High Frequency Multiplicative Component GARCH ♣*

Ananda Chanda*, Robert F. Engle** and Magdalena E. Sokalska***

Preliminary Draft: May 23, 2005

Abstract

This paper proposes a new way of modeling and forecasting intraday returns. We decompose the volatility of high frequency asset returns into components that may be easily interpreted and estimated. The conditional variance is expressed as a product of daily, diurnal and stochastic intraday volatility components. This model is applied to a comprehensive sample consisting of 10-minute returns on more than 2500 US equities.

JEL Classifications: C22, G15
Keywords: Volatility, ARCH, Intra-day Returns.

♣* The authors gratefully acknowledge that financial support for the research was provided by Morgan Stanley to New York University's Stern School of Business. The views expressed herein are solely those of the authors and not those of any other person or entity.
* Morgan Stanley, Anada.Chanda@morganstanley.com.
** Department of Finance, Stern School of Business, New York University, rengle@stern.nyu.edu
*** Department of Finance, Stern School of Business, New York University, msokalsk@stern.nyu.edu, and Warsaw School of Economics msokal@sgh.waw.pl.

1. INTRODUCTION

This paper proposes a new way of modeling and forecasting intraday returns. We decompose the volatility of high frequency asset returns into multiplicative components, which may be easily interpreted and estimated. The conditional variance is expressed as a product of daily, diurnal and stochastic intraday volatility components. This model is applied to a comprehensive sample consisting of 10-minute returns on more than 2500 US equities. We apply a number of different specifications. Namely we build models for separate companies, pool data into industries and consider various criteria for grouping returns. It turns out that results for the pooled regressions seem to be more stable. The forecasts from the pooled specifications outperform the corresponding forecasts from company-by-company estimation but there are several issues in the best way to pool.

A number of papers have presented related work on intraday returns. The most commonly cited studies include Andersen and Bollerslev (1997, 1998). The authors propose models for 5-minute returns on Deutschemark-dollar exchange rate and the S&P500 index. In the first paper, Andersen and Bollerslev build a multiplicative model of daily and diurnal volatility, and in the second they add an additional component which takes account of macroeconomic announcements. For most of their models, the intra-daily volatility components are deterministic. In contrast, the intra-daily components in our model include both deterministic (the diurnal) and stochastic (a separate intra-daily ARCH).

Conventional GARCH approaches were argued to be unsatisfactory by authors at the Olsen conference on High Frequency Data Analysis but in response see Ghose and Kroner (1997). A long memory stochastic volatility approach was applied by Deo Hurvich and Lu (2005). This paper diurnally adjusts in the frequency domain and then uses a local Whittle estimator on log of squared returns to estimate the parameters.

We expect our model to be of particular interest for derivative traders or hedge funds who seek high frequency measures of risk or time varying hedge ratios. Another potential use lies in devising the optimal trading strategy either in placement of limit orders or scheduling trades. The paper is organized as follows: Section 2 presents the model. Section 3 describes the data and gives results of estimation. Conclusions and an outline of further research are offered in Section 4.

2. THE MODEL

2.1. Notation

We use the following notation. Days in the sample are indexed by t (t =1,…, T). Each day is divided into 10 minute intervals referred to as bins and indexed by i (i =0,…, N). The current period is $\{t,i\}$. The price of an asset at the end of day t and bin i is denoted by $P_{\{t,i\}}$. The continuously compounded return $r_{\{t,i\}}$ is modeled as:

$$
\begin{aligned}
r_{\{t,i\}} &= \ln\left(\frac{P_{\{t,i\}}}{P_{\{t,i-1\}}}\right) \qquad \textit{for} \;\; i \geq 1 \\
&= \ln\left(\frac{P_{\{t,1\}}}{P_{\{t-1,N\}}}\right) \qquad \textit{for} \;\; i = 0
\end{aligned}
\tag{1}
$$

The overnight return in bin zero is deleted leading to a total number of return observations, $M=TN$

2.2. Model

We propose a new GARCH model for high frequency intraday financial returns, which specifies the conditional variance to be a multiplicative product of daily, diurnal and stochastic intraday volatility. Intraday equity returns are described by the following process:

$$
r_{\{t,i\}} = \sqrt{h_t s_i q_{\{t,i\}}}\,\varepsilon_{\{t,i\}} \quad \text{and} \quad \varepsilon_{\{t,i\}} \sim N(0,1) \tag{2}
$$

where:

h_t is the daily variance component,

s_i is the diurnal (calendar) variance pattern,

$q_{\{t,i\}}$ is the intraday variance component with mean one, and

$\varepsilon_{\{t,i\}}$ is an error term.

The daily variance component could be specified in a number of ways. Andersen and Bollerslev (1997, 1998), estimate this component from a daily GARCH model for a longer sample, going back a number of months or years. It could also be estimated based on daily realized variance as proposed by Engle(2001) and Engle and Gallo(2005). We adopt a different route, however, and utilize commercially available volatility forecasts produced daily for each company in our sample. This eliminates the need for longer series for the daily model than for the intra-daily model. With the turnover of corporate ownership, it is difficult to get consistent long series for a big universe of stocks.

The diurnal component is calculated as the standard deviation of returns in each bin after deflating by the daily volatility. To see this consider the variance of these returns:

$$\frac{r_{i,t}^2}{h_t} = s_i q_{i,t} \varepsilon_{i,t}^2$$

and

$$E\left(\frac{r_{i,t}^2}{h_t}\right) = s_i E\left(q_{i,t}\right) = s_i \tag{3}$$

Practically, we estimate the model in two stages. First we normalize returns by daily and diurnal volatility components, and then model the residual volatility as a unit GARCH(1,1) process:

$$y_{\{t,i\}} = r_{\{t,i\}} / \sqrt{\hat{h}_t} = \sqrt{q_{\{t,i\}}}\varepsilon_{\{t,i\}} \tag{4}$$

$$q_{\{t,i\}} = \omega + \alpha(r_{\{t,i-1\}} / \sqrt{\hat{h}_t \hat{s}_{i-1}})^2 + \beta\, q_{\{t,i-1\}} \tag{5}$$

The GARCH specification can be rewritten as:

$$\begin{aligned} z_{\{t,i\}} \Big| F_{\{t,i-1\}} &\sim N\left(0, q_{\{t,i\}}\right) \\ q_{\{t,i\}} &= \omega + \alpha z_{\{t,i-1\}}^2 + \beta\, q_{\{t,i-1\}} \\ z_{\{t,i\}} &= r_{\{t,i\}} / \sqrt{h_t \hat{s}_i} \end{aligned} \tag{6}$$

The unit GARCH might enforce the constraint $\omega = 1 - \alpha - \beta$ although in the empirical work this has not been done.

3. ECONOMETRIC ISSUES

In this section we will discuss statistical properties of the two-step estimator of the model outlined in the previous section. The estimation proceeds in two steps. First we specify and estimate the diurnal component. Following equation (3) we estimate the diurnal component for each bin as the variance of y in this bin. That is:

$$\hat{s}_i = \frac{1}{T}\sum_{t=1}^{T} y_{\{t,i\}}^2,\ \forall i = 1,\dots,N \tag{7}$$

The second step consists of standardizing $y_{\{t,i\}}$ by $\hat{s}_i$ and estimating parameters of the GARCH(p,q) model, which describes the dynamics of the intraday stochastic component as in (6). Such a multi-step estimation strategy is potentially misleading as errors in one stage can lead to errors in the next stage. Nevertheless it will be shown below that the estimator is consistent but that the standard errors should be adjusted.

In deriving the asymptotic properties of the estimators in this sequential procedure, we will follow Newey and McFadden (1994) (later denoted as **NM**) and cast the above steps into the GMM framework. We will consider the GMM estimator of the moment conditions stacked one on the other. We will use the following notation. Vector $\psi = \begin{pmatrix} \phi \\ \theta \end{pmatrix}$ contains both the k_1 parameters ϕ, estimated in the first step, and the k_2 parameters θ, estimated in the second step,. Let there be k_1 moment conditions $g_1(\phi)$ and k_2 moment conditions $g_2(\phi, \theta)$ comprising vector $g(\psi) = \begin{pmatrix} g_1(\phi) \\ g_2(\phi,\theta) \end{pmatrix}$. The corresponding sample sums are g_{1M} and g_{2M}, giving $g_M=(g_{1M}', g_{2M}')'$. We will consider the GMM estimator of the parameter vector

$$\hat{\psi} = \begin{pmatrix} \hat{\phi} \\ \hat{\theta} \end{pmatrix} = \arg\min g_M ' W g_M = \arg\min g_M ' g_M \tag{8}$$

Since it is a just identified system, W=I. To solve this system, ϕ must solve the first set of equations and θ must solve the second set conditional on the estimated value of ϕ. Thus it is a natural framework to analyze two step estimators of this type. Newey and McFadden (1994) (c.f. their Theorem 6.1, p. 2178), have shown that if $\hat{\phi}$ and $\hat{\theta}$ are consistent estimators of the true ϕ_0 and θ_0, respectively, and g_M satisfies a number of standard regularity conditions, the resulting GMM estimator is consistent and asymptotically normal:

$$\sqrt{M} \begin{pmatrix} \hat{\phi} - \phi_0 \\ \hat{\theta} - \theta_0 \end{pmatrix} \xrightarrow{d} N\left(0, G^{-1}\Omega G^{-1}{}'\right) \tag{9}$$

,

where $G = E\left(\frac{\partial g(\psi)}{\partial \psi'}\right)$ and $\Omega = E\left(g(\psi) g(\psi)'\right)$.

As in Hansen (1982), the above matrices can be consistently estimated by replacing expectations by sample averages and parameters by their estimates.

The NM approach is very convenient and may be applied when parameters at some steps are estimated by ML. In this case some of the GMM moment conditions are taken to be score functions. In the current two-step setting, the sample sums in the first and the second stages are:

$$g_{1,M}(\phi) = g_{1,M} = \begin{pmatrix} 1/T\sum_{\{t,1\}}^{T}(y_{\{t,1\}}{}^2 - s_1) \\ \vdots \\ 1/T\sum_{\{t,N\}}^{T}(y_{\{t,N\}}{}^2 - s_N) \end{pmatrix} \tag{10}$$

$$g_{2,M}\left(\hat{\phi},\theta\right) = g_{2,M} = 1/TN\sum_{t=1}^{T}\sum_{i=1}^{N}\nabla_\theta\left(\log(q_{\{t,i\}}) + \left(y_{\{t,i\}}{}^2 / \hat{s}_i q_{\{t,i\}}\right)\right) \tag{11}$$

$$G = \frac{1}{M}\sum\begin{pmatrix} \nabla_\phi g_1 & 0 \\ \nabla_\phi g_2 & \nabla_\theta g_2 \end{pmatrix} \quad \text{and} \quad \frac{1}{M}\sum\begin{pmatrix} g_{1,t}^2 & g_{1,t}g_{2,t} \\ g_{1,t}g_{2,t} & g_{2,t}^2 \end{pmatrix} \xrightarrow{p} \Omega \tag{12}$$

In order to apply NM's Theorem 6.1, we have to make sure that $\hat{\phi}$ and $\hat{\theta}$ are consistent estimators of the true parameter values at each stage. This is indeed the case for estimator (7). In random sampling from a stationary ergodic distribution, the sample mean is a consistent estimate of the expected value. Consistency of $\hat{\theta}$ follows from, for example, Hansen and Lee (1994) or Lumsdaine (1996). In sum, the consistency and asymptotic normality of the estimator (11) is a corollary to Theorem 6.1 (p. 2178) in Newey and McFadden (1994). The above results could, in principle, be generalized to a multi-step estimation.

4. EMPIRICAL RESULTS

4.1. DATA

Our sample consists of price data on 2721 companies obtained from the TAQ database. We analyze logarithmic returns standardized by a commercially available volatility forecast for each company and each day and the standard deviation of returns in each 10-minute bin. Data spans a three-month period in April-June 2000.

4.2. RESULTS FOR A SINGLE STOCK

Some results will be presented using a single randomly chosen NYSE-listed stock – Republic Group Inc. In the following, we will refer to the company by its ticker symbol RGC. We have its daily price data since August 1985 and tick-by-tick data during the period from January 2, 1998 to June 30, 1999. This return series has been previously used by Engle and Patton (2004) and comes from their 4^{th} decile for trading frequency. Over the examined period (1.5 years) it had almost 13 thousand trades or about 35 per day. On the basis of this ultra-high frequency data we calculate 10-minute returns based on closing midquotes, which are used in the subsequent analysis.

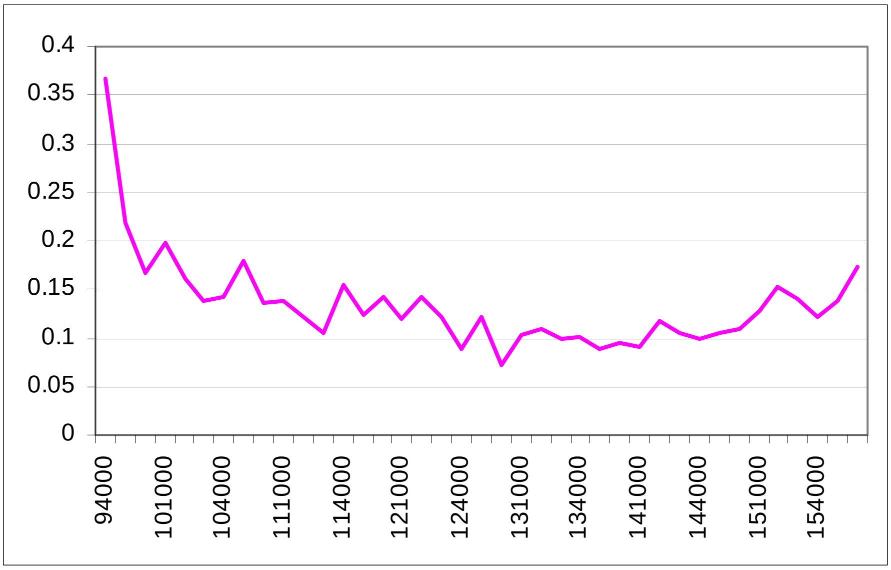
We obtain forecasts of daily volatilities h_t from a sequence of daily GARCH(1,1) models estimated from August 1985 until one day before the forecast period. This procedure is similar to the daily volatility component estimation applied by Andersen and Bollerslev (1997, 1998). Daily forecasts are computed for RGC between January 2, 1998 and June 30, 1999. Next we divide 10-minute returns by their respective daily volatility forecasts. What can be observed for these data, however, is a very pronounced diurnal volatility pattern. Figure 1 plots the standard deviation of returns in each of 40 10-minute bins. There is a pronounced increased variation in the beginning of each day, a calm period in the middle and somewhat increased variation towards the end. This diurnal pattern has been observed by many studies for all sorts of financial returns.

The sample standard deviation for each bin will be our estimate of diurnal component s_i. Hence in the second step, returns are normalized by their respective diurnal standard deviations. In order to take account of the remaining intraday dynamics, we fit a GARCH (1,1) model into returns standardized in that way. Tables 1 and 2 present results of daily and intraday GARCH estimation. Figure 2a. superimposes the three volatility components described above. We have chosen to only show two-week period (Fig 2a) and 2.5-month period (Fig.2b) for the reason of clarity.

The black rectangular line denoted shows daily volatility forecast, common for each day. The blue line represents the regular diurnal pattern, and the stochastic intraday component appears in pink. We may appreciate that this component is able to modify the regular deterministic diurnal pattern.

Figure 1

Standard deviation of return during a day for the RGC Stock



Note: The horizontal axis labels denote hours during a trading day.

Figure 2a

Volatility Components for the RGC Stock

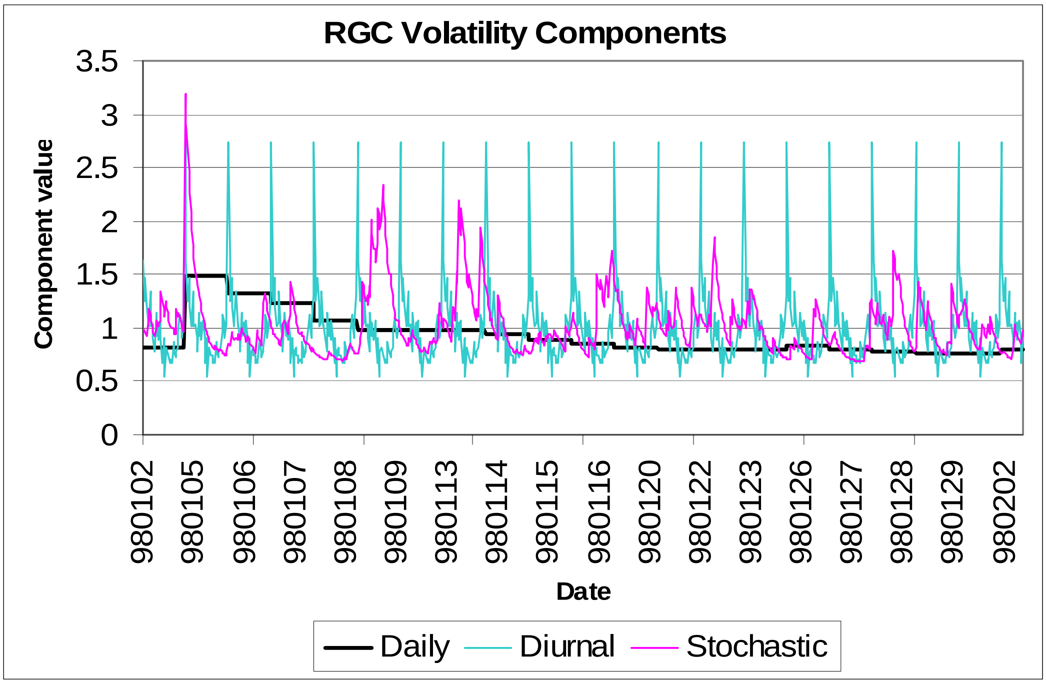


Figure 2b

Volatility Components for the RGC Stock

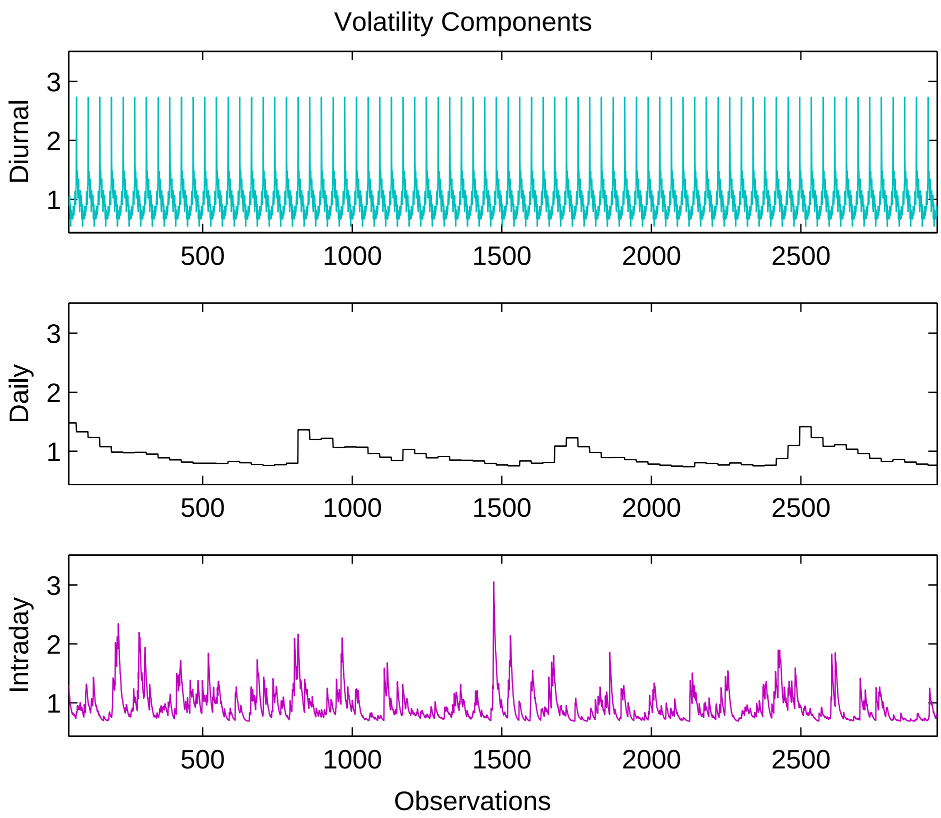


Figure 2c

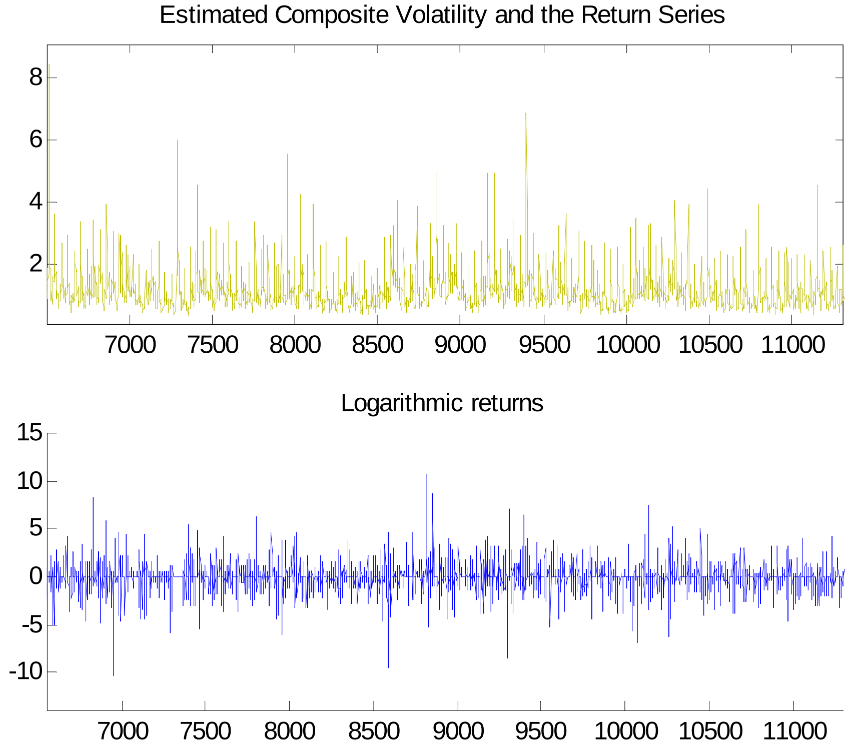
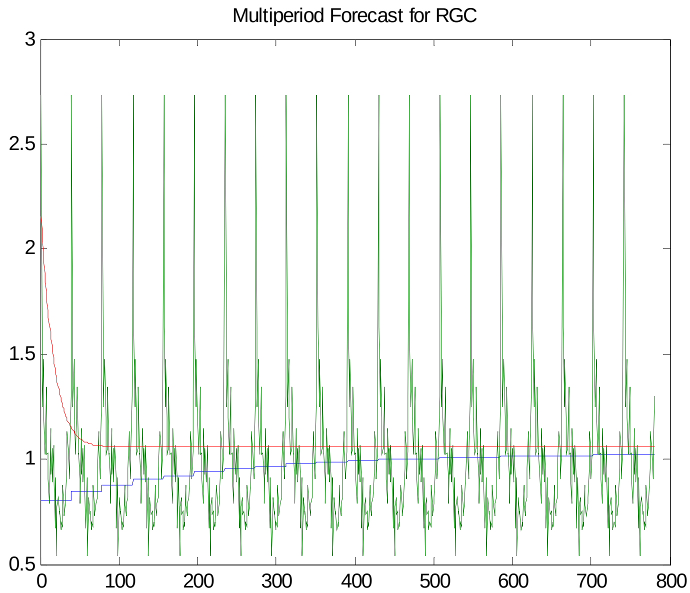


Figure 2d



4.3. RESULTS FOR A SAMPLE OF 2721 STOCKS

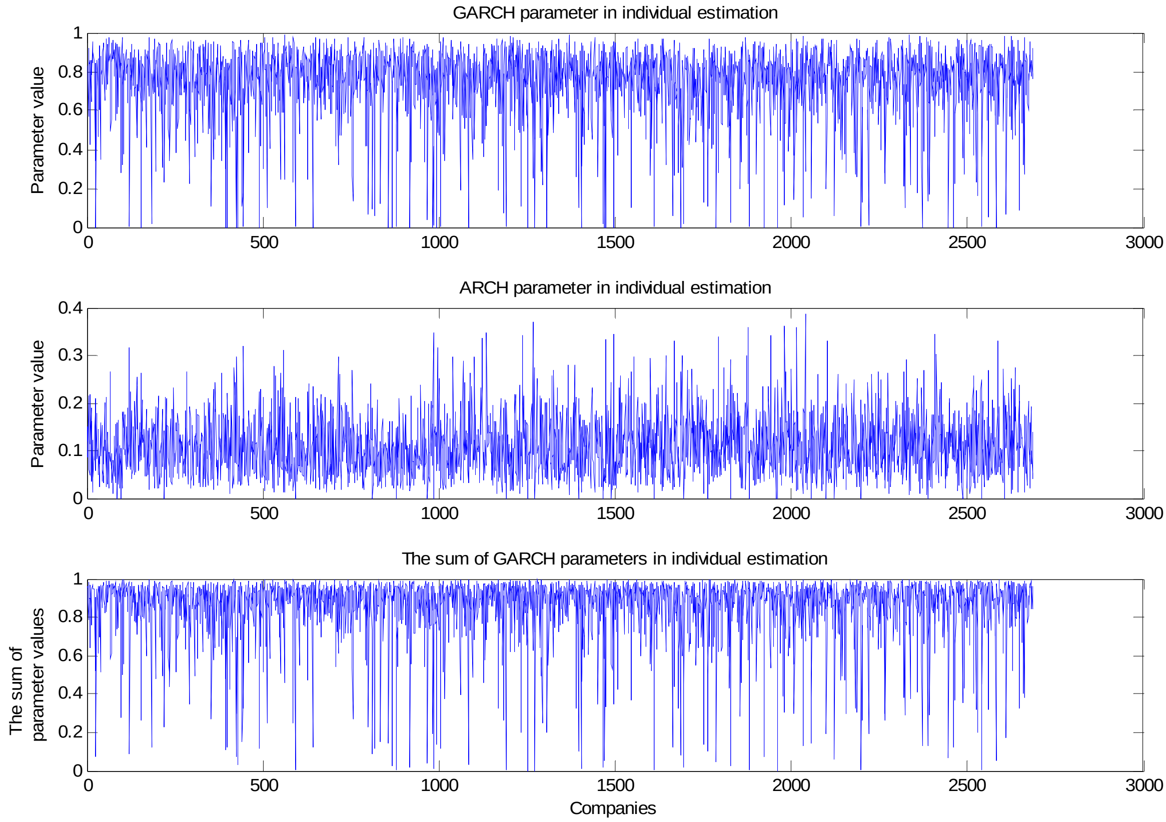
4.3.1 SEPARATE ESTIMATION RESULTS

Model (6) is estimated for 2721 US stock equity returns, which have been previously divided by a volatility forecast for a day and "diurnally adjusted" by the standard deviation for each bin. Any remaining serial correlation is eliminated by fitting an ARMA(1,1). Estimation is performed for the period April-May 2000, and the combined count of observations during this period exceeds 4.2 million data points. Since it is rather demanding to fit results of this estimation into a table of a manageable size, we report results of this procedure resorting to graphical methods. By a "GARCH parameter" and an "ARCH parameter", we refer to β and α coefficients from equation (3). The top and middle panels of Figure 3 depict β and α parameters, respectively. The bottom panel plots the sum of both parameters, thus informing us how persistent the volatility is. We may observe a fair amount of variation in the values of parameters and the measure of persistence. This result cannot be surprising taking into account that the estimation sample spans only two months. Although for 10-minute returns this translates into 1560 observations for most companies, the actual interval appears to be very short. Figure 4 sheds some more light on the nature of the observed parameter variation in separate regressions. It reports results of the very same estimations; however the sequence of companies has changed. We have sorted the companies according to the percentage of "zero" returns. This percentage could roughly be understood as a measure of intensity of trading. Substantial percentage of zero returns would indicate lack of active trading (although a zero return may result from trading at the same price).

Companies at the very left of Figure 4 almost always trade, and at the very right- hardly ever trade or trade at the same price most of the time. It can be observed that estimates' variability increases with the percentage of zero returns. Further, there is an upward trend in the GARCH parameter and a downward tendency for the ARCH parameter.

Figure 3

GARCH Estimation Results for the Intra-Day Component
2721 Separate Models



Note: The horizontal axis indicates companies.

B. Histogram of GARCH parameters

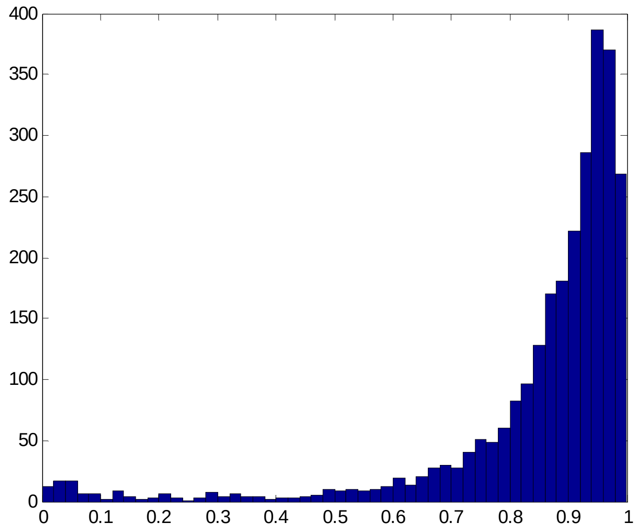
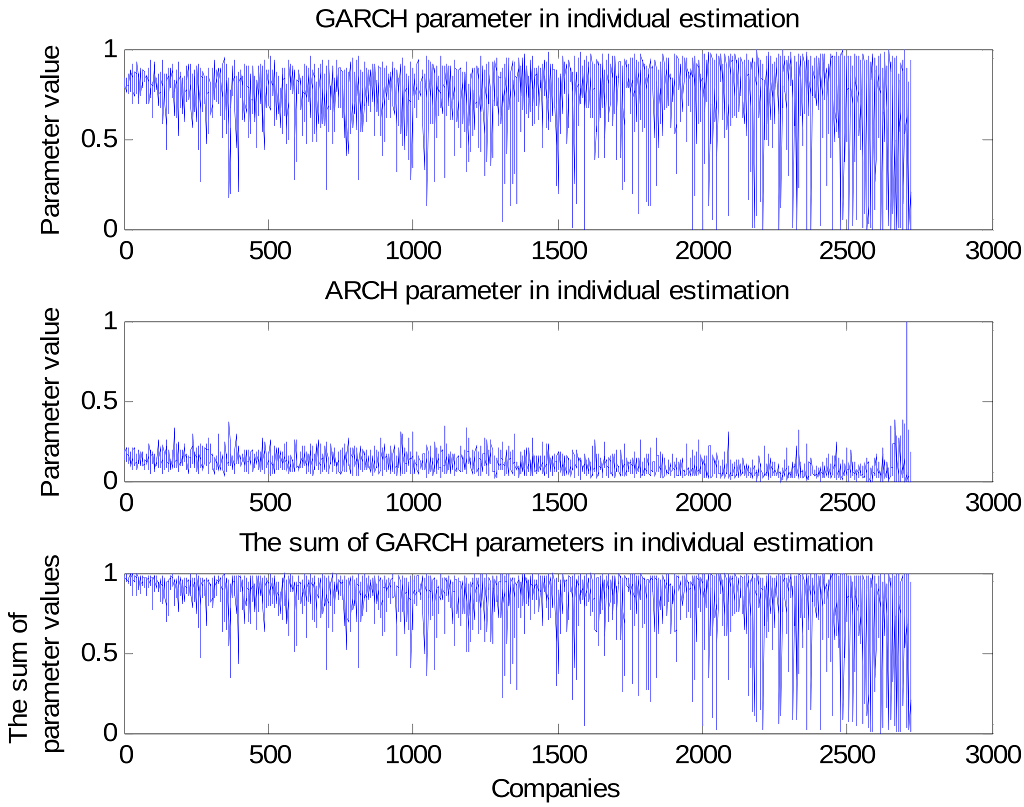


Figure 4

GARCH Estimation Results for the Intra-Day Component
2721 Separate Models, Companies Sorted According to Their Trading Intensity



Note: The horizontal axis indicates particular companies.

4.3.2 GROUPED ESTIMATION RESULTS

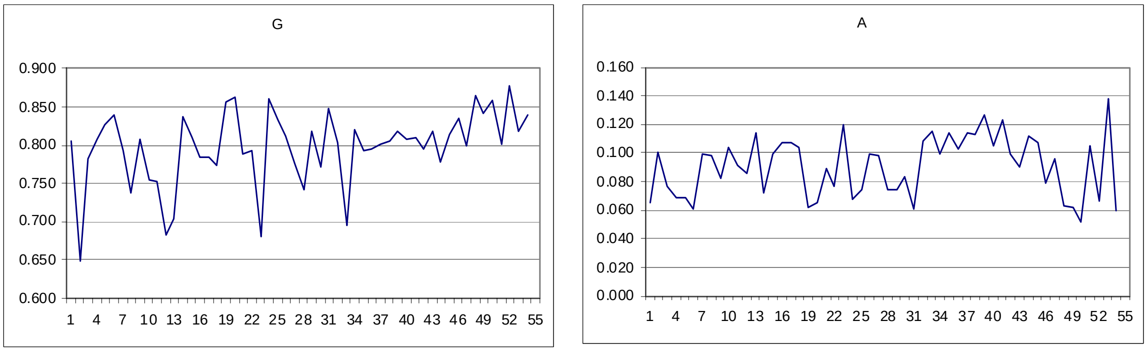
In order to improve forecasting performance of our model, we have decided to group data into 54 industries and estimate 54, instead of 2721 GARCH (1,1) models for an intraday component. Each return series has been divided by its standard deviation in order to render returns comparable across stocks. Standardized returns within each industry have been pooled into industry series. Estimation results of this step are summarized in Table 3 and parameters plotted in Figure 5. The last column indicates a very substantial reduction in persistence of volatility. Of course, industries vary in number of companies.

Table 4 reports results for ARCH estimation for the intraday component for one giant pool of all the companies, comprising over 4.2 million observations. Similarly to the industry case, this table also indicates modest persistence of intraday volatility.

A third approach to pooling stocks is to sort them by time series characteristics. This could have the benefit of reduced parameter uncertainty but preserving the natural heterogeneity across this dimension. Based on the previous results, we sorted on liquidity as measured by the percentage of time bins with zero returns. We estimated 50 categories of Liquidity for the GARCH models.

Figure 5a

GARCH Estimation Results for the Intra-Day Component
54 Industry Models



Note: G and A denote β and α parameters in model (2)-(3), respectively.

Figure 5b

GARCH Estimation Results for the Intra-Day Component
54 Industry Models, Persistence Parameter A+G, Histogram

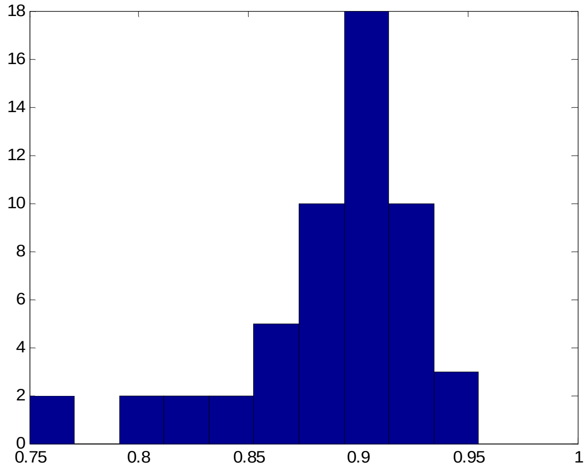
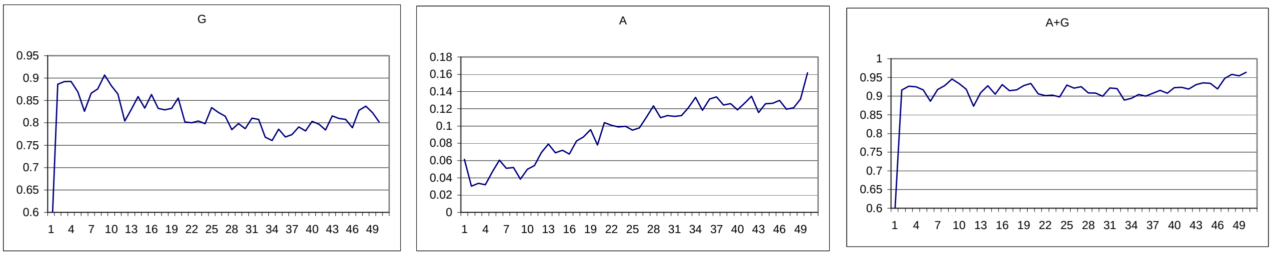


Figure 6

GARCH Estimation Results for the Intra-Day Component
50 Liquidity-Sorted Models



5. FORECASTING RESULTS

5.1. LOSS FUNCTIONS AND DESIGN

We now turn to out of sample forecast accuracy. We use the parameter estimates for the period April-May 2000, and forecast one step ahead volatilities for each bin in June 2000. Forecasts are obtained in a sequential procedure on the basis of estimated parameters and the volatility forecast calculated at previous bin. From the structure of the model, forecasts of the variance of returns are the product of the daily variance forecast, the diurnal variance and the garch variance. In this analysis, the variance that is forecast is of the return deflated by the daily vol times the diurnal vol.

It should be appreciated that forecasting volatility is connected with an additional complication since of course we do not observe the variable we want to forecast. In our forecasting evaluation we will compare our forecasts with the squared return ${z_{\{t,i\}}}^2 = r^2_{\{t,i\}} / \hat{h}_t \hat{s}_i$. This return is a random variable drawn from a distribution with a variance we are trying to estimate. We expect that the squared return will be large only when the true variance is large, however, the squared return may be small even when the variance is large. As a consequence, it is not at all clear what a sensible loss function should be. For recent discussions of forecast accuracy measures, see Granger (2003) or Patton (2004). In the following, we use two loss functions:

L1	LIK	Out-of-sample likelihood	$L_{1t} = \log q_{\{t,i\}} + \frac{{z_{\{t,i\}}}^2}{q^f_{\{t,i\}}}$
L2	MSE	Mean Squared Error	$L_{2t} = \left({z_{\{t,i\}}}^2 - q^f_{\{t,i\}}\right)^2$

The use of squared return in place of the true volatility renders many popular loss functions problematic (so does the RV measure). However, under MSE and LIK loss functions optimal forecasts unbiased (c.f. Patton, 2004).

We determine forecasts for each company separately, using parameters estimated in both separate and pooled estimations. Therefore for each time period, for each company, we obtain 5 different forecasts that will form the basis for a subsequent model evaluation and comparison.

5.2. OUT-OF-SAMPLE FORECAST COMPARISON

We have performed five different estimations for companies pooled into groups in various ways and will refer to these ways as modes. The first mode (NSTOCH) contains no stochastic component (5) at all. Mode No. 2 (UNIQUE) involves no pooling, i.e. we estimate unique GARCH models for separate companies. Mode No. 3 (INDUST) denotes a GARCH estimation for companies grouped according to their primary industry classification. In Mode No. 4 (LIQUID) we have grouped companies according to the average number of trades per day. The last mode (ONEBIG) involves estimation of a large GARCH model, for all companies pooled together to form one group.

For each of these 5 separate estimations we have calculated a series of forecast errors. These forecast errors are used to calculate accuracy measurement criteria using loss functions L_1 and L_2, $L_j = \frac{1}{\tau}\sum_{t=1}^{\tau} L_{jt}$.

Table 5 reports aggregated average results on forecast accuracy measures. It calculates average loss measures across all stocks. It also contains the ordering that each criterion assigns to the five estimation modes. Number 1 denotes the best model with the smallest error, model numbered as fifth performs the worst. The last column shows the average score, or the average ordering. According to the average score, the liquidity sorted model appears to forecast best out of sample for grouped data. Industry grouping comes as second, followed by one large GARCH mode. Please note that all the criteria assume very similar values for the three grouping modes. A clear message emerges from this table, however, that the separate estimation is inferior in terms of forecasting accuracy in comparison to the pooled modes. LIK criterion favours UNIQUE mode of NSTOCH model, whereas the MSE measure gives an opposite answer.

The results reported above are highly aggregated and call for a more detailed investigation. We will consider forecast accuracy measures separately for each of the 5 estimations and for each company, which amounts to a total of (10* 2721 =) 27210 numbers.

Table 6 reports aggregated results for pair wise forecast comparisons. It calculates the probability that the mode in the row will outperform the mode in the column or a random stock for a particular loss function. For example, the third row second column compares results of NSTOCH vs. separate estimations. Here the number 0.618 means that the specification without component (5) yields worse forecasts than the individual company by company estimation 62 % of times. The second column of the table informs us that NSTOCH estimation mode performed worse than all the other modes.

As we learn from column three, separate estimation gives worse forecasting results than INDUST, LIQUID and ONEBIG modes. The fifth column establishes forecasting inferiority of LIQUID mode in comparison to both INDUST and ONEBIG modes (as far as aggregate results are concerned). Finally the sixth and forth columns fails to reveal a clear winner in this forecasting competition. The MSE criterion marginally favours INDUST estimation, whereas the out-of-sample likelihood prefers one large GARCH estimation.

Tables 5 and 6 give somewhat conflicting answers to the question - which method of company grouping should be adopted. They only agree that the very grouping is desirable to the separate estimation. We investigate the supposed disagreement looking at liquidity issues. We limit our attention to the upper and lower tails of companies' distribution, as far as liquidity is concerned. We repeat the same calculations that were reported in Table 5, but take two separate samples of 550 most liquid and most illiquid companies. Table 8.A reports forecast accuracy measures for the subsample of least liquid stocks. The measures presented in Tables 5 and 8 are not exactly comparable, since each of them refers to a different sample/subsample. It is clear, however, that the means of error functions are bigger for least liquid stocks than the aggregate numbers in Table 5, which are in turn bigger than the means for most liquid stocks (Table 8, Panel B.). Nowhere the difference between tails is as visible as for the MSE measure. The figures differ by an order of a magnitude, and this illustrates the sensitivity of the MSE criterion to outlying observations, more frequently haunting illiquid stocks.

If on inspecting Table 8, we concentrate on the question of which estimation mode gives superior forecasts, we will notice that the conclusions for illiquid stocks are almost identical to the results for the aggregate results. Table 8.A recommends liquidity sorted GARCH model as a preferred solution.

The picture suddenly changes when we look at the results assembled in Table 8.B, which concern the most liquid stocks. Here one big GARCH solution, followed closely by industry grouping outperform both liquidity-sorted models and separate estimation. These conclusions closely resemble results reported for separate companies.

To sum up, we have seen uniformly improved forecasting performance on the basis of cross section information, i.e. when applying different methods of pooling. The exact way how we chose to group companies heavily relies on company characteristics, liquidity in this case. Most liquid stocks seem to benefit from the widest pooling possible, i.e. using all available cross section information. Most illiquid stock apparently exhibit intraday dynamics which is idiosyncratic to their particular liquidity-determined group.

6. CONCLUSION

This paper proposes a new way of modeling and forecasting intraday returns. We decompose the volatility of high frequency asset returns into components that may be easily interpreted and estimated. The conditional variance is expressed as a product of daily, diurnal and stochastic intraday volatility components. This model is applied to a comprehensive sample consisting of 10-minute returns on more than 2500 US equities.

REFERENCES

Alexander, C., Risk Management and Analysis, (New York: Wiley).

Andersen, T.G., and T. Bollerslev (1997), "Intraday Periodicity and Volatility Persistence in Financial Markets," Journal of Empirical Finance, 4, 115-158.

Andersen, T.G., and T. Bollerslev (1998), "DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies," Journal of Finance, 53, 219-265.

Blair, B.J., S.-H. Poon, and S.J. Taylor, 2001, "Forecasting S&P 100 volatility: the incremental information content of implied volatilities and high-frequency index returns", Journal of Econometrics, 105, 5-26.

Deo, R., C. Hurvich and Y. Lu,(2005) "Forecasting realized volatility using a long memory stochastic volatility model: Estimation, prediction and seasonal adjustment" Forthcoming in the *Journal of Econometrics*

Engle, R. F., "Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation," Econometrica 50(4) (1982), 987-1007.

Engle, R.F., 2002, "New Frontiers for ARCH Models", Journal of Applied Econometrics, 17, 425-446.

Engle, R.F. and G.P. Gallo(2005), A Multiple Indicators Model for Volatility Using Intra-Daily Data

Engle R. F. and A. Patton (2004), "Impacts of Trades in an Error-Correction Model of Quote Prices," Journal of Financial Markets, Vol. 7, No. 4.

Granger, C. (2003),Empirical modeling in economics.Specification and Evaluation, CUP.

Patton, A (2004), Volatility Forecast Evaluation and Comparison Using Imperfect Volatility Proxies.

Taylor S.J., and X. Xu (1997), "The incremental volatility information in million foreign exchange quotations" Journal of Empirical Finance, 4, 317-340.

Table 1. RGC Daily GARCH Results

Parameter	Value	Standard Error	T Statistic
C	0.00023757	0.00041135	0.577537
K	0.00013636	0.00001149	11.86358
GARCH(1)	0.65957	0.019085	34.5596
ARCH(1)	0.18073	0.010675	16.93021

Note: C denotes a constant in the mean equation, and K - a constant in the variance equation

Table 2. RGC Intraday GARCH Results

Parameter	Value	Standard Error	T Statistic
C	-0.0047551	0.0072439	-0.6564
K	0.079635	0.0020221	39.3828
GARCH(1)	0.8311	0.0033093	251.1402
ARCH(1)	0.097794	0.002116	46.2155

Note: C denotes a constant in the mean equation, and K - a constant in the variance equation

Table 3. Industry Results (Notes on the next page)

Industry	NoObs	LLF	C	tC	K	tK	G	tG	A	tA	BIC	1	2	5	10	20	40	A+G sum
1	92976	-128890	-0.0140	-4.61	0.132	121.48	0.805	558.34	0.065	110.66	257820	1.993	4.76	7.58	11.68	33.56	71.31	0.870
2	24102	-33598	-0.0110	-1.83	0.256	61.22	0.649	120.20	0.101	54.37	67236	0.011	0.77	1.45	2.39	3.82	13.97	0.750
3	71292	-98840	-0.0155	-4.49	0.144	134.95	0.783	533.04	0.076	92.81	197720	0.423	1.34	5.13	10.66	73.71	129.84	0.859
4	134200	-186330	-0.0074	-2.90	0.125	120.58	0.808	557.18	0.069	107.76	372700	1.930	3.74	7.80	12.97	24.02	161.73	0.877
5	100580	-139360	0.0116	3.95	0.107	118.80	0.826	649.19	0.069	98.56	278770	0.458	1.01	6.61	13.20	103.65	147.08	0.895
6	43641	-60851	-0.0017	-0.38	0.101	52.28	0.840	317.81	0.061	52.63	121740	0.643	2.22	3.66	5.13	10.05	39.47	0.901
7	55380	-76573	0.0034	0.86	0.111	64.10	0.793	292.88	0.100	62.23	153190	0.999	1.03	43.51	44.50	65.38	86.19	0.893
8	102610	-141890	0.0029	0.99	0.169	110.13	0.737	336.33	0.098	95.58	283830	0.359	2.72	4.46	15.16	30.18	69.32	0.834
9	23673	-32638	0.0101	1.74	0.118	78.44	0.808	358.65	0.082	66.39	65317	0.238	0.27	2.07	4.25	7.86	118.25	0.890
10	10647	-14695	0.0037	0.44	0.150	48.29	0.755	178.30	0.104	36.45	29426	0.445	0.45	1.45	7.01	142.95	160.76	0.859
11	37323	-51825	0.0050	1.04	0.159	68.97	0.752	254.07	0.092	81.17	103690	0.442	1.07	2.71	15.79	22.34	76.92	0.844
12	28236	-39090	-0.0024	-0.42	0.232	62.43	0.683	140.03	0.086	45.27	78222	0.001	0.20	1.55	3.39	7.48	253.32	0.769
13	33540	-46304	-0.0157	-3.19	0.187	54.69	0.703	133.14	0.114	43.30	92651	0.015	0.25	0.96	2.18	3.79	6.34	0.817
14	56978	-78622	-0.0208	-5.58	0.094	90.88	0.837	526.71	0.072	79.95	157290	2.048	2.08	9.04	11.63	28.75	57.13	0.908
15	44928	-60880	-0.0081	-1.94	0.094	90.89	0.809	449.96	0.100	80.36	121800	3.571	3.61	11.13	15.59	23.61	65.68	0.909
16	37401	-51323	0.0002	0.05	0.112	57.39	0.785	237.54	0.107	55.84	102690	1.898	2.61	6.81	12.28	36.76	47.13	0.892
17	90714	-123720	-0.0072	-2.42	0.113	146.53	0.783	562.79	0.107	100.38	247480	0.450	1.21	4.20	6.62	52.08	100.34	0.891
18	24960	-34412	-0.0038	-0.67	0.125	62.90	0.774	257.30	0.104	52.14	68864	1.693	2.48	7.71	14.57	24.90	42.96	0.878
19	84240	-115920	-0.0061	-1.95	0.083	128.23	0.856	830.67	0.062	113.08	231880	1.461	1.50	2.83	7.67	11.90	28.08	0.918
20	57330	-79145	-0.0151	-3.93	0.074	81.01	0.863	587.80	0.066	74.98	158330	2.166	2.17	9.64	20.53	26.12	43.01	0.928
21	103230	-141720	-0.0137	-4.72	0.126	153.10	0.788	601.23	0.089	104.86	283480	0.014	0.37	1.81	4.52	8.36	22.54	0.877
22	28119	-39152	0.0030	0.54	0.134	41.77	0.792	184.19	0.076	42.78	78346	0.690	1.00	5.21	7.52	16.13	51.38	0.869
23	37401	-51842	-0.0024	-0.51	0.206	50.80	0.680	129.13	0.120	60.64	103730	0.914	1.97	9.92	12.01	39.23	55.49	0.799
24	25584	-34907	-0.0061	-1.09	0.073	58.53	0.860	438.01	0.068	55.61	69854	0.473	0.50	8.97	10.34	25.12	75.87	0.928
25	48165	-66239	-0.0071	-1.69	0.095	92.02	0.832	529.08	0.075	71.41	132520	3.651	3.69	6.60	13.66	39.91	65.51	0.907
26	10920	-14976	-0.0062	-0.73	0.091	19.69	0.811	114.01	0.099	24.81	29988	3.649	3.69	5.12	12.46	19.14	48.62	0.910
27	42119	-57852	-0.0084	-1.87	0.127	63.59	0.777	237.50	0.099	55.96	115750	0.233	1.42	1.89	4.56	8.10	646.70	0.875
28	17121	-23836	0.0027	0.37	0.185	32.55	0.742	102.24	0.074	31.63	47711	0.092	0.88	0.99	2.34	7.47	32.70	0.816
29	103540	-142960	-0.0063	-2.18	0.112	142.93	0.817	689.12	0.074	116.15	285970	0.303	1.82	4.61	9.90	13.52	24.18	0.891
30	119110	-166030	0.0113	4.12	0.147	138.16	0.771	522.60	0.083	93.20	332100	5.716	6.45	23.68	28.00	118.85	153.48	0.854
31	58733	-81804	0.0121	3.16	0.093	76.81	0.848	493.08	0.061	71.20	163650	1.023	1.26	4.17	10.34	26.81	114.25	0.909
32	15600	-21405	0.0010	0.14	0.094	25.74	0.803	135.94	0.109	30.64	42850	0.055	1.09	3.27	8.50	18.50	34.50	0.911
33	31629	-43650	-0.0036	-0.70	0.196	78.15	0.695	194.59	0.116	58.08	87342	0.031	1.14	2.64	5.73	26.91	66.64	0.811
34	44694	-60711	-0.0031	-0.79	0.088	153.28	0.820	868.30	0.100	109.77	121470	2.506	3.29	5.75	8.59	12.75	74.63	0.920
35	51401	-70025	0.0006	0.14	0.097	63.02	0.794	329.90	0.114	76.98	140090	14.129	15.06	20.17	25.39	40.37	72.11	0.908
36	108810	-148910	-0.0043	-1.65	0.108	127.52	0.795	606.19	0.102	124.64	297860	4.554	5.46	11.45	18.69	22.48	45.36	0.897
37	214610	-290520	-0.0122	-6.57	0.089	161.37	0.801	891.12	0.114	188.33	581080	3.368	3.55	8.31	16.64	80.34	109.06	0.916
38	222300	-298840	-0.0130	-7.17	0.084	191.53	0.806	952.12	0.113	169.34	597720	4.439	7.57	21.20	44.04	50.19	100.56	0.919
39	148820	-197780	-0.0146	-6.82	0.061	108.20	0.818	744.25	0.126	131.09	395600	16.126	17.58	45.50	73.92	81.46	193.38	0.944
40	125260	-168830	-0.0178	-7.20	0.088	174.91	0.807	801.51	0.105	132.50	337710	0.368	0.52	3.31	5.79	8.46	632.56	0.913
41	258060	-344260	-0.0236	-14.61	0.073	228.02	0.810	1372.00	0.123	222.82	688570	16.312	17.97	39.25	60.83	68.41	160.80	0.932
42	34242	-46826	0.0016	0.32	0.109	60.39	0.795	282.90	0.100	59.50	93694	3.999	4.33	9.42	12.43	21.39	57.09	0.895
43	113020	-153890	-0.0180	-6.84	0.092	129.93	0.819	662.60	0.091	109.66	307820	0.926	2.41	10.34	18.61	24.77	127.73	0.909
44	50505	-68427	-0.0307	-7.79	0.116	123.27	0.777	467.03	0.112	86.16	136900	0.021	0.08	0.97	3.90	12.98	72.66	0.889
45	174680	-235800	-0.0157	-7.75	0.085	184.04	0.814	935.66	0.107	154.84	471660	4.365	5.84	16.64	26.94	32.91	66.18	0.920
46	85602	-117200	-0.0031	-1.04	0.089	118.16	0.835	729.52	0.079	116.18	234450	5.362	5.46	15.80	20.04	40.39	58.82	0.914
47	34865	-47537	0.0058	1.21	0.109	89.12	0.799	429.74	0.096	84.32	95115	5.988	6.95	12.86	17.93	27.10	62.16	0.895
48	96524	-132830	0.0049	1.72	0.074	169.03	0.865	1506.90	0.063	128.44	265700	14.657	14.72	17.65	24.20	30.03	212.02	0.928
49	218210	-302410	0.0067	3.48	0.098	185.07	0.842	1064.50	0.062	149.24	604880	3.776	4.97	10.41	16.23	26.80	154.44	0.904
50	48048	-66832	0.0039	0.95	0.090	74.70	0.859	533.63	0.052	71.95	133710	2.984	4.15	7.04	10.86	12.75	36.55	0.911
51	57679	-78235	-0.0029	-0.79	0.097	117.88	0.802	535.37	0.105	89.48	156510	0.370	0.85	4.39	8.74	15.65	41.16	0.907
52	87903	-119860	-0.0061	-2.02	0.060	156.60	0.876	1391.00	0.066	131.10	239760	4.911	4.99	6.84	9.02	20.79	98.63	0.943
53	79754	-104390	-0.0284	-10.46	0.050	74.06	0.817	501.03	0.137	97.43	208830	10.193	10.20	23.44	42.19	50.96	120.91	0.955
54	156160	-218270	0.0133	5.73	0.104	97.56	0.838	575.22	0.059	101.57	436580	1.455	2.40	7.01	17.43	45.70	64.92	0.898
											ARCH test critical values	**3.842**	**5.99**	**11.07**	**18.31**	**31.41**	**55.76**	

Notes to Table 3

NoObs	Number of observations
LLF	Value of the Loglikelihood function
C	Constant in the mean equation
tC	t-value for C parameter
K	Constant in the variance equation
tK	t-value for K parameter
G	beta in the variance equation
tG	t-value for G
A	alpha in the variance equation
tA	t-value for A
BIC	Schwartz Information Criterion
1	ARCH LM test for volatility clustering in residuals – Lag 1
2	Lag 2
5	Lag 5
10	Lag 10
20	Lag 20
40	Lag 40
A+G	the sum of alpha and beta parameters

Table 4 All Sample Results

Parameter	Value	Standard Error	T Statistic
C	-0.0072459	0.00043311	-16.7301
K	0.099443	0.00011118	894.4290
GARCH(1)	0.81556	0.00018085	4509.5115
ARCH(1)	0.087421	0.00011767	742.9462

Note: C denotes a constant in the mean equation, and K - a constant in the variance equation

Table 5. Forecast accuracy measures for different estimation modes

Aggregate results

	Forecast accuracy measures	**LIK**	**O**	**MSE**	**O**
NSTOCH	No stochastic intraday component	1.0001	5	3.8228	4
UNIQUE	Separate GARCH estimation	0.9402	4	3.8231	5
INDUS	Industry GARCH estimation	0.9292	2	3.8014	2
LIQUID	Liquidity-Sorted GARCH estimation	0.9264	1	3.7973	1
ONEBIG	One large GARCH estimation	0.9295	3	3.8018	3

Note: O denotes ordering from best (1) to worst (4).

Table 6 Forecast comparison

LIK	**NSTOCH**	**UNIQUE**	**INDUST**	**LIQUID**	**ONEBIG**
NSTOCH		0.382	0.275	0.245	0.262
UNIQUE	0.618		0.354	0.404	0.339
INDUST	0.725	0.646		0.572	0.445
LIQUID	0.755	0.596	0.428		0.377
ONEBIG	0.738	0.661	0.555	0.623	
MSE	**NSTOCH**	**UNIQUE**	**INDUST**	**LIQUID**	**ONEBIG**
NSTOCH		0.304	0.203	0.307	0.222
UNIQUE	0.696		0.419	0.465	0.432
INDUST	0.798	0.581		0.615	0.520
LIQUID	0.694	0.535	0.385		0.411
ONEBIG	0.778	0.568	0.480	0.590	

Fraction of companies where row beats column

Table 7 Forecast accuracy comparison for different estimation modes individual stocks

		NSTOCH	UNIQUE	INDUST	LIQUID	ONEBIG
LIK	Mean	1.0001	0.9485	0.9306	0.9271	0.9313
	Rank	5	4	2	1	3
LIK	Median	1.0002	0.9543	0.9430	0.9432	0.9410
	Rank	5	4	2	3	1
MSE	Mean	3.4930	3.4835	3.4668	3.4661	3.4666
	Rank	5	4	3	1	2
MSE	Median	2.9765	2.9571	2.9495	2.9498	2.9478
	Rank	5	4	2	3	1

Table 8. Forecast accuracy measures for different estimation modes

A. Least Liquid Stocks						
	Forecast accuracy measures	**LIK**	**O**	**MSE**	**O**	**Av. Score**
NSTOCH	No stochastic intraday component	1.0002	5	5.8425	2	3.5
UNIQUE	Separate GARCH estimation	0.9779	4	5.8871	5	4.5
INDUST	Industry GARCH estimation	0.9548	2	5.8481	3	2.5
LIQUID	Liquidity-Sorted GARCH estimation	0.9405	1	5.8287	1	1
ONEBIG	One large GARCH estimation	0.9629	3	5.8541	4	3.5
B. Most Liquid Stocks						
	Forecast accuracy measures	**LIK**	**O**	**MSE**	**O**	**Av. Score**
NSTOCH	No stochastic intraday component	0.9999	5	2.7439	5	5
UNIQUE	Separate GARCH estimation	0.9679	4	2.7309	4	4
INDUST	Industry GARCH estimation	0.9293	2	2.7084	2	2
LIQUID	Liquidity-Sorted GARCH estimation	0.9339	3	2.7172	3	3
ONEBIG	One large GARCH estimation	0.9274	1	2.7044	1	1

Note: O denotes ordering from best (1) to worst (4). Average Score is calculated as the mean of ordering measures in each row.

Table 9. Forecast accuracy comparison Diebold-Mariano test – t-values

A. Least Liquid Stocks					
LIK	NSTOCH	UNIQUE	INDUST	LIQUID	ONEBIG
NSTOCH		-1.603	-8.911	-4.618	-2.524
UNIQUE	1.603		- 1.739	-1.911	-0.731
INDUST	8.911	1.739		-0.967	0.476
LIQUID	4.618	1.911	0.967		1.465
ONEBIG	2.524	0.731	-0.476	-1.465	
B. Most Liquid Stocks					
LIK	NSTOCH	UNIQUE	INDUST	LIQUID	ONEBIG
NSTOCH		-26.822	-36.386	-10.686	-12.407
UNIQUE	26.822		-9.411	-0.621	-1.832
INDUST	36.386	9.411		0.758	-0.408
LIQUID	10.686	0.621	-0.758		-1.174
ONEBIG	12.407	1.832	0.408	1.174	