Emerging Markets Group

EMG Working Paper Series

WP-EMG-19-2009

'MICEX vs RTS, Battle of Exchanges: Who Wins the Order Flow Supremacy'

Slawa Roschkow, Ian Marsh and Natasa Todorovic

February 2009

Emerging Markets Group
Cass Business School
City University
106 Bunhill Row
London
EC1Y 8TZ
UK

www.cass.city.ac.uk/emg/

MICEX vs RTS

Battle of Exchanges: Who wins the order flow supremacy?

Slawa V. Roschkow[1]

Ian Marsh[2]

Natasha Todorovic[3]

Abstract

Using high frequency data for seven stocks traded on the Russian MICEX and RTS exchanges, we apply established methodology and various time frequencies to reconstruct the order book and time series of transactions, as well as examine the price discovery process for both exchanges. The evidence shows a common stochastic trend amongst the RTS and MICEX pricing, however, the overall result seems to indicate a clear relationship. Further, although the causality relationship is asymmetrical, the price discovery occurs more in the denser MICEX market, with a yet also statistically significant role for the competing RTS market.

This version: February 2009

[1] Doctoral student, Cass Business School, London. Corresponding author. Contact: s.roschkow@city.ac.uk

[2] Professor of Finance, Cass Business School, London

[3] Senor Lecturer in Investment Management, Cass Business School, London

1 Introduction

With advancement of globalisation of financial markets, increasing number of firms has been cross- listing their shares both overseas and on multiple domestic markets. This has implied some of the key questions that have been the focus of specialised research performed in the market microstructure of multi- market trading such as: Has the price determination been affected by the new competition for order flow from multiple markets trading the shares? Has the increased trading in the overseas (alternative domestic) or at home market permanent or just a transitory effect of the listing event itself? Has the information that arises in the new overseas (alternative domestic) markets contributed to price discovery? On the contrary, have the new markets contributed to greater fragmentation instead and thereby generated systematic deviations from price parity or opportunities for arbitrage? Has this automatically lead to a more liquid trading environment for the shares? And has this listing represented a zero sum game with increased trading in the overseas market being offset by reduced trading in the home market?

The nature of cross- listed stock makes the market dynamics process is more complicated by enabling market- makers from more than one market to compete for order flow. It appears that the home market plays a major role in price discovery however some studies support the notion that, new markets- especially these which are abroad playing an increasingly important role. The trading environment itself is an important consideration for cross-sectional variation across markets, which could be formulated as a stylised fact: the contribution of a market to price discovery is greater if the fraction of global trading volume which takes place in the new market is higher. However, whether this relationship has been characterised by more permanent or just a transitory effect in the markets, the causality of that relationship still remains an open question, particularly in emerging markets. With the exception to Asian emerging markets, none of the above reviewed research has investigated the cross-listed securities on Russian exchanges.

It is well known that the stocks of Russian companies are cross- listed locally as well as abroad in the form of American Depositary Receipts (ADRs). Yet, the pricing behaviour of the stocks traded across domestic and foreign markets are not necessarily identical. Generally, there are a large number of papers in cross- listing field of microstructure literature, which examines domestic as well as international markets. However there is a

debate in conclusions based upon the vast range of data sampling frequencies. According to the existing research literature there is evidence that the denser home market may have the determining position. Although the inferences made are firm, the samples of the reviewed literature hardly touched the subject of emerging markets especially Russian domestic market.

The objective of this paper is to examine the price discovery process across two Moscow stock exchanges: Russian Trading System (RTS) and Moscow Interbank Currency Exchange (MICEX). The contribution to the cross- listing market microstructure research area would be justified by the unique data set and the limit order books reconstruction methodology. This study ought to provide insight into information flow and transmission between the competing domestic, by examining most liquid cross- listed shares of Russian companies traded in order driven mode on MICEX and RTS stock exchanges, namely: RAO UES, MMC Norilisk Nickel, Lukoil, Rostelecom, Sibneft, Surgutneftgaz and Tatneft. The research is based on limit order book reconstructed from intraday databases. It is hypothesised that the price discovery would occur to a larger extent in the most liquid market.

2 Literature Review

There has been a vast research literature devoted to the issues of the cross listed securities, where recent work by Karolyi (2006) may be used as the primary reference to detailed review of this growing sphere of literature. The market micro structure research field in cross listed securities could be categorised into five main areas: 1) multi- market trading, 2) price discovery, 3) order book reconstruction, 4) arbitrage opportunities and 5) liquidity. However, given the nature of this research, here we review the first three of the aforementioned areas. Nevertheless, none of the studies below has investigated the relationship between the Russian stock markets while employing a unique intraday data set taken directly from the underlying exchanges.

2.1 Multi- market trading

In the context of multi- market trading, the key contributions were inspired by theoretical models of Kyle (1985), Admati and Pfleiderer (1988) and Chowdhry and Nanda (1991), who examine the interaction among market makers, private information based smart traders and information less noise liquidity traders. When the markets are “thick” with other liquidity traders, information based smart traders seek to mask their information by timing their trading. These theories suggest explanations about the price impact of these trades by

examining volume- volatility relationships in these special circumstances, about clustering of trading volumes in some markets and while not in others and about clustering of trading volume around market opens and closes.

The Stulz (1999) critique is a recent catch up to the research interest in multi- market trading and liquidity. Technological developments such as more rigorous research methodologies at the researchers disposal and higher frequency data available for more markets around the world, provide a gaining insight into liquidity such as spreads, volume and volatility changes for newly cross- listed firms that may be related closely to changes they have involved in the information environment, the corporate governance systems and ownership structure of a firm. After all, despite the progress made, open question remains about the causality relationship between market intermediaries such as brokers and traders and the liquidity that arises in the competing markets for their shares, particularly in the emerging markets.

2.2 Price discovery

In the area of price discovery, there are numerous papers that attempted to address the issue which of the markets contributes to the price discovery dominantly or submissively on average, i.e. is it the domestic versus foreign or most liquid versus less liquid market? The available literature seems not to reach a consensus because of different methodologies and quality of data utilised. The field of price discovery has been pioneered by studies of Harris et al. (1995, 2002) and Hasbrouck (1995) which have examined the relative contribution to the price discovery between NYSE and regional exchanges of domestic stocks trading on these exchanges. Although using different methodologies, the above mentioned studies reveal significant price discovery in both the home and foreign market.

The early price discovery evidence from low- frequency daily data, as seen for example in Kato et al., (1991), Wahab et al., (1992), Park and Tabakkol, (1994), Miller and Morey, (1996), indicates that the issue of price discovery for cross- listed shares remains rather unresolved. Generally, these studies are not truly conducting tests for price discovery in the lead- lag relationship sense, however are examining links in pricing across markets. Inferences may have been drawn, however the results are mixed and some studies find the US prices to be dominant, majority of low- frequency evidence indicates that the home market being the dominant source of pricing.

A more recent study, Baruch et al (2005), have provided a theoretical model and empirical support for trading volume of cross- listed firms to be concentrated in the market with the highest correlation of cross- listed asset returns with other asset returns in that market. Harris et al (2003) provide a connection between liquidity, information, and home bias in international investment. Domestic investors may be better informed about and better able to monitor local firms than foreign firms.

Grammig et al (2004, 2005) have investigated this issue and found that price discovery occurs largely in the home market, while the impact of the exchange rates on price discovery seems to be insignificant. The results are similar to Phylaktis and Korczak (2004, 2007 and 2005). Similarly to Eun and Sabherwal (2003), these studies show that the liquidity of US trading discovery is positively related to the extent of US price by exploiting a large cross-section of stocks. These papers confirm that the price discovery largely occurs in the home market despite the globalisation trends of stock markets.

In terms of determining the lead/lag relationship between home and foreign markets, the most recent research in the area of price discovery suggests that stocks traded on the home market prices lead relative to their foreign listed "derivative" securities market. The evidence of this can be found in Kato et al (1990), Lau and Diltz (1994) and Lieberman, et al (1999). All these studies use domestic currency prices which have been converted into dollar prices using a daily exchange rate except Wang et al (2002), who examine a group of Hong Kong stocks which are also traded in London however the exchange rate has not been incorporated into their analysis. Their findings indicate a bi- directional causality for local market returns between the two markets however with Hong Kong being the dominant market.

A potential research question which could be applied to Russian stock market is to identify the factors contributing to location migration of the trading activity (and their likelihood), since the proportion of actual trading activity that takes place across competing markets seems to be significantly related to price discovery. For example, it has been observed that the trading volume has migrated from RTS to MICEX over the course of a decade. Pulatkonak and Sofianos (1999) investigate this issue by including global trading data on 254 NYSE cross- listed foreign stocks. They found that on average, one third of trading takes place on the NYSE and that the time-zone seems to be the most important determinant of this as firms with home markets that have been traded around the US time- zone have higher probability to be more active on US markets.

Bacidore and Sofianos (2002) have found that in general, the cross-listed stocks have wider spreads with less market depth than US stocks. It has been advocated that such results could be attributed to adverse selection risks and higher information asymmetry for which additional compensation is required by market makers and other liquidity providers. Phylaktis and Korczak (2004, 2007) and Moulton and Wei (2005) have taken this analysis further by evaluating the specialist trading activity of cross- listed stocks in order to demonstrate how the forces of competition have shaped market quality intraday patterns. Moulton and Wei (2005) show that when the home markets is trading, there has been availability of substitutes for investors, while realised spreads, effective spreads and quoted spreads had been significantly lower for the cross- listed stocks. Phylaktis and Korczak (2004) uncover evidence that concentration of stocks from a given country in an individual specialist tends to increase the share of US in price discovery by the means of the reduction in information asymmetries.

2.3 Limit Order Book Reconstruction

Although this literature field is still in infancy stages, there have been several attempts to reconstruct the limit order book. However, recent technological advancement in computational power enables the high frequency specialty data to be more readily available and processable for researchers. It seems that each stock market requires its own complex and unique reconstruction procedure. Following with pioneering works of Hasbrouck (1991) and Harris (1995), Kavajecz (1999) and Goldstein and Kavajecz (2000, 2004) construct an estimate of the NYSE order book using TORQ database in four steps: identification of limit orders at the time the sample started in pre- book; adding the current orders to the pre- book; matching order records with execution records and matching order records with cancellations. An snapshots at 30 minute intervals is then obtained as result.

More recently, the attempts to reconstruct order book have been carried out mostly on European exchanges: Auguy and Le Saoût (1999, 2001) and Auguy et al. (2000), briefly describe the book analysis for Paris Bourse; De Winne and D'Hondt (2005) reconstruct the order book very accurately for eighty two selected stocks from Paris, Brussels and Amsterdam; Beltran et al. (2005) and Frey and Grammig (2005) do the order book reconstruction for DAX 30 stocks with the same dataset from XETRA system. In the similar fashion, Hall and Hautsch (2004) reconstruct the complete order book by replicating the electronic trading at the Australian Stock Exchange at real time frequency.

In sum, numerous securities, market combinations and methodologies have been considered in the above mentioned studies, however none of have investigated the cross-listed securities on Russian exchanges. Furthermore, besides the research gap on the Russian market, only few studies employ the specialty ultra high frequency data overall. There is no reconciliation in the literature which data type i.e. quotes or transaction prices are optimal for the price discovery. Also, the optimal sampling frequency is a unique feature of each sample therefore remains an unanswered question. The straight forward framework following in the next sections of this paper aims to close these research gaps by bringing new insight into information flow and transmission between the underlying Russian markets with existing state of the art methodologies.

3 RTS and MICEX Markets: Facts

3.1 General comparisons

Russian stock markets could be divided into three main segments: Russian Trading System (RTS), Moscow Interbank Currency Exchange (MICEX) and American Depository Receipts (ADRs). Table 1 exhibits the turnover of the stocks markets competing for order flow. Most stocks are still traded in RTS yet the trading volume has migrated from RTS to MICEX and foreign exchanges such as LSE and NYSE. MICEX is a more liquid market than RTS. The trading volume in December 2006 was in ratio of 20: 1. Although trading on MICEX is concentrated in blue chips, 82 percent are Gazprom, Lukoil, Norilsk Nikel and RAO UES, the RTS index still remains the main benchmark. The other less visible however not less important common competitor, is over- the counter (OTC) which has comparable trading volumes.

Table 1: The turnover of the Russian exchanges 2003- 2006, bln USD

	2003	2004	2005	Jan- Apr 2006
MICEX SE	99.1	151.2	225.6	164.4
% of total market	88.20%	85.30%	85.60%	83.20%
RTS	13.3	26.1	38	33.1
% of total market	11.80%	14.70%	14.40%	16.80%

(Source: MICEX)

The Table 2 summarises the main differences between MICEX and RTS exchanges. The stock exchanges in Moscow are fully automated electronic and are quite similar to the developed markets equivalents.

Table 2: Comparison between the two Russian exchanges

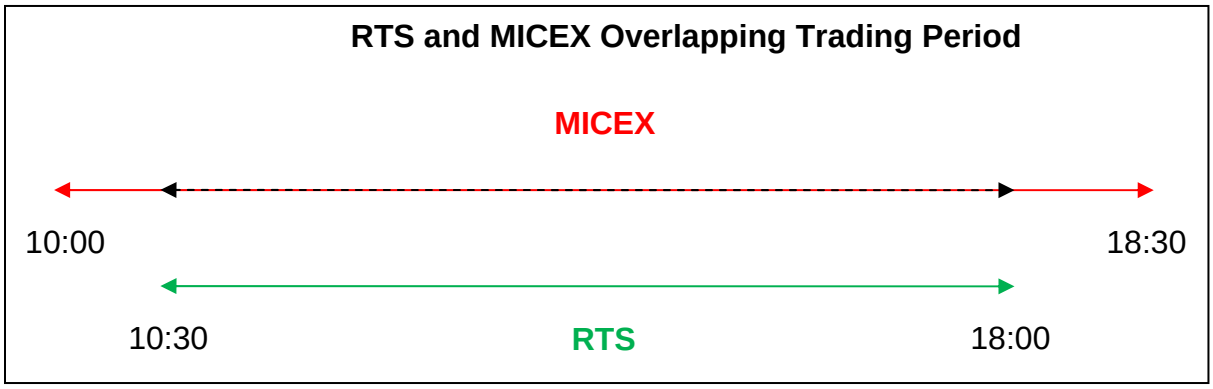
RTS	MICEX
SInce 1995 as a stock market	Since 1992 as FX market, since 1997 as a stock market
Hybrid dealership system	Order driven system
Quotes in US dollars	Quotes in Roubles
230 listed stocks	130 listed stocks
40 actively traded stocks	80 actively traded stocks
250 brokers/dealers	530 brokers/dealers
Minimum trading: lot of $5000	No minimum trading
Continuous trading since opening	Continuous trading since opening
No overnight trading	No overnight trading
Settelement delayed by 4 days	Settlement not delayed
No short selling	No short selling
No market orders de juro	No market orders de juro
No order revision	Limited order revision
No hidden orders	No hidden orders

The reconstruction process requires the knowledge about the trading rules and mechanism. These rules vary across markets and all of the facts from Table 2 should be considered when working with the orders databases and order books on MICEX and RTS.

3.2 Continuous Trading Time Overlap

As can be seen from Figure 4 both Moscow exchanges trading overlaps almost the entire trading day. MICEX starts to trade from 10:00 and closes 18:30 Moscow time. RTS start to trade at 10:30 and closes at 18:00 Moscow time. Although MICEX covers RTS trading time Figure 1 indicates that the intersection of the continuous trading hours of both exchanges is from 10:30- 18:00 Moscow time- that is continuous seven and one half hours per day.

Figure 4: Trading times on the RTS and MICEX markets



Besides the competition of these two markets, one should not ignore that fact that international markets, where the stock ADRs are also listed, do compete with each other, as well as other ADRs are available such as GDR, are also internationally traded. MICEX and RTS trading times do overlap with the other exchange such as Frankfurt Stock Exchange (FSE) and London Stock Exchange (LSE), which directly compete with the home markets, yet they also contributing to some extent to the price discovery process, which is beyond the scope of this paper.

The focus of this empirical study is on continuously traded overlapping period of MICEX and RTS exchanges. Trading of Russian stocks occurs in US dollars on RTS yet not on MICEX, where trading is only permitted in Russian Rouble. Changes in exchange rates require a change on RTS and MICEX stock prices in order to preserve the law of one price and avoid arbitrage opportunities. However Russian economy operates under capital flow restrictions and therefore the Rouble exchange rate is fixed over night based on central bank policy and closing results on MICEX.

The MICEX market trading is "thicker" than RTS trading. MICEX set has on average twenty thousand transactions per trading day. With much lower liquidity, RTS has less than 500 transactions per trading day. That is where the major discrepancy lies.

4 Data

The data employed in this study originated from the databases which have been obtained directly from both Moscow exchanges. The database sample covers the period of one trading year between January, the 10th and December, the 27th of the year 2006. The MICEX and the RTS databases contain both quotes and transaction prices for seven Russian blue chip stocks

comprised of RAO UES, NorNickel, LUKoil, Rostelecom, Sibneft, Surgutneftgaz, Tatneft. The list of stocks and their respective codes on the exchanges are provided in Table 3.

Table 3: List of cross- listed stocks by their trading venues

Name	LSE CODE	Currency	MICEX CODE	Currency	RTS CODE	Currency
RAO UES	UESD	USD	EESR	RUB	EESR	USD
LUKOIL	LKOD	USD	LKOH	RUB	LKOH	USD
SURGUTNEF	SGGD	USD	SNGS	RUB	SNGS	USD
MMC NORN	MNOD	USD	RU14GMKN0XXX	RUB	GMKN	USD
ROSTELEKC	RKMD	USD	RTKM	RUB	RTKM	USD
TATNEFT	ATAD	USD	RU14TATN3006	RUB	TATN	USD
SIBNEFT	SIF	USD	SIBN	RUB	SIBN	USD

4.1 Databases Contents

There are two database types stemming from the exchanges: databases based on quotes, which contain quotes entered into the systems and trades databases consisting in resulted transactions. From Table A1 in the appendix, it can be seen that raw order databases consist of ten columns, which display the name of the stock, the ID number, date, time to the nearest second, type such as buy or sell, action such as accept, cancel, split, quantity and validity of the order accept, match or cancel. If an order is split into pieces, the part at the original price level stays with the same order ID and keeps its priority while the new parts are given new IDs and lose priority. Raw transaction data shown in Table A2 in the appendix is composed of nine columns showing the name of the stock, the date, time, session, quantity, price and ID of the transaction as well as the IDs of the buy and sell orders which have been matched.

4.2 Sample Description

The total number of observations in the original databases is very large. For the MICEX quotes of the seven stocks, just the number of rows is about 36 million, which is only taking into account the number in the seven and half overlapping trading hours where both exchanges compete with each other. The total number of overlapping trading days of the year 2006 is 246. The summary observations for MICEX database are shown in the Table 4.

Table 1: Summary observations for MICEX Quotes Database during trading times

Name	RAO UES	LUKOIL	SURGUTNEFT	MMC NORILSK			ROSTELEKC	TATNEFT	SIBNEFT
MICEX CODE	EESR	LKOH	SNGS	GMKN	RU14GMK	RU14GMKN	RTKM	RU14TATN3	SIBN
January	362,503	480,513	137,070	178,309	10,526	167,783	166,617	38,745	36,586
February	939,422	453,527	134,857	170,445	11,182	159,263	446,470	59,512	39,789
March	932,919	629,720	196,627	196,566	10,034	186,532	496,309	102,826	37,810
April	961,707	531,977	209,776	354,578	17,503	337,075	305,340	103,417	51,960
May	956,179	626,609	179,543	619,531	26,252	593,279	293,696	109,603	68,554
June	1,318,323	773,029	207,036	402,749	25,796	376,953	339,554	137,248	61,694
July	1,005,348	730,433	210,151	471,014	17,031	453,983	183,328	295,826	65,025
August	1,111,520	631,158	205,723	491,321	26,227	465,094	171,199	242,265	78,072
September	966,605	546,912	255,315	440,608	19,514	421,094	94,610	252,551	67,743
October	659,486	836,122	326,594	612,474	23,750	588,724	95,543	269,849	74,887
November	923,686	623,945	281,649	450,470	16,646	433,824	131,674	213,322	70,033
December	733,676	408,699	263,049	333,989	20,950	313,039	170,096	178,663	77,293
Total DB Qu	**10,871,374**	**7,272,644**	**2,607,390**	**4,722,054**	**225,411**	**4,496,643**	**2,894,436**	**2,003,827**	**729,446**
Average	905,948	606,054	217,283	393,505	18,784	374,720	241,203	166,986	60,787

4.3 Order Book Reconstruction Methodology

The methodology presented in this section contributes to the literature filed of order book reconstruction by generating new variables, thus developing understanding of the complex behaviour of trading, through an application to MICEX and RTS. This approach proposes a unique method to link trades to orders, i.e. instant market orders constituting the transaction book to the stock of limit orders of the order book, which is just more than simply reconstructing order book.

Limit order book reconstructing is very rewarding as it allows the creation of lossless information variables which are difficult to access for general public, yet are essential to understand the functioning of order- driven markets, constituting complementing alternative insight to the existing knowledge. While general methodology of reconstruction remains the same for all order driven markets, exchange specific trading rules such as absence of de juro market orders would require modifications.

The order book reconstruction has several advantages over solely reproduction of a book of trades based on transaction data. First, there is an advantage that orders are posted more frequent than transaction occur because a transaction is a result of an executed trade based on a match of bid and ask orders. Second, order data provides a more complete picture about the behaviour of the market participants. This eliminates the problem which arises with transaction data, namely the so called observations “data holes”. Different missing data filling methods may lead to different and sometimes ambiguous inference results. Sometimes, the

dataset has to be reduced to lower frequency sampling periods inevitably, due to asynchronous and infrequent quoting. This seems to be a solution with the trade off between scaling to larger intervals would provide less accuracy, however would also require less interpolation to fill the time "data holes" which could be sometimes a costly compromise which may result in complete ambiguity. Third and the main advantage is opportunity to analyse the bid- ask spread and the associated indirect costs of trading, risk, inter- trade duration and many other important aspects of financial research interests.

4.4 Reconstruction Process

The objective of reconstruction is to convert the order and transaction data into the order book in such a way that the information about the variables becomes observably reproduced in their original order similar to the way traders would have observed the trading process uninterruptedly live in front of the screens. The general criterion is to create panel and time series with best prevailing bid and ask quotes, eliminating market orders which cause bid- ask quote crossovers and reconstructing just limit orders with minimum inside spread. The rebuild data should comply with further criteria such as being accurate, consequent, chronological, consistent, synchronous and scalable across exchanges. Moreover the reconstruction contains an option of choosing the sampling frequency between real time 1 second and inter-temporally aggregated 30 and 60 seconds intervals. This results in a new dataset containing useful variables for financial analysis.

Firstly, we link the database files of order and transaction data. Then, queries, as shown in Table A3 in appendix, are created, which sort the raw data chronologically the by security IDs, sorting bid and ask quotes and filtering the unnecessary noise values out for instance orders with prices zero. Then, we use the information of overlapping entry and exit times of orders, for which the two database files (orders and transactions) were required to choose the right values according to selection criteria (maximum for bids and minimum for asks), processing each order according to the priority rules and fills in the order book one by one allocating each to the specific time point in the chosen sampling intervals. If orders match and then are executed, they are first considered in their time interval being in the trading system and then eliminated from the order book. Each row may indicate the status of an order such as new, revision, split or cancellation if these last two satisfy the rules. The major challenge appears to be dealing with bid- ask order crossovers. The case of bid- ask crossover means that the counterparties matched in quotes, and the transaction IDs is executed fully or

partially with contemporary equilibrium price and quantity at the given time point reported by the system. However, the difficulty lies in differentiating between the numerous case scenarios, such as full match, partial match, cancellation or split. To prevent the crossing over of bids and asks, the buying orders and selling orders in the quotes data are matched to the execution orders in the transaction data in order to eliminate orders from the current book in case of matching. As a last step of reconstruction, the new variables such as time series of midpoint prices between bid and ask are created corresponding to the chosen sampling frequency.

A distinct feature, which differentiates the reconstructed order book from the observations made on screens of traders, is that the marketable orders which resulted in transaction were derived from the book. In order to detect whether an order is a limit order or marketable order, we check the actual bid and ask quotes. If the new coming buy or sell order price equals or overlaps ask or bid quote respectively, the order is identified as a marketable order. When a marketable order arrives, instead of directly taking out the matching order from the book, it is initially considered by the order book in order to see the exit moment from the system, graphically, the bid- ask spread in this case becomes zero or causes a crossover for an instant, and it is then removed. As a consequence, because of the chosen methodological subject, crossovers caused by such quotes are eliminated. Such an approach separates marketable and limit orders into a separate baskets while reflecting the demand or supply for stock more accurately.

Once the limit order book is rebuilt accurately, it becomes possible to follow the evolution of behaviour of a market at each moment. Several variables can now be deduced at each point of time including bid and ask prices, quantities and order IDs of the bid and ask sides as well as the balance of bid- and ask variables. The example of reconstruction results for MICEX database for 30 seconds intervals are depicted in Table A4 in the appendix.

Since the main research interest is focused on dynamic relationship, the order book reconstruction has to be applied to both exchanges, creating therefore two order books; each of them requires idiosyncratic approach. Furthermore, the condition for analysis of the dynamics across markets the order books must be synchronous to each other with common overlapping trading period. In order to synchronise the overlapping continues trading time between the exchanges, the order books were reduced to common denominator of time and were used as an input for newly created database containing the common overlapping time

and midpoint time series. The Rouble denominated prices of MICEX stocks were converted into US dollar prices as already quoted on RTS with officially set daily exchange rates obtained from MICEX. The combined time series are exhibited in Table A5 in the Appendix.

Yet, prior to the creation of a common database, one should consider some idiosyncratic peculiarities of traded stocks. Each of them requires a unique approach in choosing an optimal sampling frequency because of the differences in information expressed in frequency of initial orders posted for each security across markets. Alternative sampling frequencies over the choice of five minutes as being suitable relative to higher frequencies such as one minute, ten or five seconds, are also considered in the initial analysis. Due to inter- temporal aggregation, the accuracy diminishes, at lower frequencies, so that the correlation dynamics between the variables dissolves, indicated by falling significance in their coefficients. At higher sampling frequencies than thirty seconds there would probably be no gain in terms of increasing the significance of correlation between the variables, however there has been a trade off with issues such as microstructure noise, signal and non synchronism of observations, which may lead to conclusion that the thirty seconds interval as being the most optimal.

4.5 Time Series Reconstruction Methodology

An alternative to the proposed order book reconstruction method is a creation of a book based on transactions data as an additional robustness test. The MICEX and RTS sample may contain high number of transactions during one sampling interval, however may have none during the next. In case of RTS data sample such an observation “hole” could last for about more than twenty plus minutes depending on given liquidity. In order to be able to do any modelling and to make any inferences, the overlapping trading times between the RTS and MICEX have to meet following criteria: they have to be continues and synchronous. So the number of observation has to be equal. Equalisation requires interpolation and various extrapolation methods to be employed, which are comprised of the following steps: 1) Create a query with filtering for a security ID from the RTS and MICEX databases; 2) Allocate the prices to each given moment of time intervals, separating prices into buy and sell orders; 3) Interpolate the missing data with last known price ceteris paribus in order to make the series continuous and 4) Merge quotes and trades databases into one as an option.

The less liquid RTS data set has been interpolated to fill the “data holes” by assumption that the last observed price would hold constant until the next observation. This is the most

conservative method since it assumes that the prices would not change over unobservable period, which is not entirely true with respect to theoretical price levels derived from the quotes of an order book. Given that MICEX sample is more liquid, the interpolation there has been ignorable.

Furthermore, order to synchronise the overlapping continuous trading time between the two exchanges, the differences in liquidity expressed in number of observations, required the same solution as in case of order book reconstruction. This seems to be a rational scaling because smaller interval would provide more accuracy, however would also require more interpolation to fill the time "data holes". Moreover, from previous studies, there is evidence that these data holes provided that they are kept in reasonable proportions do not substantially affect the quality of inferences[4].

The aggregation method suggested here has a property of being inter- temporally self adjusting. The self adjustment takes place automatically depending on frequency of data available because there is a simple pattern which is always followed- this method uses only the last observation. No matter how high the frequency of an interval is i.e. containing more or less observations, the outcome of the aggregation is always the last one observation. It is possible to create other observation patterns such as taking the first the first or the median observation however there would be a problem with even number of observations in the given intervals.

The use of executed prices data has its advantages and disadvantages. The advantage of using it eliminates the possibility that quotes are being revised just on one side of the market or the other to reflect positions. It is unlikely that the choice of execution versus midpoint prices will have any essential effect on the results. The main disadvantage of transaction prices however is that the time series by their nature contain less prices per observed interval than tick bid ask price series. This is expressed in first more frequent and larger, in terms of time, observation gaps and second more discreteness. Since the time lead- lag relationship in price discovery is very sensitive to the way the datasets are manipulated, different price extraction methods could be used, which may lead to different results and hence conclusions.

4.6 Time Series Construction Results

The Table 5 exhibits the expected number of rows for the year 2006.

[4] ibid 1, pp. 8- 9

Table 2: Expected number of observations with range of frequencies

Frequency (sec)	1	5	10	15	30	60	300
sec	60	60	60	60	60	60	60
min	60	60	60	60	60	60	60
hours	7.5	7.5	7.5	7.5	7.5	7.5	7.5
days	250	250	250	250	250	250	250
obs/day	**27,000**	**5,400**	**2,700**	**1,800**	**900**	**450**	**90**
obs/anno	**6,750,000**	**1,350,000**	**675,000**	**450,000**	**225,000**	**112,500**	**22,500**
securities	7	7	7	7	7	7	7
markets	2	2	2	2	2	2	2
Total	**94,500,000**	**18,900,000**	**9,450,000**	**6,300,000**	**3,150,000**	**1,575,000**	**315,000**

The Figures 1 and 2 display one reconstructed MICEX and RTS trading day respectively. The degree of accuracy is high: The bid and ask quotes neither touch each other nor cross. Micro bid- ask bounce movements are clearly visible.

Figure 1: Bid- Ask time series of one MICEX trading day (30sec frequency)

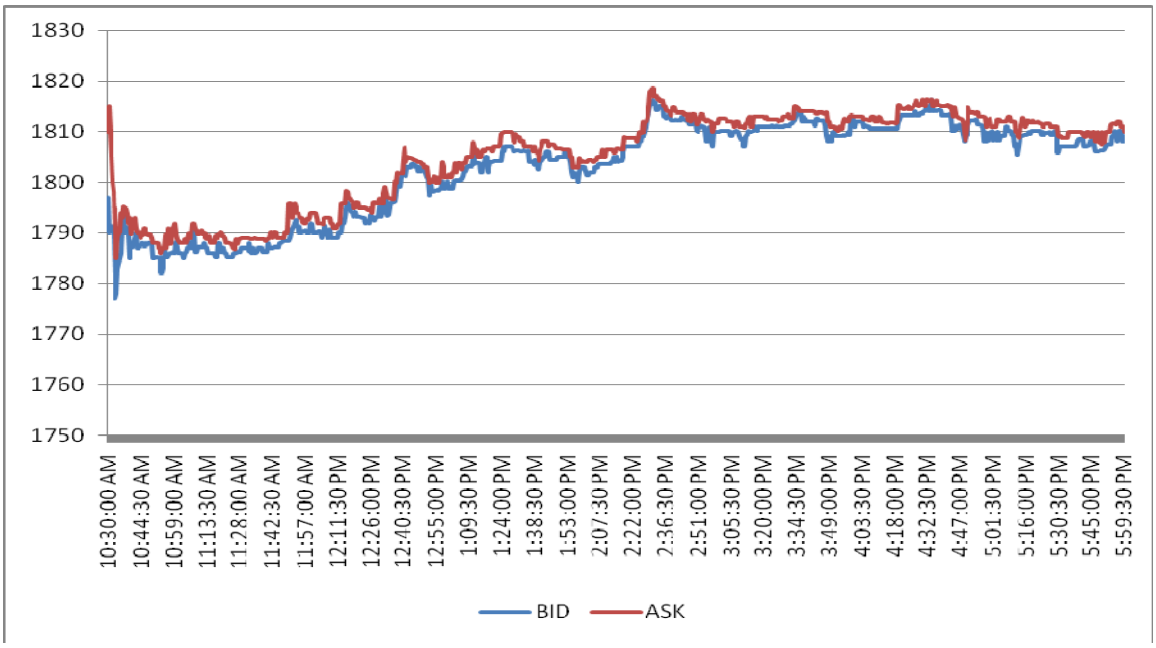
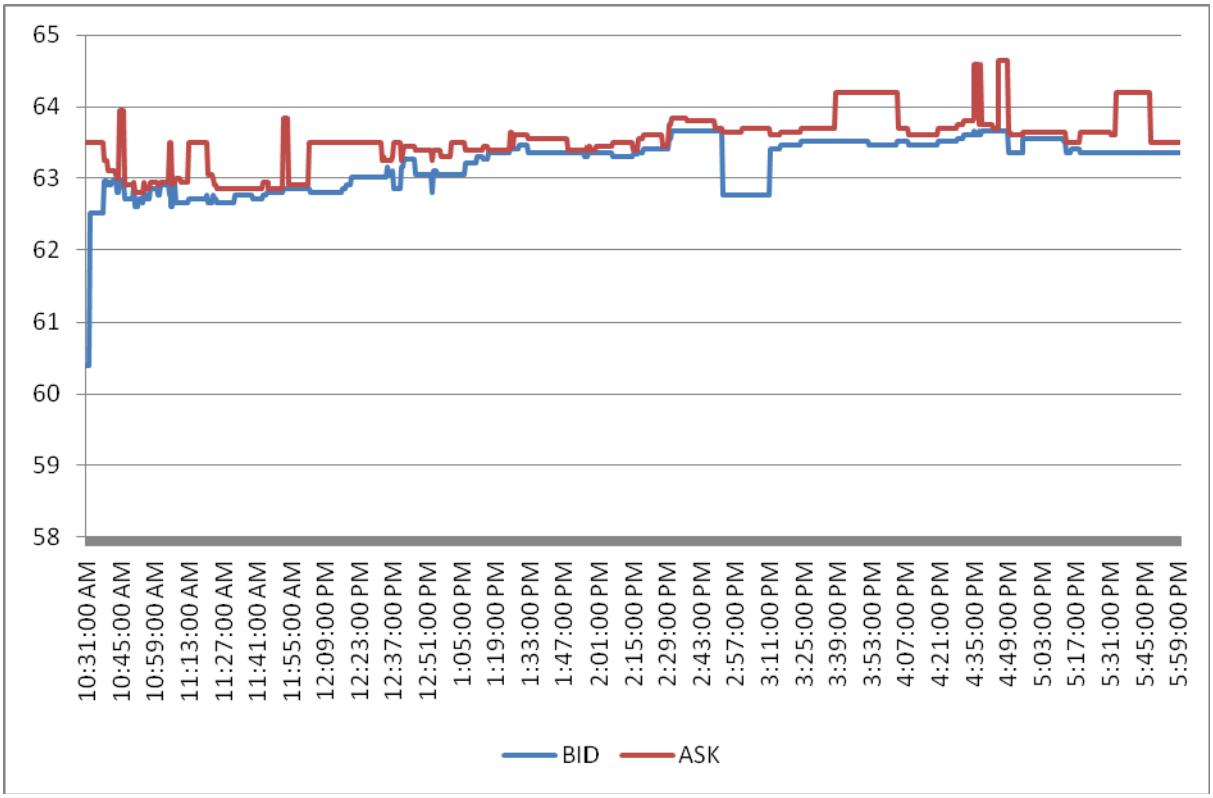
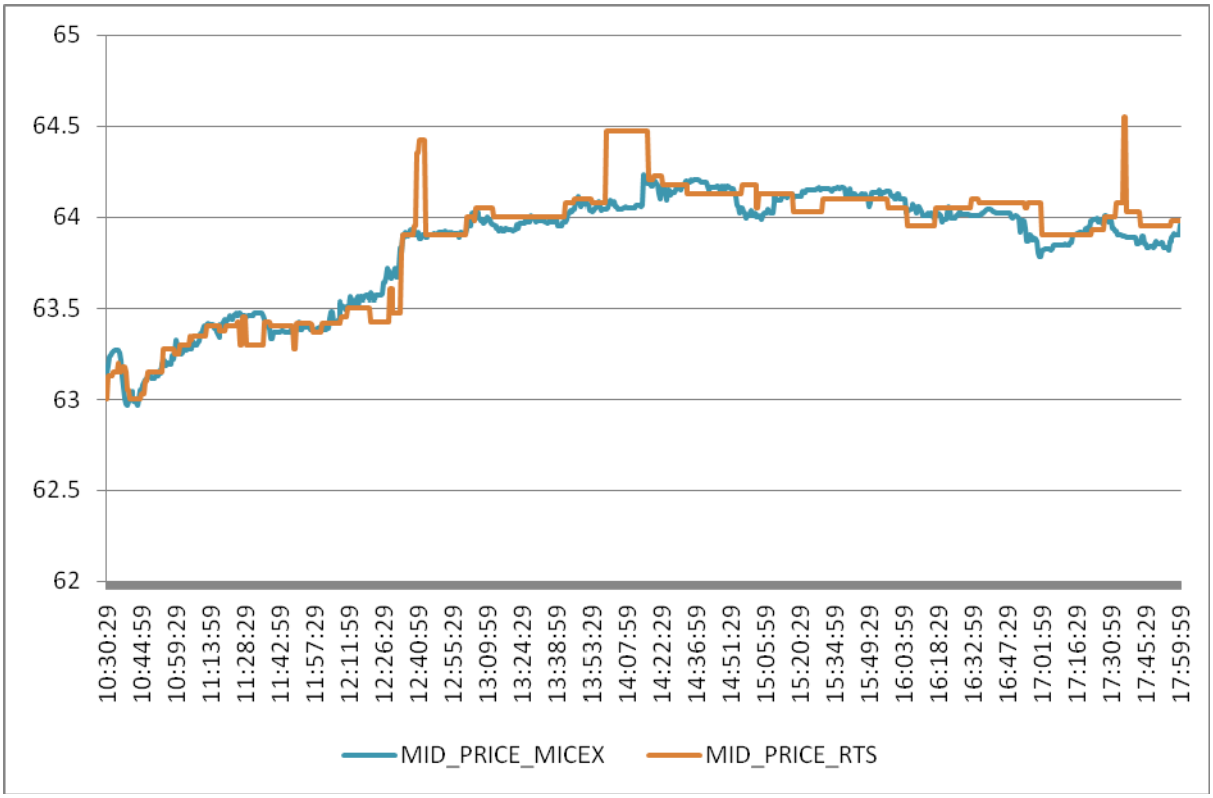


Figure 2: Bid- Ask time series of one RTS trading day (30sec frequency)



If looked at the constructed time series graph of both markets combined as in Figure 3, it becomes visible that there is high degree of accuracy since both market waves seem to go close together almost in resonating fashion.

Figure 3: Midpoint price series of one MICEX- RTS trading day (30sec frequency)



5 Methodology

In order to investigate on the subject of the dynamic relationship between securites and their dynamics, this study proposes to estimate models of cointegration/ error- correction models (ECM) in Engle and Granger (1987), Johansen (1988) cointegration/ vector error- correction (VECM) frameworks, by applying Harris (1995) and Hasbrouck (1995) methodologies. The main purpose of these models is to provide insight into the relationship of the markets, whether MICEX and RTS, are co- moving or cointegrated, whether there is a lead- lag relationship as well as their information flow concentration and dispersion.

The investigation is organised as follows: Firstly, the markets are examined for their order of integration and whether there is a common stochastic trend in the long run. Secondly, MICEX or RTS are modelled to investigate the price discovery relationship. Lastly, the market contribution to the price discovery is measured through their content of information shares.

5.1 Models of Information and Price Discovery

We apply two methods for testing for information and price discovery: 1) the unconditional Engle-Granger model, and 2) Johansen (1988) vector autoregressive (VAR) in conjunction with maximum likelihood estimation method for consistent model dynamic specification and in order to measure the contributions of markets to price discovery.

5.1.1 Unconditional ECM

The unconditional Engle- Granger error correction model (ECM) does not directly involve residuals of the levels of the log prices as the conditional ECM does, however involves the series of differences between log prices i.e. as one step procedure and the actual series spread is treated as a variable instead. The unconditional ECM regression model is estimated as follows:

$$\Delta X_t = \alpha + \Delta X_{t-1} + \Delta Y_{t-1} + \lambda \mu^*_{t-1} + \eta_t \qquad (1)$$

and the analogous vice versa

$$\Delta Y_t = \beta + \Delta Y_{t-1} + \Delta X_{t-1} + \theta \lambda^*_{t-1} + \pi_t \qquad (2)$$

Where market spread between MICEX and RTS vice versa is defined as $\mu^*_t = X_t - Y_t$ and $\lambda^*_t = Y_t - X_t$.

The parameters λ and θ are the indicators of long run relationship similar to the representations in the previous ECM. These parameters rather describe the speed of adjustment if the disequilibrium occurs. A positive significance and higher degree of value in one these parameters is expected for a lagging market and vice versa. For instance if RTS or lags MICEX, the prices on RTS are expected to catch up MICEX prices at higher speed then in vice versa case.

5.1.2 VAR and Vector Error Correction Models (VECM)

An alternative and more consistent methodology to model cointegration and error correction would be the Johansen (1988) procedure.

Johansen's approach is based on maximum likelihood estimation (MLE) of the VECM, by step- wise eliminating the parameters out i.e., maximizing the likelihood function over a

subset of parameters, treating the other parameters as known, given the number of cointegrating vectors, where the matrix β^T is the last to be concentrated out.

Let X to be an 1x n vector of unit- root processes where it is assumed that there exist 1−n cointegrating vectors, which would imply an existence of single common stochastic trend according to Stock and Watson (1988). Based on Engle and Granger (1987), the series have the following vector autoregressive VAR (n) representation:

$$\Delta X_t = \alpha + \Pi X_{t-1} + \sum_{k=1}^{q-1} \Gamma \Delta X_{t-k} + \varepsilon_t \qquad (3)$$

$$\varepsilon_t \sim NID(0, \sum \varepsilon)$$

$$\text{where } \Pi_k = I + A_1 + \ldots + A_i; \quad \Pi = I - A_1 - A_2 - \ldots - A_q$$

are the matrices of coefficients determining the cointegration relationship between the variables.

Once the cointegrating relationships have been confirmed by likelihood ratio (LR) test, Π can now be substituted with factorised matrix $\Pi = \alpha\beta^T$, where β and α are n x (n-1) matrices of rank (n-1). This transformation would result in following VECM:

$$\Delta X_t = \alpha_0 + \alpha_\perp \beta^T X_{t-1} + \sum_{k=1}^{q-1} C \Delta X_{t-k} + \varepsilon_t \qquad (4)$$

The columns of β consist of the (n-1) cointegrating vectors and each column of α consists of adjustment coefficients, which define the speed of adjustment back to equilibrium similar to bivariate ECM. The matrix Π is decomposed in such a way that $\beta^T X_t$ consists of (n-1) vector of stationary series. The covariance matrix of the error term is given by $E = \left[\varepsilon_t \varepsilon_t^T\right] = \Omega$.

5.1.3 Gonzalo- Granger (GG) Permanent- Transitory Measure

This approach specifies the proportion of information shares discovered in either market. The basic idea is to decompose the cointegrated system Y_t into a permanent the common factor component $A_1 f_t$ and a transitory stationary component $\tilde{Y}_t$:

$$Y_t = A_1 f_t + \tilde{Y}_t \qquad (5)$$

where A_1 represents a factor loading matrix. By using the identifying restrictions that f_t is a linear combination of Y_t and that the transitory component $\tilde{Y}_t$ does not Granger- cause Y_t in the long- run Gonzalo and Granger (1995) definition, the dynamics above can be decomposed as:

$$Y_t = \beta_\perp \left(\alpha_\perp ' \beta_\perp\right)^{-1} \alpha_\perp ' Y_t + \alpha \left(\beta ' \alpha\right)^{-1} \beta Y_t \qquad (6)$$

Since f_t is given $\alpha_\perp ' Y_t$ the elements of $\alpha_\perp$ are the common factor weights of the variables driving the cointegrated system. Once $\alpha_\perp$ has been normalised so that its elements sum to unity, it measures the fraction of system innovations attributable to each variable. In a bivariate system, the Gonzalo- Granger measure is defined as a ratio of coefficient of errors over the difference in coefficients of errors α_1 and α_2 from (4) and (6) between both markets:

$$GG_Y = \frac{\alpha_1}{\alpha_1 - \alpha_2} \qquad (7)$$

$$GG_X = \frac{\alpha_2}{\alpha_2 - \alpha_1} \qquad (8)$$

For instance, assuming MICEX being a dependent variable from equations (7) and (8), if the proportion of GG is higher than 50 percent, then MICEX contains higher proportion of information share in price discovery, in this case more than half, indicating that RTS or LSE is leading or a more dominant market. In vice versa case with lower than 50 percent market share, RTS or LSE would indicate that MICEX is lagging since their share of information increases.

5.1.4 Hasbrouck Information Share

The method of measuring the price discovery is starting with the information share method proposed by Hasbrouck (1995). Hasbrouck Information Share (IS) is following Stock and Watson (1988), Hasbrouck (1995) transforms equation (3) into the following vector moving average (VMA) representation:

$$\Delta X = \psi(L)\varepsilon_t \qquad (9)$$

alternatively, $X_t = X_0 + \psi(1)\sum_{i=1}^{t} \varepsilon_t + \psi^*(L)\varepsilon_t$ (10)

Since the two series are cointegrated, the Engle- Granger representation theorem implies the following:

$$\beta^t \psi(1) = 0 \quad \text{and} \quad \alpha \psi(1) = 0 \tag{11}$$

Therefore, there is $\psi(1) = \alpha_{\perp}^{T} \beta_{\perp}$ where α⊥ and β⊥ are orthogonal vectors to α and β

respectively. Equation (3) could be expressed as:

$$X_t = X_0 + \beta_{\perp} \alpha_{\perp}^{T} \psi(1) \sum_{i=1}^{t} \varepsilon_t + \psi^*(L) \varepsilon_t \tag{12}$$

The term $\alpha_{\perp}^{T} \sum_{i=1}^{t} \varepsilon_t$ represents the common stochastic trend component, which follows a random walk process. The $\psi(1)\varepsilon_t$ term represents the long- run impact of innovation on price. It is also clear that existence of n- 1 cointegrating vectors implies that the impact matrix $\psi(1)$, which is the sum of the moving average coefficients, has rank 1. Ψ represents the identical row of $\psi(1)$. Hasbrouck (1995) mentions that $\psi \varepsilon_t$ constitutes the long- run impact of the innovations on each of the prices and suggests the following measure of information share (IS) of market j for the case where the covariance matrix Ω is diagonal i.e., the innovations are independent:

$$S_j = \frac{\psi_j^2 \Omega_{jj}}{\psi \Omega \psi^T} \tag{13}$$

Here, the j ψ is the j- th element of the identical row of the impact matrix $\psi(1)$. The information share measure when the covariance matrix is not diagonal is given by

$$S_j = \frac{\left([\psi F]_j \right)^2}{\psi \Omega \psi^T} \tag{14}$$

where F is the Cholesky factorisation of Ω and $[\psi F]_j$ represents the j- th element of the row vector Fψ. Since the Cholesky factorisation depends on the ordering, equation (14) would

provide an information share for a particular ordering. By considering all possible orderings we can compute the upper and lower bounds on IS.

6 Findings

This study supports the notion that price discovery occurs in the most liquid domestic market with yet also statistically significant results for less liquid competitor. The results are consistent with the hypothesis that the more liquid market is the primary market and the less liquid one is the secondary market. These findings are also in line with studies of Hasbrouck (1995) and Harris (1995). There is evidence which indicates that the null hypothesis should not be rejected. On the one hand, the thicker MICEX market seems to be a dominant price discoverer, containing the major information share. On the other, both markets contribute to price discovery significantly, yet the RTS market plays a supportive role.

The findings are preliminary. The estimation results are based on maximum of fifteen trading days of January 2006. The results are approximately true for six out seven stock pairs, which have been sampled with frequency range of one, thirty and sixty seconds resulting in about 200 thousand observations per stock. Despite the rise of the market density in the course of the year on both markets, unless there is structural break in the data, the results of further research are expected to be consistent for the remainder of the year due to the fair robustness and consistency of formulated models.

6.1 VECM Estimation Results

The focus of the research has been the estimates of the restricted VECM model and the associated information shares. The initial observation each day for each stock is determined by the first sampling interval following the RTS opening containing quotes in both markets. The presented results are based on the highest sampling frequency i.e. one second- real time interval, which is the most accurate sampling resolution, since lower frequencies results were almost perfectly in line with them.

The more parsimonious choice of lag length for VECM has been finally determined by the Schwarz Information Criteria (SIC). The initial lag length is 30 lags, which represents 30 seconds in a sample with observations at one second intervals. Then, using the same set of observations that was used for the estimation of the model with 30 lags, the VECM is estimated at each shorter lag length down to one lag to determine the lag structure that minimises the SIC. Lag lengths range from 1 to 8 for all stocks.

Table 6 clearly indicates the dominance of the MICEX market price in price discovery. The information shares for RTS price innovations are seen to be somewhat of an inverse image of the MICEX information shares.

Table 6: VECM cointegrating vectors estimation Results

1sec						
Cointegrating Eq:	**EESR**	**LKOH**	**RTKM**	**SIBN**	**SNGS**	**TATN**
LOG_MICEX(-1)	1	1	1	1	1	1
LOG_RTS(-1)	-1.06971	-1.04137	-1.09142	-0.95417	-1.03509	-0.97117
	-0.02443	-0.00598	-0.00828	-0.01908	-0.00917	-0.01099
	[-43.7864]	[-174.017]	[-131.872]	[-50.0137]	[-112.845]	[-88.3984]
30sec						
Cointegrating Eq:	**EESR**	**LKOH**	**RTKM**	**SIBN**	**SNGS**	**TATN**
LOG_MICEX(-1)	1	1	1	1	1	1
LOG_RTS(-1)	-1.00292	-1.03924	-1.07133	-1.04379	-1.03604	-0.9713
	-0.00081	-0.00795	-0.02982	-0.02145	-0.00986	-0.01131
	[-1240.89]	[-130.652]	[-35.9256]	[-48.6658]	[-105.094]	[-85.8574]
60sec						
Cointegrating Eq:	**EESR**	**LKOH**	**RTKM**	**SIBN**	**SNGS**	**TATN**
LOG_MICEX(-1)	1	1	1	1	1	1
LOG_RTS(-1)	-1.00292	-1.03924	-1.07133	-1.04379	-1.03604	-0.9713
	-0.00081	-0.00795	-0.02982	-0.02145	-0.00986	-0.01131
	[-1240.89]	[-130.652]	[-35.9256]	[-48.6658]	[-105.094]	[-85.8574]

The higher the information share of the MICEX price innovations in explaining MICEX price, the lower the RTS information shares.

The hypothesis that the MICEX market is the primary market and the RTS the supportive market would be consistent with a larger role for price discovery in the MICEX market than on the RTS. Tables 7-9 in the sections that follow indicate that this is clearly true for all the stocks in the sample.

However, five stocks have a considerable information share greater than 10 percent role for RTS price discovery and however none of the stocks display a larger information share for RTS price innovations than MICEX market price innovations. The interesting question of what explains the differences across stocks will be addressed in the cross- section analysis below.

6.2 Conditional and Unconditional ECMs Estimation Results

In accordance to Engle and Granger representation theorem, since the cointegration condition has been fulfilled, the study proceeds with the evaluation of the formulated ECMs. The results for various frequencies are shown in Tables 7, 8 and 9.

Table 7: ECM estimates Summary 1sec frequency

Dependent Variable: LOG_DMICEX												
	EESR		LKOH		RTKM		SIBN		SNGS		TATN	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
LOG_DRTS(-1)	-0.001204	0.2916	2.30E-04	0.7793	1.46E-03	0	0.003675	0.0097	0.001739	0.0978	0.001584	0.2445
LOG_DRTS(-2)	0.001704	0.1353	5.36E-04	0.5133	-2.11E-04	0.1195	0.001405	0.3226	0.000874	0.4056	0.001784	0.1898
LOG_DRTS(-3)	0.001243	0.2763	4.04E-03	0	-2.67E-04	0.0488	0.00411	0.0038	0.002486	0.018	-3.84E-05	0.9775
LOG_DRTS(-4)	0.00135	0.237	1.05E-03	0.1996	-2.06E-04	0.1288	0.004354	0.0022	-0.001808	0.0853	0.000813	0.5505
LOG_DRTS(-5)	0.002063	0.0706	2.78E-03	0.0007	-2.09E-04	0.1227	0.003497	0.0138	0.001409	0.1798	0.001428	0.294
LOG_DRTS(-6)	0.002053	0.072	-6.63E-04	0.4187	-9.94E-05	0.4628	9.85E-04	0.488	0.000335	0.7499	0.001021	0.4532
LOG_DRTS(-7)	0.000301	0.792	3.08E-03	0.0002	-1.73E-04	0.2024	0.001644	0.247	2.30E-05	0.9825	0.006023	0
LOG_DRTS(-8)	-0.000819	0.4732	1.55E-03	0.0585	-1.26E-04	0.3536	2.06E-03	0.148	0.001059	0.3133	-0.001819	0.1813
LOG_DRTS(-9)	0.001203	0.2917	-1.08E-04	0.8953	-1.66E-04	0.2217	0.002057	0.1475	0.000637	0.5441	0.001838	0.1768
LOG_DRTS(-10)	-0.001281	0.2618	2.06E-04	0.8018	-1.63E-04	0.2298	2.26E-03	0.1124	0.003067	0.0035	0.000549	0.6864
LOG_DMICEX(-1)	0.007949	0.0001	0.001093	0.6269	-6.35E-03	0.0002	-8.73E-03	0	-0.004848	0.0023	-0.003291	0.031
LOG_DMICEX(-2)	-0.005597	0.0051	-0.008948	0.0001	-0.012445	0	-0.012869	0	-0.021802	0	-0.007257	0
LOG_DMICEX(-3)	-0.015315	0	-1.52E-02	0	-1.76E-02	0	-3.78E-02	0	-0.020153	0	-0.009071	0
LOG_DMICEX(-4)	-0.011308	0	-1.64E-02	0	-1.39E-02	0	-2.30E-02	0	-0.017008	0	-0.010361	0
LOG_DMICEX(-5)	-0.004237	0.034	-2.21E-03	0.3272	-1.17E-02	0	-2.10E-02	0	-0.016492	0	-0.015794	0
LOG_DMICEX(-6)	-0.003797	0.0575	5.40E-03	0.0164	-4.32E-03	0.0107	-1.43E-02	0	-0.01179	0	-0.012414	0
LOG_DMICEX(-7)	-0.005493	0.006	1.13E-03	0.6162	-5.92E-03	0.0005	-9.98E-03	0	-0.005061	0.0015	-0.009353	0
LOG_DMICEX(-8)	0.000548	0.784	-2.82E-03	0.2107	-4.16E-03	0.0139	-1.27E-02	0	-0.013754	0	-0.001905	0.2116
LOG_DMICEX(-9)	0.002545	0.2028	1.39E-03	0.5367	-3.24E-03	0.0555	-7.51E-03	0	-0.002901	0.068	-0.005709	0.0002
LOG_DMICEX(-10)	-0.002406	0.2287	1.44E-02	0	-6.41E-03	0.0002	-1.28E-02	0	-0.003522	0.0267	-0.005552	0.0003
SPREAD(-1)	0.000187	0	2.09E-04	0.0001	2.28E-04	0	5.90E-05	0.0211	0.000177	0	0.00019	0
C	-2.82E-07	0.2044	4.76E-07	0.109	-4.16E-07	0.0781	2.92E-07	0.3166	5.49E-07	0.0232	5.82E-07	0.0644
Dependent Variable: LOG_DRTS												
	EESR		LKOH		RTKM		SIBN		SNGS		TATN	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
LOG_DRTS(-1)	-0.000205	0.9184	7.12E-05	0.9749	0.007826	0	0.001151	0.4531	-0.001753	0.2701	-0.006602	0
LOG_DRTS(-2)	0.000433	0.8285	0.002893	0.1981	0.008701	0	0.001467	0.3388	0.002978	0.061	-0.001324	0.3852
LOG_DRTS(-3)	-0.000599	0.7642	-0.001247	0.579	0.008627	0	0.002151	0.1607	0.000665	0.6755	1.91E-05	0.99
LOG_DRTS(-4)	0.001594	0.425	0.001352	0.5476	0.008813	0	-0.026677	0	0.001829	0.2498	-0.005379	0.0004
LOG_DRTS(-5)	0.000304	0.8792	0.001767	0.4318	0.008728	0	0.001695	0.2691	-1.31E-05	0.9934	0.001476	0.333
LOG_DRTS(-6)	-0.001312	0.5113	0.001442	0.5213	0.008167	0	-0.000698	0.6492	0.000932	0.5578	-0.001712	0.2615
LOG_DRTS(-7)	-0.004952	0.0132	0.00099	0.6597	0.011246	0	-0.01811	0	0.00073	0.6462	-8.14E-05	0.9574
LOG_DRTS(-8)	0.005318	0.0078	0.001315	0.5587	0.008306	0	-0.00032	0.8348	0.000754	0.6353	-0.008346	0
LOG_DRTS(-9)	0.000704	0.7246	-0.001128	0.6158	0.008267	0	-0.003107	0.0428	1.39E-06	0.9993	0.005168	0.0007
LOG_DRTS(-10)	0.002331	0.2434	-0.00031	0.8904	0.008605	0	0.000572	0.709	-0.003818	0.0163	0.000696	0.6483
LOG_DMICEX(-1)	-0.002209	0.5277	0.009954	0.1063	-0.002508	0.9054	-0.000627	0.7048	0.001185	0.6223	0.000105	0.9512
LOG_DMICEX(-2)	0.002523	0.4708	0.016195	0.0086	-0.009043	0.6682	-0.002395	0.1481	0.003015	0.21	0.001285	0.4522
LOG_DMICEX(-3)	-0.000629	0.8573	-0.003683	0.5503	-0.002046	0.9228	-2.62E-05	0.9874	-0.000929	0.6994	0.001756	0.3043
LOG_DMICEX(-4)	0.000557	0.8734	-0.021875	0.0004	0.000893	0.9662	0.009524	0	0.00172	0.4747	0.005305	0.0019
LOG_DMICEX(-5)	0.016246	0	0.008371	0.1749	-0.01063	0.6145	-0.000375	0.8213	0.000999	0.6779	0.001067	0.5325
LOG_DMICEX(-6)	0.005675	0.1048	0.003999	0.5168	-0.010458	0.6202	0.00011	0.9473	0.001398	0.5611	0.003243	0.0577
LOG_DMICEX(-7)	0.017358	0	-0.008448	0.1708	0.266961	0	0.00358	0.0308	0.001478	0.5388	0.001599	0.3494
LOG_DMICEX(-8)	-0.0194	0	0.01331	0.0309	-0.00328	0.8765	0.000835	0.6143	0.002448	0.3087	-0.001528	0.3713
LOG_DMICEX(-9)	-0.013275	0.0001	0.004637	0.4521	-0.003225	0.8785	-0.003971	0.0165	0.002731	0.2561	0.001996	0.2429
LOG_DMICEX(-10)	-0.00667	0.0566	-0.005131	0.4054	-0.002133	0.9195	-9.40E-05	0.9547	0.000149	0.9507	-0.002281	0.182
SPREAD(-1)	-0.000711	0	-0.002187	0	-0.017505	0	-0.000164	0	-0.000735	0	-0.000361	0
C	1.67E-06	0	5.22E-06	0	3.71E-05	0	6.84E-07	0.0296	3.56E-07	0.3303	3.30E-07	0.35

Table 8: ECM estimates Summary 30sec frequency

Dependent Variable: LOG_DMICEX

	EESR		LKOH		RTKM		SIBN		SNGS		TATN	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
LOG_DRTS(-1)	0.028583	0	0.00682	0.1903	0.013684	0.0104	-0.001233	0.2454	0.002087	0.6947	-0.00258	0.7121
LOG_DRTS(-2)	0.014019	0.0322	-0.003596	0.4908	0.002668	0.617	0.003308	0.0018	0.016458	0.002	0.02933	0
LOG_DRTS(-3)	-0.008266	0.2067	0.002695	0.616	0.022485	0	-0.000973	0.3597	0.000311	0.9534	-0.010294	0.1419
LOG_DRTS(-4)	0.006822	0.2972	0.001246	0.8166	0.014774	0.0057	-0.001727	0.1038	-0.014608	0.006	0.022845	0.0011
LOG_DRTS(-5)	0.00856	0.1907	0.004807	0.371	0.005052	0.344	-0.000238	0.8225	-0.006792	0.2011	-0.002596	0.711
LOG_DRTS(-6)	0.004662	0.4494	0.000559	0.9169	0.004831	0.3655	0.000466	0.6607	0.009582	0.0713	-0.000366	0.9583
LOG_DRTS(-7)	-0.001849	0.7639	0.000607	0.9093	0.002312	0.6649	0.000106	0.9202	-0.000904	0.8647	0.005228	0.4553
LOG_DRTS(-8)	0.01591	0.0095	-0.001379	0.7927	-0.002021	0.7049	0.001406	0.1853	-0.000417	0.9374	-0.003284	0.639
LOG_DRTS(-9)	0.016252	0.008	-0.007019	0.1644	-0.006119	0.2509	0.001183	0.2654	0.008071	0.1286	0.013988	0.0458
LOG_DRTS(-10)	-0.006117	0.3179	0.002145	0.6683	-0.004792	0.3686	0.002024	0.0541	-0.004815	0.3627	-0.010304	0.1392
LOG_DMICEX(-1)	0.033228	0.0017	0.141668	0	-0.031822	0.0006	-0.022744	0.0087	-0.02232	0.0108	-0.053726	0
LOG_DMICEX(-2)	0.040982	0.0001	0.115016	0	0.018803	0.0417	0.028563	0.001	0.028171	0.0013	0.042749	0
LOG_DMICEX(-3)	0.008878	0.4025	-0.00548	0.6652	0.001757	0.8491	0.027029	0.0018	0.027643	0.0016	0.006384	0.4488
LOG_DMICEX(-4)	0.009817	0.3534	0.01854	0.1428	0.004882	0.5966	0.032429	0.0002	0.035772	0	0.024982	0.003
LOG_DMICEX(-5)	0.003181	0.7637	0.029317	0.0203	0.00288	0.7548	0.01354	0.1185	0.017806	0.0423	-0.033253	0.0001
LOG_DMICEX(-6)	0.011815	0.2648	-0.033588	0.0078	0.000837	0.9276	0.020789	0.0167	0.022104	0.0118	0.004911	0.5599
LOG_DMICEX(-7)	0.003427	0.7463	0.004673	0.7102	-0.045133	0	0.004525	0.5937	0.005047	0.5654	-0.007429	0.3774
LOG_DMICEX(-8)	0.006055	0.5676	0.032568	0.0094	-0.006511	0.4801	0.001514	0.8614	0.002462	0.7792	0.013152	0.1183
LOG_DMICEX(-9)	-0.021732	0.0389	-0.015693	0.2095	-0.00364	0.6929	0.009294	0.2834	0.008718	0.3204	-0.00087	0.9176
LOG_DMICEX(-10)	-0.008272	0.4308	-0.018242	0.1407	0.0022	0.8112	0.004202	0.6274	0.004054	0.6437	-0.016016	0.0565
SPREAD(-1)	**0.006855**	**0**	**0.008312**	**0.0002**	**0.003815**	**0.0001**	**0.000506**	**0.0771**	**0.00528**	**0**	**0.00434**	**0**
C	-9.03E-06	0.1572	7.95E-06	0.3576	-6.35E-06	0.3437	1.19E-05	0.0707	1.19E-05	0.0698	1.60E-05	0.0621

Dependent Variable: LOG_DRTS

	EESR		LKOH		RTKM		SIBN		SNGS		TATN	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
LOG_DRTS(-1)	-0.009467	0.3743	-0.075507	0	0.010273	0.2659	-0.000862	0.918	-0.063683	0	-0.085639	0
LOG_DRTS(-2)	-0.05066	0	-0.26959	0	-0.025021	0.0067	0.00263	0.7533	0.024561	0.005	0.002561	0.7606
LOG_DRTS(-3)	-0.004806	0.6525	-0.082973	0	-0.011489	0.2131	0.002551	0.7515	0.008332	0.3433	0.003867	0.6628
LOG_DRTS(-4)	-0.005967	0.5759	-0.100317	0	-0.000745	0.9356	0.000341	0.9675	-0.024914	0.0045	-0.003609	0.6678
LOG_DRTS(-5)	0.017683	0.0973	-0.053309	0	0.016985	0.0657	0.002363	0.7778	-0.007532	0.3901	-0.002798	0.7393
LOG_DRTS(-6)	-0.008977	0.3716	-0.05085	0.0001	-0.003917	0.6713	0.001955	0.8154	-0.009269	0.2901	0.001778	0.8324
LOG_DRTS(-7)	0.00792	0.43	-0.049631	0.0001	0.00457	0.6205	0.00012	0.9886	-0.059195	0	0.001509	0.8575
LOG_DRTS(-8)	-0.02557	0.0105	-0.035205	0.0059	0.001715	0.8525	0.001851	0.825	-0.004091	0.6407	-0.001347	0.8727
LOG_DRTS(-9)	0.001904	0.8488	-0.020458	0.0964	-0.033916	0.0002	0.001774	0.8322	0.005505	0.5298	-0.015272	0.0691
LOG_DRTS(-10)	-0.002703	0.7866	-0.026767	0.0283	-0.007545	0.413	0.003262	0.6937	0.002998	0.7312	-0.007443	0.3733
LOG_DMICEX(-1)	0.055755	0.0013	0.256608	0	0.03297	0.0388	-0.054962	0.4211	0.062575	0	0.052345	0
LOG_DMICEX(-2)	0.024141	0.1626	0.125384	0	0.03872	0.0153	0.031081	0.649	0.028811	0.0461	-0.007885	0.4352
LOG_DMICEX(-3)	0.058151	0.0008	0.036049	0.2429	-0.007825	0.6242	-0.135354	0.0476	0.042029	0.0036	0.008131	0.4215
LOG_DMICEX(-4)	0.048873	0.0046	0.092838	0.0026	0.001613	0.9195	-0.078119	0.2531	0.036111	0.0125	0.004019	0.6908
LOG_DMICEX(-5)	0.155289	0	-0.016579	0.5903	0.002067	0.8969	-0.125448	0.0644	0.005767	0.69	0.022211	0.028
LOG_DMICEX(-6)	-0.010471	0.5443	0.081442	0.0082	0.039002	0.0145	-0.020803	0.7612	0.052949	0.0003	0.018764	0.0634
LOG_DMICEX(-7)	0.03754	0.0297	0.047415	0.122	0.023734	0.1365	-0.164888	0.0158	0.005033	0.7282	0.023077	0.0223
LOG_DMICEX(-8)	0.038602	0.0254	0.116984	0.0001	0.005289	0.6932	-0.148974	0.0293	0.012665	0.3819	-0.003703	0.714
LOG_DMICEX(-9)	-0.000412	0.9808	-0.02778	0.3621	0.013563	0.3948	0.05759	0.3224	0.005592	0.6992	0.055965	0
LOG_DMICEX(-10)	-0.001823	0.9152	0.095645	0.0015	-0.013662	0.3909	0.095351	0.1624	0.003767	0.7944	0.050231	0
SPREAD(-1)	**-0.01136**	**0**	**-0.026672**	**0**	**-0.016093**	**0**	**-0.069801**	**0**	**-0.017868**	**0**	**-0.008218**	**0**
C	3.36E-05	0.0013	7.28E-05	0.0006	3.17E-05	0.0063	1.58E-05	0.762	8.13E-06	0.454	9.75E-06	0.3427

Table 9: ECM estimates Summary 60sec frequency

Dependent Variable: LOG_DMICEX

	EESR		LKOH		RTKM		SIBN		SNGS		TATN	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
LOG_DRTS(-1)	0.008704	0.3755	0.001022	0.9056	0.012547	0.0966	0.011417	0.245	0.009293	0.2245	0.016585	0.1126
LOG_DRTS(-2)	-0.009013	0.3589	0.014152	0.1166	**0.037255**	**0**	**0.038565**	**0.0001**	-0.014691	0.0548	0.018205	0.0814
LOG_DRTS(-3)	0.010639	0.2779	0.001458	0.8727	0.01151	0.1278	-0.007067	0.4718	0.009266	0.2255	0.0093	0.373
LOG_DRTS(-4)	**0.043728**	**0**	0.000143	0.9874	-0.001527	0.8396	0.001628	0.8684	-0.002711	0.7228	0.005696	0.5846
LOG_DRTS(-5)	-0.009041	0.3497	0.001638	0.8564	-0.006323	0.4015	0.00465	0.6331	-0.006978	0.3616	0.011186	0.2829
LOG_DRTS(-6)	0.000946	0.922	0.003317	0.7123	-0.002696	0.7202	**0.020817**	**0.0326**	0.004249	0.5785	-0.016218	0.1175
LOG_DRTS(-7)	**-0.068546**	**0**	-0.000873	0.9221	0.003119	0.6784	-0.008053	0.4112	0.004358	0.5673	-0.007677	0.4585
LOG_DRTS(-8)	-0.011371	0.2175	0.009735	0.2683	**0.024951**	**0.0009**	0.013452	0.1691	-0.01049	0.1675	-0.009014	0.3837
LOG_DRTS(-9)	**-0.018771**	**0.0399**	**0.020389**	**0.0166**	0.001315	0.8612	-0.003881	0.6914	0.005491	0.4693	-0.00411	0.6911
LOG_DRTS(-10)	0.003401	0.709	0.002857	0.7206	-0.00194	0.796	-0.006882	0.4811	0.005625	0.4574	0.004639	0.6537
LOG_DMICEX(-1)	0.062662	0	0.193423	0	0.004082	0.7593	-0.038729	0.0013	0.043052	0.0005	0.026553	0.0265
LOG_DMICEX(-2)	0.030962	0.0413	0.036566	0.0458	0.013345	0.3163	0.000639	0.9577	0.065392	0	0.020905	0.0806
LOG_DMICEX(-3)	0.017471	0.2504	-0.006875	0.7074	-0.000747	0.9551	-0.00796	0.5088	0.025386	0.0418	-0.033437	0.0052
LOG_DMICEX(-4)	0.001611	0.9156	0.04189	0.0223	-0.050916	0.0001	0.003045	0.8003	0.010824	0.3854	0.018223	0.1279
LOG_DMICEX(-5)	-0.008651	0.5686	-0.040647	0.0266	-0.015396	0.2469	0.038217	0.0015	0.011316	0.3642	-0.019357	0.1057
LOG_DMICEX(-6)	-0.010766	0.4778	-0.039679	0.0305	-0.014025	0.2914	-0.00077	0.949	0.013072	0.2945	-0.003573	0.766
LOG_DMICEX(-7)	-0.028635	0.0574	-0.01265	0.4899	-0.031168	0.0189	-0.025983	0.0309	-0.03194	0.0105	-0.0074	0.5376
LOG_DMICEX(-8)	-0.02045	0.1728	-0.018103	0.322	-0.039861	0.0027	-0.007662	0.5248	0.006227	0.6176	-0.011282	0.3468
LOG_DMICEX(-9)	0.02259	0.1322	-0.00962	0.5947	0.048639	0.0002	-0.016397	0.1728	0.00656	0.5987	-0.013385	0.2641
LOG_DMICEX(-10)	0.040173	0.0068	0.033495	0.0589	0.023377	0.0777	-0.000555	0.963	-0.015339	0.2181	-0.024412	0.0417
SPREAD(-1)	**0.012917**	**0**	**0.01206**	**0.0011**	**0.0066**	**0.0012**	**0.001566**	**0.2306**	**0.009732**	**0**	**0.008061**	**0**
C	-2.07E-05	0.1088	1.85E-05	0.3183	-1.03E-05	0.4417	1.78E-05	0.2161	2.35E-05	0.0666	3.22E-05	0.052

Dependent Variable: LOG_DRTS

	EESR		LKOH		RTKM		SIBN		SNGS		TATN	
	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.	Coefficient	Prob.
LOG_DRTS(-1)	-0.065704	0	-0.341915	0	-0.018454	0.1663	0.011121	0.3547	-0.003947	0.7516	-0.00179	0.8809
LOG_DRTS(-2)	-0.030326	0.0488	-0.182228	0	-0.013572	0.3088	0.039182	0.0011	-0.021486	0.0849	0.006125	0.6079
LOG_DRTS(-3)	0.011537	0.4525	-0.149598	0	0.012616	0.3445	-0.069303	0	-0.00761	0.5415	-0.003258	0.7849
LOG_DRTS(-4)	-0.032569	0.0332	-0.110547	0	0.001006	0.9398	-0.025959	0.0308	-0.055984	0	0.003491	0.7694
LOG_DRTS(-5)	-0.020216	0.1818	-0.098714	0	-0.043326	0.0011	0.001365	0.9088	0.008321	0.5047	-0.019929	0.0942
LOG_DRTS(-6)	-0.154258	0	-0.094725	0	-0.020897	0.1158	-0.111627	0	0.039966	0.0014	0.010189	0.3895
LOG_DRTS(-7)	-0.096975	0	-0.102912	0	-0.036182	0.0065	-0.031983	0.0077	0.017863	0.1504	0.010055	0.3957
LOG_DRTS(-8)	-0.017089	0.2366	-0.062795	0.0012	-0.023548	0.076	0.020063	0.0936	0.000945	0.9392	-0.008147	0.491
LOG_DRTS(-9)	-0.015023	0.2937	-0.049967	0.0075	0.012345	0.3523	-0.03087	0.0099	0.011928	0.335	-0.00237	0.8411
LOG_DRTS(-10)	-0.00443	0.7563	-0.020596	0.2404	-0.048593	0.0002	-0.004651	0.6971	0.009214	0.4553	0.001859	0.875
LOG_DMICEX(-1)	**0.076851**	**0.0012**	**0.266432**	**0**	0.039005	0.0974	0.046206	0.0017	**0.080419**	**0.0001**	0.008152	0.5511
LOG_DMICEX(-2)	**0.117261**	**0**	**0.12954**	**0.0013**	0.018154	0.44	0.009475	0.5205	0.027515	0.1752	**0.028685**	**0.0359**
LOG_DMICEX(-3)	**0.135075**	**0**	**0.120192**	**0.0028**	**0.075681**	**0.0013**	0.025189	0.0874	**0.065724**	**0.0012**	**0.028798**	**0.0352**
LOG_DMICEX(-4)	**0.104089**	**0**	**0.145144**	**0.0003**	0.003068	0.8959	0.017608	0.2319	0.035216	0.0833	0.01451	0.2889
LOG_DMICEX(-5)	0.020239	0.3944	**0.135872**	**0.0007**	-0.008405	0.7203	0.029254	0.0467	-0.015163	0.4559	**0.102956**	0
LOG_DMICEX(-6)	0.023238	0.3279	0.062835	0.1186	0.005401	0.818	**-0.050526**	**0.0006**	0.016572	0.415	0.016233	0.2368
LOG_DMICEX(-7)	-0.00552	0.815	**0.116595**	**0.0038**	-0.005175	0.8252	**0.051566**	**0.0005**	0.00821	0.6865	**-0.058599**	0
LOG_DMICEX(-8)	0.036919	0.1161	**0.092634**	**0.021**	0.01532	0.5131	0.019712	0.1811	0.012157	0.55	0.020391	0.1368
LOG_DMICEX(-9)	0.013026	0.5793	-0.028612	0.4711	-0.006448	0.7826	**-0.043519**	**0.0031**	**-0.055048**	**0.0068**	0.02681	0.0504
LOG_DMICEX(-10)	0.045134	0.0523	**0.092315**	**0.0177**	0.043892	0.0606	**0.04217**	**0.004**	-0.018526	0.3616	-0.007798	0.5692
SPREAD(-1)	**-0.013426**	**0.0006**	**-0.033588**	**0**	**-0.027268**	**0**	**-0.00716**	**0**	**-0.036492**	**0**	**-0.014675**	**0**
C	4.53E-05	0.0251	0.000115	0.0045	6.00E-05	0.0115	3.78E-05	0.0322	1.46E-05	0.485	1.74E-05	0.3583

The serial correlation misspecification seems to be minimised with the inclusion of a lag length comprised of ten lags on both time series variables. The most crucial explanatory variable of long run error correction relationship spread between RTS and MICEX from ECM of both models is negative and highly significant. That means that RTS returns variations are significantly, inversely related to the error correction term and the lagged residual spread returns term variable in the alternative model. In other words, if the cointegrated difference of logs of RTS and MICEX prices were positive in one period, the RTS price would fall to restore equilibrium in the next period and vice versa. Subsequently, in the long run perspective RTS prices would converge with MICEX prices and vice versa, where RTS lagging behind MICEX meaning that if the causality is unidirectional then price discovery could be mainly occurring in the thicker MICEX market, which would be a confirmation to the primary hypothesis. However different sampling frequencies and

inconsistent estimation results might have some undermining power. There might be a possibility that RTS might lead since the chosen sample frequency is a signal- nose trade off.

However, two fundamental conditions have been fulfilled, that is that the time series cointegrate, while the errors in prices converge correcting in the long run and display significant and negative coefficients resulting in markets having mutual explanatory powers. However, the thesis would be insufficiently supported by solely by these two conditions since there is no proof that the causality is unidirectional. There is a possibility that there is bidirectional causality. In order to prove that MICEX really leads, more evidence is required, which would prove that, when looking from other market perspective the error correction and residual terms coefficients are mutually and significantly excluding otherwise there would be some ambiguity in causality direction terms. Here, the estimated coefficients indicate a clear direction for all models. Additionally, MICEX and RTS significance does not differ across the different sampling frequencies employed.

Next to the most crucial long rung explanatory variable in ECM are the lag structure of short run variables. Opposite to long run variable there is no such persistence of evidence that in the short run one market is leading continuously the time however there is no clear direction of corrections. No contradictions arise with the estimation methods employed and with the sample size restrictions. If looked from the ECM short run perspective then the majority of RTS variables seem to lag with significant and negative coefficients.

In sum, the results do not significantly differ across the scope, sample size and their frequency of sampling. Given that, there is sufficient evidence to support the original thesis treated as a null hypothesis that overall, the information and price discovery seem to largely occur in the most liquid market. Overall, the thesis remains uncontested: Since there is not a single period in the whole sample where long run relationship coefficient of MICEX leadership is insignificant for all models, from the opposite direction the results are not undermined by the positive significance of the RTS resulting in unidirectional relationship, the question moves towards: What is the extent of the differences in information contents between MICEX and RTS?

6.3 Cross- sectional Information Shares in Price Discovery

Ganzalo and Granger (GG) information measure does provide additional insight in explaining the information shares of either market. This is very intuitive measure utilises the factor loadings of VECMs, while the values sum up to one. The Hasbrouck information shares (HIS) measures are more consistent than GG shares however provide only upper and lower bound of the information shares. The HIS measure is not reported due to the reason of time constraint.

The information shares of Ganzalo- Granger are confirming the inferences made that MICEX is the dominant price discoverer. The MICEX information measures are significantly above 50 percent for all six stocks as shown in Table 10.

Table 10: VECM Error Correction and GG shares Summary

Stock	Error Corr	D(LOG_MIC	D(LOG_RT	D(LOG_MIC	D(LOG_RT	D(LOG_MIC	D(LOG_RT
		1sec		30sec		60sec	
EESR	**CointEq1**	-0.000186	0.000739	-0.005965	0.008446	-0.011205	0.016553
		-4.20E-05	-7.30E-05	-0.00122	-0.00198	-0.00246	-0.00384
		[-4.46506]	[10.1619]	[-4.88875]	[4.26828]	[-4.56098]	[4.30723]
	GG	0.7989189	0.2010811	0.58608008	0.4139199	0.5963326	0.4036674
LKOH	**CointEq1**	-0.000247	0.002327	-0.007674	0.041137	-0.014539	0.062664
		-5.80E-05	-0.00016	-0.00172	-0.00419	-0.00373	-0.00822
		[-4.29033]	[14.7739]	[-4.45615]	[9.81018]	[-3.89890]	[7.62774]
	GG	0.9040404	0.0959596	0.84278134	0.1572187	0.8116783	0.1883217
RTKM	**CointEq1**	-0.000205	0.016374	-0.004171	0.016308	-0.00786	0.031213
		-2.30E-05	-0.00029	-0.00093	-0.00161	-0.00185	-0.00325
		[-8.83580]	[56.6905]	[-4.47443]	[10.1300]	[-4.24905]	[9.59440]
	GG	0.987635	0.012365	0.79632795	0.2036721	0.7988381	0.2011619
SIBN	**CointEq1**	-7.93E-05	0.000179	5.33E-05	0.043235	-0.002631	0.009656
		-2.70E-05	-3.00E-05	-0.00022	-0.00175	-0.00137	-0.00169
		[-2.88639]	[6.03915]	[0.24212]	[24.7089]	[-1.92226]	[5.71669]
	GG	0.6929926	0.3070074	1.00123432	-0.001234	0.7858712	0.2141288
SNGS	**CointEq1**	-0.000185	0.000783	-0.004239	0.021458	-0.009591	0.039594
		-4.00E-05	-6.10E-05	-0.0011	-0.00181	-0.00216	-0.00353
		[-4.60975]	[12.9272]	[-3.85641]	[11.8329]	[-4.43509]	[11.2187]
	GG	0.8088843	0.1911157	0.83503911	0.1649609	0.8050015	0.1949985
TATN	**CointEq1**	-0.000201	0.000359	-0.005247	0.00994	-0.010033	0.017907
		-3.70E-05	-4.20E-05	-0.00103	-0.00123	-0.00197	-0.00226
		[-5.35849]	[8.58691]	[-5.11783]	[8.08219]	[-5.08980]	[7.91823]
	GG	0.6410714	0.3589286	0.65450714	0.3454929	0.6409091	0.3590909

One could only presume that the reason for clear values far from half is a relatively very large MICEX liquidity relative to RTS and hence information share is contained in MICEX market, which is consistent to the findings of Harris et al (2003).

The striking question that emerges from the results reported in Figure is why stocks differ so much in terms of price discovery shares between MICEX and RTS with a range of frequencies. With higher frequencies the GG information share shows a higher information share of MICEX where on contrast with lower frequencies RTS share seems to increase by values between ten and one percent. The MICEX market information shares for market prices range from about 99 percent for RTKM to about 59 percent for EESR. The associated RTS information shares for home market prices range from less than 1 percent to about 41 percent, respectively. In between these extremes, there are some cases, where there is a substantial role for RTS price innovations in MICEX price discovery while in other cases, there is however a small role.

6.4 Restricted VECM and Unrestricted VAR

A general to specific modeling strategy leads to similar results with the presence of cointegration being robust to the number of lags. Initially, the optimal lag structure for unrestricted VAR has been identified according whatever is highest of the following information criteria of Akaike (AIC), Schwarz (SIC) and Hannan- Quinn (HQ). The optimum lag range has been between twenty and four lags for thirty seconds intervals. For real time frequency the lag structure tends to increase whereas for lower frequencies to decrease. For VECM, the determinant of lag length choice has been the SIC.

The performed Johansen cointegration tests revealed results which clearly support the hypothesis of one cointegrating vector among the two variables and hence one common stochastic trend as observed (see Table 12).

With the variables ordered as midpoint RTS and MICEX log market prices, the estimated cointegrating vectors are close to the vector $\beta^T = (1, -1)$ indicated by theory. Whether the practise has deviated significantly from the theory has been tested on later stage in VECM by imposing restrictions.

The estimates of the cointegration models determining the cointegrating rank, the factorisation results in the matrix of one cointegrating vector, β^T and the weighting matrix

α estimates of error correction adjustments for six stocks in the sample are more consistent estimates then the ECMs of Engle and Granger however are in line with stated inferences.

Almost all cointegrating vectors showed closeness to the theoretical $\beta^T = (1, -1)$ as in Table 6. However, most of the vectors deviated fairly from theoretical ideal. Therefore the restrictions for the VECM of the cointegrating vectors (1, – 1)' were imposed and tested whether the practise is significantly different from theory. The results are shown in Table 11.

Table 11: Restricted VECM cointegrating vectors test Results

1sec						
Cointegration Restrictions:	**EESR**	**LKOH**	**RTKM**	**SIBN**	**SNGS**	**TATN**
B(1,1)=1,B(1,2)=-1						
Chi-square(1)	7.88442	42.38366	114.3841	4.833948	14.08333	6.177694
Probability	0.004986	0	0	0.027905	0.000175	0.012937
30sec						
Cointegration Restrictions:	**EESR**	**LKOH**	**RTKM**	**SIBN**	**SNGS**	**TATN**
B(1,1)=1,B(1,2)=-1						
Chi-square(1)	1.72527	21.37958	5.576523	3.039437	12.75676	5.76537
Probability	0.189016	0.000004	0.018203	0.081264	0.000355	0.016345
60sec						
Cointegration Restrictions:	**EESR**	**LKOH**	**RTKM**	**SIBN**	**SNGS**	**TATN**
B(1,1)=1,B(1,2)=-1						
Chi-square(1)	1.72968	13.63942	4.848947	3.674659	11.99179	5.713086
Probability	0.188452	0.000221	0.027663	0.055246	0.000534	0.016839

The test results of the restricted models for stocks such as Lukoil displayed a rejection of null that there is no significance. This finding is not really surprising given the anomaly of the first day and relative short observation period.

Although the cointegrating vector is expected to become near theoretical perfect, it would be quite interesting to investigate if it that will not be the case since one would not expect a significantly constant equilibrium gap between the markets. A possible explanation for the existence of the market price gap could be the institutional difference of market venues, which are characterised by besides the difference of currencies quoted also by idiosyncratic trading rules. Furthermore, one should consider that the exchange rate variable has not been modelled. Yet, there is a currency risk, which might be attributed to the risk premium or discount contained in the relative large gap of the first trading day and persistent in cointegrating equilibrium. Therefore, the exchange rate variable might not be exogenous as previously assumed. Given such a discovery, a question that arises consequently is whether there are any persistence of such as bias in for the remaining period of 2006, which may lead

potential arbitrage opportunities question on the domestic or if applicable for cross- border markets however that would be the focus of the final chapter of the thesis.

Summarising the results so far, price discovery for most firms occurs largely in the MICEX market with a smaller, however statistically and economically significant role for RTS stocks. This is consistent with the MICEX market being the primary market for most stocks with RTS trading following the MICEX market. However, the RTS has a significantly less than half information share for all stocks and has more than a 10 percent information share for five stocks.

6.5 Model Misspecifications

The MICEX and RTS time series seem to have a long memory and autocorrelation bias as suggested in results reported in earlier sections. The quantity of significantly negative short run dynamics variables of MICEX out weights RTS. According to the Durbin- Watson statistic, the MICEX model seems to be rather positively serially correlated, while RTS seems to be more negatively. Therefore, the OLS standard errors estimates, which contribute to determination of significance of the dependent variables, in this case MICEX returns, is in inefficient and inconsistent relative to true standard errors. The consequence of presence of such a problem can lead to overstatement or understatement of true variability of MICEX as well as RTS significance of regressors, keeping in mind that inefficiency of standard error increases the probability of type II error, where null hypothesis is zero, not rejecting the null hypothesis which could in fact be false for MICEX variables.

Besides the presence issues of standard errors inefficiency, there is strong evidence that the residuals of examined samples are not normally distributed. The return distribution of the residuals is highly non normal. The main characteristic of this non normality is high kurtosis which is caused by a number of extreme values and outliers. Despite higher observation number the kurtosis and skewness values tend to worsen.

Consequently, overall inferences made on ECMs are questionable given the violation of OLS estimator consistency, which makes the estimates inconsistent unless the distribution is assumed to be normal only under the assumptions central limit theorem, that with infinitely large sample size the given series distribution would converge to normal distribution. However, in practise this assumption does not usually hold especially as in this case of market microstructure research. So, in order to achieve some model consistency Generalised

Methods of Moments (GMM), Maximum Likelihood or QML estimation methods have to be applied. Therefore, this study seeks remedy by employing Johansen (1988) VAR approach.

6.6 Stationarity and Order of Integration

Augmented Dickey- Fuller test rejected the null in first differences therefore the test indicating unit roots in the log of all seven stock time series for most sampling frequencies. All variables were identified as being integrated of order one at the sampling frequency of thirty seconds. Initially, the real time frequency revealed no unit root for RTS stocks. This could be probably be explained by the low bid- ask bounce which is caused by the fact that RTS market is significantly thinner than MICEX. In other words, RTS market does have enough neither signal nor noise relative to MICEX. Despite that the results are all in line with expectations.

6.7 Cointegration Test Results

The fundamental condition for the lead- lag relationship analysis is the cointegration between the log price variables. If the residuals of the level regression between markets time series are not stationary unlike their corresponding price series then the regression results are spurious. However, despite some issues with real time frequency, the ADF test coefficients displayed significance at least in 95 percent confidence level meaning that the null hypothesis, that there is a unit root, has been rejected for all stocks and frequencies. Therefore, the residuals of the level regression are concluded to be stationary. The results are reported in the Table 12.

Table 12: Johansen Cointegration Test Results at 1sec frequency

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)					
Stock	**Hypothesized**		**Trace**	**0.05**	
	No. of CE(	**Eigenvalu**	**Statistic**	**Critical Va**	**Prob.****
EESR	None *	0.00551	51.79832	12.3209	0
	At most 1	0.000119	1.090662	4.129906	0.3447
LKOH	None *	0.001224	242.7571	15.49471	0.0001
	At most 1	1.38E-10	2.73E-05	3.841466	0.9981
RTKM	None *	0.009002	3209.155	15.49471	1
	At most 1	1.08E-05	3.832261	3.841466	0.0503
SIBN	None *	0.000116	51.45111	15.49471	0
	At most 1	3.41E-06	1.47181	3.841466	0.2251
SNGS	None *	0.0005	200.2568	15.49471	0.0001
	At most 1	2.41E-06	0.960583	3.841466	0.327
TATN	None *	0.000251	109.0516	15.49471	0.0001
	At most 1	1.44E-06	0.62069	3.841466	0.4308

Additionally, Johansen cointegration tests have been performed with different options allowing for intercept and deterministic trend in the assumptions of data. The results clearly support the hypothesis of one cointegrating vector among the two variables amongst all stock pairs and sampling frequencies. With the variables ordered as midpoint log RTS and MICEX market prices, the estimated cointegrating vectors are close to the vector $\beta^T = (1, -1)'$ as indicated by theory.

6.8 Direction of Dynamics

From Tables 7-9, the interpretation of the regression output is divided into estimation models, where the causality is directional i.e. where long run coefficient of both markets is significant while having mutually exclusive signs hence indicating unambiguous causality relationship. In the given estimated model, the non rejection of the null hypothesis would be comprised of the number of confirmations, which is complemented by the alternative regression model. This for instance means that any given estimated model is confirmed by the alternative model then that such example would indicate that there is some degree of robustness. Given that, it is likely that null hypothesis should not be rejected, indicating that the thicker, MICEX market would be the primary leading price discoverer. In other words the more confirmations there are, the higher the robustness of the estimation results and hence the higher the likelihood that RTS lags MICEX.

Similar conclusions should arise if the order based data rebuild samples are compared to samples which have been derived from transactions data. However, there might be a possibility that the degree of Granger- causality could increase both with the range of sampling frequencies and with sampling methodologies. The order based sample may not reject the thesis while the transaction based method might reject the one way causality of the thesis. The order book based sample is expected to comply fully with the thesis, having one sided negative long run and positive short run significant coefficients. RTS market could take the short run price discovering position, which is still possible with lagging in the long run. Such inferences might start changing with different samples and frequencies employed, so where the positive and significant coefficient of the short run explanatory variable is likely to change the direction. When the lower frequencies are considered the directionality may disappear in contemporaneous correlation while both markets could have significance in the discovery process anyway. Ambiguity may also arise by the presence of equal information shares presented in the measurements methods of Ganzalo and Granger (1995) and

Hasbrouck (1995). However it, it is unlikely that all alternative methods, would contribute to ambiguity expressed in a both sided significances of the proposed thesis.

7 Conclusions

The preliminary findings of the first chapter provide evidence of a common stochastic trend amongst the RTS and MICEX pricing. The overall result indicates a clear relationship between the two. The evidence, though exhibits that the causality relationship is asymmetrical, which is expected for lead- lag relationship, with conclusions, which strongly suggest MICEX market leadership independent of the sampling methods applied to reconstruct the order books from given data samples. The results also indicate that the direction of causality is not unidirectional. Furthermore, the asymmetry is consistent to the notion of non frictionless markets and inefficiency. The overall conclusion which can be drawn is following: Price discovery seem to occur more in the denser MICEX market with a yet also statistically significant role for the competing RTS market.

References

Admati, A. and Pfleiderer, P. (1988), "A theory of intraday patterns: Volume and price variability", *Review of Financial Studies* 1, 3–40.

Auguy, M., Le Saoût, E., (1999), "La liquidité cachée durant la séance de cotation a la Bourse de Paris: une étude du règlement mensuel" Working Paper, U of Aix-Marseille III and U of Paris I.

Auguy, M., Le Saoût, E., (2001) "Analyse de l'effet compensation" Working Paper, U of Aix- Marseille III and U of Paris I.

Bacidore, J. M. and Sofianos G. (2002) Liquidity provision and specialist trading in NYSE-listed non-U.S. stocks, *Journal of Financial Economics* 63, 133–158.

Baruch, S., A. Karolyi, and M.L. Lemmon, (2005) "Multi-Market Trading and Liquidity: Theory and Evidence" working paper, Ohio State University

Beltran, H., Grammig, J., Menkveld, A.J., (2005) "Understanding the Limit Order Book: Conditioning on Trade Informativeness" Working Paper, Catholic U of Louvain, U of Tuebingen and U of Amsterdam.

Chowdhry, B. and Nanda, V. (1991) "Multimarket trading and market liquidity", *Review of Financial Studies* 4, 623–656.

De Winne, R., D'Hondt, C., (2005), "Market Transparency and Traders' Behavior: An Analysis on Euronext with Full Order Book Data" Working Paper, Catholic U of Mons.

Engle, R.F. and C.W.J. Granger (1987), "Cointegration and Error-Correction: Representation, Estimation, and Testing" Econometrica 55, pp. 251-276

Eun, C. and Sabherwal, S. (2003), "Cross-border listings and price discovery: Evidence from US-listed Canadian stocks", *Journal of Finance* 58, 549–574.

Frey, S., Grammig, J., (2005), "Liquidity supply and adverse selection in a pure limit order book market" Working Paper, U of Tuebingen.

Gonzalo, J. and C. W. J. Granger (1995), "Estimation of common long-memory components in co- integrated systems," Journal of Business and Economic Statistics, 13, 27–36.

Grammig, J. Melvin, M. and C. Schlag (2005) "Internationally cross- listed stock prices during overlapping trading hours: price discovery and exchange rate effects" Journal of Empirical Finance, 12, 1, pp. 139- 164

Goldstein, Michael A. & A. Kavajecz, Kenneth, (2000). "Eighths, sixteenths, and market depth: changes in tick size and liquidity provision on the NYSE," Journal of Financial Economics, Elsevier, vol. 56(1), pages 125-149, April.

Goldstein, Michael A. & Kavajecz, Kenneth A., (2004). "Trading strategies during circuit breakers and extreme market movements," Journal of Financial Markets, Elsevier, vol. 7(3), pages 301-333, June

Hall, A.D., Hautsch, N., (2004), "Order Aggressiveness and Order Book Dynamics" Working Paper, U of Technology, Sydney and U of Cophenagen.

Harris, F. H., McInish, T. H., Shoesmith, G. L., and Wood, R. A. (1995), "Cointegration, error correction, and price discovery on informationally linked security markets", *Journal of Financial and Quantitative Analysis* 30, 563– 580.

Harris, F. H., McInish, T. H., Shoesmith, G. L., and Wood, R. A. (2002), "Security price adjustment across exchanges: An investigation of common factor components for Dow stocks", *Journal of Financial Markets* 5, 277– 308.

Harris, L, (2003) "Trading and Exchanges", Oxford: Oxford University Press.

Hasbrouck, J. (1995), "One security, many markets: Determining the contributions to price discovery", *Journal of Finance* 50, 1175–1199.

Johansen, S. (1988),"Statistical analysis of cointegration vectors," Journal of Economic Dynamics and Control, 12, 231–254.

Karolyi, G. A., (2006), "The World of Cross-Listings and Cross- Listings of the World: Challenging Conventional Wisdom" pp.1-54

Kato, K., S. Linn, and J. Schallheim, (1991), "Are There Arbitrage Opportunities in the Markets for American Depository Receipts?" Journal Of International Financial Markets, Institutions and Money, 1, 13-32

Kavajecz, K., A., (1999), "A Specialist's Quoted Depth and the Limit Order Book" Journal of Finance 54, 747-771.

Kyle, A. (1985), "Continuous auctions and insider trading", *Econometrica* 53, 1315–1335.

Lau, S.T., and J.D. Diltz, (1994) "Stock returns and the transfer of information between the New York and Tokyo stock exchanges", Journal of International Money and Finance 13, pp. 211- 222

Lieberman, O., Ben- Zion, U. and S. Hauser, (1999) "A characterization of the price behaviour of international dual stocks: an error correction approach" Journal of International Money and Finance, Elsevier, 18(2), pp. 289- 304

Miller, P. D., and R. M. Morey, (1996), “The Intraday Pricing Behavior of International Dually Listed Securities” Journal of International Financial Markets, Institutions and Money, 6, pp. 79- 89

Moulton, P. and Wei, L. (2005), “A tale of two time zones: Cross-listed stocks liquidity and the availability of substitutes”, working paper, NYSE, New York, NY.

Park, J. and A. Tavakkol, (1994), “Are ADRs A Dollar Translation of their Underlying Securities?: The Case of Japanese ADRs” Journal of International Financial Markets, Institutions & Money, 4, pp. 77 – 87

Phylaktis, K. and Korczak, P. (2004), “Specialist trading and the price discovery process of NYSE listed Non-U.S. stocks”, working paper, City University of London, London, England.

Phylaktis, K. and Korczak, P. (2005), “Price discovery process in international cross-listings: Evidence from U.S.-listed British and French companies”, working paper, City University of London, London, England.

Phylaktis, K. and Korczak, P. (2007) "Specialist Trading and the Price discovery Process of NYSE- Listed Non- US Stocks" Working Paper Europa-University Viadrina Frankfurt.

Pulatkonak, M. and Sofianos, G. (1999), “The distribution of global trading in NYSE-listed non-U.S. stocks”, working paper, NYSE 99-03, New York, NY.

Stulz, R. M. (1999), “Globalization, corporate finance, and the cost of capital”, *Journal of Applied Corporate Finance* 12, 8–25.

Wahab, M. and A. Khandwala, (1993), “Why Diversify Internationally With ADRs?” Journal Of Portfolio Management, 20, pp. 75-82

Wahab, M., M. Lashgari, and R. Cohn, (1992), “Arbitrage in the American Depository Receipts Market Revisited” Journal of International Markets, Institutions and Money, 2, pp. 97-128

Wang, S.S., O.M. Rui, and M. Firth, (2002) “Return and volatility behavior of dually-traded stocks: the case of Hong Kong” Journal of International Money and Finance, 21, pp. 265-291

Appendix

Order Book Reconstruction

Table A1: Raw MICEX order based database

ORDERNO	ENTRY DATE	ENTRY TIME	SECURITY ID	BUY SELL	STATUS	PRICE	QUANTITY	BALANCE	AMENDTIME	PERIOD	VAL
236499086	10-Jan-06	101500	EESR	S	C	13.05	100	100	184502	O	1305
236499087	10-Jan-06	101500	SNGS	S	M	32.37	2000	0	0	O	64738
236499088	10-Jan-06	101500	RU14TATN3006	S	M	97.3	300	0	0	O	29190
236499089	10-Jan-06	101501	EESR	S	M	12.17	100	0	0	O	1217
236499092	10-Jan-06	101501	EESR	B	C	11.92	100	100	184502	O	1192
236499093	10-Jan-06	101502	LKOH	S	M	1720	40	0	0	O	68800
236499094	10-Jan-06	101502	RU14TATN3006	B	W	102.4	5000	5000	103027	O	5E+05
236499095	10-Jan-06	101502	LKOH	B	W	1705	1000	1000	103244	O	2E+06
236499099	10-Jan-06	101502	EESR	S	M	12.52	100	0	0	O	1252
236499103	10-Jan-06	101503	EESR	S	M	12.6	100	0	0	O	1260
236499104	10-Jan-06	101503	SNGS	B	W	32.2	5000	5000	103443	O	2E+05
236499105	10-Jan-06	101503	EESR	S	C	13.09	15000	15000	184502	O	2E+05

Table A2: Raw MICEX transaction based database

TRADENO	BUY SELL	TRADEDATE	TRADETIME	ORDERNO	SECURITY ID	PRICE	QUANTITY	VAL	PERIOD
215089410	B	10-Jan-06	103000	236503652	LKOH	1801	150	270150	N
215089410	S	10-Jan-06	103000	236505524	LKOH	1801	150	270150	N
215089411	B	10-Jan-06	103001	236505539	RU14GMKN0507	2489	10	24885	N
215089411	S	10-Jan-06	103001	236502315	RU14GMKN0507	2489	10	24885	N
215089412	B	10-Jan-06	103001	236505541	RTKM	66.3	300	19890	N
215089412	S	10-Jan-06	103001	236503005	RTKM	66.3	300	19890	N
215089413	B	10-Jan-06	103001	236503652	LKOH	1801	100	180100	N
215089413	S	10-Jan-06	103001	236505542	LKOH	1801	100	180100	N
215089414	B	10-Jan-06	103001	236503652	LKOH	1801	292	525892	N
215089414	S	10-Jan-06	103001	236505547	LKOH	1801	292	525892	N
215089415	B	10-Jan-06	103001	236505051	LKOH	1801	418	752818	N
215089415	S	10-Jan-06	103001	236505547	LKOH	1801	418	752818	N
215089416	B	10-Jan-06	103001	236505548	RU14GMKN0507	2489	10	24885	N

Table A3: Query ordered MICEX database

ORDERNO	ENTRY DATE	ENTRY TIME	SECURITY ID	BUY SELL	STATUS	PRICE	QUANTITY	BALANCE	AMENDTIME	PERIOD	VAL
236505534	10-Jan-06	103001	LKOH	S	W	1855	90	90	105238	N	166950
236505542	10-Jan-06	103001	LKOH	S	M	1801	100	0	0	N	180100
236505547	10-Jan-06	103001	LKOH	S	M	1801	710	0	0	N	1E+06
236505567	10-Jan-06	103002	LKOH	B	M	1792	1	0	0	N	1792.4
236505580	10-Jan-06	103003	LKOH	B	W	1776	300	300	103133	N	532803
236505584	10-Jan-06	103003	LKOH	S	M	1808	10	0	0	N	18077
236505591	10-Jan-06	103004	LKOH	B	W	1770	100	100	104214	N	177000
236505592	10-Jan-06	103004	LKOH	S	M	1740	5	0	0	N	8700
236505621	10-Jan-06	103008	LKOH	B	C	1710	300	300	184502	N	513000
236505622	10-Jan-06	103008	LKOH	B	C	1720	200	200	184502	N	344000
236505623	10-Jan-06	103008	LKOH	B	C	1763	234	234	184502	N	412542

Table A4: MICEX reconstructed order book in 30sec intervals

ORDERNO B	ORDERNO A	SECURITY ID	ENTRY DATE	ENTRY TIME	BID	ASK	MID_PRICE	QUANTITY	BALANCE
236505534	236511445	LKOH	10/01/2006	10:30:00	1797	1810	1803.5	90	90
236506051	236511873	LKOH	10/01/2006	10:30:30	1790	1815	1802.49	500	500
236506569	236512425	LKOH	10/01/2006	10:31:00	1791.5	1810	1800.74	2	0
236507084	236512940	LKOH	10/01/2006	10:31:30	1790.2	1805	1797.595	500	500
236507597	236513417	LKOH	10/01/2006	10:32:00	1790	1800	1795.015	649	649
236508079	236513842	LKOH	10/01/2006	10:32:30	1790.1	1798	1794.05	292	0
236508593	236514181	LKOH	10/01/2006	10:33:00	1781	1789	1785	500	0
236509077	236514572	LKOH	10/01/2006	10:33:30	1777	1795	1785.995	10	0
236509522	236515048	LKOH	10/01/2006	10:34:00	1778	1785	1781.485	5000	0
236510001	236515404	LKOH	10/01/2006	10:34:30	1783	1789	1786	20	0
236510516	236515766	LKOH	10/01/2006	10:35:00	1784	1790	1786.995	7	0

Table A5: MICEX and RTS time series combined in 30sec intervals

SECURITY ID	DATE	TIME	BID_MICEX	ASK_MICEX	BID_RTS	ASK_RTS	MID_PRICE_MICEX	MID_PRICE_RTS
LKOH	11-Jan-06	10:31:00	63.0922579	63.1975873	62.5	63.5	63.1449226	63
LKOH	11-Jan-06	10:31:30	63.0922579	63.2843084	62.75	63.5	63.18828317	63.125
LKOH	11-Jan-06	10:32:00	63.1975873	63.2843084	63	63.25	63.24094782	63.125
LKOH	11-Jan-06	10:32:30	63.2165465	63.2678068	63	63.25	63.24217667	63.125
LKOH	11-Jan-06	10:33:00	63.232697	63.2906281	63	63.25	63.26166259	63.125
LKOH	11-Jan-06	10:33:30	63.2439322	63.2906281	63.05	63.25	63.26728015	63.15
LKOH	11-Jan-06	10:34:00	63.2502519	63.2906281	63.05	63.25	63.27044003	63.15
LKOH	11-Jan-06	10:34:30	63.2513052	63.2906281	63.05	63.25	63.27096668	63.15
LKOH	11-Jan-06	10:35:00	63.2513052	63.2906281	63.05	63.25	63.27096668	63.15
LKOH	11-Jan-06	10:35:30	63.2509541	63.2906281	63.05	63.25	63.27079113	63.15
LKOH	11-Jan-06	10:36:00	63.2509541	63.2906281	63.1	63.3	63.27079113	63.2
LKOH	11-Jan-06	10:36:30	63.2330481	63.2906281	63.1	63.2	63.26183814	63.15
LKOH	11-Jan-06	10:37:00	63.1940763	63.2323459	63.1	63.2	63.2132111	63.15
LKOH	11-Jan-06	10:37:30	63.0606592	63.193023	63.1	63.25	63.12684107	63.175