data.

Based on the result the acceptance of H5 is the following:

H5/a: I cannot reject the hypothesis that the price impact of the bid and ask side is symmetric.

H5/b: I reject the hypothesis that there is a trend in the vPIF time series.

H5/c: I cannot reject the hypothesis that there are cycles in the vPIF time series.

H5/d: I cannot reject the hypothesis that the vPIF is a mean reverting process.

H5/e: I cannot reject the hypothesis that effect of shocks on the price impact lasts longer.
4. References

Acerbi, C. [2010]: The value of liquidity – Can it be measured? RBC Dexia Investor Services
Barra, I. [2008]: Analysis of market liquidity based on transaction durations. Thesis, Corvinus University of Budapest
Bélyácz, I. [2009]: *Befektetési döntések megalapozása*. Aula Kiadó Kft., Budapest
Berkowitz, J. [2000]: Incorporating liquidity risk into value at risk models. Working paper, University of California, Irvine


Csávás, Cs. & Erhart, Sz. [2005]: Likvideke-a magyar pénzügyi piacok? – A deviza- és állampapír-piaci likviditás elméletben és gyakorlatban. MNB study 44.


Darvas, Zs. [2004]: Bevezetés az idősorelemzés fogalmáiba. Lecture notes. Corvinus University of Budapest.


Kerékgyártó, Gy. & Mundruczó, Gy. [1995]: *Statisztikai módszerek a gazdasági elemzésben*. Aula kiadó, Budapest


5. Own publications

Publications in Hungarian in the topic of the dissertation:

Reviewed journal:

Publications in English in the topic of the dissertation:

Reviewed journal:
Szűcs, Balázs Árpád & Váradi, Kata [2012]: Measuring and managing liquidity risk in practice. Pénzügyi Szemle, accepted tender material.

Working papers in English:

Articles in conference proceedings:

Abstracts in conference proceedings:
Publications in English in other topic:

Reviewed journal:

Articles in conference proceedings:

Abstracts in conference proceedings: