ON MEASURING THE COST OF LIQUIDITY IN THE LIMIT ORDER BOOK

M.M. Dyshaev

Chelyabinsk State University, Chelyabinsk, Russia
Mikhail.Dyshaev@gmail.com

The work is devoted to the elaboration and demonstration of a method for measuring the cost of illiquidity in the delta hedging of futures-style options on the Moscow Exchange. Illiquidity is usually measured per unit of asset or money. However, given the specifics of futures, this article proposes to use the total volume of the initial margin and the state of the limit order book. In this case, it becomes possible to compare the cost of illiquidity for various futures. The most liquid futures that are traded on the Moscow Exchange are analyzed and the cost of liquidity is compared.

Keywords: limit order book, illiquidity, hedging, initial margin.

Introduction

Recently, a number of works have been published on the issue of taking into account the influence of the state of the limit order book (LOB) during delta hedging of options [1–12].

At hedging of options, a trader buys or sells a certain amount of a basic asset at various points in time. In most of papers, devoted on issues of the delta hedging of options, the market price typically uses for accounting. This is the price, where the underlying asset is bought or sold in models. The main approach is to consider the market price as the average price between the best bid price $s_{\text{bid}}$ and the best offer price $s_{\text{ask}}$, i.e.

$$\bar{s} = \frac{s_{\text{bid}} + s_{\text{ask}}}{2}.$$  

However, as the practice shows, if the order amount $h$ (for example, the buy order) is greater than the amount of an asset at the best price $s_{\text{ask}}$, the trader will pay an amount almost always greater than $\bar{sh}$. The value $\bar{sh}$ often called the mark-to-market value. In fact, the trader buys at weight price

$$s^f(h) = \frac{1}{h} \sum_{i=1}^{N(h)} a_i s_i, \quad \sum_{i=1}^{N(h)} a_i = h,$$

where $a_i$ is bought amount of the underlying asset at price $s_i$ in the limit order book. In practical terms, if a trader submits market orders for delta hedging options without taking into account this price difference, he or she will have some loss of money.

There is an extensive literature on the optimal execution of large orders. We list only a few of them [13–17]. In these works, the effects of the temporary price impact and the permanent price impact are studied. Sometimes these two types are complemented by the third type, transient price impact.

The reported study was funded by RFBR according to the research project no 19-01-00244.
On the other hand, the case with absolute resiliency is of interest. It corresponds to the situation when, after a trader’s market order fills, the limit order book quickly returns to its previous state. There is no the permanent price impact in this case. The trader pays the costs associated with the fact that he or she lacked the underlying asset at the best price.

Since the difference between the mark-to-market value and the realized cost is a consequence of a not perfect liquidity, the main purpose of this work is to demonstrate and compare two methods for accounting illiquidity. These methods were chosen because they allow to visually show the cost of the liquidity in the case of the strong resiliency. And they can be easily applied to practical issues. The first method bases on the article of L.C.G. Rogers and S. Singh (2010) [7]. The second method of the measurement of the illiquidity presented in the work of P. Malo and T. Pennanen (2012) [18].

For practical purposes, an estimate will be made of the cost of the illiquidity in transactions with futures on the Moscow Exchange. Therefore, unlike earlier works, this article additionally studies the measurement of the illiquidity, depending on the required initial margin [19] actually paid by the trader to the exchange when buying or selling a futures contract. When a trader buys or sells a futures contract, the exchange or the broker blocks the initial margin on the trader’s account. Therefore, a trader cannot gain an infinitely large position. This practice is aimed at reducing risks in the financial system. The initial margin depends of many parameters, such as exchange rates, interest rate curves and volatility curves (see [19] for details). Since the initial margin already takes into account most market parameters, the use of these units will allow a more accurate comparison of various futures contracts.

The following section provides a brief description of the both methods. After that, the application of the above methods is considered for most liquid futures contracts that are traded on the Moscow Exchange. Main results and the direction of future research are presented in the last section.

1. How can take the illiquidity into account?

A brief description is given of how to measure the illiquidity of an existing limit order book (LOB) in the case of the strong resiliency. Suppose a trader needs to buy \( h \) units of an underlying asset to hedge an option. A trader submits a buy market order to the limit order book. A market order implies sequentially filling the limit orders available in the LOB until the required amount of \( h \) is collected. LOB quickly fills new orders after the trader’s transaction is completed.

1.1. The measurement of illiquidity in the model of L.C.G. Rogers and S. Singh

The authors describe the liquidity as the non-linear transaction cost. A market is considered without a permanent impact on the price. A trader buys \( h \) units of an underlying asset

\[
h = \frac{s^m(h)}{\rho(s)} ds,
\]

where \( s^m(h) = s_N(h)/s \) is the marginal relative price when the trader buys \( h \) units of the underlying asset. Thus, \( s^m(h) \) is the last, maximal (for the buy market order) relative
price. The total cost that the trader will pay to depend upon $h$, is

$$T(h) = hs^f(h) = \bar{s} \int_1^{s^m(h)} s \rho(s) ds,$$

where $\rho(s)$ is the density of the underlying asset orders volume in the LOB. The difference between the mark-to-market value and the total cost $T(h) = hs^f$ is the cost of the illiquidity:

$$T(h) - h\bar{s} = \bar{s} \int_1^{s^m(h)} s \rho(s) ds - h\bar{s} = \bar{s} \int_1^{s^m(h)} (s - 1) \rho(s) ds.$$

Thus, the cost of the illiquidity, expressed in the amount of money, takes the form

$$T(h) - h\bar{s} = \bar{s} \int_1^{s^m(h)} (s - 1) \rho(s) ds = \bar{s} l(h),$$

where

$$l(h) = \int_1^{s^m(h)} (s - 1) \rho(s) ds. \quad (1)$$

The authors of [7] used the cost function $l(h)$ for $h$ units of an underlying asset in the form

$$l(h) = \frac{1}{2} \varepsilon h^2,$$

where $\varepsilon$ is a small parameter. For numerical solutions, the authors take $\varepsilon = 0.006$, based on practical measurements. This form of $l(h)$ was chose for reasons of the tractability of the HJB equation (see [7, Remark and Equation (3.12)]).

1.2. The measurement of illiquidity in the model of P. Malo and T. Pennanen

As monetary measure of the illiquidity authors [18] suggested to use the function, that in our notations has the form

$$r(h\bar{s}) = \ln(s^m(h)). \quad (2)$$

The function $r(\cdot)$ is called the relative price impact curve. It depends from the mark-to-market value of the market order for $h$ units of the underlying asset. If at the best bid or ask price in the limit order book there are many units of an asset, then $r = 0$. If trader can not to buy a needed volume at the best ask price, then $r > 0$ and this corresponds to the presence of the illiquidity. The relative price impact curve $r$ is a non-decreasing, piecewise constant function of the units of an underlying asset.

Next, $s^m(h)$ can be expressed in terms of the mark-to-market value as $s^m(h) = e^{r(h\bar{s})}$. And then the total cost may be presented (in our notations) as

$$T(h) = \bar{s} \int_0^h s^m(z) dz = \bar{s} \int_0^h e^{r(z\bar{s})} dz,$$

and the cost of the illiquidity, expressed in the amount of money, takes the form

$$T(h) - h\bar{s} = \bar{s} \int_0^h e^{r(z\bar{s})} dz - h\bar{s} = \bar{s} \left( \int_0^h e^{r(z\bar{s})} dz - h \right).$$
The authors considered the function $r$ as $r(hs) = \beta hs$, where $\beta > 0$ characterizes the illiquidity in the LOB. A smaller $\beta$ corresponds to a greater liquidity. Then the total cost is (cf. [7, Example 2.1 (2.2)])

$$T(h) = \bar{s} \int_{0}^{h} e^{\beta z} dz = \frac{e^{\beta h} - 1}{\beta}.$$

Thus, the cost of the illiquidity, expressed in the amount of money, takes the form

$$T(h) - \bar{h}s = \frac{e^{\beta h} - 1}{\beta} - \bar{h}s = \frac{1}{\beta} (e^{\beta h} - \beta \bar{h}s - 1).$$

Since

$$\lim_{\beta \to 0} \left( \frac{1}{\beta} (e^{\beta h} - \beta \bar{h}s - 1) \right) = 0,$$

the cost of the illiquidity also shrinks to zero. This corresponds to the horizontal constant line $r$ and the infinity of the liquidity.

The authors find values of $\beta$ for the bid and ask sides of the LOB separately for many stocks, that trades on the Copenhagen Stock Exchange.

1.3. Empirical data

300 LOB «snapshots» were used as primary data for the cash-settled futures on Brent oil (BR-3.20, BRH0, expiration 02 March 2020), the cash-settled RTS Index futures (RTS-3.20, RIH0, expiration 19 March 2020), and cash-settled futures on US dollar — Russian ruble (USD/RUB) exchange rate (Si-3.20, SIH0, expiration 19 March 2020). The LOB dynamics is studied in the range from 15:36 to 16:05 MSK at 17/02/2020. On the both sides of the LOB 50 best price levels are presented. The average time interval between fixing of the LOB state is about 5 seconds. Most of parameters are presented in Table 1 (see [20] for details).

<table>
<thead>
<tr>
<th>Table 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>The parameters for various futures contracts on the Moscow Exchange on 17/02/2020</td>
</tr>
<tr>
<td>Contract size (lot)</td>
</tr>
<tr>
<td>Quotation</td>
</tr>
<tr>
<td>Price tick</td>
</tr>
<tr>
<td>Value of price tick</td>
</tr>
<tr>
<td>Initial margin (buy) $m_b$</td>
</tr>
<tr>
<td>Initial margin (sell) $m_s$</td>
</tr>
</tbody>
</table>

From data of the LOB snapshots the empirical values of $h$, $s^m(h)$ and $l(h)$ are obtained. In the model of L.C.G. Rogers and S. Singh, the illiquidity is measured in the amount of an asset that a trader buys or sells ($h_b > 0$ or $h_s < 0$). In P. Malo and T. Pennanen model, the determining factor is the mark-to-market value $\bar{sh}$. Since this article takes a similar approach to assessing the illiquidity for futures, the initial margin ($m_b$ or $m_s$) is taken as a variable in the measurement of the illiquidity.

Instead of units of assets $h$ or the mark-to-market value $\bar{sh}$ we used the value of the total initial margin for $h$ units adjusted to US dollar:

$$mh = \begin{cases} 
\frac{m_b}{USD/RUB} h, & \text{if } h > 0; \\
\frac{m_s}{USD/RUB} h, & \text{if } h < 0;
\end{cases}$$
where USD/RUB is the current average spot exchange rate. Here it would be more correct to also take into account the difference between the USD/RUB bid and ask prices, but for the simplicity USD/RUB = 63.40 is used.

2. Main results

2.1. The shape of the limit order book

First of all, let see on the function \( l(h) \) from equation (1) on Figure 1. Three functions present the LOBs for futures on the Brent oil, the RTS Index and the USD/RUB exchange rate. If to assume that the dependence is quadratic in the form of \( l(h) = \frac{1}{2} \varepsilon(mh)^2 \) as in [7], then we obtain the values for \( \varepsilon \), which are presented in Table 2.

![Figure 1](image)

**Fig. 1.** The function \( l(h) \) for the futures on Brent oil, the RTS Index and the USD/RUB exchange rate. For the graphs, the mean values of \( \varepsilon \) are used here. The widest line corresponds to the futures on the RTS index, which has the least liquidity. The thinnest line shows the futures for the USD/RUB exchange rate, which is the most liquid. The horizontal axis presents the total volume of the USD-adjusted initial margin \( mh \) (in US dollar millions).

The average empirical value of \( \varepsilon, \beta \) and \( \gamma \) for various futures contracts on the Moscow Exchange on 17/02/2020. The significance level is \( \alpha = 0.05 \).

<table>
<thead>
<tr>
<th></th>
<th>Brent</th>
<th>RTS index</th>
<th>USD/RUB</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \varepsilon )</td>
<td>((3.90 \pm 0.45) \times 10^{-11})</td>
<td>((1.07 \pm 0.02) \times 10^{-11})</td>
<td>((5.01 \pm 0.25) \times 10^{-11})</td>
</tr>
<tr>
<td>( \beta )</td>
<td>((2.83 \pm 0.16) \times 10^{-9})</td>
<td>((2.54 \pm 0.04) \times 10^{-9})</td>
<td>((0.79 \pm 0.09) \times 10^{-9})</td>
</tr>
<tr>
<td>( \gamma )</td>
<td>((1.12 \pm 0.13) \times 10^{-8})</td>
<td>((1.66 \pm 0.03) \times 10^{-8})</td>
<td>((2.51 \pm 0.13) \times 10^{-8})</td>
</tr>
</tbody>
</table>

Another methods to evaluate the illiquidity in the LOB are based on the study of the relative price impact curve \( r(s^m) \) from equation (2). On Figure 2 the function \( r(h) \) for the futures on the Brent oil, the RTS Index and the USD/RUB exchange rate are depicted. The horizontal axis presents the total volume of the USD-adjusted initial margin \( mh \) in US dollar millions. The relative price impact curves are piecewise constant increasing functions. For the good visualization, the figure shows only the right points of the constant sections.

Interpreting the curve \( r(\bar{sh}) \) in the same way as in [18], the following points can be noted. Based on the slope of the \( r \) curve, we can conclude that the most liquid underlying asset is USD/RUB futures. The least liquid is the futures on the RTS index. It should be noted that the calculated value of \( \beta \) for the Brent futures is slightly larger than for...
the RTS futures (see Table 2). This is the result of applying the least squares method to all data for Brent futures. If only central values of $mh \in (-1.5, 1.5)$ are taken into account, the data can be successfully approximated by a straight line. The second point is that there is a difference between the bid and ask side of the LOB.

If to assume that the dependence is linear in the form of $r(\bar{sh}) = \beta mh$ as in [18], then we obtain the values for $\beta$, which are presented in Table 2.

2.2. The total cost of illiquidity

On Figure 3 the total costs of the illiquidity $T(h) - \bar{sh}$ for the futures on the Brent oil, the RTS Index and the USD/RUB exchange rate are depicted. The horizontal axis presents the total volume of the USD-adjusted initial margin $mh$ in US dollar millions. This functions allow to compare values of the cost of the illiquidity for each futures. As shown in Figure 4, for deals in $h$ units that require less than 0.1 mio US-dollar for initial margins, the cost of the liquidity is approximately the same for the futures in question. The majority of retail investors operate in this interval.

Again, if to assume that the dependence $T(h) - h\bar{s} = \gamma(mh)^2$ for example, then we obtain the values for $\gamma$, which are presented in Table 2.

3. Conclusion

Two methods of assessing the insufficient liquidity were used in the work. They gave the same qualitative results, allowing to estimate in advance the total cost of the insufficient liquidity for a trader. This is important, as in most models it is believed that a trader hedges his position on the mark-and-market value. In this case, the cost of the illiquidity, of course, is zero. This is the discrepancy with the practice, especially for large orders.
Large participants use such methods of selling or acquiring an underlying asset as an iceberg or splitting a large order into small lots. This allows to reduce the cost of the illiquidity. Small traders often make trades within the spread of the best bid and ask prices.

Thus, our study is suitable for a mid, not large and not small, trader. On the one hand, the funds of such a trader are not enough to place iceberg orders and there are no other resources to perform hedging that is uniform in time or volume. On the other hand, the volume of her or his order exceeds the cumulative volume of the orders at the best bid and ask prices in the limit order book.
The paper suggests using the initial margin to compare liquidity costs for different futures. The initial margin can also be the basis for comparing other parameters, since many market indicators are taken into account in the calculation.

References
Работа посвящена разработке и демонстрации метода измерения стоимости неликвидности при дельта-хеджировании маржируемых опционов (futures-style options) на Московской бирже. Неликвидность обычно измеряется на единицу актива или денежных средств. Однако, учитывая специфику фьючерсов, в данной статье предлагается использовать общий объём начальной маржи и состояние лимитной книги ордеров. В этом случае становится возможным сравнить стоимость неликвидности для различных фьючерсов. Анализируются наиболее ликвидные фьючерсы, которые торгуются на Московской бирже, и сравнивается стоимость неликвидности при операциях с ними.

Ключевые слова: книга лимитных ордеров, неликвидность, хеджирование, начальная маржа.

Поступила в редакцию 19.01.2020
После переработки 28.02.2020

Сведения об авторе

Дышаев Михаил Михайлович, кандидат физико-математических наук, заведующий научно-исследовательской лабораторией финансового моделирования, Челябинский государственный университет, Челябинск, Россия; e-mail: Mikhail.Dyshaev@gmail.com.

Исследование выполнено при финансовой поддержке РФФИ в рамках научного проекта 19-01-00244/19.