Bivariate mixture model for pair of stocks: evidence from developing and developed markets

Stanislav Anatolyev* New Economic School

Alexander Varakin New Economic School

Abstract

We extend the Modified Mixture of Distribution model of Andersen (1996) to the case of a pair of assets whose return volatilities and trading volumes are driven by own latent information variables, with the shocks to the two being correlated. The model allows one to reveal what fraction of information flows is due to news that may be common for the whole market, common for the industry, common for a particular exchange where the stocks are traded, etc. We estimate the model using modifications of the GMM procedure, and data from the Russian stock market represented by two exchanges and a small number of stocks traded on both, and from the American stock market represented by one exchange and stocks from a few industries. The results indicate that the information flows are more highly correlated in the Russian market for a number of reasons, while at the American market the common component seems to be negligible, except when the two companies belong to the same industry.

Key words: Return volatility; Trading volume, Information flow, Mixture of Distribution Hypothesis, Generalized method of moments, Stock market.

*Corresponding author. Address: Stanislav Anatolyev, New Economic School, Nakhimovsky Prospekt, 47, Moscow, 117418 Russia. E-mail: sanatoly@nes.ru. We thank all members of the NES research project "Dynamics in Russian and Other Financial Markets" for intensive discussions.

1 Introduction

The relationship between return volatility and trading volume has been the focus of theoretical and empirical research for a long time. Along with univariate models for return volatilities, bivariate models for returns and trading volumes have been developed under a variety of approaches. Within the ARCH framework, Lamoureux and Lastrapes (1990) inserted the volume directly in the GARCH process for the return volatility, and found that the volume was strongly significant while the past return shocks were insignificant, which confirmed that the trading volume is driven by the same factors that generate the return volatility. Another approach was taken by Gallant, Ross and Tauchen (1992) who used semi-nonparametric estimation of the joint density of price changes and trading volumes conditional on past price changes and trading volumes. Tauchen, Zhang and Liu (1996) used a semi-nonparametric framework and impulse response analysis to investigate the relationship between return volatility, trading volume, and leverage. Tauchen and Pitts (1983) put forth a structural approach called the "Mixture of Distribution Hypothesis" (MDH) to modeling the joint distribution of returns and trading volumes conditional on an underlying latent variable that proxies information flowing to the market. The MDH paradigm was improved upon in several respects by Andersen (1996) and Liesenfeld (1998, 2001); see Section 2.

In this paper, we extend the Modified MDH model of Andersen (1996) to the case of a pair of assets whose return volatilities and trading volumes are each driven by its own latent information variable. The shocks to the two information variables are allowed to be correlated, with the corresponding correlation coefficient being of primary interest. Such modeling allows one to reveal what fraction of information flows is caused by news that may be common for the whole market, common for the industry, common for a particular exchange where the stocks are traded, etc., after cross-comparison of results for a variety of asset pairs. We estimate the model using data from the developing Russian stock market represented by two exchanges and a small number of stocks from few industries traded on both exchanges, and data from the developed American stock market represented by one exchange and stocks from many more industries. The results indicate that the information flows are more highly correlated in the Russian market due to high political and economic risks, more highly correlated when the companies belong to the same industry, and more highly correlated when the stocks are traded at the same exchange, although the correlation is nearly perfect for the same stocks traded at different exchanges. At the American market, the common component of information flows seems to be negligible except when the two companies belong to the same industry, although some relatively high correlations exist for some pairs of companies from different industries.

From an existing variety of estimation methods usually applied to bivariate mixture models we choose the GMM framework also used by Richardson and Smith (1994) and Andersen (1996). To cope with the problem of not so big sample sizes, we apply several modifications of the GMM – the continuously updating GMM of Hansen, Heaton and Yaron (1996) and the downward testing algorithm of selecting correct moment restrictions described in Andrews (1999); for details, see Section 3. The GMM diagnostic tests attest that the exploited features of the model do provide a good fit to the data even though the model as a whole may not account for all observed features of the joint distribution of return volatilities and trading volumes.

A study close in goals to this paper is Spierdijk, Nijman, and van Soest (2002) which tries to identify commonality of information and distinguish sector and stock specific news for a pair of assets using ultra-high frequency data. The authors apply their bivariate model for trading intensities to transaction data of stocks of several NYSE-traded US department stores. They conclude that there is a large amount of common information in information flows, although it is not completely clear if this is due to common industry news, or common exchange news, or news common for the entire market.

The present paper is organized as follows. Section 2 briefly overviews the history of bivariate mixture models, and presents an extension of Andersen's (1996) model to the case of two stocks. Section 3 contains the discussion of estimation methods. The description of the data is given in Section 4. The results are reported and analyzed in Section 5, and Section 6 concludes.

2 Model

2.1 Bivariate mixture models

The structural approach to analyzing the relationship between return volatility and trading volume based on information arrivals was first put forth by Tauchen and Pitts (1983). In their framework, the asset market passes through a sequence of equilibria driven by arrivals of new information to the market. The changes in prices and trade volumes aggregated across traders are approximately normally distributed; when aggregated throughout the day t having I_t information arrivals the daily return r_t and daily trading volume V_t are also approximately normal conditional on I_t which is random:

$$\begin{aligned} r_t|I_t &\sim N(0, \sigma_r^2 I_t), \\ V_t|I_t &\sim N(\mu_V I_t, \sigma_V^2 I_t). \end{aligned}$$

This model is termed the *Mixture of Distribution Hypothesis (MDH).* The dynamic behavior of the return and trading volume depends on the dynamics of the latent variable I_t. Richardson and Smith (1994) estimate and test this model without restrictions placed on the form of the process the latent information variable follows using the GMM procedure. They find out that the latent information variable has positive skewness and large kurtosis and exhibits underdisperion. While many standard distributional assumptions for this variable can be rejected, Richardson and Smith (1994) find that parameter restrictions passing the tests are close to those implied by a log-normally distributed information variable. Other authors have attempted to impose a dynamic structure on the information variable, typically an autoregressive process of low order in logarithms or another transformation, to identify the parameters of its dynamics, primarily the degree of persistence.

Liesenfeld (2001) proposes an alternative *Generalized Mixture of Distribution Hypothesis (GMH)* where the parameters measuring the sensitivity of traders' reservation prices are time varying and directed by a common latent variable J_t measuring the general degree of uncertainty. As a result, the returns and volumes are driven by two latent variables, I_t and J_t:

$$\begin{aligned} r_t|I_t, J_t &\sim N\left(0, (\sigma_{r,1}^2 J_t^{\alpha_1} + \sigma_{r,2}^2 J_t^{\alpha_2}) I_t\right), \\ V_t|I_t, J_t &\sim N\left(\mu_{V,1} + \mu_{V,2} J_t^{\alpha_2/2} I_t, \sigma_V^2 J_t^{\alpha_2} I_t\right), \end{aligned}$$

where I_t and J_t follow autoregression-type processes in logarithms. By estimating the MDH and GMH for IBM and Kodak stocks using the SML procedure Liesenfeld (2001) finds that the MDH is clearly rejected against the GMH. One of conclusions is that due to low persistence in return volatility in the estimated MDH and some other aspects the baseline MDH model cannot capture some important aspects of the volatility dynamics adequately.

Andersen (1996) develops another alternative model using the theoretical framework of Glosten and Milgrom (1985). In his modification, there are two types of trading volume that are due to informed traders and uninformed traders. The uninformed component is governed by a time invariant Poisson process with constant intensity m_0, while the informed volume has a Poisson distribution with parameter $m_1 I_t$ conditional on the number of news arrivals. Hence the daily trading volume, being a sum of informed and uniformed components, is distributed as Poisson too:

$$V_t | I_t \sim Po(m_0 + m_1 I_t).$$

The bivariate distribution in the Andersen (1996) *Modified Mixture of Distribution Hypothesis (MMH)* model is

$$\begin{aligned} r_t | I_t &\sim N(\bar{r}, I_t), \\ V_t | I_t &\sim c \cdot Po(m_0 + m_1 I_t), \end{aligned}$$

where the parameter σ_r^2 is set equal to 1 because the model is invariant to a scale transformation of the information variable. The parameter c in the conditional distribution of volume comes out from the process of detrending (for details, see Andersen, 1996), and allows to distinguish the conditional mean and variance of volumes. The coefficients cm_0 and $cm_1 EI_t$ characterize the average uninformed and informed parts of volume respectively, so one can easily find the corresponding shares of volumes of uninformed and informed trades. Note also that for greater flexibility the conditional distribution of returns has a nonzero mean in contrast to the previous discussion.

As can be seen the volume may take only positive values so this feature can be considered as the advantage of this model over the MDH, which is an obvious advantage over previous specifications. Using the GMM procedure without restrictions placed on the dynamics of the information variable Andersen (1996) estimates both the MDH and MMH for several NYSE-traded stocks. He finds that the MMH is an adequate model for these assets while the MDH is clearly rejected. Furthermore, he imposes a restriction on the process for the information variable in the form

$$I_t^{1/2} = \omega + \beta I_{t-1}^{1/2} + \alpha I_{t-1}^{1/2} u_t, \quad u_t \sim i.i.d.\ (1, \sigma_u^2), \quad u_t > 0.$$

Considering different distributions of u_t (with σ_u^2 being some known constant) he estimates the MMH together with the univariate mixture model for returns. One of main conclusions is that there is a significant reduction in the measure of volatility persistence when the univariate model for returns is expanded to encompass data on trading volumes. The full MMH model passes all diagnostic tests.

Liesenfeld (1998) is more pessimistic about the adequacy of the MMH model. He obtains similar results using the data on four major German stocks. He estimates the univariate model for returns, the MDH, and the MMH using the SML procedure, with the information variable following an AR(1) process in logarithms

$$\ln I_t = \alpha + \beta \ln I_{t-1} + u_t, \quad u_t \sim i.i.d.\ N(0, \sigma_u^2).$$

Liesenfeld (1998) finds that while the MMH is generally more preferred than the MDH the estimates of the persistence of the information variable in both models are still lower than in the univariate model for returns, so he doubted the validity of bivariate models. He proposed a formal test to show that there is an additional source of persistence in return volatility which is not captured by the information variable; this test reveals the presence of such source.

The literature has other examples of criticism of the MDH paradigm. Interestingly, Luu and Martens (2003) argue that rejections of the MDH obtained within the ARCH framework may be caused by an imprecise measure of volatility.

2.2 Two-stock MMH model

We formulate the MMH model for a pair of stocks by extending the MMH model considered in Andersen (1996) except that the logarithm of the information variable follows a Gaussian AR(1)-process:

$$\begin{aligned} &r_t | I_t \sim N(\bar{r}, I_t) \\ &V_t | I_t \sim c \cdot Po(m_0 + m_1 I_t) \\ &\ln I_t = \alpha + \beta \ln I_{t-1} + u_t, \quad u_t \sim i.i.d.\ N(0, \sigma_u^2), \end{aligned} \tag{1}$$

and r_t and V_t are independent conditional on I_t. In choosing the form of the conditional distribution of the information variable, we are driven by the following two reasons. First, Richardson and Smith (1994) found that estimates of various moments of the information variable were close to those implied by its being log-normally distributed. The second reason is a relative simplicity of formulating the set of moment conditions when we consider the extension of this model. As a guard against possible misspecifications of the conditional distributions and/or form of dynamics we use an estimation procedure robust to the presence of such misspecifications (see Section 3).

The key idea in extending this framework to a pair of stocks is that the dynamics of the return volatility and trading volume of each stock is driven by the dynamics of its own information variable that characterizes the amount of news coming to the market during the day and concerning this particular stock. At the same time, the flows of information concerning different stocks may interact with each other. This interaction can be allowed and analyzed via the correlation coefficient between shocks to the information variables for the two stocks. Hence, for two stocks labelled 1 and 2, the model is

$$\begin{aligned}
r_{j,t}|I_{1,t}, I_{2,t} &\sim N(\bar{r}_j, I_{j,t}), \quad j \in \{1,2\}, \\
V_{j,t}|I_{1,t}, I_{2,t} &\sim c_j \cdot Po(m_{j,0} + m_{j,1} I_{j,t}), \quad j \in \{1,2\}, \\
\ln I_{j,t} &= \alpha_j + \beta_j \ln I_{j,t-1} + u_{j,t}, \quad j \in \{1,2\}, \\
\begin{pmatrix} u_{1,t} \\ u_{2,t} \end{pmatrix} &\sim i.i.d.\ N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix} \right),
\end{aligned} \tag{2}$$

and $r_{1,t}$, $r_{2,t}$, $V_{1,t}$ and $V_{2,t}$ are independent conditional on $I_{1,t}, I_{2,t}$. It is expected that an estimate of σ_{12} will be positive due to the information common for the two stocks. This common information can have several sources. First, it may be common for the whole market, resulting from overall political or economic news that have an effect on the stock market. This kind of information may have especially significant effects on decisions of traders and investors in emerging markets in developing and transition countries, while the amount of such information is presumably lower in the developed countries due to lower political risks. Second, if the stocks belong to companies from the same industry, the information concerning this industry may have an effect on these companies simultaneously, and traders may change their decisions concerning stocks of companies from this industry. A primary example of information common for the industry is changes in world prices of energy sources. Third, the correlation between information variables belonging to seemingly unrelated stocks may be high in concentrated markets with only few liquid assets when the traders body contains few big players who can invest or withdraw funds into or from different assets simultaneously.

A key variable of interest thus is the correlation coefficient

$$\rho_{12} = \frac{\sigma_{12}}{\sigma_1 \sigma_2} \tag{3}$$

which can be tested for equality to zero (corresponding to the assumption of independence between the two information variables) but is expected to be positive. Indeed, suppose that the daily shocks ($u_{j,t}$) for the two information variables are divided into two parts: one part (u_t) is a common shock (common for the whole market or for these particular two stocks), and the other part ($\tilde{u}_{j,t}$) contains shocks that are unique for each stock, independent of each other and the common shock:

$$
\begin{aligned}
u_{j,t} &= u_t + \tilde{u}_{j,t}, \quad j \in \{1,2\}, \\
u_t &\sim i.i.d.\ N(0,\sigma^2), \\
\begin{pmatrix} \tilde{u}_{1,t} \\ \tilde{u}_{2,t} \end{pmatrix} &\sim i.i.d.\ N\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} \tilde{\sigma}_1^2 & 0 \\ 0 & \tilde{\sigma}_2^2 \end{pmatrix}\right).
\end{aligned}
$$

It is easy to see that $\sigma_j^2 = \sigma^2 + \tilde{\sigma}_j^2$, $j \in \{1,2\}$, and $\sigma_{12} = \sigma^2 > 0$, hence $\rho_{12} > 0$ too. However, we do not impose this condition during estimation in order to let the data determine the sign of this correlation.

3 Estimation issues

The key problem in estimating the mixture models is that the information variable that drives the dynamics of returns and volumes is latent. Three major methods of estimation of such models are the Generalized Method of Moments (GMM) used by Andersen (1996) and Richardson and Smith (1994), Simulated Maximum Likelihood (SML) used by Liesenfeld (1998, 2001) and Liesenfeld and Richard (2002), and Bayeasian Markov Chain Monte Carlo applied in Watanabe (2000, 2003). In this paper, we take the GMM approach because it is simpler and less computer-intensive (which especially matters in the two stocks model), and in addition is able to handle models with elements containing misspecification (see below). That is, not exploiting all distributional features of the model, which Liesenfeld (1998, 2001) attributes to drawbacks of the GMM, we instead regard as an advantage.

In the next few subsections, we give details on how we run the GMM procedure. Many of the modifications to the baseline GMM that we apply in this paper are motivated by relatively smaller sample sizes than those used in this literature when the GMM is employed.

3.1 Continuously updating GMM

In the present context, the classical GMM procedure (Hansen, 1982) is based on the minimization of the quadratic distance between the sample moments and analytical moments using the efficient weighting matrix that is inversely proportional to the (long-run) variance of the sample moments. Using the specified conditional distributions of returns and trading volumes one can find unconditional moments of returns and volumes. These moments are certain closed-form functions of the deep parameters of the model (these functions are derived in Appendices A and B). In our model formulation, the deep parameters are the parameters figuring to the conditional distributions and the law of motion for the information variable.

It is widely recognized that in situations when theoretical and empirical moments are matched, the classical GMM estimator may be severely biased in samples that are not large (see, for example, Andersen and Sørensen, 1995). In this paper, we apply a modification of the GMM estimation called the *continuously updating GMM (CU).* This method presumes simultaneous optimization of the GMM criterion function over parameters both in the moment function and in the weighting matrix. The CU was introduced in Hansen, Heaton and Yaron (1996) where the CU estimator was shown to exhibit smaller biases than classical GMM estimators in time series applications when sample sizes are not large. An intuitive explanation for such behavior was provided by Donald and Newey (2000), while Newey and Smith (2004) showed the presence of such tendencies by appealing to second order asymptotic properties.

3.2 GMM weighting matrix

The weighting matrix in the GMM or CU procedure is chosen so that it minimizes the asymptotic variance of the estimated parameters. In our problem when moment functions are serially correlated with unknown, possibly infinite, order, the form of the inverse to the efficient weighting matrix is the long-run variance of the moment function

$$\sum_{j=-\infty}^{+\infty} E\left[m(Z_t,\theta)m(Z_{t-j},\theta)'\right],$$

where $m(Z_t,\theta)$ is the moment function (the difference between sample moments and analytical moments), whose arguments are the vector of data Z_t that includes returns and volumes together with their lags, and the parameter vector θ. A (positive definite) estimate of this matrix can be obtained in the Newey and West (1987) form:

$$\frac{1}{T}\sum_{j=-b}^{b}\left(1-\frac{|j|}{b+1}\right)\sum_{t=\max(1,1+j)}^{\min(T,T+j)}\left(m(Z_t,\theta)-\bar{m}(\theta)\right)\left(m(Z_{t-j},\theta)-\bar{m}(\theta)\right)',$$

where T is the sample size, b is a positive lag truncation parameter, and

$$\bar{m}(\theta)=\frac{1}{T}\sum_{t=1}^{T}m(Z_t,\theta).$$

Notice that when not all moment conditions are satisfied (see below), it is important to subtract the average of the moment function $\bar{m}(\theta)$ (see Andrews, 1999, and Hall, 2000).

It is important to choose carefully lag truncation for the Newey–West estimator. It has been argued in the literature that when the sample is large the number of lags in the estimate of the weighting matrix should be sufficiently large too. Andersen and Sørensen (1995) suggested the following formula: $b \approx \lfloor \gamma T^{1/3} \rfloor$, where γ is some constant which varies between 0.6 and 5, and equals 1.2 for most experiments in their work. Andersen (1996) for the sample of about 4,700 observations used 75 lags in the estimation of the weighting matrix. Our samples are three times as small. Following Andersen and Sørensen (1995), we set the lag truncation parameter equal to $\lfloor 1.2T^{1/3} \rfloor$; for our sample sizes (about 1300 observations) it equals 13.

3.3 Moment selection

As mentioned before, because of a tight distributional specification of the model, we use estimation robust to the presence of possible misspecifications. This means that we run CU on a set of moment conditions (of which the model implies an infinite number) that result from a consistent procedure of moment selection. To this end, we use the "downward testing algorithm" described in Andrews (1999) applied to an initial set of (most reliable) moment restrictions to end up with a set of only "right" ones. The downward testing algorithm, along with the upward testing algorithm and the selection algorithm based on information criteria has been proved to be consistent and well behaving in finite samples (see Andrews, 1999). The idea of this algorithm is the following: moments are successively removed from the set of moment restrictions until it is not possible to reject the model using the J-test with the 5%-significant level, with the model having minimal J-statistic being preferred among models with the same number of moment restrictions.

The starting set of moment conditions is formed according to the following principles. Because of relatively small samples that are used in this study the number of moment restrictions should not be very large (see Andersen and Sørensen, 1995). It is also known that it is harder to estimate moments of higher order from such samples. Therefore, we confine ourselves only to moments of order not higher than two, and run simulation experiments to be certain that such moments can be accurately estimated using samples of sizes we have (such experiments indicate, in particular, that third and fourth order moments exploited, e.g., in Andersen, 1996, are estimated rather imprecisely). This stone also kills a second bird: by using only lower order moments we refrain from using relationships between low and high order moments implied by the posited distributions. In forming the set of moments, we abstain from using moments conditions that are less likely to be satisfied in data (e.g., the implicit zero skewness in returns, or the implicit zero covariance of returns of different assets). Similarly, to guard ourselves against misspecifications in the dynamics of information variables, we include only first three lags of dynamic moments (Andersen, 1996, used up to 20 lags, but the sample was far larger). In addition, in the two-stock model the cross-moments enter symmetrically: if, for example, $E[r_{1,t}V_{2,t}]$ is included, then $E[r_{2,t}V_{1,t}]$ is included too until such moments have identical expressions via parameters. Eventually, the outlined strategy resulted in the following starting set of moment restrictions for the one-stock model:

$$\begin{gathered} E[r_t], \quad E[|r_t-\bar{r}|], \quad E\left[(r_t-\bar{r})^2\right], \quad E[V_t], \\ E[|r_t-\bar{r}|\,|r_{t-k}-\bar{r}|], \quad E\left[\left(V_t-\bar{V}\right)^2\right], \quad E\left[\left(V_t-\bar{V}\right)\left(V_{t-k}-\bar{V}\right)\right], \end{gathered} \tag{4}$$

where $k \in \{1,2,3\}$, and $\bar{V} = E[V_t]$, so initially there are 11 moments and 7 parameters. For the two-stock model, the starting set of moments is

$$\begin{gathered} E[r_{i,t}], \quad E[|r_{i,t}-\bar{r}_i|], \quad E\left[(r_{i,t}-\bar{r}_i)^2\right], \quad E[V_{i,t}], \\ E[|r_{i,t}-\bar{r}_i|\,|r_{i,t-k}-\bar{r}_i|], \quad E\left[\left(V_{i,t}-\bar{V}_i\right)^2\right], \quad E\left[\left(V_{i,t}-\bar{V}_i\right)\left(V_{i,t-k}-\bar{V}_i\right)\right], \\ E[|r_{i,t}-\bar{r}_i|\,r_{j,t}], \quad E[V_{i,t}V_{j,t}], \quad E[V_{i,t}r_{j,t}], \quad E[|r_{i,t}-\bar{r}_i|\,V_{j,t}], \end{gathered} \tag{5}$$

where $i,j \in \{1,2\}$, $i \neq j$, $k \in \{1,2,3\}$, and $\bar{V}_i = E[V_{i,t}]$, so initially there are 29 moments and 15 parameters. The analytical expressions for these moments are derived in Appendices A and B. In the course of applying the downward testing algorithm, we remove only moment restrictions that figure in the second line in (4) and in the second and third lines in (5); the moments that figure in the first lines in (4) and (5) are always regarded right. Interestingly, for no stock or pair of stocks did we have to exclude more than two restrictions, in the majority of cases removing only one moment or not removing at all. These facts give a rather convincing empirical support to the MMH model in both original and modified forms.

3.4 Estimation algorithm

To summarize, the estimation is run in the following steps. First, the one-stock model is estimated by the continuously updating GMM using the downward testing algorithm.

The shares of volumes of uninformed and informed trades are calculated. Then, using the obtained estimates as starting values (the starting value for σ_{12} is set to zero), the two-stock model is estimated by the continuously updating GMM using the downward testing algorithm. Finally, the correlation coefficient ρ_{12} is computed using the estimates of σ_1^2, σ_2^2 and σ_{12}, and its standard errors are constructed by the delta-method.

4 Data

We use data from the developing Russian stock market which is of primary interest to us, and in addition data from the developed American stock market. The Russian market is represented by two exchanges and a small number of stocks traded on both exchanges and belonging to three industries. In contrast, the American market is represented by one exchange and stocks from many more industries. In order to make more fair comparisons, we use samples of approximately equal size.

The organized stock market in Russia is composed of several stock exchanges, two of which, MICEx (short for "Moscow Interbank Currency Exchange") and RTS (short for "Russian Trading System"), account for more than 95 percent of trade turnover, with the share of MICEx being near 80 percent. A brief introduction to the Russian stock market can be found in Ostrovsky (2003); details on the MICEx and RTS are available in English at `www.micex.com` and `www.rts.ru/?tid=2`, respectively. The assets are traded in rubles at the MICEx, but in US dollars at the RTS. The players primarily represent Russian investors; the percentages of American and European investors are relatively small. On each exchange more than a hundred equity stocks are transacted along with corporate and government bonds and other assets. Most of stocks are traded very rarely, but several blue chips are traded at a frequency of up to 6,000 transactions a day. The MICEx and RTS are evidently quite active for an Eastern European market compared, for example, with the Czech stock market, with most liquid stocks being traded at 67 trades per day (Hanousek and Podpiera, 2003). There is a universal perception in the Russian financial market that market prices of traded equities do not reflect their underlying fundamental values. Dividends on blue chips are extremely rarely paid; when paid, they constitute a tiny fraction of the market price. Capitalization figures also have little to do with the fundamental value; they are inherited from Soviet era bookkeeping, and are said to be underestimated. Hence, price fluctuations reflect more the dynamics of overall economic and political factors than changes in fundamental values.

So, the first sample covers the period from March 1, 1999 (when the normal trading regime started), to June 4, 2004, composed of 1,311 trading days, and contains daily closing prices and number of lots for four Russian corporations whose common stocks were most frequently traded at the both exchanges during the whole period. Among these four companies, two, SurgutNefteGaz (SNGS) and Lukoil (LKOH) are oil extractors, Unified Energy System of Russia (EESR) is the largest electricity producer, and RosTeleKom (RTKM) is a leading Russian telecommunications company. We do not consider some important stocks whose trading history does not go so far back as well as belonging to companies that were subject to government attacks during this period. One of leading Russian oil extractors Yukos falls into both categories. The data are taken from `www.finam.ru`, `www.micex.ru`, and `www.rts.ru`.

The second sample covers the period from January 4, 1999, to April 30, 2004, composed of 1,332 trading days, and contains daily closing prices corrected for dividends, and daily number of traded shares for the common stocks of British Petroleum (BP), Chevron-Texaco (CVX), Ford Motor (F), DaimlerChrysler AG (DCX), International Business Machines (IBM), Hewlett–Packard (HPQ), Verizon Communications (VZ), SBC Communications (SBC), Merck&Co (MRK), GlaxoSmithKline (GSK), McDonald's (MCD), Yum! Brands (YUM) at the New York Stock Exchange (NYSE). These stocks represent six different industries with two stocks in each industry: Oil & Gas Integrated (BP and CVX), Auto & Truck Manufacturers (F and DCX), Computer Hardware (IBM and HPQ), Communications Services (VZ and SBC), Major Drugs (MRK and GSK), and Restaurants (MCD and YUM). Such choice allows us to see if there is higher correlation between the information variables of two stocks that belong to the companies from the same industry. These data are taken from `www.finance.yahoo.com`.

The daily return r_t is the log-difference of closing stock prices, $r_t = \ln p_t - \ln p_{t-1}$. The daily observed volume series V_t^O is the number of traded shares or lots. The summary statistics for the returns are presented in Tables **??** and 2 for the Russian and American stocks, respectively. As one can see, the returns on Russian stocks are larger and slightly more volatile. The distributions of returns are non-normal, with no or positive skewness (with an exception of two restaurants) for both Russian and American stocks, and comparable kurtosises (with an exception of HPQ). Interestingly, stocks for the same companies traded at different Russian exchanges have more similar characteristics than stocks for different companies. The values of the Ljung-Box statistics indicate that there is a significant autocorrelation in squared and absolute returns much varying across stocks.

The observed volume series for all stocks have a trend component that should be removed. It been argued that if trading volume is strongly trended the estimation results for the bivariate mixture model may be very misleading (Tauchen and Pitts, 1983); in addition, it is important to have stationary series to use GMM (Andersen, 1996). We follow a simple procedure similar to one used by Liesenfeld (1998) and Watanabe (2000) to remove the exponential trend from volumes, but we also take care of the effects of holidays and weekends as Andersen (1996) reports the existence of such effects. We regress logarithm of trading volume on a constant, the time trend t and the variable $nontr_t$ that equals the number of non-trading days preceding the current trading day t:

$$\ln V_t^O = c_1 + c_2 t + c_3 nontr_t + error_t.$$

The detrended volume V_t is the exponent of the residuals from this regression. Summary statistics for the detrended trading volumes are presented in the same tables. The variability in degrees of skewness and kurtosis across the stocks in the Russian market is amazing. Some of it (e.g., large skewness and kurtosis for SNGS at the RTS) is driven by few instances when an unusually huge volume was transacted. There is also a significant difference in the kurtosis across the stocks traded at the NYSE. The values of the Ljung-Box statistics indicate very high autocorrelation in detrended trading volumes. Finally, there is significant positive contemporaneous correlation between return volatility and volume, as confirmed by correlation coefficients between the volume and squared return.The link between returns and trading volumes is evidently weaker in the Russian market, but is still strong for the bivariate mixture model to work.

5 Empirical results

5.1 One-stock MMH model

The estimation results for the one-stock MMH model (1) with the stocks traded at the MICEx and RTS are presented in Table **??**, with the stocks traded at the NYSE – in Table 4. The estimates of persistence of the information variable (β) are consistent with the evidence in Andersen (1996), Liesenfeld (1998) and Watanabe (2003), and are on average higher for the Russian market. This means that the news coming today has lower effect on tomorrow's decisions of traders at the NYSE than at the MICEx or RTS. This may be caused by the different nature of information coming to different markets: the information concerning an overall political or economic situation may have a longer effect on traders and investors, than the information concerning the company whose stock is traded. The variance of the information shock is quite variable across stocks, but the figures are comparable in size in the two markets. Interestingly, for two Russian stocks, EESR and LKOH, this variance is much smaller when these stocks are traded at the MICEx than when they traded at the RTS, and the other way round for the other two Russian stocks, RTKM and SNGS. There is an impression that stock-specific news have a tendency to appear in a particular exchange rather than in the whole market. The average share of the uninformed volume in the total trading volume (S_V^u) is shown in the last columns of the tables. These shares are quite high in both markets fluctuating near 50%, and, most importantly, the numbers are comparable and even similar across the markets.

5.2 Two-stock MMH model

The estimation results for the two-stock MMH model (2) with the stocks traded at the MICEx, RTS and NYSE are presented in Table **??**, 6, and 7, respectively (in the latter case the results only for 6 pairs are reported). We also estimate the model (the results are not shown to save space) for all pairs of stocks where one stock is drawn from the MICEx, and the other – from the RTS. The parameter estimates differ from those obtained for the one-stock MMH model, but the differences are consistent with the reported standard errors. The estimates of the parameter σ_{12} are nonnegative for all pairs of stocks, reported and unreported (except for the DCX–MCD pair where it is negative but insignificant and close to zero), in spite of the fact that we do not impose any restrictions on this parameter during estimation. This confirms the story behind the positive correlatedness of information flows given below (3).

As discussed previously, the key parameter in the two-stock model is the correlation coefficient ρ_{12} between the shocks of information variables computed as (3). Tables **??** and 9 report estimates of this parameter. Let us first consider correlations in the Russian market. The northwest quadrant in Table **??** shows those for stocks traded at the MICEx, the southeast quadrant – for stocks traded at the RTS, and the northeast quadrant shown cross-correlations between shocks to information variables for stocks traded at the two exchanges. All estimated correlations are highly significant. Correlations for different stocks traded at the same exchange vary from about 0.3 to about 0.8, while those for different stocks at the different exchanges vary from about 0.2 to about 0.7, i.e. are somewhat smaller but not appreciably. There is a tendency to companies from the same industry to have higher correlated information variables: the highest correlation at the MICEx, 0.680, belongs to the two oil companies LKOH and SNGS, and so does the highest correlation at the RTS, 0.782. The between-exchanges correlations for these two companies, 0.631 and 0.623, are also high, although somewhat smaller. In contrast, the lowest correlation at the MICEx, 0.306, belongs to the pair RTKM (telecommunications industry) and SNGS (oil extraction), and so does the lowest correlation at the RTS, 0.316. The between-exchanges correlations for these two companies, 0.217 and 0.252, are also the lowest. In the northeast quadrant there is some weak evidence of a symmetry relative to the diagonal, although, for example, the high correlation for the SNGS stock from the MICEx and the EESR stock from the RTS does not repeat itself for the EESR stock from the MICEx and the SNGS stock from the RTS (that high correlation, 0.721, seems to be an exception from many other tendencies). If one compares the same-exchange correlations between stocks of two companies to the cross-exchange correlations, one can see that the cross-exchange correlations tend to be lower than the maximal same-exchange correlation for these two stocks, and most often lower than the minimal of them. Interestingly, the cross-exchange correlations for stocks of the same company are very close to unity. For the EESR, the most heavily traded stock, the point estimate even exceeds unity (recall that we do not restrict $|\rho_{12}|$ to be lower than unity during estimation); it is also very high for RTKM and SNGS, and a bit lower for LKOH. This points at an almost free information mobility between the two exchanges. The fact that the same-exchange correlations are generally larger than the cross-exchange correlations if the stocks are not of the same company but of the same industry indicates that there is some specialization of traders to work with securities at a particular exchange.

The fact that the lowest estimated same-exchange correlation equals 0.306 at the MICEx and 0.316 at the RTS, while the lowest cross-exchange correlation equals 0.217, indicates that some of the correlation is due to overall political and economic risk factors and some is due to the commonality of the trading platform, i.e. due to exchange specialization. Further, the commonality of the industry drives the correlations up appreciably from the average same-exchange or cross-exchange correlations. This sharply contrasts with the evidence from the NYSE presented in Table 7. The lowest correlations for the American market are so close to zero that it is reasonable to assume that the political and economic risks have practically zero effect; the diversity of assets and liquidity are so high that common information can arise only from industry-wide news. Remember that the NYSE-traded stocks are chosen from six industries with two stocks in each. The empirical evidence confirms the hypothesis that the correlation of shocks of information variables for stocks of same-industry companies is higher than of those from different industries, although not perfectly. The correlation is indeed high for the pairs BP and CVX (0.64), F and DCX (0.59), VZ and SBC (0.79), IBM and HPQ (0.47), but it is lower for MRK and GSK (0.29), and much lower for MCD and YUM (0.16). There is also quite high correlation for the stocks of companies from different industries, for example for SBC and IBM (0.44), MCD and BP (0.43), YUM and DCX (0.39). In some cases it is easy to understand what kind of information may be common for the industry in order to have effect on the dynamics of both stocks from that industry. For example, for the Oil & Gas Integrated industry it may be the world oil prices, for the Communications Services and Major Drugs industries it may be advents of new technologies crucial for the development of these industries, but it is hard to imagine what kind of information may be common for the Restaurants industry.

6 Conclusion

The proposed natural extension of the Modified Mixture of Distribution model of Andersen (1996) does provide interesting evidence about interconnection of information flows associated with different assets. Of course, inferring which fractions of common information are due to different factors from a large number of pairwise comparisons is far from perfect. Hence, the model can be potentially extended to a larger number of assets, and possibly introduce more complex lead–lag relationships for information flows, provided that the span of data is long enough. A specification similar to those used in the panel data analysis is a possibility.

References

Andersen, T.G. (1996) Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility, *Journal of Finance* 51, 169–204.

Andersen, T.G. and B.E. Sørensen (1995) GMM estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Business & Economic Statistics* 14, 328–352.

Andrews, W.K. (1999) Consistent moment selection procedures for generalized method of moments estimation. *Econometrica* 67, 543–564.

Donald, S. and W.K. Newey (2000) A jacknife interpretation of the continuous updating estimator. *Economics Letters* 67, 239–243.

Gallant, A.R., P.E. Rossi, and G.E. Tauchen (1992) Stock prices and volume. *Review of Financial Studies* 5, 199–242.

Glosten, L.R. and P.R. Milgrom (1985) Bid, ask, and transaction prices in a specialist market with heterogeneously informed traders. *Journal of Financial Economics* 14, 71–100.

Hall, A.R. (2000) Covariance Matrix Estimation and the Power of the Overidentifying Restrictions Test. *Econometrica* 68, 1517–1527.

Hanousek, J. and R. Podpiera (2003) Informed trading and the bid-ask spread: evidence from an emerging market. *Journal of Comparative Economics* 31, 275–296.

Hansen, L.P. (1982) Large sample properties of generalized method of moments estimators. *Econometrica* 50, 1029–1054.

Hansen, L.P., Heaton, J., and A. Yaron (1996) Finite-Sample Properties of Some Alternative GMM Estimators. *Journal of Business & Economic Statistics* 19, 262–280.

Lamoureux, C.G. and W.D. Lastrapes (1990) Heteroskedasticity in stock return data: Volume versus GARCH effects. *Journal of Finance* 45, 221–229.

Liesenfeld, R. (1998) Dynamic bivariate mixture models: modeling the behavior of prices and trading volume. *Journal of Business & Economic Statistics* 16, 101–109.

Liesenfeld, R. (2001) A generalized bivariate mixture model for stock price volatility and trading volume. *Journal of Econometrics* 104, 141–178.

Liesenfeld, R. and J.-F. Richard (2002) The estimation of dynamic bivariate mixture models: Comments on Watanabe (2000). *Journal of Business & Economic Statistics* 21, 570–576.

Luu, J.C. and M. Martens (2003) Testing the mixture of distributions hypothesis using "realized" volatility. *Journal of Futures Markets* 23, 661–679.

Newey, W.K. and R.J. Smith (2004) Higher order properties of GMM and generalized empirical likelihood estimators. *Econometrica* 72, 219–255.

Newey, W.K. and K.D. West (1987) A Simple, Positive Semi-definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix. *Econometrica* 55, 703–708.

Ostrovsky, A. (2003) From chaos to capitalist triumph. *Financial Times (UK)*, Oct 9, pg. 4.

Richardson, M. and T. Smith (1994) A direct test of the mixture of distributions hypothesis: Measuring the daily flow of information. *Journal of Financial and Quantitative Analysis* 29, 101–116.

Spierdijk, L., Nijman, T.E., and A.H.O. van Soest (2002) Modeling Comovements in Trading Intensities to Dinstinguish Sector and Stock Specific News. Manuscript, Tilburg University.

Tauchen, G., Zhang, H., and M. Liu (1996) Volume, volatility, and leverage: A dynamic analysis. *Journal of Econometrics* 74, 177–208.

Tauchen G. and M. Pitts (1983) The price variability-volume relationship on speculative markets. *Econometrica* 51, 485–505.

Watanabe, T. (2000) Bayesian analysis of dynamic bivariate mixture models: Can they explain the behavior of returns and trading volume? *Journal of Business & Economic Statistics* 18, 199–210.

Watanabe, T. (2003) The estimation of dynamic bivariate mixture models: Reply to Liesenfeld and Richard comments. *Journal of Business & Economic Statistics* 21, 577–580.

A Derivation of moment conditions

We express the moments as functions of models parameters and $E\left[I_t^a\right]$ and $E\left[I_{i,t}^a I_{j,t-k}^b\right]$ that are in turn can be expressed as functions of model parameters as shown in Appendix B. For one stock, the static moments are

$$\begin{aligned}
E\left[r_t\right] &= \bar{r}, \\
E\left[\left|r_t-\bar{r}\right|\right] &= \sqrt{\frac{2}{\pi}} E\left[I_t^{1/2}\right], \\
E\left[\left(r_t-\bar{r}\right)^2\right] &= E\left[I_t\right], \\
E\left[V_t\right] &= cm_0+cm_1 E\left[I_t\right] \equiv \bar{V}, \\
E\left[\left(V_t-\bar{V}\right)^2\right] &= c\bar{V}+(cm_1)^2(E\left[I_t^2\right]-(E\left[I_t\right])^2),
\end{aligned}$$

and the dynamic moments for $k \geq 1$ are

$$\begin{aligned}
E\left[\left|r_t-\bar{r}\right|\left|r_{t-k}-\bar{r}\right|\right] &= \frac{2}{\pi} E\left[I_t^{1/2} I_{t-k}^{1/2}\right], \\
E\left[\left(V_t-\bar{V}\right)\left(V_{t-k}-\bar{V}\right)\right] &= (cm_1)^2\left(E\left[I_t I_{t-k}\right]-(E\left[I_t\right])^2\right).
\end{aligned}$$

For two stocks i and j, $i \neq j$, and $k \geq 1$,

$$E\left[\left|r_{i,t}-\bar{r}_i\right| r_{j,t}\right] = \sqrt{\frac{2}{\pi}} \bar{r}_j E\left[I_{i,t}^{1/2}\right],$$

$$
\begin{aligned}
E\left[V_{i,t}V_{j,t}\right] &= c_i m_{i,0}\bar{V}_j + c_i m_{i,0} c_j m_{j,0} E\left[I_{i,t}\right] + c_i m_{i,1} c_j m_{j,1} E\left[I_{i,t} I_{j,t}\right], \\
E\left[V_{i,t} r_{j,t}\right] &= \left(c_i m_{i,0} + c_i m_{i,1} E\left[I_{i,t}\right]\right)\bar{r}_j \equiv \bar{V}_j \bar{r}_j, \\
E\left[V_{i,t}\left|r_{j,t}-\bar{r}_j\right|\right] &= \sqrt{\frac{2}{\pi}}\left(c_i m_{i,0} E\left[I_{j,t}^{1/2}\right] + c_i m_{i,1} E\left[I_{i,t} I_{j,t}^{1/2}\right]\right).
\end{aligned}
$$

B Derivation of moments of information variable

Denote $\lambda_{i,t} = \ln I_{i,t}$. We need to express $E\left[I_{i,t}^a I_{j,t-k}^b\right] = E\left[\exp\left(a\lambda_{i,t} + b\lambda_{j,t-k}\right)\right]$ for $k \geq 0$ via the parameters of processes the information variables follow. From the dynamics of the information variables we have:

$$
a\lambda_{i,t} + b\lambda_{j,t-k} = a\frac{\alpha_i}{1-\beta_i} + a\sum_{l=0}^{\infty}\beta_i^l u_{i,t-l} + b\frac{\alpha_j}{1-\beta_j} + b\sum_{l=0}^{\infty}\beta_j^l u_{j,t-k-l},
$$

from which it follows that

$$
a\lambda_{i,t} + b\lambda_{j,t-k} \sim N\left(a\frac{\alpha_i}{1-\beta_i} + b\frac{\alpha_j}{1-\beta_j}, a^2\frac{\sigma_i^2}{1-\beta_i^2} + b^2\frac{\sigma_j^2}{1-\beta_j^2} + 2ab\beta_i^k\frac{\sigma_{12}}{1-\beta_i\beta_j}\right).
$$

Hence,

$$
E\left[I_{i,t}^a I_{j,t-k}^b\right] = \exp\left(a\frac{\alpha_i}{1-\beta_i} + b\frac{\alpha_j}{1-\beta_j} + \frac{a^2}{2}\frac{\sigma_i^2}{1-\beta_i^2} + \frac{b^2}{2}\frac{\sigma_j^2}{1-\beta_j^2} + ab\beta_i^k\frac{\sigma_{12}}{1-\beta_i\beta_j}\right).
$$

In particular,

$$
\begin{aligned}
E\left[I_{i,t}^a I_{i,t-k}^b\right] &= \exp\left(\left(a+b\right)\frac{\alpha_i}{1-\beta_i} + \left(\frac{a^2}{2} + \frac{b^2}{2} + ab\beta_i^k\right)\frac{\sigma_i^2}{1-\beta_i^2}\right), \\
&E\left[I_{i,t}^a\right] = \exp\left(a\frac{\alpha_i}{1-\beta_i} + \frac{a^2}{2}\frac{\sigma_i^2}{1-\beta_i^2}\right).
\end{aligned}
$$

Table 1: Summary statistics for returns and detrended volumes of stocks traded at the MICEx and RTS

	EESR MICEx	LKOH MICEx	RTKM MICEx	SNGS MICEx	EESR RTS	LKOH RTS	RTKM RTS	SNGS RTS
Returns								
mean, $\times 10^{-3}$	1.40	1.50	0.91	1.62	1.31	1.20	0.69	1.37
stdev, $\times 10^{-2}$	2.97	3.86	3.88	3.61	3.80	3.03	3.71	3.56
skew	-0.025	0.342	0.776	-0.028	0.260	-0.092	0.522	0.026
kurt	5.71	6.03	11.9	6.63	7.08	7.46	9.94	7.12
$Q_{30}(r)$	51.08	45.62	46.31	34.47	51.78	63.35	63.85	48.30
$Q_{30}(r^2)$	387.5	280.4	106.3	240.5	291.0	416.2	150.0	261.1
$Q_{30}(\|r\|)$	466.1	487.3	606.4	316.6	648.3	726.9	749.7	407.6
Detrended volumes								
mean	1.37	1.15	1.40	1.23	1.26	1.32	1.66	1.50
stdev	2.06	0.61	1.21	0.97	0.87	1.00	1.81	2.51
skew	10.98	1.10	2.42	5.10	1.84	1.88	2.67	17.21
kurt	161.5	4.37	13.28	58.43	8.90	8.58	12.94	438.8
$Q_{30}(V)$	253.7	10021.	3974.	1956.	2219.	513.	1304.	1313.
Correlations								
ρ_{V,r^2}	0.118	0.234	0.134	0.313	0.185	0.188	0.184	0.320
$\rho_{V,\|r\|}$	0.151	0.313	0.276	0.378	0.277	0.273	0.283	0.312

Note: Ljung-Box statistic $Q_{30}(.)$ is distributed as χ^2_{30}, with 5% critical value being 43.77.

Table 2: Summary statistics for returns and detrended volumes of stocks traded at the NYSE

	BP	CVX	F	DCX	IBM	HPQ	VZ	SBC	MRK	GSK	MCD	YUM
Returns												
mean, $\times 10^{-3}$	0.253	0.239	-0.336	-0.49	-0.003	0.072	-0.102	-0.414	-0.207	-0.271	-0.218	0.359
stdev, $\times 10^{-2}$	1.78	1.60	2.65	2.27	2.44	3.56	2.26	2.38	1.97	1.95	2.09	2.44
skew	-0.086	0.059	0.198	-0.379	-0.087	1.645	0.095	-0.017	-0.011	-0.016	-0.147	-0.397
kurt	4.77	4.84	6.49	6.32	8.21	27.06	5.65	4.96	5.15	4.82	6.62	11.81
$Q_{30}(r)$	38.11	37.18	80.26	35.42	44.29	39.94	55.30	30.30	45.75	49.58	26.41	33.04
$Q_{30}(r^2)$	321.6	391.5	208.7	177.7	98.7	5.4	188.9	108.4	166.1	175.0	78.8	53.0
$Q_{30}(\|r\|)$	363.1	455.1	253.2	435.1	452.2	148.9	363.1	209.1	296.4	283.2	131.0	396.8
Detrended volumes												
mean	1.11	1.06	1.12	1.15	1.09	1.11	1.08	1.07	1.08	1.12	1.10	1.16
stdev	0.568	0.407	0.642	0.660	0.537	0.609	0.507	0.448	0.477	0.591	0.565	0.811
skew	2.27	2.25	3.243	2.38	4.75	3.61	4.50	1.96	2.73	2.14	2.83	4.68
kurt	12.08	14.46	21.92	14.51	59.90	25.61	45.56	9.42	16.83	11.36	16.47	43.06
$Q_{30}(V)$	3006.	2250.	1745.	965.	1420.	1090.	1403.	2352.	1121.	1024.	762.	1546.
Correlations												
ρ_{V,r^2}	0.335	0.382	0.478	0.373	0.631	0.239	0.554	0.464	0.499	0.441	0.563	0.533
$\rho_{V,\|r\|}$	0.358	0.385	0.506	0.450	0.601	0.462	0.516	0.433	0.486	0.411	0.549	0.515

Note: Ljung-Box statistic $Q_{30}(.)$ is distributed as χ^2_{30}, with 5% critical value being 43.77.

Table 3: Estimation results for the one-stock model using the MICEx and RTS data

	$\bar{r}$	cm_0	cm_1	c	α	β	σ_u^2	J-test	S_V^u
MICEx									
EESR	0.00145 (0.00115)	0.589 (0.069)	385.5 (58.5)	0.045 (0.005)	-0.427 (0.082)	0.938 (0.012)	0.084 (0.017)	4.088 (0.394)	0.513 (0.056)
LKOH	0.00152 (0.00084)	0.576 (0.121)	938.4 (182.2)	2.744 (0.872)	-0.749 (0.295)	0.899 (0.040)	0.123 (0.051)	0.958 (0.916)	0.418 (0.079)
RTKM	0.00080 (0.00118)	0.544 (0.097)	597.3 (84.3)	0.239 (0.049)	-0.905 (0.205)	0.871 (0.030)	0.214 (0.053)	3.565 (0.468)	0.392 (0.067)
SNGS	0.00169 (0.00097)	0.586 (0.119)	499.9 (131.9)	0.231 (0.057)	-0.980 (0.243)	0.862 (0.034)	0.214 (0.053)	3.717 (0.446)	0.489 (0.105)
RTS									
EESR	0.00125 (0.00115)	0.581 (0.130)	503.9 (140.4)	0.259 (0.037)	-0.969 (0.270)	0.861 (0.039)	0.170 (0.059)	5.704 (0.127)	0.465 (0.101)
LKOH	0.00128 (0.00082)	0.734 (0.109)	686.3 (190.6)	0.458 (0.050)	-1.287 (0.372)	0.828 (0.050)	0.242 (0.081)	7.388 (0.117)	0.561 (0.076)
RTKM	0.00021 (0.00118)	0.679 (0.155)	865.7 (165.7)	1.364 (0.164)	-0.483 (0.487)	0.932 (0.068)	0.092 (0.100)	4.592 (0.204)	0.400 (0.096)
SNGS	0.00117 (0.00094)	0.698 (0.357)	543.1 (318.7)	0.558 (0.359)	-0.176 (0.512)	0.975 (0.072)	0.035 (0.104)	1.863 (0.761)	0.518 (0.319)

Note: Standard errors for parameters and p-values for J-tests are in parentheses.

Table 4: Estimation results for the one-stock model using the NYSE data

	$\bar{r}$	cm_0	cm_1	c	α	β	σ_u^2	J-test	S_V^u
BP	0.00044 (0.00042)	0.511 (0.076)	2027.5 (332.4)	0.095 (0.014)	-0.867 (0.257)	0.896 (0.031)	0.089 (0.032)	4.615 (0.329)	0.462 (0.073)
CVX	-0.00002 (0.00038)	0.596 (0.063)	1936.8 (327.1)	0.054 (0.009)	-1.570 (0.262)	0.817 (0.031)	0.142 (0.039)	4.732 (0.316)	0.565 (0.056)
F	-0.00030 (0.00066)	0.418 (0.098)	1050.2 (197.6)	0.021 (0.022)	-1.860 (0.292)	0.756 (0.038)	0.255 (0.053)	4.343 (0.362)	0.373 (0.085)
DCX	-0.00051 (0.00057)	0.444 (0.113)	1447.2 (305.9)	0.175 (0.028)	-1.446 (0.473)	0.816 (0.060)	0.136 (0.067)	3.948 (0.413)	0.390 (0.096)
IBM	-0.00020 (0.00059)	0.670 (0.050)	739.6 (118.2)	0.056 (0.024)	-1.578 (0.291)	0.800 (0.037)	0.262 (0.062)	3.335 (0.503)	0.624 (0.042)
HPQ	-0.00032 (0.00082)	0.577 (0.094)	445.5 (111.1)	0.047 (0.031)	-2.504 (0.462)	0.650 (0.065)	0.426 (0.149)	6.876 (0.143)	0.535 (0.088)
VZ	-0.00011 (0.00051)	0.631 (0.057)	849.6 (130.6)	0.058 (0.017)	-1.689 (0.442)	0.785 (0.056)	0.204 (0.069)	2.672 (0.614)	0.599 (0.053)
SBC	-0.00071 (0.00056)	0.574 (0.052)	877.4 (100.2)	0.043 (0.007)	-1.515 (0.282)	0.804 (0.037)	0.177 (0.040)	4.161 (0.385)	0.537 (0.050)
MRK	-0.00058 (0.00048)	0.521 (0.074)	1476.1 (237.3)	0.046 (0.014)	-2.426 (0.376)	0.701 (0.046)	0.240 (0.058)	4.162 (0.385)	0.483 (0.063)
GSK	-0.00021 (0.00045)	0.539 (0.069)	1568.9 (227.7)	0.111 (0.015)	-1.855 (0.380)	0.773 (0.047)	0.195 (0.054)	5.395 (0.249)	0.485 (0.058)
MCD	0.00006 (0.00055)	0.555 (0.070)	1308.0 (217.0)	0.064 (0.020)	-2.604 (0.496)	0.679 (0.062)	0.293 (0.072)	6.873 (0.143)	0.518 (0.060)
YUM	0.00056 (0.00061)	0.654 (0.072)	836.0 (155.6)	0.117 (0.064)	-2.829 (0.746)	0.645 (0.094)	0.524 (0.173)	5.610 (0.230)	0.591 (0.054)

Note: Standard errors for parameters and p-values for J-tests are in parentheses.

Table 5: Estimation results for the two-stock model using the MICEx data

	$\bar{r}$	cm_0	cm_1	c	α	β	σ_u^2	σ_{12}	J-test
LKOH	0.00167 (0.00076)	0.492 (0.114)	994.2 (178.3)	1.752 (0.670)	-0.788 (0.266)	0.894 (0.036)	0.133 (0.045)	0.038	14.68
EESR	0.00160 (0.00097)	0.516 (0.065)	462.4 (61.8)	0.044 (0.005)	-0.451 (0.083)	0.934 (0.012)	0.076 (0.015)	(0.012)	(0.401)
LKOH	0.00141 (0.00073)	0.497 (0.125)	1039.1 (197.1)	2.245 (0.690)	-0.693 (0.329)	0.907 (0.044)	0.103 (0.051)	0.079	20.61
RTKM	0.00047 (0.00105)	0.535 (0.083)	566.0 (75.8)	0.223 (0.051)	-0.847 (0.209)	0.881 (0.030)	0.215 (0.060)	(0.023)	(0.112)
LKOH	0.00095 (0.00075)	0.485 (0.130)	1142.1 (221.9)	2.291 (0.739)	-0.942 (0.288)	0.874 (0.039)	0.146 (0.048)	0.109	18.83
SNGS	0.00138 (0.00086)	0.605 (0.094)	536.3 (110.9)	0.183 (0.049)	-0.956 (0.331)	0.867 (0.046)	0.177 (0.062)	(0.027)	(0.172)
EESR	0.00127 (0.00093)	0.546 (0.066)	424.1 (53.0)	0.045 (0.005)	-0.432 (0.078)	0.937 (0.011)	0.076 (0.015)	0.070	18.63
RTKM	0.00073 (0.00110)	0.573 (0.083)	518.9 (63.0)	0.234 (0.047)	-0.752 (0.223)	0.893 (0.032)	0.184 (0.058)	(0.016)	(0.180)
EESR	0.00075 (0.00092)	0.612 (0.054)	379.9 (49.7)	0.042 (0.006)	-0.423 (0.093)	0.939 (0.013)	0.082 (0.018)	0.059	21.21
SNGS	0.00187 (0.00088)	0.609 (0.079)	468.1 (81.6)	0.193 (0.053)	-0.914 (0.288)	0.873 (0.040)	0.195 (0.061)	(0.019)	(0.096)
RTKM	0.00009 (0.00098)	0.602 (0.084)	536.0 (75.8)	0.242 (0.052)	-1.037 (0.243)	0.854 (0.034)	0.267 (0.068)	0.075	18.73
SNGS	0.00173 (0.00085)	0.537 (0.082)	551.7 (93.9)	0.219 (0.051)	-1.084 (0.257)	0.849 (0.036)	0.227 (0.051)	(0.023)	(0.175)

Note: Standard errors for parameters and p-values for J-tests are in parentheses.

Table 6: Estimation results for the two-stock model using the RTS data

	$\bar{r}$	cm_0	cm_1	c	α	β	σ_u^2	σ_{12}	J-test
LKOH	0.00085	0.559	940.8	0.383	-1.842	0.754	0.264	0.142	20.528
	(0.00072)	(0.125)	(245.8)	(0.047)	(0.381)	(0.051)	(0.083)		
EESR	0.00050	0.527	552.6	0.208	-1.330	0.809	0.184	(0.037)	(0.114)
	(0.00091)	(0.120)	(131.4)	(0.035)	(0.376)	(0.054)	(0.064)		
LKOH	0.00107	0.367	1379.7	0.414	-1.769	0.764	0.169	0.126	23.664
	(0.00072)	(0.223)	(431.0)	(0.052)	(0.407)	(0.054)	(0.078)		
RTKM	-0.00023	0.415	1055.3	0.856	-1.565	0.782	0.300	(0.038)	(0.050)
	(0.00102)	(0.172)	(211.7)	(0.139)	(0.330)	(0.046)	(0.086)		
LKOH	0.00104	0.596	921.4	0.414	-1.776	0.763	0.256	0.132	16.365
	(0.00071)	(0.124)	(239.0)	(0.049)	(0.416)	(0.056)	(0.083)		
SNGS	0.00120	0.608	680.3	0.349	-0.783	0.890	0.111	(0.050)	(0.292)
	(0.00079)	(0.162)	(181.2)	(0.305)	(0.673)	(0.094)	(0.104)		
EESR	0.00111	0.432	724.7	0.228	-1.274	0.819	0.132	0.077	21.235
	(0.00093)	(0.170)	(190.9)	(0.035)	(0.419)	(0.059)	(0.056)		
RTKM	0.00026	0.715	757.2	1.263	-0.536	0.925	0.106	(0.032)	(0.068)
	(0.00101)	(0.121)	(131.1)	(0.157)	(0.486)	(0.068)	(0.106)		
EESR	0.00092	0.584	476.3	0.218	-1.060	0.847	0.164	0.104	13.867
	(0.00095)	(0.097)	(105.1)	(0.032)	(0.379)	(0.054)	(0.061)		
SNGS	0.00150	0.579	642.5	0.566	-0.895	0.874	0.152	(0.042)	(0.460)
	(0.00082)	(0.160)	(176.9)	(0.341)	(0.540)	(0.076)	(0.099)		
RTKM	-0.00013	0.486	911.0	0.834	-1.493	0.790	0.324	0.059	20.478
	(0.00104)	(0.128)	(142.3)	(0.146)	(0.316)	(0.044)	(0.085)		
SNGS	0.00118	0.178	1092.7	0.645	-0.725	0.898	0.108	(0.022)	(0.116)
	(0.00086)	(0.262)	(293.3)	(0.337)	(0.594)	(0.083)	(0.089)		

Note: Standard errors for parameters and p-values for J-tests are in parentheses.

Table 7: Estimation results of the two-stock model using the NYSE data

	$\bar{r}$	cm_0	cm_1	c	α	β	σ_u^2	σ_{12}	J-test
BP	0.00015 (0.00032)	0.551 (0.080)	1861.5 (385.9)	0.066 (0.015)	-1.959 (0.689)	0.769 (0.081)	0.160 (0.066)	0.083	22.07
CVX	0.00023 (0.00029)	0.541 (0.089)	2456.1 (547.4)	0.052 (0.008)	-1.631 (0.330)	0.812 (0.038)	0.105 (0.041)	(0.033)	(0.08)
BP	0.00056 (0.00037)	0.546 (0.069)	1873.7 (313.6)	0.085 (0.013)	-0.815 (0.295)	0.903 (0.035)	0.0851 (0.036)	0.064	16.88
MCD	0.00041 (0.00045)	0.473 (0.089)	1600.5 (303.1)	0.054 (0.022)	-2.998 (0.558)	0.632 (0.069)	0.260 (0.079)	(0.017)	(0.26)
BP	0.00040 (0.00040)	0.502 (0.084)	2120.9 (392.9)	0.082 (0.013)	-0.911 (0.325)	0.892 (0.039)	0.082 (0.035)	0.036	17.53
YUM	0.00056 (0.00060)	0.493 (0.104)	1331.7 (275.0)	0.099 (0.069)	-3.679 (0.821)	0.541 (0.103)	0.453 (0.122)	(0.022)	(0.23)
CVX	0.00022 (0.00032)	0.598 (0.053)	1978.5 (284.2)	0.053 (0.009)	-1.608 (0.264)	0.813 (0.031)	0.150 (0.037)	0.070	14.51
MCD	0.00020 (0.00049)	0.484 (0.081)	1549.4 (268.0)	0.056 (0.023)	-3.018 (0.541)	0.629 (0.067)	0.283 (0.078)	(0.015)	(0.41)
CVX	-0.00002 (0.00035)	0.583 (0.061)	2072.9 (322.3)	0.053 (0.009)	-1.542 (0.245)	0.820 (0.029)	0.131 (0.035)	0.057	12.71
YUM	0.00033 (0.00059)	0.590 (0.080)	980.9 (187.7)	0.080 (0.064)	-3.430 (0.730)	0.571 (0.092)	0.574 (0.120)	(0.026)	(0.55)
MCD	0.00009 (0.00045)	0.557 (0.066)	1332.4 (212.3)	0.066 (0.019)	-2.490 (0.476)	0.694 (0.059)	0.271 (0.069)	0.057	17.82
YUM	0.00060 (0.00060)	0.636 (0.067)	840.3 (156.9)	0.128 (0.053)	-2.575 (0.728)	0.678 (0.091)	0.447 (0.143)	(0.031)	(0.22)

Note: Standard errors for parameters and p-values for J-tests are in parentheses.

Table 8: Correlations of shocks of information variables at the MICEx and RTS

	EESR MICEx	LKOH MICEx	RTKM MICEx	SNGS MICEx	EESR RTS	LKOH RTS	RTKM RTS	SNGS RTS
EESR MICEx	1	0.373 (0.077)	0.594 (0.051)	0.462 (0.135)	1.048 (0.051)	0.552 (0.061)	0.541 (0.097)	0.429 (0.109)
LKOH MICEx		1	0.528 (0.081)	0.680 (0.106)	0.420 (0.073)	0.845 (0.080)	0.541 (0.070)	0.623 (1.066)
RTKM MICEx			1	0.306 (0.087)	0.533 (0.057)	0.471 (0.076)	0.881 (0.082)	0.252 (0.086)
SNGS MICEx				1	0.721 (0.140)	0.631 (0.077)	0.217 (0.091)	0.914 (0.128)
EESR RTS					1	0.643 (0.079)	0.652 (0.203)	0.655 (0.144)
LKOH RTS						1	0.562 (0.068)	0.782 (0.197)
RTKM RTS							1	0.316 (0.115)
SNGS RTS								1

Note: Standard errors are in parentheses.

Table 9: Correlations of shocks of information variables at the NYSE

	BP	CVX	F	DCX	IBM	HPQ	VZ	SBC	MRK	GSK	MCD	YUM
BP	1	0.64 (0.09)	0.04 (0.07)	0.15 (0.07)	0.25 (0.08)	0.25 (0.07)	0.21 (0.09)	0.46 (0.08)	0.14 (0.08)	0.24 (0.08)	0.43 (0.09)	0.19 (0.10)
CVX		1	0.20 (0.06)	0.16 (0.10)	0.29 (0.08)	0.20 (0.09)	0.39 (0.10)	0.41 (0.07)	0.23 (0.08)	0.19 (0.10)	0.34 (0.06)	0.21 (0.08)
F			1	0.59 (0.10)	0.13 (0.07)	0.10 (0.06)	0.23 (0.04)	0.19 (0.05)	0.22 (0.07)	0.18 (0.08)	0.22 (0.06)	0.08 (0.07)
DCX				1	0.07 (0.07)	0.19 (0.09)	0.18 (0.11)	0.16 (0.08)	0.24 (0.08)	0.31 (0.08)	-0.02 (0.07)	0.39 (0.13)
IBM					1	0.47 (0.07)	0.40 (0.10)	0.44 (0.06)	0.28 (0.07)	0.28 (0.08)	0.21 (0.07)	0.16 (0.08)
HPQ						1	0.27 (0.09)	0.26 (0.06)	0.15 (0.05)	0.21 (0.07)	0.20 (0.06)	0.12 (0.06)
VZ							1	0.79 (0.09)	0.29 (0.07)	0.30 (0.09)	0.22 (0.09)	0.27 (0.10)
SBC								1	0.45 (0.06)	0.29 (0.06)	0.28 (0.06)	0.29 (0.09)
MRK									1	0.29 (0.06)	0.17 (0.06)	0.13 (0.06)
GSK										1	0.21 (0.09)	0.17 (0.08)
MCD											1	0.16 (0.08)
YUM												1

Note: Standard errors are in parentheses.