2016/Issue 1 (Volume 3): Economy and Finance

THE MEASUREMENT OF MARKET RISK WITH A FORWARD–LOOKING LAVAR MODEL

László Madar – Bettina Tálos – Ádám Kocsis

Liquidity risk is closely related to market risk. The VaR (Value-at-Risk) based general model, which is used as the best practice for the quantification of market risks, tends to underestimate the risks in the Central and Eastern European and Middle Eastern region investigated in our analysis. The reason for this is the inadequate extent and quality of the available data. In the countries of the region with a more developed capital market (higher capitalisation, liquidity), the availability of data is adequate, but there is only a limited selection of traded instruments. This problem is aggravated by a stylised fact of the equity markets, namely the concentration of liquidity. This is where the liquidity is concentrated in a few instruments, while in the case of the smaller securities the lack of depth of the order book and available price levels have a substantial impact on price in some transactions. Even with a small quantity this can represent a problem if there are gaps in the order book, and with a large quantity the dynamic of the recovery of the order book brings the actual price impact of the transaction into question. For this reason, the risk management model needs to manage the difference between the market and liquidation value of the portfolio. This study is concerned with the measurement of this effect, and the development of an alternative, liquidity-adjusted VaR estimate.[1]

JEL codes: C01, C18, C24, C52, C53, G21

Keywords: LAVaR, liquidity adjustment, market risk

1. BACKGROUND

A tool commonly used in the course of risk management is the Value at Risk (VaR) calculation, and specifically what is referred to as the delta-normal VaR, because this is simple and quick to calculate, and it can be used to assess complex portfolios. The VaR tells us, at a given significance level (α) and assuming a given

1 The study was conducted in the course of the research project entitled "*Research into innovative mathematical models for the measurement of bank risks under Basel and quantification of the capital requirement, with respect to market, operational, liquidity and secondary risk; and the behaviour-based forecasting of change in the prices of financial products*", funded under the New Széchenyi Plan (project number PIAC_13-1-2013-0073).

holding period (T), the maximum loss that can be suffered on a given position, either in terms of an amount denominated in forint, or as a percentage. The significance level for market risks is typically set at 95–99.9%, while the holding period is usually a one-day or ten-day time frame, although under Pillar 2 of the Basel Accord it may be as long as one year. The holding period is usually determined in terms of the liquidity of the market concerned; that is, based on how quickly the position can be offloaded in the market. The delta-normal VaR presupposes a perfect liquidity situation.

The traditional VaR calculation does not encompass the full market risk, because it does not take the liquidity risk into account. The traditional VaR calculation works on the assumption that the instrument can be traded at the median price within a fixed time frame, which is not necessarily the case under actual market conditions. For this reason, in the VaR calculation it is necessary to allow for the fact that we are not always able to trade the instrument at the median price; in other words, the liquidity also needs to be quantified.

Several families of model have emerged in the field of market risk that also aimed to manage the liquidity aspect of market risk. Typically these methodologies developed further the classic delta-normal framework, which eases the process of interpreting these and comparing them with the logic of the delta-normal calculation that serves as the basic model. Market liquidity adjustment models are typically labelled with their English acronym: LAVaR (Liquidity Adjusted Value at Risk), and can be divided into two main groups: 1) models based on the order book data, and 2) models based on optimal execution.

Liquidity risk, however, can be broken down by two more factors: exogenous and endogenous liquidity risk. Exogenous liquidity risk originates from market processes, and is the same for all market players. This type of liquidity risk is not affected by the activities of any of the market players (although it may be influenced by the joint activity of the players). Exogenous liquidity risk cannot be measured, for example, in terms of the size of the bid-ask spread, the turnover, or the quantity of orders available at best order. Endogenous liquidity risk is the liquidity risk that can be influenced by the participants, as its extent can be influenced by the decisions of the market players. An example of this is a liquidity wave caused by an attempt to execute a large position.

In liquid markets, the bid-ask spread takes on a relatively stable and low value, while the quantity of orders available at best order shows a relative high, and also stable, value. Besides this, in liquid market turnover is also high. In contrast to this, in illiquid markets such as those of developing countries, the value of the bid-ask spread is highly variable, and higher than in the case of liquid markets. In addition to this, the value of the quantity offered at best order is also changeable, and moreover there are often few orders, and turnover also falls considerably short of the developed markets.

In our analysis we chose the endogenous model for investigating this aspect of market risk. This is mainly because, on the one hand, these models take into account both endogenous and exogenous risk; and on the other, this methodology is also suitable for modelling illiquid markets, as they are sufficiently complex and capable of isolating the market risk depending on the holding period, but they are not influenced by the market at every moment.

In order to develop the appropriate model, we conducted a factor analysis of the liquidity indicators derived from the data of the selected stocks, to ensure that the input parameters of the model for measuring market risk are easy to interpret due to the ways in which the procedure concentrates various dimensions of liquidity.

In order for the model to determine as effectively as possible the potential loss resulting from market liquidity, it is advisable to use as many variables as possible for the principal component analysis (PCA). The redundancy of variables causes no problems, as the matching types of variable are separate from the other types of variable, and together they add to the explanatory power of the principal component. The objective is for the principal components to describe the highest possible percentage of the variance.

2. LITERATURE REVIEW

The traditional Value at Risk calculation is based on the assumption that the financial instrument can be traded at the median price in every case, so it does not take liquidity risk into account. The studies examined in this section, however, have shown that liquidity risk accounts for a not insignificant share of the overall market risk.

In the course of their research *Lawrence* and *Robinson* (1996, in: *François-Heude* and *Van Wynendaele*, 2001) observed that a VaR calculation that omits liquidity risk leads to a 30% underestimation of market risk. *Bagnia* et al. (1998), in respect of developing countries, found that models which leave out liquidity tend to underestimate market risks by 25-30%. *Stange* and *Kaserer* (2009a) arrived at a similar result in their study of data from Deutsche Börse AG: the traditional market models underestimate market risk by 25%.

Bagnia et al. (1998) divide market risk into two main parts: price risk and liquidity risk. The price risk is the risk of a substantial shift in the median price in response to market trends, while the liquidity risk is the risk that we will be unable to trade at the median price. They break down the liquidity risk further into two components: exogenous liquidity risk and endogenous liquidity risk.

2.1. Models based on exogenous liquidity risk

In 1998 *Bagnia* et al described a VaR model, adjusted for exogenous liquidity that can be easily used by participants in the market, thus enabling them to incorporate liquidity risk into their Value at Risk calculations. This model, known as the BDSS model after the names of the authors, formed the basis for many subsequent lines of thought.

The BDSS model only takes the bid-ask spread into account in its calculations, so the LaVaR value is obtained by adding together the traditional VaR and the liquidity risk calculated form the bid-ask spread.

$$LaVaR = Pmid_t \left[(1 - e^{\mu - \alpha\sigma}) + \frac{1}{2}(\bar{S} + \alpha'\tilde{\sigma}) \right], \quad (1)$$

where $Pmid_t$ is the instrument's price at t time, μ is the log yield, α is the predetermined percentage of the log yield's distribution, σ is the standard deviation of the log yield, $\bar{S}$ is the average relative spread is the distribution of the relative spread, and $\tilde{\sigma}$ is the standard deviation of the relative spread (*Bagnia* et al., 1998, p. 8.).

The bid-ask spread is easily accessible to the market player, but the model is based on the spread's normal distribution, and experience shows that this is not true of the spread's distribution. Due to the trends, the edge is wider and more skewed than the normal distribution. It may also be the case that the distribution has several modes. Because the model only takes into account the exogenous variables (and not the endogenous risk), it underestimates the actual liquidity risk.

Ernst et al. (2008) set out to correct the error in the BDSS model that assumes a normal distribution, by taking the skewness and pointedness of the distribution into account; but this approach still does not remedy the other flaws in the model.

The inclusion of endogenous risk in the model can be corrected with the models presented in the next subsection, while estimating the correlation from actual market date would correct the error resulting from the assumption of a perfect correlation between the exogenous liquidity risk and the price risk.

Radnai and *Vonnák* (2009) proposed a method similar to the BDSS model. Based on the bid-ask spread they present a supplementary capital requirement for less liquid instruments, as a form of penalty for the bank's failure to transfer the illiquid assets to the banking book. The authors take the position that the bid-ask spread is good tool for measuring liquidity because supply and demand tend to diverge as liquidity decreases. "A possible solution is to determine the capital requirement as a linear function of the spread, besides which the banks must be given the opportunity – provided that they meet the quantitative and qualitative requirements for the estimate – to calculate the supplementary capital requirement through internal modelling based on the historical distribution of the spread." (*Radnai–Vonnák*, 2009, p. 252.)

2.2. Models based on endogenous liquidity risk

Models that also take endogenous liquidity risk into account give a more precise result with respect to the size of the liquidity risk, as they encompass both the exogenous and the endogenous liquidity risk.

The first model of this type is attributed to *François-Heude* and *Van Wynendaele* (2001), who use intraday data. In this way, the accessibility of the intra-day data facilitates a more accurate LAVaR calculation, which does not only use a sample of the observations to describe the events of a whole day. The model developed by them applies the foundations of the BDSS model, but takes into account the best five orders in the order book, in contrast to the BDSS model, which only examines the best order. Owing to the new approach, they are capable of examining the price impact of several transactions of various sizes, executed at the best five levels. The model can be described with the following formula:

$$LaVaR = Pmid_t\left[\left(1-\left(1-\frac{\overline{Sp}(Q)}{2}\right)*\left(e^{-\alpha\sigma}\right)\right)+\frac{1}{2}\left(Sp_t(Q)-\overline{Sp}(Q)\right)\right], \qquad (2)$$

where *Pmidt* is the median price at time t, $\overline{Sp}(Q)$ is the average spread given quantity Q, $Sp_t(Q)$ is the size of the spread given quantity Q at time t, α is the percentage given for the median price yield distribution, and σ is the yield's standard deviation (*François-Heude* and *Van Wynendaele*, 2001, p.10.).

The basis of the model of *Giot* and *Gramming* (2005, in: *Váradi*, 2012) is the price impact arising upon the sale and purchase of a specified instrument. "This price impact; that is, at what price the given order will be executed for a market player giving a market order, depends on the state of the order book at any given time. The two authors names this measure the CRT (Cost of Round Trip)." (*Váradi*, 2012, p. 99.)

The model can be expressed with the following formula:

$$LaVaR = 1 - \exp\left(\mu_{rnet(q)} + \alpha\sigma_{rnet(q)}\right), \qquad (3)$$

where $rnet(q)$ indicates the net yield, $\mu_{rnet(q)}$ is the expected value of the net yield, α is the given percentile of the distribution of the net yield, and $\sigma_{rnet(q)}$ represents the standard deviation of the net yield (*Váradi*, 2012, p. 99.).

The principal flaw in the module is that it uses t-distribution instead of empirical distribution. A solution to the problem is provided by the work of *Stange* and *Kaserer* (2009a), who use empirical distribution in their calculation. The model is not capable of allowing for the possibility of differing liquidity on the bid and ask side. The latter error is corrected by the work of *Qi* and *Ng* (2009), who also work with intra-day data, but calculate the liquidity risk separately on the bid and ask side, because the market shifts asymmetrically upwards and downwards. The authors calls this approach the LAIVAR (liquidity adjusted intraday VaR) model. The model is capable of measuring the liquidity risk depending on the market position, determining more precisely the Value at Risk.

3. METHODOLOGY

During the analysis we developed a LAVaR model family that assists institutions in determining their market risks using a liquidity-adjusted VaR methodology. The methodology belongs in the family of endogenous models, since it attempts to evaluate whether, based on the liquidity indicators of a given institution, a more accurate liquidity model can be created in comparison with the delta-normal VaR model. Our objective was to analyse whether a family of liquidity models can be developed that is applicable at the micro level, and which gives a more accurate, liquidity-adjusted analysis in the case of a given institution or company. Another objective of our research was to investigate how predictive such a model could be.

We tested the developed framework on stocks traded in liquid and illiquid markets, and compared the results with those of a similar model that does not take market liquidity indicators into account.

In the analysis, we studied the intraday data of three stocks. Essentially we focused on the Central and Eastern European market, but for the purposes of comparison we also included a developed-market liquid stock, to ensure that the differences between the various liquidity levels can be seen. It was using this database that development of the liquidity-adjusted VaR methods, and the principal component analysis, took place.

In the course of the devellopment we elaborated a liquidity-adjusted VaR model that, based on our tests, is capable of determining the market risks more accurately, albeit only slightly, than the classic delta-normal VaR model.

The specification of the model is as follows:

$$\mathbf{LAVaR} = \mathbf{G^{-1}}(\boldsymbol{\alpha}) \times \boldsymbol{\sigma} \times \sqrt{\mathbf{T}} \times (\mathbf{bx} + \mathbf{c}), \tag{4}$$

in contrast to the classic VaR specification:

$$\mathrm{VaR} = \mathrm{G}^{-1}(\alpha) \times \sigma * \sqrt{\mathrm{T}} \quad \text{, where} \tag{5}$$

$G^{-1}(\alpha)$ is the inverse of the standard normal distribution at the probability level,
σ is the standard deviation of the log yield of the analysed financial instrument,
T is the time in which the dimension of the standard deviation and log yield exists (in practice a day),
bx+c is a liquidity adjustment regression component that describes the covariance of the liquidity variables generated from the intraday data and the development of the following log yield for the following day.

The regression component must be modelled on the basis of the historical data, building on the liquidity variables. Due to possible market changes the reviewing of the model (that is, the liquidity regression) is unavoidable, but as in the case of all other Basel II models it is sufficient to perform this annually.

The input data on which the regression component is based can be real-world data, or the principal components obtained as the result of the principal component analysis. The process of determining the parameters of the model, and its analysis, are presented in detail in the following section.

4. DATA

The input data for the model can be divided into two groups:

- Data necessary for the VaR calculation: the yields of shares, their standard deviation and their correlations with each other.
- Data necessary for the quantification of liquidity: the most comprehensive possible determination of the individual dimensions of liquidity with various mono or multidimensional indicators.

In what follows we review the expected input data parameters.

4.1. Parameters of the VaR calculation

In our analysis we use the delta-normal method for the quantification of market risks. Therefore, two parameters of the normal distribution are needed: the expected value and the standard deviation. In this case we work on the assumption that the prices of the investigated shares have a lognormal distribution, and thus the yields follow a normal distribution. Since the portfolio is a linear combination of the instruments it contains, the portfolio will also have a normal distribution (*Jorion*, 1999). The first input parameter of the VaR calculation, therefore, will be the log yields of the individual instruments:

$$y_{i,T,t} = \log\left(\frac{P_T}{P_t}\right) \tag{6}$$

where $y_{i,T,t}$ the log yield of the portfolio's i-th security, between T and t. The term "log" denotes the natural logarithm, P_T the price of the share in T, and P_t the price of the share in t. The use of log yields has the additional advantage that the log yields can be added together, thus $y_{i,T,t_0} = \sum_{j=1}^{T} y_{i,j,j-1}$.

The distribution of the log yields, based on the central limit distribution, approaches the normal distribution with an increase of time frame. This phenomenon is observable mainly in liquid instruments on a one-year horizon; on a shorter horizon than this, however (for example intraday, overnight, weekly) it does not necessarily apply. On such a horizon the log yields are more leptokurtic; that is, flatter, and the probability of extreme cases is greater. Since the VaR is concentrated on the left-hand edge of the yield distribution, by attributing a lower probability to the extreme negative events it underestimates the risk. In our model, however, we attribute a more important role to the positive impact of the assumption of normal distribution, and accordingly we use the delta-normal method.

The log yields can be quantified in respect of different time frames depending on the available data. The calculation of overnight, monthly or annual yields is obvious, but in the case of intraday data the transactions are numerable, specifically between the market prices. For the calculation of market risk, as an input parameter of the VaR model the log yields are usually determined at daily level, as was the case in our analysis.

4.2. Parameters of the liquidity indicators

The input parameters of the market risk management model will include the variables, derived from the factor analysis, that concentrate the various dimensions of liquidity. In order for the model to determine as effectively as possible the potential loss arising from market liquidity, it is advisable to use as many variables as possible for the principal component analysis. The redundancy of variables causes no problems, as the matching types of variable are separate from the other types of variable, and together they add to the explanatory power of the principal component. The objective is for the principal components to describe the highest possible percentage of the variance.

The question of liquidity has already been studied by many; *Gyarmati* et al. (2010) listed the five dimensions of liquidity: tightness, depth, breadth, resiliency and immediacy.

The first three are static, and the last two dynamic dimensions. This grouping can be augmented with diversity, which measures the heterogeneity of the market.

Static dimensions interpret liquidity at given moment of the order book. Tightness quantifies the transaction cost of trading (for example the bid-ask spread). Depth shows the quantity of best offers on the ask and bid side, while the breadth accounts for the quantity of all offers above and below the market price. The former is customarily approached from the angle of market turnover, and the latter is determined through the quantification of price sensitivity.

Unlike the static dimensions, the dynamic dimensions show the change in liquidity during a given period. Resiliency captures the speed of the smoothing out of price fluctuations resulting from trading, while immediacy expresses the time in which a given portfolio can be sold or bought (*Gyarmati* et al., 2010). With respect to the dimensions of liquidity there are indicators that quantify a single dimension (mono-dimension indicators) and others that concentrate several dimensions of liquidity into a single indicator (*von Wyss*, 2010).

The quantification of as many of the dimensions of liquidity as possible, and their inclusion among the variables of the principal component analysis, should be supported. The scope for quantification of the variables depends on the available market data. In general terms one can say that the static dimensions are typically quantifiable and publicly available, but in most cases the use of the dynamic dimension requires a sophisticated database, which is not available to the majority of market participant, or possibly to anyone at all. One such example is the Budapest Liquidity Rate (BLM), which builds on the XLM developed by the Deutsche Böerse Group, but is only published in the form of monthly data, and is not quantified anywhere else apart from the German, Slovenian and Hungarian markets, and thus it cannot constitute a part of the principal component analysis.

4.3. Transaction data necessary for the reviewed parameters

For the trading volume, transaction data is needed with a depth that includes the times of the transactions related to the given instrument, and the number of shares involved in the concluded transaction. The turnover data can be calculated together with the prices of the transactions. The number of transactions, the liquidity rates and the flow rate will also be quantifiable given such data. It should be noted that we set out to elaborate a model that does not require a full knowledge of the order book, so of the five dimensions, we are able to measure the breadth and resiliency.

In respect of the shares under investigation, therefore, we need a specified period in which to access the data. In respect of this period the times of all the transactions, the traded quantity, and the traded price, need to be provided.

In practice the actual shares will be determined by the institution's share portfolio; for the building and testing of the model, however, it is necessary to set up a hypothetical portfolio. An expectation in this regard is that the portfolio should be made up of a sufficient quantity of different securities for the potential weak points of the model (for example the noise of the covariance matrix) to be revealed in the event of the quantification of the portfolio's risk, so that the parameter setting of the model receives feedback from the model testing, thereby improving its performance. A further expectation in respect of the securities is that the defined indicators should be quantifiable. The Hungarian market also has some illiquid stocks which have not been involved in trades for many years. These instruments should be avoided for the purposes of creating the model, because when selecting the securities of the hypothetical portfolio we make it a condition that the market capitalisation of the shares constituting parts of the portfolio should exceed HUF 1 billion.

4.4. Data search

The three stocks chosen by us were Zwack Unicum Zrt., supplemented with MOL and Tesla shares. The logic of our selection was as follows:

We are looking for a fairly liquid security, taking care not to choose one that has not been traded for years, since this would introduce considerable distortion to the model. Furthermore, due to their total lack of liquidity they would not have any data on which we could perform calculations. We decided that a Hungarian tock that is not included in the BUX index would be suitable for this purpose.

Our next step was to designate a share with liquidity that can perhaps be regarded as middling by global standards. Here the choice was the shares of the MOL group, the most important stock in the BUX index alongside OTP Bank. MOL is one of the favoured blue chips in Hungary. Although recent events related to oil prices and the oil industry have had a substantial impact on the company's shares too, we nevertheless chose this company over OTP because the factors influencing the latter's interests in the region represent a greater distorting force, because they do not relate to the sector as whole.

The next task was to find a share with very high liquidity. We chose Tesla. This was mainly for reasons of convenience. Apple, as the most traded stock in the world, has such a quantity of transactions that it exceeds the capacities of Excel even in a daily breakdown, so the data collection alone would have been very complicated and time consuming, which conflicts with our objectives. In Tesla's case we were able to retrieve the data in a monthly breakdown.

We retrieved the intraday data series necessary for the modelling using the Bloomberg system. This service provided by Bloomberg only stores the detailed information going back six months, so we do not have a longer horizon that this for the construction and testing of the model. The selected time frame related to deals made between 11/11/2013 and 09/05/2014.

Taking the ability to obtain the data into consideration, we introduced the following liquidity indicators:

Trading volume: $\boldsymbol{Q_t = \sum_{i=1}^{N_t} q_i}$, where

Nt the number of transactions in the period between *t*–1 and *t*,

qt the number of shares in the *i*-th transaction.

The trading volume shows what the total number of traded shares is for a given stock on the trading day in question. The higher the volume, the more liquid the security might be.

Turnover: $\boldsymbol{V_t = \sum_{i=1}^{N_t} p_i q_i}$, where

Nt the number of transactions in the period between *t*–1 and *t*,

qt the number of shares in the *i*-th transaction,

pt price in the *i*-th transaction.

The turnover shows the total value in which transactions were concluded with the given share on the trading day in question; in other words, it is nothing other than an auxiliary variable for calculation of the daily average price presented at the time of the examination of the basic data. The greater the turnover, the more liquid the instrument. The variable takes on a value of zero if no transaction takes place on the trading day in question.

Number of transactions:

Nt the number of transactions in the period between *t*–1 and *t*.

The number of transactions signifies how many transactions were concluded on a given day with a given security, regardless of the number of securities involved in the deal, or their price. The more deals were made on a given day, the higher the liquidity that this variable indicates. Accordingly, the data can only be natural numbers that is either zero or positive whole numbers greater than zero.

Liquidity rate 1: $\boldsymbol{LR1_t = \frac{V_t}{|r_t|} = \frac{\sum_{i=1}^{N} p_i q_i}{|r_t|}}$, where

Nt the number of transactions in the period between *t*–1 and *t*,

qt the number of shares in the *i*-th transaction,

pt price in the *i*-th transaction,

rt the yield in the period between *t*–1 and *t*.

The liquidity rates already take the impact of the price shift into account for the quantification of liquidity. Liquidity rate 1 is the quotient of the turnover and the absolute value of the daily log yield. Because both the turnover and the absolute value of the log yield can only be positive or zero, the variable may also only take on such values. Because the turnover can be very high, and, when generating the log yield we use a 0.01% log yield on days when the average price matches the average price of the previous day – and thus the yield would be zero – the Liquidity rate 1 can also take on a high value; in other words, the range is high. The relative range; that is, the range divided by the mean, is high at an average of 46. The relative standard deviation also supports this, as the average value of the variable for the 119 shares is 4.8. It can be concluded, therefore, that Liquidity rate 1 deviates significantly.

Liquidity rate 3: $LR3_t = \frac{\sum_{i=1}^{N}|r_i|}{N_t}$, where

Nt the number of transactions in the period between *t*–1 and *t*,

rt the yield in the period between *t*–1 and *t*.

Liquidity rate 3, similarly to Liquidity rate 1, uses the log yields in order to measure the performance of another liquidity ratio – in this case the number of transactions – not in absolute terms, but in relative terms. Because here the yields have been put into the numerator, the variable can take on low values. Both the numerator and the denominator can have a zero or positive value, so the value of the variable is also restricted. The lower limit is zero and the upper limit is 1, because when generating the log yields we set 1 as the value of the variable at which no transaction took place. In this way, unlike Liquidity rate 1, the variable can only vary within a small range. This was confirmed by the statistical test: as expected, in absolute terms 1 is the greatest range indicator, but what is more important is that in comparison to the other variable, in relative terms – divided by the mean – the variable's standard deviation and range are also far lower.

Flow rate: $FR_t = N_t \times V_t$, where

Nt the number of transactions in the period between *t*–1 and *t*,

Vt turnover (*Dömötör–Marossy*, 2010).

The flow rate is the product of two previously examined liquidity indicators, turnover and the number of transactions. Because both can take on a large value in a daily breakdown, the variable derived from them can also be large, but on low-turnover days the value of the variable could remain low. Accordingly, the flow rage may have a large range and standard deviation. Among the examined shares, the lowest flow rate was 5, and the highest 35 685 billion, so the total range really is very large. Again, dividing the values by the mean flow rate of the individual stocks gives us more manageable data. The relative range in the variable is higher than those previously examined, but still manageable. The relative standard deviation remained below 4 in the case of most examined shares.

For the modelling, therefore, we need the shares that we are including in the analysis; we need to determine which period to examine, and in respect of this period we need to provide the necessary input data, such as the times of all the transactions, the traded quantity, and the traded price. This makes it possible to quantify the defined liquidity indicators, but at the same time it will not be necessary to give too much information in order to obtain an effective model that requires less computation and data, but which nevertheless functions well.

4.5. Descriptive data analysis

In the tables below we present the descriptive statistics of the collected data:

Table 1
Descriptive statistical properties of the Zwack share data

	Price	Trading volume	Turn-over	Number of trans-actions	Liquidity rate 1	Liquidity rate 3	Flow rate
Average	14 262	106	1 403 525	5	741×10⁶	0.266%	13×10⁶
Standard deviation	817	123	1 518 660	4	1 229×10⁶	0.428%	23×10⁶
Median	13 745	55	738 880	4	259×10⁶	0.093%	3×10⁶
Skewness	1	2	1	1	3	2	3
Minimum	13 295	0	0	0	1×10⁶	0.000%	13 695
Maximum	15 980	555	5 304 180	15	5654×10⁶	1.776%	132×10⁶

Table 2
Descriptive statistical properties of the Tesla share data

	Price	Trading volume	Turn-over	Number of trans-actions	Liquidity rate 1	Liquidity rate 3	Flow rate
Average	57 314	4 801 683	275×10⁹	26 589	61×10¹²	0.0001%	8 849×10¹²
Standard deviation	3288	2 202 284	127×10⁹	12 661	124×10¹²	0.0000%	9 995×10¹²
Median	56 822	4 321 052	247×10⁹	23 716	16×10¹²	0.0001%	5 923×10¹²
Skewness	0	2	2	2	3	0	4
Minimum	49 896	1 156 386	66×10⁹	6 371	5×10¹²	0.0000%	424×10¹²
Maximum	68 034	14 094 310	886×10⁹	79 836	554×10¹²	0.0002%	70 708×10¹²

Table 3

Descriptive statistical properties of the MOL share data

	Price	Trading volume	Turnover	Number of transactions	Liquidity rate 1	Liquidity rate 3	Flow rate
Average	11 966	106 442	1 316×10⁶	535	686×10⁹	0.002%	1 228×10⁹
Standard deviation	795	101 975	1 402×10⁶	407	5 016×10⁹	0.001%	3 293×10⁹
Median	11 836	74 662	888×10⁶	416	125×10⁹	0.002%	367×10⁹
Skewness	2	3	3	3	11	1	6
Minimum	10 713	17 716	204×10⁶	169	26×10⁹	0.000%	38×10⁹
Maximum	14 980	721 405	9 998×10⁶	2 868	57 046×10⁹	0.006%	28 647×10⁹

Based on the data, the differences between the three shares are clear. The trading volume is exceptionally high in the case of Tesla. It is apparent that Zwack also had days when there were no deals involving the share, so its liquidity is also likely to be far lower than that of the other two shares. For Liquidity rate 1 the difference between the three shares is clearly observable, and this is also true of Liquidity rate 3. The Tesla shares stand out in terms of their liquidity compared to the other two, and it is clear that in respect of these indicators Zwack produces the lowest properties. In the case of the Liquidity rate 3 variable Zwack shows the highest values and Tesla the lowest, which leads to the conclusion that Zwack has the lowest liquidity, while Tesla has the highest liquidity among the three examined shares. The skewness of the data is positive in every case; in other words the data is concentrated on the left and does not have normal distribution. This indicator also supports our earlier liquidity hypothesis on the basis of the statistical properties.

5. METHOD OF UTILISING THE LIQUIDITY INDICATORS

Once we had determined the five liquidity factors defined as the input data, it was possible to begin the analysis of the data for each stock where we intended to use these factors. Our modelling, as we have mentioned already, builds on the delta-normal VaR model, making use of its advantageous properties. We have modified this to take into account the real-world situations in which the market is not liquid, and therefore it is not always possible to trade at the median price.

5.1. Principal component analysis of the variables

We processed the created database using the IBM SPSS software suite. We will perform the principal component analysis in respect of the six liquidity variables, but we have transplanted the database, and thus the ticker, the transaction date and price, in its entirety. Since we are performing the principal component analysis on daily data, in the course of the price calculation the database contains the previously calculated daily average price, and, in the case of the liquidity indicators, the daily data.

After the import we set all six liquidity variables up to be measured on a ratio scale, and set the appropriate format for the date and the variables. The database contained data for 14 414 records.

After setting up the analysis, we examined the correlation of the variables with each other, since the objective of the principal component analysis is, through a process of orthogonal transformation from variables that correlate with each other by pairs at some level, to generate uncorrelated principal components, where the first few principal components account for a large enough proportion of the total variance of the variables (*Kovács*, 2011). The correlation test is important therefore, but at the same time this could be set as an option in SPSS during the principal component analysis.

Based on the linear correlation coefficients, we conclude that the correlation between the individual variables is weak, and there is only a strong correlation, of close to 1, between the turnover and the flow rate, due to the fact that the flow rate is the quotient of the turnover and the number of transactions. On this basis, we expect that the variables will be clearly distinguishable based on the individual factors, but the flow rate will not have a significantly greater explanatory power than the turnover. Due to the low correlations, it is expected that only several principal components will be capable of ensuring sufficient explanatory power; we could only expect one or two variables to have a high variance explanatory power in the event of high correlation. We anticipate, therefore, that there will be no single clearly definable latent factor, but that the variables will determine the liquidity inherent in the 119 shares from several angles.

Table 4
Linear correlation analysis of the indicators

	Traded quantity	**Turn-over**	**Number of trans-actions**	**Liquidity rate 1**	**Liquidity rate 3**	**Flow rate**
Traded quantity	1	0.374	0.117	−0.001	−0.016	0.298
Turnover	0.374	1	0.139	−0.001	−0.017	0.946
Number of transactions	0.117	0.139	1	−0.006	−0.105	0.130
Liquidity rate 1	−0.001	−0.001	−0.006	1	−0.002	−0.001
Liquidity rate 3	−0.016	−0.017	−0.105	−0.002	1	−0.12
Flow rate	0.298	0.946	0.130	−0.001	−0.12	1

As already discussed, when examining the system of relationships between a given number of variables that correlate with each other, if we transform the original variables into uncorrelated variables, then this is a principal component analysis.

Performing a standard deviation breakdown of the observed variables, the following three components can be differentiated:

Total variance = common variance + individual variance + error variance (7)

Here the common variance shows that there is a common factor underlying the several variables, the individual variance indicates that there is a single factor underlying one variable, and the error variance is a measuring error. In a principal component analysis, we explain the common and individual variances together (*Kovács*, 2011).

Earlier we wrote that for p variant and n observation, as a rule of thumb, $n \geq 5p$ should be achieved. Since a minimum of 14 030 cases and 6 variables are available, sufficient data is available for the principal component analysis.

The variables differ in their units of measurement, so first either the variables have to be standardised, or instead of the covariance matrix we need to start out from the correlation matrix during the process, and this needs to be broken up into eigenvalues and eigenvectors. Another problem is that not only do the units of measurement differ between the variables, but there are also differences of scale within the variables themselves. Therefore it is important for the variance of the variables to be almost identical, because any variable with a high standard deviation would dominate the principal component. For this reason, we performed the principal component analysis on the correlation matrix instead of the covariance matrix.

We performed the PCA so as to ensure that we retained eigenvalues greater than 1. Based on this, we were able to choose three eigenvalues, which together explain 71.1% of the variance of the variables. The first principal component explains

36.4%, the second 18% and the third 17% of the variance. The first eigenvalue is the dominant one, but the other two eigenvalues of around 1 are significant for capturing the full explanatory power. This bears out the hypothesis that the information inherent in the liquidity variables can be captured using several principal components. The remaining three eigenvalues explain 29% of the total variance; two of them are around 0.8, while the third is close to 0. These contain the proportion of the information that we disregard in order to reduce the dimensions.

Table 5
Process of the principal component analysis

Principal component	Initial eigenvalue			Used eigenvalue		
	Value	Variance%	Cumulative%	Value	Variance%	Cumulative%
1	2.185	36.411	36.411	2.185	36.411	36.411
2	1.081	18.016	54.427	1.081	18.016	54.427
3	1.000	16.671	71.098	1.000	16.671	71.098
4	0.881	14.675	85.774			
5	0.803	13.387	99.161			
6	0.050	0.839	100.000			

The most important result of the study is shown by *Table 6*, which quantifies the linear correlation of the liquidity variables with the principal components. It is clear that each variable shows strong correlation with one principal component, while the exposure to the other principal components is low, typically around 0. The Traded quantity, the Turnover and the Flow rate show covariance with the 1st principal component; the Number of transactions and the Liquidity rate with the 2nd principle component; and the Liquidity rate 1 with the 3rd principal component. The Traded quantity displays the lowest correlation (0.562), while the other variables correlate more strongly with the three principal components.

Table 6
Result of the principal component analysis

Variable	Principal component		
	1	2	3
Traded quantity	**0.562**	–0.002	0.000
Turnover	**0.958**	0.110	0.004
Number of transactions	0.268	**–0.648**	–0.024
Liquidity rate 1	–0.003	0.021	**0.999**
Liquidity rate 3	–0.055	**0.796**	–0.047
Flow rate	**0.936**	0.121	0.004

As the final result of the analysis we conclude that the greater the value of the principal component in the given share, the more liquid the instrument. This is true of all three principal components, as the correlations are positive and a high value for the variable means high liquidity. The exception to this is Liquidity rate 3, but the correlation with the principal component is negative; in other words, a large principal component also represents liquidity.

The variables of liquidity can be captured with a total of three significant principal components in the PCA; however, there is no principal component that does not significantly build onto one of the input variables. Accordingly, the principal component analysis has become a form of regression, with every single input factor important in the process of describing the liquidity position; the principal component analysis has combined the correlating variables.

5.2. Model specification

Our liquidity-adjusted model is a delta-normal specification, assisted by an endogenous regression component. Of the model specifications found in literature recommendations, this contains the methodology best suited to the collected data, and compared to the traditional delta-normal specification this has the smallest deviance; it is the most suitable for analysing, all other factors being equal, the additional impact of the liquidity adjustment.

The model is supplemented in the following manner, using the delta-normal specification:

$$LAVaR = G^{-1}(\alpha) \times \sigma \times \sqrt{T} \times (bx + c) \tag{8}$$

where G is the distribution function of the standard normal distribution, σ is the standard deviation of the yield, T is the holding period, and $(bx+c)$ is the result of the liquidity measurement regression, specifying a higher correction for the illiquid market.

The regression estimate can be specified through the changes in price, the price impact, with a given holding period. What needs to be measured is the extent of the actual change in standard deviation with the delta-normal specification (in respect of the input parameter of the delta-normal model), assuming given input variables. That is

$$\frac{S_T}{\sigma} = bx + c, \tag{9}$$

where S_T is the standard deviation measured during period T, $\sigma \times \sqrt{T}$ is the time-adjusted standard deviation parameterized on the basis of the full period, the average relative standard deviation expected given the measured price change/VaR logic. Normally, the relevant variables of liquidity will explain the change in standard deviation, anticipating the likelihood of a higher standard deviation in illiquid periods. All the available data (all securities) must be used for the parametrizing of the regression, on order to ensure a robust estimate.

5.3. Results of the LAVaR model

In our single-variable analysis we established that the variables have a high standard deviation, and are very volatile, so in the interests of ensuring better manageability and usability, we substituted the original variables with the time line calculated from the 10 and 20-day moving average. Without exception, we can say that the data calculated with the 10-day moving average also represent a considerable improvement in the usability of the data, but the time lines calculated with the 20-day moving average, most of the time, come close to the linear trend that is necessary in order for the linear regression model to ensure a good result. The liquidity indicators, in their original form, do not yield an appreciably better result.

The moving average is also warranted because, for the purpose of the VaR calculation also, the calculation specification performs a form of historical averaging; the standard deviation of (recent) past data provides the framework for the VaR calculation.

Once the single-variable analysis of the data had taken place, we performed the 10 and 20-day regression estimate.

The goodness of the regression and the goodness of fit are given by the adjusted R^2. R^2 in itself shows the percentage by which the independent variables explain the dependent variable. In our case this means the extent, expressed in terms of a percentage, in which the established liquidity indicators explain the relative standard deviation that we have used as the indicator of liquidity. The adjusted R^2 also takes into account the number of variables. In the case of the Zwack stock, in the first, 10-day model the adjusted R^2 is 0.655; while in the second model, built up with a 20-day moving average, this value is 0.755. This means that the explanatory variables explain the relative standard deviation in an extent of 65% in the first case, and 75% in the second case. This is not surprising, as in the course of the individual data analysis we saw that with the 20-day moving average we were better able to bring the data closer to a trend, the linear trend, which is one of the prerequisites for linear regression.

Where the Zwack stock is concerned, with the first model we observe a standard error of 0.04, while for the second this figure is 0.02. This again shows that the second model is better than the first. The standard error refers to how much fluctuation is shown by a parameter obtained in our sample, for reasons attributable to the sampling. This figure tends to decrease as the size of the sample grows, but in our case this figure is almost half even for the second model, which contains fewer observations, so this model is far more favourable.

Based on an examination of the significance of the included variables, in both cases we find that the trading volume, as a variable, can be omitted from our models, because in both cases the p-value associated with this variable is higher than 0.05, which is the customary limit, and with a p-value in excess of this we no longer regard the indicator as significant. The other explanatory variables can be declared significant on the basis of the same values.

Liquidity rate 3 stands out from among the other explanatory variables, as based on both models this has the greatest impact on the independent variable. If Liquidity rate 1 increases by one percent, the relative standard deviation increases by 23.36 percent based on the first model, and by 20.56 percent based on the second model, this being the extent of the increase in the security's lack of liquidity.

We performed the estimate with only the significant variables and the principal components included in the model, but in this way we did not succeed in increasing the explanatory power of the model.

5.4. Development of a forward-looking model

The model developed by us, containing multi-dimensional liquidity indicators, gives a good estimate for the purpose of estimating the standard deviation given in the model specification. This is extremely important, because this enables us to adjust the VaR models for liquidity, so the model takes into account those real-world market situations when trading at the median price is not possible, and thus the Value at Risk of our portfolio changes due to the liquidity of the securities.

In this stage we altered the model to make it capable of determining as accurately as possible the change in distribution, and thus the higher exposure, not only retrospectively, but where possible in advance. This is an important criterion in order for the model to be a tool that is suitable for everyday use. We built up our model so as to makes the most effective possible use of the available six-month data series. The model took on the task of estimating the next day's distribution on the basis of the momentary data.

In this section, we present the results of regression using the variables calculated from the 10 and 20-day moving average of the three stocks that we examined. In the case of Zwack, we also performed these regressions, in keeping with the sensitivity analysis, on the various Liquidity rate 3 data within the 20-day model.

The R^2 indicator of the adjusted model came to 39%, which in this model is decidedly low, which in turn draws attention to the difficulties of forecasting, because while the model used the distribution in a non-predictive way, it had good explanatory power. The explanatory power of the model is low, so it is only capable to a small degree of estimating the changes in the relative standard deviation that we intend to model.

Using the versions prepared with a 20-day moving average, we observe that, in the case of the regression, the value at which we cap the Liquidity rate 3 variable, as a measure of liquidity, has no effect on the goodness of the model. The variable showing the goodness of the model only improves by 0.02% in response to this. However, in comparison to the 10-day model the explanatory power of the model increased, with an R^2 value of 58%, which is because this method allows us to span a longer time frame, and as we have seen, the variables conform better to longer-term trends in an individual analysis too.

5.5. Analysis of the results

The last phase of the analysis was the use of the results achieved thus far, in order to compare the delta-normal VaR model and the LAVaR model specified by us. In the case of both models we determined the model error for the VaR and LAVaR models, and also quantified the extent of the overruns in every case.

For the final comparison of the models, we first calculated the VaR and LAVaR values with the delta-normal method. The basic delta-normal VaR calculation was the following:

$$\text{VaR} = G^{-1}(\alpha) \times \sigma \times \sqrt{T}, \tag{10}$$

where G is the standard normal distribution at confidence level α, and standard deviation is the forward-looking, 60-day standard deviation calculated form the daily log yield of the given stock. We chose this in order to improve the estimation ability of the model, and given the availability of the appropriate quantity of data, this choice is satisfactory. In our case a very short, six-month time line is available to us, but our aim is to test the goodness of the model, so at present we will disregard this negative factor. In our model the value of α is 1%, so we observe the lower side. The reason for this is simply that, in our opinion, this makes it easier to visually perceive the difference between the two models on the diagrams. Naturally we also took into account the fact that we are not dealing with a daily distribution when calculating the time, as well, so we brought this into line with the 60-day observation too.

We modified the formula shown above as described earlier, in order to ensure that the less liquid period are also taken into account in the calculation:

$$\text{LAVaR} = G^{-1}(\alpha) \times \sigma \times \sqrt{T} \times (bx + c) \tag{11}$$

The expanded term also contains the results of the previously performed regression component for the individual models.

After calculation of the VaR and LAVaR, we determined the model error in the following manner: if the received VaR and LAVaR value was greater than the log yield for the same day, then this represents an error in the model, because the Value at Risk is higher than the real-world change in value (yield). We indicated the model error with a binary code: if we received a larger value we worked with an error indication, and the data was given a code of 1, otherwise a code of 0 was used. By totalizing the ones we received the number of model errors in the individual cases.

Following this we also quantified the extent of the overruns, thus ensuring an indicator for the goodness or error of the finished model. We determined this data by taking the difference of the calculated VaR and LAVaR, and the log yields, thereby quantifying the extent of the overrun, and then multiplied this with the binary code used earlier, to ensure that it is only calculated for the erroneous data. By totalizing these overruns we ascertain the total extent of the overrun produced by the model as a whole.

First we compare the 10-day moving average-based models in the table below, then we also display the obtained results in the form of charts.

Table 7
Model errors of 10-day moving average models

	Model error (VaR)	Number of model errors (LAVaR)	Extent of overrun (VaR)	Extent of overrun (LAVaR)
Zwack	6	8	3.57%	3.37%
MOL	6	4	2.98%	2.72%
Tesla	9	8	13.31%	13.41%

It is also clear from the summary table that in the case of the MOL and Tesla shares the LAVaR model results in fewer errors, but in the case of the very illiquid Zwack stock, the delta-normal VaR model gives fewer errors. The extent of the overrun for both Zwack and MOL is greater on the basis of the delta-normal VaR calculation than in the case of the model developed by us, while for the Tesla shares our model produces a 0.1% higher overrun.

A summary of the 20-day models is shown in the table below:

Table 8
Model errors of the 20-day moving average models

	Model error (VaR)	Number of model errors (LAVaR)	Extent of overrun (VaR)	Extent of overrun (LAVaR)
Zwack	2	2	0.53%	0.59%
MOL	4	2	1.06%	1.24%
Tesla	2	2	4.09%	4.24%

With these models – disregarding the MOL shares – with regard to model error there was no difference between the VaR and LAVaR model; however, the extent of the overrun is smaller in every case when the Value at Risk is calculated with the delta-normal VaR model, than it is in the case of the model developed by us.

6. CONCLUSIONS

In our analysis we examined the aspect of market risk that relates to the liquidity of the portfolio and to the individual stocks. We augmented the commonly used, easily manageable delta-normal VaR method with a regression term with which, using the liquidity indicators defined and used by us, we intended to incorporate market liquidity into our model.

In the first phase of the model building we fitted the model to existing data; in other words, we obtained a model that fits onto the already known and integrated data, and measures the strength of covariance. In these cases the results were encouraging; the explanatory power of the models was above 80%, and the model's error relatively low. We found that the models using a 20-day moving average gave a more accurate estimate than the models which only used a 10-day moving average. This is due to the better linearity.

After obtaining favourable results in the dimensions that we had analysed, we decided that a forward-looking model would be far more favourable, and therefore in this phase of the modelling we utilised the data with the aim of predicting the non-integrated and utilised data. In this case too, we remained with the 10 and 20-day moving average-based models. This concept, however, was limited significantly by the fact that we were only able to extract a six-month data series from the Bloomberg terminal. Through the use of the moving average our data set was shortened, and the fact that we were developing a forward-looking model led to a further decrease.

The tables showing the results also confirm that, with a forward-looking model, in the case of an illiquid stock the LAVaR model estimates market risk with a greater error, which is precisely the opposite of the outcome that we set out to achieve. In the case of liquid stocks, based on the modelling performed by us, we can state that it entails fewer errors; that is, it is capable of giving a more accurate VaR, than the delta-normal method. As a result of this, banks could determine their regulatory capital more accurately, thereby reducing the opportunity cost of funds that are not lent out.

Although the model is not capable of correctly forecasting insufficient liquidity, it does effectively show the current risk. Naturally this begs the question of what causes this effect – presumably the liquidity level is capable of changing very rapidly for every type of stock, so the indicators of liquidity are only suitable for describing a concurrent relationship, and are incapable of providing a forecast of how the institution's liquidity will develop even over the short period ahead, in the next 10-20 days.

The model can be developed further, insofar as data of this depth is available for a longer time frame of at least 4-5 years, and a sufficient quantity of data is on hand for both the development and testing of the model. At the same time, it is important to note that the liquidity and the distribution examined by us, in our experience, is a concurrent phenomenon, so it is conceivable that even on a longer time line a forward-looking model would not yield a significantly better result, and in this case where liquidity-adjusted VaR calculations are concerned we must turn to a different logic for the generation of capital reserves, as the existing endogenous LAVaR methods are not suitable for this.

REFERENCES

Bangia, Anil – Diebold, Francis X. – Schuermann, Til – Stroughair, John D. (2002): *Modeling Liquidity Risk, with Implications for Traditional Market Risk Measurement and Management. Risk Management: The State of the Art.* The New York University Salomon Center Series on Financial Markets and Institutions, Vol. 8, pp. 3–13.

Dömötör, B. – Marossy, Z. (2010): A likviditási mutatószámok struktúrája. [The structure of liquidity indicators.] *Hitelintézeti Szemle* 9 (6), pp. 581–603.

Ernst, C. – Stange, S. Kaserer, C. (2008): Accounting for non-normality in liquidity risk. CEFS Working Paper No. 2008-14, http://ssrn.com/abstract=1316769.

François-Heude, A. – Van Wynendaele, P. (2001): Integrating Liquidity Risk in a Parametric Intraday VaR Framework. Working paper, IST EIN (Istituto Einaudi), http://www.istein.org.

Giot, P. – Gramming, J. (2005): How large is liquidity risk in an automated auction market? *Empirical Economics*, 30 (4), pp. 867–887.

Gyarmati, Á. – Michaletzky, M. – Váradi, K. (2010). The Budapest Liquidity Measure and its application – Liquidity risk in VaR measure. Working paper, Budapest Stock Exchange.

Jorion, P. (1999). *A kockázatott érték* [Value at Risk]. Budapest: Panem.

Kovács, E. (2011): *Pénzügyi adatok statisztikai elemzése* [Statistical analysis of financial data]. Budapest: Tanszék Pénzügyi Tanácsadó és Szolgáltató Kft.

Lawrence, C. – Robinson, G. (1997). Liquidity, dynamic hedging and value at risk. *Risk Management for Financial Institutions* 1 (9), pp. 63–72.

Michaletzky, M. (2010): Pénzügyi piacok likviditása. [Liquidity of financial markets.] PhD dissertation, Budapesti Corvinus Egyetem.

Qi, J. – Ng, W. L. (2009): Liquidity adjusted intraday value at risk. Proceedings of the World Congress on Engineering, Vol. II. WCE, July 1-3, London, U. K.

Radnai, M. – Vonnák, Dzs. (2009): Likviditási kockázat az Európai Tőkemegfelelési Direktíva tervezett módosításában. [Liquidity risk in the planned amendment of the European Capital Requirements Directive.] *Hitelintézeti Szemle* 8 (3), pp. 248–256.

Sajtos, L. – Mitev, A. (2007): *SPSS kutatási és adatelemzési kézikönyv.* [*SPSS research and data analysis manual*] Budapest: Alinea.

Stange, S. – Kaserer, C. (2009a): Why and how to integrate liquidity risk into VaR- framework. CEFS working paper, http://ssrn.com/abstract=1292289.

Stange, S. – Kaserer, C. (2009b): Market liquidity risk – An overview. Working Paper Series, Center for Entrepreneurial and Financial Studies, http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1362537.

Váradi, K. (2012): Likviditási kockázat a részvénypiacokon. [Liquidity risk in the equity markets.] PhD dissertation, Budapesti Corvinus Egyetem.

Wyss, V. R. (2004): Measuring and Predicting Liquidity in the Stock Market. PhD Dissertation, Universität St. Gallen.