

Intraday Dynamics of Volatility and Duration: Evidence from Chinese Stocks

Chun Liu
School of Economics and Management
Tsinghua University

John M. Maheu*
Dept. of Economics
University of Toronto
& RCEA

September 2011

Abstract

We propose a new joint model of intraday returns and durations to study the dynamics of several Chinese stocks. We include three U.S. stocks for comparison. Flexible innovation distributions are used for durations and returns, and the total variance of returns is decomposed into different volatility components associated with different transaction horizons. The new model provides strong improvements in density forecasts for duration and returns and only modest gains for points forecasts of the variance of returns. The conditional hazard functions are non-monotonic and there is strong evidence for different volatility components. Although diurnal patterns, volatility components, and market microstructure implications are similar across the markets, there are interesting differences. Durations for lightly traded Chinese stocks tend to carry more information than heavily traded stocks. Chinese investors usually have longer investment horizons, which may be explained by the specific trading rules in China.

Key words: Market microstructure; Transaction horizon; High-frequency data; ACD; GARCH

JEL: C22, C11, G10

*We are grateful for the helpful comments from the editor S. Ghon Rhee, an anonymous referee, Charles Cao, Feng Gao, Chuan Goh, Raymond Kan, Li Gan, Gael Martin, Alex Maynard, Matthew Pritsker, Tom McCurdy, Angelo Melino and participants of China International Conference in Finance, Beijing. Liu thanks Tsinghua Humanities and Social Sciences Revitalization Fund for financial support (Grant No. 2010WKYB004). Maheu thanks the Social Sciences and Humanities Research Council of Canada for financial support.

1 Introduction

As China moves toward its potential output and commands an increasing share of world output, trade and commerce, it is important to study the properties of its domestic asset market. Despite the importance that China has on the world stage, there is little to no research on the time-series properties of intraday Chinese stock returns.¹ The purpose of this paper is to begin to fill that gap. We propose a new joint model of intraday returns and duration to study the dynamics of several Chinese stocks. Besides using flexible innovation distributions the conditional variance is made of multiple components based on transaction time which contributes to large improvements in density forecasts compared to existing models. We include three U.S. stocks for comparison purposes and highlight the similarities and differences.

The final frontier in the time-series analysis of asset returns is intraday transactions. This is the highest frequency that data is available and the structure of this data generating process determines all lower frequency dynamics. Lower frequency returns are derived from the high frequency process. Hence the time series features of this process are critical to understanding the characteristics of daily, and weekly returns. Compared with the traditional low frequency analysis, high frequency data have an immense number of observations, a pronounced seasonal structure and random time between trades.

Research on the duration process, or time between trades, is large and began with the seminar paper of Engle and Russell (1998) which proposes the univariate Autoregressive Conditional Duration (ACD) model. Many extensions have been considered, such as Bauwens and Giot (2000), Lunde (1999), Grammig and Maurer (2000), Zhang, Russell and Tsay (2001), Ghysels, Gouriou and Jasiak (2004) and Bauwens and Veredas (2004).

Among all the characteristics associated with high-frequency transaction records, the trading time and the price are the most important. The bivariate dynamics of returns and durations are taken up in Engle (2000). The joint model of return and duration allows for density forecasts in the duration to the next trade and the associated return distribution. A joint model is necessary to provide a clear picture of the dynamics of the volatility process. Durations have an important effect on the volatility of the returns, see for example, Dufour and Engle (2000), Grammig and Wellner (2002), and Ait-Sahalia and Mykland (2003). A joint model can help discriminate market microstructure theories. Specifically, the waiting time between events plays a key role in understanding the process of private and public information in financial markets.

Besides Engle (2000), bivariate models of returns and duration include Engle and Sun (2005), and Ghysels and Jasiak (1998). They use a GARCH-type model for the volatility and an ACD model for the duration. However, the standard GARCH model has problems capturing the strong persistence in the volatility process, which is an important property in high-frequency data in many studies. In this paper, we propose a new joint model to describe the dynamics of the high-frequency data. Specifically, we model the volatility dynamics

¹There is unrelated work by Zhou and Zhu (2009) who construct jump component from high-frequency data, and Tian and Guo (2007) who compare interday and intraday volatility. Lee et al. (2008) and Chong and Su (2006) investigate the differences in asset classes due to ownership restrictions. Cai et al. (2006) study which trades move prices. Finally, Cheung et al. (1994) investigate the intraday market return volatility of the Hong Kong stock market.

by a component approach, where the volatility components are associated with different transaction horizons.² The components are essentially a *realized volatility* measure calculated according to trading time and not calendar time.³ The use of volatility components is motivated by the Heterogeneous Market Hypothesis (HMH) of Muller et al. (1997). The HMH recognizes the existence of heterogeneity in markets. Different types of traders trade for different purposes, and thus decide to execute their transactions in different market situations. They create different volatilities.

In the daily time horizon, Corsi (2009) proposes the Heterogeneous Autoregressive (HAR) model based on the HMH. The HAR model is popular in modeling daily realized volatility.⁴ Other research which decomposes daily volatility includes Calvet and Fisher (2008), Engle and Lee (1999), and Maheu (2005). These papers deal with fixed-interval volatility and ignore the information from the duration process.

Besides the component model for the conditional variance, return innovations follow a t-distribution and duration innovations follow a Burr distribution. The new joint model provides strong improvements in density forecasts for duration and returns and only modest gains for points forecasts of the variance of returns. Compared to existing models it provides better in-sample performance and displays improved out-of-sample forecasting power.

The new model is estimated for three stocks from the Chinese market. We also include three U.S. stocks to provide a comparison with and to gauge the differences. Chinese stocks are segmented into two groups according to investor entrance permission. The first, share-A stock, is dominated by the Chinese Yuan, and open only to domestic investors⁵. The second is share-B stock, dominated by foreign currency (U.S. Dollar in Shanghai Stock Exchange and Hong Kong Dollar in Shenzhen Stock Exchange) and open to both domestic and foreign investors. In addition, we are also interested in the different performance of stocks with different market capitalization in the Chinese market. We randomly choose one stock from each of the groups: the Share-A large stock group, the Share-A small stock group, and the Share-B group.

We optimally select the number of volatility components according to model fit. They include the squared instantaneous return from the last transaction and the average of those ranging from several minutes up to one hour. Information beyond one hour has no effect on intraday volatility for any of the stocks considered. Contemporaneous durations have important effects on the volatility process. The longer the duration, the lower is the conditional variance.⁶

We find a number of common properties that Chinese markets share with developed

²The idea is also related to the realized volatility literature such as Andersen, Bollerslev, Diebold and Labys (2001), Andersen, Bollerslev, Diebold and Ebens (2001) and Barndorff-Nielsen and Shephard (2002a, 2002b) since we compute realized volatility over various transaction horizons.

³For a recent review of volatility including realized volatility see Andersen et. al. (2009).

⁴Recent literature using HAR models includes Andersen, Bollerslev and Diebold (2007), Andersen, Bollerslev, and Huang (2006), Bollerslev, Kretschmer, Pigorsch and Tauchen (2009), Corsi, Kretschmer, Mittnik, and Pigorsch (2005), Forsberg and Ghysels (2007), and Maheu and McCurdy (2011).

⁵In 2002, the Qualified Foreign Institutional Investor (QFII) system was introduced in China to allow selected foreign institutional investors to invest in Share-A market directly. 86 overseas investors had been granted QFII status as of August, 2010.

⁶The results support the Easley and O'Hara (1992) formulation in which the long waiting time means no information.

ones. Like the U.S. market, trade durations and the conditional variance of returns display strong serial dependence. Diurnal patterns of durations and the volatility are similar. There are more transactions and larger price changes at the beginning of a day and prior to the close of the market, while less transactions and more stable prices around lunch time.⁷ The probability of a trade, as a function of the time from the last trade, increases the first few seconds but then decreases thereafter. This inverted “U” shape of the conditional hazard function is inconsistent with the exponential distribution that is popular in the literature.

However, we find some interesting differences. The autocorrelation function of returns and our model estimates indicate a lower degree of dependence in volatility measures compared to U.S. stocks. Durations for lightly traded Chinese stocks tend to carry more information than heavily traded U.S. stocks. In the Chinese market, the trading frequency is much lower. Long-term volatility components have a larger effect than the short-term components. One explanation is the existence of longer investment horizon in China. This may be related to particular institutional trading rules in China and the overall shorter market open time per day.

The remainder of this paper is structured as follows. Section 2 discusses our new model along with existing benchmark specifications. Estimation issues are presented in Section 3. Section 4 reviews the institutional features of the Chinese market and the data. Section 5 presents empirical results. The last section summarizes.

2 Models

The high-frequency transaction data contain two types of observations. One is the time of the transaction. The other is a vector of the quantities, called the marks, observed at the time of the transaction. The marks include price, volume, and spread in addition to other characteristics. The time information is summarized by duration x_i which is defined as the time between two transactions. Denoting t_i as the time associated with the transaction i , the duration is then $x_i = t_i - t_{i-1}$. In our paper, we deal with two type of the durations: The first type is defined as the time between adjacent trades, while the second one is defined by the arrival time of the transactions occurring at a new price, or the thinned point prices by Engle and Russell (1998).

The return is the difference of log prices, $\tilde{r}_i = p_i - p_{i-1}$. Engle (2000) argues that the natural measure of volatility is the variance per unit of time. Since the variances are connected to the squared returns, we construct our model based on the return per square root of time, which is defined as $r_i = \tilde{r}_i/\sqrt{x_i}$. Therefore, the data we are dealing with are a sequence of joint observations of the duration and return denoted by $\{(x_i, r_i), i = 1, \dots, T\}$, where T is the total number of observations.⁸

⁷The Chinese market closes for a 90 minute lunch period.

⁸In this paper, we are dealing with returns per square root of time except when specified.

2.1 ACD Models for Duration

We assume the duration process follows the ACD(1,1) model proposed by Engle and Russell (1998)

$$x_i = \psi_i z_i \text{ with } \psi_i = \omega_a + \alpha_a x_{i-1} + \beta_a \psi_{i-1} \quad (1)$$

where ψ_i is the expectation of the duration conditional on the last period's information, and the duration innovation z_i follows independent identical distribution $f_z(\cdot)$ with a non-negative support. The standard ACD model assumes that $f_z(\cdot)$ is an exponential distribution, which is called an EACD model.

In an EACD model, the conditional hazard function $h(t) = \psi_i^{-1}$ is constant. Bauwens and Veredas (2004), Lunde (1999), Hamilton and Jorda (2002) and Zhang et al. (2001) question whether imposing this restriction on the hazard function is appropriate, and propose specifications that offer greater flexibility. Bauwens, Giot, Grammig and Veredas (2004) compare most of the popular duration models, and suggest the ACD specification with a more flexible innovation distribution, such as a Burr or a Generalized Gamma distribution.

2.2 Burr ACD (BACD) Model

Grammig and Maurer (2000) use the Burr distribution, denoted as $\text{Burr}(\mu, \kappa, \varpi^2)$ to provide a more flexible innovation distribution, z_i . Coupled with the time-series dynamics of x_i in (1) we have the BACD model. It allows for different forms of the hazard function with different parameter values. Under the restriction $\mu = 1$, the duration innovation z_i is

$$z_i = x_i / f(\psi_i) \quad (2)$$

where

$$f(\psi_i) = \psi_i \frac{(\varpi^2)^{\left(1 + \frac{1}{\kappa}\right)} \cdot \Gamma\left(\frac{1}{\varpi^2} + 1\right)}{\Gamma\left(1 + \frac{1}{\kappa}\right) \cdot \Gamma\left(\frac{1}{\varpi^2} - \frac{1}{\kappa}\right)} \quad (3)$$

with $0 < \varpi^2 < \kappa$. The density for z_i is a Burr distribution and can be written as

$$g(z_i) = \frac{\kappa z_i^{\kappa-1}}{(1 + \varpi^2 z_i^\kappa)^{(1/\varpi^2)+1}}, \quad z_i \geq 0. \quad (4)$$

The conditional hazard function is

$$h(x_i) = \frac{f(\psi_i)^{-\kappa} \cdot \kappa \cdot x_i^{\kappa-1}}{1 + \varpi^2 \cdot f(\psi_i)^{-\kappa} \cdot x_i^\kappa} \quad (5)$$

which is non-monotonic with respect to duration for $\kappa \geq 1$ and $\varpi^2 > 0$.

The BACD reduces to a Weibull ACD model when $\varpi^2 \rightarrow 0$. The standard EACD model is also a special case if both $\varpi^2 \rightarrow 0$ and $\kappa = 1$.

2.3 Benchmark GARCH-BACD Model

Engle (2000) uses a GARCH-ACD model for the joint process of returns and duration. Duration follows an ACD model and the conditional variance has a GARCH-type structure.

There is usually a pronounced market microstructure effect for high frequency data. For example, an i.i.d. noise process that is uncorrelated with the latent price would induce an MA(1) structure in the observed returns. In a more general case of dependent noise an ARMA(1,1) process is usually adopted for the observed return process.⁹ Using Engle's model as a starting point we consider

$$r_i = \rho r_{i-1} + u_i + \phi u_{i-1} \quad (6)$$

where the innovation u_i is

$$u_i = \sqrt{q_i} \zeta_i \text{ with } \zeta_i \stackrel{iid}{\sim} t_\nu(0, 1), \quad (7)$$

and q_i is the variance of the return conditional on past volatility and current duration information. In contrast to Engle who assumes normal innovations¹⁰, $t_\nu(0, 1)$ denotes a Student-t distribution with mean 0, scale parameter 1, and degree of freedom ν to account for the fat tails in high-frequency returns.

Engle (2000) suggests several versions of the GARCH model to describe the dynamics of the conditional variance. The most successful one with current duration information is

$$q_i = \omega_g + \alpha_g u_{i-1}^2 + \beta_g q_{i-1} + \gamma_1 x_i^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \psi_i^{-1} + \gamma_4 \xi_i. \quad (8)$$

Here, three parts affect the conditional variance q_i . The first part is a GARCH(1,1)-type effect including the previous squared innovation u_{i-1}^2 and the last conditional variance q_{i-1} . The second part captures the interaction between the volatility and the duration. It includes the reciprocal duration x_i^{-1} , the reciprocal of expected duration ψ_i^{-1} , and the duration surprise x_i/ψ_i . The third part is the long-run volatility variable ξ_i , which captures the long-term persistence of the conditional variance. It is computed by exponentially smoothing squared returns r^2 with a preset parameter 0.995, as in $\xi_i = .005r_{i-1}^2 + .995\xi_{i-1}$.

This generalized GARCH-BACD model is our benchmark specification in our empirical work.

2.4 HAR-BACD Model

In addition to including the more flexible Burr distribution for durations and the Student-t for returns, our main contribution from a modelling perspective is to improve the conditional variance specification.

We apply a component approach to high-frequency volatility. The basic idea is motivated from the Heterogeneous Market Hypothesis (HMH) of Muller et al. (1997), which recognizes the presence of heterogeneity in traders.¹¹ Market participants have different time horizons, and therefore perceive, react to, and cause different types of volatility. For example, market makers and intraday speculators have very short time horizons and focus on the tick-by-tick data; while, on the other end, the central banks and some pension funds may only be concerned with the long-term performance of the markets. In a heterogeneous market, different

⁹Refer to Campbell, Lo and Mackinlay (1997) for a review.

¹⁰He uses this in quasi-maximum likelihood estimation.

¹¹A close and related idea is advocated by Andersen and Bollerslev (1998a) where they show different fractions of return volatility are associated with different information flows.

types of traders trade for different purposes and thus decide to execute their transactions in different market situations, hence they create different volatilities. Motivated by HMH, Corsi (2009) proposes the Heterogeneous Autoregressive (HAR) model for the daily volatility. He shows that although the HAR-type model is not a true long-memory model, it does provide a good approximation to the dynamics of long memory which is a stylized fact of high frequency data.¹²

We apply this idea in the high-frequency framework, and decompose the conditional volatility into components which are associated with different transaction horizons in business time. Specifically, the M -component HAR-BACD model includes (1), (2), (6) and (7) along with the following conditional variance,

$$q_i = \beta_0 + \sum_{m=1}^M \beta_m VC_{i-1, h_m} + \gamma_1 x_i^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \psi_i^{-1}. \quad (9)$$

M is the total number of the components, and VC_{i-1, h_m} is the m -th realized volatility component defined as

$$VC_{i-1, h_m} = \frac{u_{i-1}^2 + \dots + u_{i-h_m}^2}{h_m} \quad (10)$$

where h_m is the number of ticks (transactions) associated with each component. Therefore, components are associated with transaction times and not calendar time as in Corsi (2009). Oomen (2006) shows that transaction time sampling leads to more efficient estimates of volatility.

We will discuss how to choose M and h_m later. When $h_m > 1$, VC_{i-1, h_m} is the realized volatility constructed from squared return innovations during the time period from transaction $i - h_m$ to transaction $i - 1$. When trading is light (heavy) the calendar time over which VC_{i-1, h_m} is computed will be large (small) and in general will vary throughout the day.

Compared with the conditional volatility equation (8) the HAR-BACD decomposes the total volatility into different volatility components VC_{i-1, h_m} for $m = 1, \dots, M$. According to HMH, each component corresponds to a group of market participants with a transaction horizon h_m . Short-run components are captured by small h_m , and medium to long-run components are captured by larger h_m . Of course the components are relative to the trading activity in the market. By summing up all the M terms, the dynamics of the total volatility process is just the aggregate of the different market volatility components. Since we have included volatility components with a variety of transaction time horizons, the HAR-BACD model takes into account any long-run volatility component naturally, thus we exclude the term ξ_i found in the GARCH model (8). The remaining interaction terms between the duration and volatility are the same.

¹²For evidence of strong temporal dependence in high frequency data, see for example Engle and Russell (2009).

3 Econometric Issues

3.1 The Likelihood

The information set up to observations s is $I_s = \{(x_i, r_i), i = 1, \dots, s\}$. Given the model parameter θ , denote the sample data $Y_T = \{y_i\}_{i=1}^T = \{(x_i, r_i)\}_{i=1}^T$, then the likelihood of the T observations is

$$p(Y_T|\theta) = \prod_{i=1}^T f(r_i, x_i|I_{i-1}, \theta). \quad (11)$$

The conditional distribution of each observation (r_i, x_i) can be written as

$$f(r_i, x_i|I_{i-1}, \theta) = f(x_i|I_{i-1}, \theta) f(r_i|I_{i-1}, x_i, \theta) \quad (12)$$

where $f(x_i|I_{i-1}, \theta)$ is the marginal density of the duration and $f(r_i|x_i, I_{i-1}, \theta)$ is the density of the return conditional on current duration. The marginal density of duration is calculated from (1) as

$$f(x_i|I_{i-1}, \theta) = \left(\frac{1}{\psi_i}\right) f_z\left(\frac{x_i}{\psi_i} | I_{i-1}, \theta\right), \quad (13)$$

where $f_z(\cdot)$ is the Burr density function for duration innovation z_i . The density of the return conditional on current duration can be calculated from equation (6) as

$$\begin{aligned} f(r_i|I_{i-1}, x_i, \theta) &= f_u(u_i|I_{i-1}, x_i, \theta) \\ &= f_u(r_i - \rho r_{i-1} - \phi u_{i-1}|I_{i-1}, x_i, \theta) \end{aligned} \quad (14)$$

where $f_u(\cdot)$ refers to the probability density function for u_i which is a Student-t with variance q_i and degree of freedom ν .

3.2 Bayesian Estimation

We estimate the models in the Bayesian framework. According to Bayes rule, the posterior distribution $p(\theta|I_T)$ is proportional to the product of the likelihood and the prior density, $p(\theta|I_T) \propto p(Y_T|\theta)p(\theta)$. Since neither the posterior nor the conditional posterior distributions have a known distribution, we cannot use the Gibbs sampling algorithm. Instead, we adopt the random walk Metropolis-Hastings (M-H) algorithm to simulate from the posterior. Good introductions to MCMC methods for Bayesian estimation can be found in Koop (2003) and Geweke (2005).

We iterate over the following steps to jointly sample all parameters at once.

1. Given the current value of the parameter vector θ , propose a new parameter vector according to

$$\theta' = \theta + V, \quad V \sim N(0, \tau\Sigma). \quad (15)$$

2. Accept θ' with probability

$$\eta = \min \{p(\theta'|Y_T)/p(\theta|Y_T), 1\}. \quad (16)$$

Otherwise retain θ as the current draw from the chain.

τ is a tuning constant set to achieve an acceptance frequency between 0.3-0.5. Σ is the sample covariance matrix estimate calculated from an initial first run using a single-move version of the above algorithm. The single-move samples from each conditional posterior density $p(\theta_j|\theta_{-j}, Y_T)$ until each component of the parameter vector is updated.

After dropping an initial set of burn-in draws from step 1–2 we collect the remaining N draws, $\{\theta^{(j)}\}_{j=1}^N$ which are then used in posterior inference. For instance, the posterior mean of θ can be consistently ($N \rightarrow \infty$) estimated as

$$\hat{\theta} = \frac{1}{N} \sum_{j=1}^N \theta^{(j)}. \quad (17)$$

3.3 Model Comparison

The Bayesian approach allows for the comparison and the ranking of nested and non-nested models by Bayes factor or posterior odds. The Bayes factor for model M_0 versus M_1 is defined as $BF \equiv p(Y_T|M_0)/p(Y_T|M_1)$, which is the ratio of marginal likelihoods and summarizes the evidence for model M_0 against M_1 . The marginal likelihood (ML) for model M_j , $j = 0, 1$ is defined as

$$p(Y_T|M_j) = \int p(Y_T|\theta, M_j)p(\theta|M_j)d\theta, \quad (18)$$

where $p(Y_T|\theta, M_j)$ is the likelihood and $p(\theta|M_j)$ the prior for model M_j . This is a measure of the success the model has in accounting for the data after the parameter uncertainty has been integrated out. Model comparison by Bayes factors penalizes highly parametrized models that do not deliver improved predictive content. For a discussion on the advantages of Bayes factors for model comparison see Koop and Potter (1999).

For the marginal likelihood we use the method of Gelfand and Dey (1994) adapted by Geweke (2005) (Section 8.2.4). This estimate is based on

$$\frac{1}{N} \sum_{i=1}^N \frac{g(\theta^{(i)})}{p(Y_T|\theta^{(i)}, M_j)p(\theta^{(i)}|M_j)} \rightarrow p(Y_T|M_j)^{-1} \quad \text{as } N \rightarrow \infty, \quad (19)$$

where $g(\theta^{(i)})$ is a truncated multivariate Normal. $\theta^{(i)}$ is a MCMC draw from the posterior. Note that the prior, likelihood and $g(\theta)$ must contain all integrating constants. Finally, to avoid underflow/overflow we use logarithms in this calculation.

3.4 Out-of-Sample Density Forecasts

The comparison of out-of-sample forecasting power is also very straightforward in Bayesian framework. As argued in Geweke and Whiteman (2006), the predictive likelihood (PL) evaluates the out-of-sample prediction of a model, making it the central quantity of interest for model comparison. Specifically, the predictive likelihood (Geweke (1995, 2005)) is defined for data y_s, \dots, y_t , $s < t$ and model M_j as

$$p(y_s, \dots, y_t|I_{s-1}, M_j) = \int p(y_s, \dots, y_t|\theta, I_{s-1}, M_j)p(\theta|I_{s-1}, M_j)d\theta \quad (20)$$

and it is the predictive density evaluated at the realized outcome y_s, \dots, y_t . The integration is performed with respect to the posterior distribution based on the information set I_{s-1} . Specially, if $s = 1$, this is the marginal likelihood we defined above. Since

$$p(y_s, \dots, y_t | I_{s-1}, M_j) = p(y_1, \dots, y_t | M_j) / p(y_1, \dots, y_s | M_j), \quad (21)$$

the log(PL) for the out-of-sample data can be calculated by taking the difference between the log(ML) for the full sample and the Log(ML) for the in-sample data.

3.5 Out-of-Sample Point Forecasts

To evaluate the accuracy of the conditional variance of returns for a model we compare them to high frequency squared returns r_i^2 . The out-of-sample performance over observations $s, s+1, \dots, t$, $s < t$, is measured by mean absolute error (MAE), and root mean squared error (RMSE) defined as

$$MAE = \frac{1}{t-s+1} \sum_{i=s}^t |\text{Var}(r_i | I_{i-1}, x_i) - r_i^2|, RMSE = \sqrt{\frac{1}{t-s+1} \sum_{i=s}^t (\text{Var}(r_i | I_{i-1}, x_i) - r_i^2)^2}. \quad (22)$$

$\text{Var}(r_i | I_{i-1}, x_i)$ is the 1-step ahead conditional variance of the return given current information from equation (8) or (9) for GARCH-ACD and HAR-ACD models respectively. The out-of-sample conditional variance is computed by averaging over the relevant quantities from the MCMC output based on the current information set. As in the last section, the model is re-estimated at each stage in the out-of-sample period to compute the variance forecast. r_i^2 is the squared return and serves as an estimate of ex-post variance.

4 Institutional Features and Data

According to World Federation of Exchanges¹³, at the end of August 2009, the New York Stock Exchange had a market value of 10,842 billion dollars with total share turnover of 12,158 billion. The London Stock Exchange had a market value 2,560 billion and turnover of 2,321 billion. The two stock exchanges in China, the Shanghai Stock Exchange and Shenzhen Stock Exchange had a total market value of 2,739 billion and turnover of 5,017 billion. The Chinese markets are comparable with developed markets.

There are some unique features of the Chinese stock market. First, there is a ‘‘T+1’’ rule, which means that if an investor buys a stock today, it cannot be sold until tomorrow. Second, short sales are forbidden.¹⁴ Third, there is a limit move rule, where the daily price change of an individual stock cannot exceed 10%. These rules discourage short-term transactions since a lot of intraday trading opportunities (day traders) are curtailed. As a consequence, the investment horizon tends to be longer, and the trading intensity can be expected to be lower in Chinese markets.

¹³<http://www.world-exchanges.org/statistics/>

¹⁴China began the trial margin trading and short selling scheme on March 31, 2010, and approved six brokerages as the first batch of pilot firms to take part in it.

There are two separate stock markets in China: Share A and Share B markets. The Share A market is open only to domestic investors, while Share B market is open to both domestic and foreign investors.

We select three stocks from the Chinese stock markets and the U.S. market respectively. Specifically, we classify Chinese stocks into 3 groups, and randomly choose one stock from each group: 1) The China Petroleum and Chemical Corporation (Sinopec) from the Share-A large market capitalization group; 2) Xinfu Pharmaceutical Co. Ltd (XFPC) from the Share-A small market capitalization group; 3) China Wanke Co. Ltd B share (WKB) from the Share-B group.¹⁵ In the U.S. market, we select 3 stocks from different industries. 1) IBM from the Technology sector; 2) Exxon Mobil Corporation (XOM) from the Basic Materials sector; 3) Pfizer Inc. (PFE) from the Health-care sector. They all are heavily traded stocks.

All Chinese stock market data are provided by the China Finance Online Company. We include all the normal trading days from March 1, 2006 to May 31, 2006. There are 60 valid days.¹⁶ The Chinese stock market opens 4 hours a day, which is from 9:30am to 11:30am, and 1:00pm to 3:00pm. We use the observations within this period. Since records in high frequency data usually contain many errors or redundant information, we filter them first.

To summarize, the filters we apply with corresponding summary statistics in Table 1 are:

- Filter 1 deletes error records.
- Filter 2 integrates all the transactions occurring at the same time into a single transaction with a volume weighted average price
- Filter 3 drops all observations with no price change.

Table 1 displays the reduction in data due to filtering. The error records are usually less than 1%. Transactions with the same time stamp are around 5% of the total observations. The average number of the valid observations per hour is 310 for Sinopec, 117 for XFPC and 89 for WKB. There are a lot of transactions with no price change, ranging from 48.47% (XFPC), 50.45% (WKB) to 55% (Sinopec) of the total observations. The average number of the price change observations per hour is 128 for Sinopec, 57 for XFPC and 42 for WKB. For the univariate ACD models, we use all the sample. When exploring our joint model, we reserve the first approximate 1 month of data as startup values as the maximum lag length that we consider in the conditional variance is 1 month. We then divide the remaining observations into the in-sample period and the out-of-sample period, with the out-of-sample data extending roughly 3 weeks. These details are listed in the bottom panel of Table 1.

The U.S. transaction data are obtained from the Trade and Quotes (TAQ) database. We choose the same sample period as in Chinese stock market, which is from March 1, 2006 to May 31, 2006 (64 days). Keeping only those records within normal trading hours (9:30 am

¹⁵Sinopec (Code: 600028) is one of the major petroleum companies in China. It is a component of the local major stock indexes. XFPC (Code: 002019) is a manufacturer and supplier of Vitamin B5, which was established in November 1994 and listed in Shenzhen Stock Exchange as a high-tech enterprise in July 2004. China Wanke Company Limited (Code: 200002) is the largest residential real estate developer in China. Its stocks are traded in both Share-A market and Share-B market. Here we choose its B Shares. The three stocks are all heavily-traded stocks in their markets.

¹⁶There are 61 normal trading days in this period. Because of the shareholder meeting, each stock trading is closed for 1 day: the Sinopec on May 24, XFPC on May 22 and WKB on May 30.

to 4:00 pm), we filter the data in a similar way.¹⁷ The filter result is reported in the last 3 columns of Table 1. Compared with the Chinese market data, U.S. stocks have much heavier trading intensity. They have very high percentage of observations happening at the same time (30.63% for IBM, 50.26% for XOM and 58.54% for PFE). There are around 500 – 800 transactions each hour, compared to the most active Chinese stock, Sinopec, which has around 128 observations each hour. The statistics of the duration for the original and the thinned process, the absolute return and the return for all the six stocks are summarized in Table 2. Average duration are much larger for the Chinese stocks.

The autocorrelation function of returns and absolute returns for Sinopec and IBM are presented in Figure 1.¹⁸ In those diagrams, the two dotted horizontal lines are the Bartlett standard error bands. Most of the autocorrelation function for the returns are within the confidence bands except for the first lag. In contrast, most autocorrelations of the absolute returns exceed the bounds even at 500 lags. The standard GARCH functional form will have problems capturing this, while our HAR-BACD model is designed to deal with this.

4.1 Diurnal Adjustments

Intraday data typically contain a very strong diurnal pattern. The diurnal adjustments for the durations and the returns follow Engle and Russell (1998). We regress the durations on the time of day using a cubic spline specification, and then we take the ratios of the durations and their fitted values to obtain *diurnally adjusted* durations. The internal knots are set on each hour. Since the transaction frequency drops quickly at the end of the day, we add an extra knot in the last half hour (knots at 10:00, 11:00, 11:30 (1:00), 2:00 and 2:30 for Chinese stocks, and at 10:00, 11:00, 12:00, 1:00, 2:00, 3:00 and 3:30 for U.S. stocks). Recall that the Chinese markets close from 11:30 to 1:00pm. For observed returns \tilde{r}_i , we first divide them by the square root of actual durations as $r_i = \tilde{r}_i / \sqrt{x_i}$, and then regress the absolute values of r_i on the time of day in the same way as durations. *Diurnally adjusted* returns are obtained by taking the ratios of r_i and their corresponding fitted values.

The daily spline estimates for duration and the absolute return are displayed in Figure 2.¹⁹ The top panel is for Sinopec data and bottom panel for IBM. Both series have very similar daily patterns. The daily spline for durations has an inverse “U” shape. Durations are very short after the open and prior to the close of the market, indicating more transactions during these periods. The durations are much longer around the lunch time in the U.S. market. In Chinese market the transaction durations are still very long before and after the lunch break. The diurnal pattern for the absolute returns is also quite similar among stocks. We can see a peak at the open, flat during most of the day until it increases again before the close. These daily patterns are consistent with daily information flow. Investors adjust their positions

¹⁷In U.S. market, the errors are removed as follows. Only records with correction indicator = 0 or 1, and the sale condition is blank or “E” are retained. According to TAQ documentation, correction indicator =0 or 1 signal those trades as good trades where 0 means “Regular trade that was not corrected, changed, or signified as cancel or error”, and 1 means “trade which was later corrected”. The blank sell condition means a trade made without any stated conditions. “E” stands for “the high-speed electronic connection for immediate automatic execution”.

¹⁸The autocorrelation functions for XFPC and WKB are not reported but display a similar pattern.

¹⁹We only report Sinopec and IBM as representative stock. Other stocks are very similar.

intensively at the beginning of the day to incorporate new information from the overnight period. However, for the Chinese data, there is a small peak around mid-day because price volatility jumps as the market is re-opened after the lunch closure.

5 Empirical Results

For all posterior simulations we use a burn-in sample of 10,000. The number of draws for the first single-move sampler is 10,000 and for the joint block sampling step is 40,000. These latter draws are used for all posterior inference in this paper. We investigate running the chain from different starting values and compute convergence diagnostics such as Geweke (1992). The results show that our posterior draws mix well and the chain converges quickly.

The priors for the following parameters are independent normal $N(0, 100)$: ρ and ϕ in the ARMA equation, $\omega_g, \alpha_g, \beta_g, \beta_0$, all β_m and $\gamma_i, i = 1, 2, 3$, in the conditional variance equation. To guarantee the non-negativity of the duration, we set priors for the ACD parameters: ω_a, α_a and β_a to be truncated $N(0, 100)$ with positive supports. The stationary condition $\alpha_a + \beta_a < 1$ is imposed. All priors are very uninformative. When the degree of freedom ν is larger than 30, the t -distribution is close to the normal, therefore we select gamma prior $\nu \sim \text{Gamma}(8, 0.5)$, that favors fat-tails with the restriction $\nu > 2$ to ensure the variance exists. This puts most weight on the region $(2, 40)$. The priors for the Burr distribution parameters κ and ϖ^2 are set to be truncated $N(0, 100)$ with positive supports and the restriction $\kappa > \varpi^2$.

5.1 ACD Model for Full Sample

Table 3 reports estimation results for all the stocks using Burr-ACD model for the full sample of data which includes durations associated with no price change (Filter 2 data). This provides a check on our bivariate model estimates which are based on data that drops zero price change observations (filter 3 data). The posterior means and the standard deviations of the coefficients are presented. All the estimates are in the 95% posterior density intervals that exclude 0. Both China and U.S. markets have a very high persistence (high β). The significance of both the κ and ω^2 parameters suggests the validity of the Burr distribution. The parameter estimates and implied dynamics of duration are in line with the results for the bivariate model which is discussed below.

5.2 Which Components are Important?

To study the joint model of return and duration, we now work on the thinned process with price change (Filter 3 data). The conditional variance of returns in (9) requires the selection of the number of components and the number of transactions that enter into a component.²⁰

²⁰According to the HMM, there are different market components, but there is no specific rule on how to determine either the optimal number of the components M or the time horizon associated with each component. For example, Muller et al. (1997) study the half-hourly time series in ϑ -time, which is transformed data in a time scale in where no intraday seasonalities exist. They select $M = 7$ market components in ϑ -time. Corsi (2009) uses 3 components, daily, weekly and monthly volatility to forecast daily realized volatility.

In principle we could include M and h_m in our main estimation procedure, however, due to the large number of observations this would increase the computation costs substantially. Therefore, we select optimal values of these parameters based on a first step of estimation using Bayes factors.

Although the components are associated with transaction time an approximate correspondence with calendar time can nevertheless be established for each stock depending on the trading intensity. The candidate volatility components are listed in Table 4. We classify all the volatility components into 5 groups according to their approximate time horizons: instantaneous components (with the time horizon $H < 1$ minute), short-term components ($1 \text{ minute} \leq H \leq 10 \text{ minutes}$), medium-term components ($10 \text{ minutes} < H \leq 1 \text{ hour}$), long-term components ($1 \text{ hour} < H \leq 1 \text{ day}$), and very-long term components ($H > 1 \text{ day}$). The number of ticks we use to construct each component is also listed. For example, for Sinopec stock, the average duration for 1 tick is around 30 seconds, so 30 seconds is its volatility component for the minimum time horizon.

Our method begins with a model with no components ($M=0$). We set this model as a benchmark and record its marginal likelihood (ML). We add the first component and compute the ML. The ML from new models are directly comparable by Bayes factors as discussed in Section 3.3. If the component improves the ML (log-Bayes Factor is greater than 0) we have an improvement in the specification. However, if the ML from the benchmark is larger (log-Bayes Factor is less than 0), the method stops, assuming the benchmark ML is larger for the next two specifications with $M+1$ and $M+2$ components. Otherwise, we choose the new model with the larger ML and set it as the new benchmark. The evaluation process is repeated until no remaining component terms can improve the ML. This is done separately for each stock. The selected terms in the optimal model are fixed and used throughout the remainder of the paper.

The chosen volatility components are listed in Table 4 with **. The Chinese stocks display a similar structure. The volatility dynamics are best described by four components. Among them, the volatility of the previous trade is the most important, as all the components associated with the previous tick are selected. Short-term components are also important, as 2, 3 and 10 ticks are selected. Middle-term components have some presence. At least one component is from this group. All the long-term and very-long term components are found to be of no importance. The coefficients on terms which have time horizons more than 1 hour are very close to 0, and including them in the conditional variance equation does not improve the ML. This suggests that when considering the intraday volatility behavior, the information beyond 1 hour has little to no effect on current price change.

The best models in the U.S. market have four or five components. Similar to the Chinese stocks, the short-term and middle-term components are important, as all three stocks have at least one component from the short and the middle term groups. The long-term components are negligible. However, as U.S. stocks have higher trading frequencies, the instantaneous (less than 1 minute) components tend to affect volatility dynamics more. The volatility of the previous trade is included in the best models and IBM and XOM have an extra component associated with 30 seconds.

5.3 Model Comparison

Table 5 reports the model comparison for the specifications GARCH-BACD, HAR-EACD and HAR-BACD. The HAR-EACD model combines the new conditional variance with the less flexible exponential distribution for durations. Panel A of the table reports the log-Bayes factors ($\text{Log}(\text{BF})$) using the in-sample data listed in Table 1.²¹ Panel B presents the results for log-predictive Bayes factors ($\text{log}(\text{PBF})$) calculated from the out-of-sample data listed in the bottom of Table 1. Panel C and Panel D report the out-of-sample root mean squared error (RMSE) and mean absolute error (MAE) as in equation (22).

Looking at the first log-Bayes factors calculated from the marginal likelihood, we see a huge improvement on model fit in moving from the exponential distribution to the Burr distribution for duration innovations. For instance, for Sinopec the HAR-BACD model is $\exp(3758)$ times better at describing the data compared to the HAR-EACD specification. This evidence is true for all stocks. The second log-Bayes factor compares the GARCH functional form with the new variance specification. Here again the evidence is very strong in favor of the component model for the conditional variance. For instance, the log evidence for the HAR-BACD is 80 for Sinopec, 48 for XFPC, 77 for WKB while the evidence for the model is greater from the U.S. stocks. This means we have decisive support for our new specification. Taking into account all the transaction data, the new conditional variance specification provides large improvements in fitting the data. For example, the HAR is $\exp(80)$ times better for Sinopec and $\exp(48)$ times better for XFPC in accounting for the data as compared to the GARCH functional form.

The log-predictive likelihood measures how the specifications perform in the out-of-sample data. The results continue to rank the models in exactly the same way. For example, the log-predictive Bayes factor of our new specification compared to the GARCH model is 44 for Sinopec, which means when we forecast using out-of-sample data, the new specification fits the data $\exp(44)$ times better than the GARCH model. The main difference between the marginal likelihood and the predictive likelihood is that the latter minimizes any impact of the prior distributions. Since the cumulative log-predictive likelihood measures the quality of out-of-sample density forecasts with parameter uncertainty integrated out, the HAR-BACD provides the best forecast performance.

Panels C and D report the accuracy of out-of-sample point forecasts of the conditional variance. We see marginal improvements in RMSE and MAE. In all cases except one, moving from the GARCH model to the HAR model improves variance forecasts. However, the gains as measured against squared returns are small.²²

By all measures, over different portions of the data sample, for Chinese and U.S. stocks, the new HAR-BACD model is very competitive.

²¹We have also compared models with a normal distribution for return innovations and found it to be dominated by the Student-t distribution used in this paper.

²²A drawback of using daily squared returns as a proxy for ex post variance is that it is noisy which can make it difficult to discriminate among models (Andersen and Bollerslev 1998b). It is not clear how important this noise is for high-frequency intraday squared returns r_t^2 .

5.4 Parameter Estimates

The estimation results for our best joint model of durations and returns are presented in Table 6. As discussed in Section 5.2, the conditional volatility equation of the HAR-BACD model consists of 4 volatility components for Chinese stocks, and 5 or 4 components for U.S. stocks. We report the posterior means of the coefficients and their associated standard deviations. The cells with stars have 0.95 posterior density intervals that exclude 0.

Panel A of this table reports the coefficients for the return equation. We see evidence of market microstructure dynamics and the typical strong persistence in the conditional variance of returns as well as fat-tails in the innovation distribution. All stocks have a strong and negative AR coefficient ρ , which is consistent with the presence of market microstructure dynamics. The degree of freedom parameter in the t-distribution ν is estimated around 7 for all stocks except for Sinopec and XOM. The smaller ν , the more fat-tailed is the return distribution relative to a normal.

Panel B lists the parameter estimates for the duration process. The duration displays the GARCH-type property of clustering. Short (long) trading durations tend to follow short (long) durations. All the coefficients are significant. The coefficient β_a is large and around 0.9, while the effect of the last duration α_a is small. All the estimated parameters of the Burr distribution have $\kappa > 1$ and $\varpi^2 > 0$. This shows that the Burr distribution is much more appropriate for duration innovations than the exponential distribution which restricts the hazard function to be constant. The implied conditional hazard functions for all the stock durations are unimodal with an inverse “U” shape, which implies that the probability of a transaction increases first with elapsing time and then decreases as time goes by. The conditional hazard functions for Sinopec and IBM are displayed in Figure 3, with other stocks have similar patterns.

Panel C describes the effects of the duration on the conditional variance. All the coefficients are different from 0, indicating that contemporaneous duration terms are very important. The signs of the coefficients γ_1 and γ_3 are the same across all the stocks, suggesting the similar interactions between trading times and price movements. Specifically, the coefficient for the inverse of the duration $1/x_i$ is positive, which means a longer contemporaneous duration is associated with a lower volatility. This is supportive for Easley and O’Hara (1992) in which long durations are interpreted as having no information so that volatility decreases. The coefficients for the duration surprise x_i/ψ_i are decreasing with trading frequency. They are positive for IBM and all Chinese stocks, but negative for XOM and PFE which are the most active stocks. This means when the actual duration is large relative to the expected duration, the transaction tends to be associated with a larger price change for lightly traded stocks, but decreasing for heavily traded stocks. For the heavily traded stocks, more investors may be noise traders²³ and they may trade for reasons other than the new information, e.g., because of liquidity or hedging considerations. As a consequence, if there is no news, transactions based on information tend to halt but the noise traders will continue trading. Short durations and duration surprises, which have information content, tend to change the return volatility by a smaller amount for heavily traded stocks.

²³According to Dow and Gorton (1997), professional traders and money managers are the main noise trader. And they usually allocate more assets in those heavily traded stocks because those markets have higher liquidity and market capitalization.

The estimated coefficients for the conditional variance are listed in the last panel. All the volatility components are positive and the majority of them have 0.95 density intervals that exclude 0. However, the volatility components with different time horizons have different effects on current volatility. Recent volatility, as measured by u_{i-h}^2 , has a larger impact on the conditional variance. The average derivative of the conditional variance q_i with respect to u_{i-h}^2 is displayed in Figure 4 for Chinese stocks and U.S. stocks respectively. h denotes the lagged transactions. The curve for Chinese stocks is above the curve for U.S. stocks for majority of the time. This means in transaction time the Chinese stocks show a slower decay in the effect of past return innovations as compared to U.S. stocks.

The differences in volatility dynamics suggest investors in the Chinese stock market focus on longer-term information and more investors behave this way than in the U.S.. There are several explanations for why investors have longer horizons in China. There is the “T+1” rule, no short sales and the “limit move” that was discussed in Section 4. These institutional rules largely discourage short-term transactions. As a consequence, the investment horizon tends to be longer, and the trading intensity is less in China.

6 Summary

In this paper, we study the intraday dynamics of three stocks from the Chinese stock market, and include three stocks from the U.S. market for comparison. We propose a new joint model of volatility of returns and the duration between trades. Specifically, we apply a component approach to the conditional variance and construct a HAR-type model of volatility at the highest available frequency. Components are based on transaction time and not calendar time. Applying our model to the Sinopec, XFPC, WKB, IBM, XOM and PFE tick-by-tick data, we find a dramatic improvement over the traditional GARCH-ACD model. The total volatility is decomposed into 4 or 5 volatility components. The common components are the stock volatility from the last transaction and the average volatility from several minutes up to one hour. Information past 1 hour does not help in the modeling or the forecasting of the high frequency data.

The new model provides strong improvements in density forecasts for duration and returns and only modest gains for points forecasts of the variance of returns.

The stocks have similar diurnal patterns and the same market microstructure implications across the markets. Compared to U.S. stocks, the Chinese stocks display much lower trading activity. For lightly traded stocks, duration carries more information content and is a more important determinant of volatility dynamics. The long-term volatility components have a larger effect than the short-term components. We attribute this to a longer investment horizon in China, which may be the consequence of specific trading rules.

References

- [1] Ait-Sahalia, Y. and P. A. Mykland (2003). “The effects of random and discrete sampling when estimating continuous-time diffusions”, *Econometrica* 71, 483-549.
- [2] Andersen T. G. and T. Bollerslev (1998a): “DM-Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer Run Dependencies,” *Journal of Finance*, 53, 219-65.
- [3] Andersen T. G. and T. Bollerslev (1998b): “Answering the Skeptics: Yes, Standard Volatility Models do Provide Accurate Forecasts,” *International Economic Review*, 39, 885-905.
- [4] Andersen, T.G., T. Bollerslev, and F.X. Diebold (2007): “Roughing It Up: Including Jump Components in the Measurement, Modeling and Forecasting of Return Volatility,” *Review of Economics and Statistics*, 89(4), 701-20.
- [5] Andersen, T., T. Bollerslev, F. X. Diebold (2009), “Parametric and Nonparametric Volatility Measurement”, in *Handbook of Financial Econometrics*, Vol 1, Eds Ait-Sahalia and Hansen, Elsevier.
- [6] Andersen T. G., T. Bollerslev, F. X. Diebold and H. Ebens (2001): “The distribution of realized stock return volatility,” *Journal of Financial Economics*, 61, 43-76.
- [7] Andersen T. G., T. Bollerslev, F. X. Diebold and P. Labys (2001): “The distribution of exchange rate volatility,” *Journal of the American Statistical Association*, 96, 42-55.
- [8] Andersen, T. G., T. Bollerslev and X. Huang (2006): “A Semiparametric Framework for Modelling and Forecasting Jumps and Volatility in Speculative Prices.” Working paper, Duke University.
- [9] Barndorff-Nielsen O. E. and N. Shephard (2002a): “Econometric analysis of realized volatility and its use in estimating stochastic volatility models,” *Journal of the Royal Statistical Society, Series B*, 64, 253-280.
- [10] Barndorff-Nielsen O. E. and N. Shephard (2002b): “Estimating quadratic variation using realised variance,” *Journal of Applied Econometrics*, 17, 457-77.
- [11] Bauwens, L. and P. Giot (2000): “The logarithmic ACD model: an application to the bid-ask quote process of three NYSE stocks”, *Annales d’Economie et Statistique*, 60, 117-149.
- [12] Bauwens, L., P. Giot, J. Grammig J and D. Veredas (2004): “A comparison of financial duration models via density forecasts,” *International Journal of Forecasting*, 20, 589-609.
- [13] Bauwens, L. and D. Veredas (2004): “The Stochastic Conditional Duration Model: A Latent Factor Model for the Analysis of Financial Durations”, *Journal of Econometrics*, 119(2), 381-412.

- [14] Bollerslev, T., U. Kretschmer, C. Pigorsch, and G. E. Tauchen (2009), “A Discrete-Time Model for Daily S&P 500 Returns and Realized Variations: Jumps and Leverage Effects”, *Journal of Econometrics*, 50(2), 151-66.
- [15] Cai, B., C. Cai and K. Keasey (2006): “Which Trades Move Prices in Emerging Markets?: Evidence from China’s Stock Market” *Pacific-Basin Finance Journal* 14(5), 453-66.
- [16] Calvet L. E. and A. J. Fisher (2008) *Multifractal Volatility: Theory, Forecasting and Pricing*, Academic Press, New York.
- [17] Campbell, J.Y., Lo, A.W. and MacKinlay, A.C. (1997): *The Econometrics of Financial Markets*, Princeton University Press.
- [18] Cheung, Y., R. Ho, P. Pope, P. Draper (1994): “Intraday stock return volatility: The Hong Kong evidence” *Pacific-Basin Finance Journal*, 2(2-3), 261-76.
- [19] Chong, T. and Q. Su (2006): “On the Comovement of A and H Shares” *Chinese Economy* 39(5), 68-86.
- [20] Corsi, F. (2009): “A simple long memory model of realized volatility,” *Journal of Financial Econometrics*, 7(2), 174-196.
- [21] Corsi, F., U. Kretschmer, S. Mittnik and C. Pigorsch (2005): “The Volatility of Realized Volatility”, *Econometric Reviews*, 27, 46-78.
- [22] Dow, J., and G. Gorton (1997): “Noise Trading, Delegated Portfolio Management, and Economic Welfare,” *Journal of Political Economy*, 105, 1024-50.
- [23] Dufour, A. and R. F. Engle (2000): “Time and the Price Impact of a Trade” *Journal of Finance*, 55, 2467-2498.
- [24] Easley, D. and M. O’Hara (1992): “Time and the Process of Security Price Adjustment”, *The Journal of Finance*, 19, 69-90.
- [25] Engle, R. F. (2000): “The Econometrics of Ultra High Frequency Data”, *Econometrica*, 68, 1-22.
- [26] Engle, R. F., and G. G. J. Lee (1999): ” A Long-Run and Short-Run Component Model of Stock Return Volatility,” in *Cointegration, Causality, and Forecasting*, ed. by Engle, and White. Oxford University Press.
- [27] Engle, R. F. and J. R. Russell (1998): “Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data”, *Econometrica*, 66, 1127-62.
- [28] Engle, R.F. and J. R. Russell (2009): “Analysis of High Frequency Data”, *Handbook of Financial Econometrics*. Ait Sahalia, Y., Hansen, L.P. (Eds.).
- [29] Engle, R. F. and Z. Sun (2005): “Forecasting Volatility Using Tick by Tick Data”, working paper, Department of Finance, New York University.

- [30] Forsberg, L., and E. Ghysels. (2007). “Why do Absolute Returns Predict Volatility so well”, *Journal of Financial Econometrics*, 5(1), 31-67.
- [31] Gelfand A. E. and D. Dey (1994): “Bayesian model choice: asymptotic and exact calculations,” *Journal Royal Statistical Society*, B, 56, 501-14.
- [32] Geweke, J. (1992): “Evaluating the accuracy of sampling-based approaches to calculating posterior moments (with discussion)”. In *Bayesian Statistics 4* [edited by J.M. Bernardo, J.O. Berger, A.P. Dawid and A.F.M. Smith], Oxford University Press, Oxford, 169-93.
- [33] Geweke J.(1995): “Bayesian comparison of Econometric Models,” Working paper, Research Department, Federal Reserve Bank of Minneapolis.
- [34] Geweke J. (2005): *Contemporary Bayesian Econometrics and Statistics*, John Wiley & Sons Ltd.
- [35] Geweke J. and C. Whiteman (2006) “Bayesian Forecasting”, *Handbook of Economic Forecasting (I)*, Edited by Elliott, Granger and Timmermann, North-Holland.
- [36] Ghysels, E., C. Gouriou and J. Jasiak (2004): “Stochastic volatility duration models,” *Journal of Econometrics*, Elsevier, vol. 119(2), 413-33.
- [37] Ghysels, E. and J. Jasiak (1998): “GARCH for Irregularly Spaced Financial Data: The ACD-GARCH Model,” *Studies in Nonlinear Dynamics and Econometrics*, 2, 133-49.
- [38] Grammig, J. and K.-O. Maurer (2000): “Non-monotonic hazard functions and the autoregressive conditional duration model,” *Econometrics Journal*, 3(1), 16-38.
- [39] Grammig, J., and M. Wellner (2002): “Modeling the Interdependence of Volatility and Inter-Transaction Duration Processes.” *Journal of Econometrics*, 106, 369-400.
- [40] Hamilton, J. D. and O. Jorda (2002): “A Model for the Federal Funds Rate Target,” *Journal of Political Economy*, 110, 1135-67.
- [41] Koop G. (2003): *Bayesian Econometrics*, John Wiley & Sons Ltd.
- [42] Koop G and Potter S. (1999): “Bayes factors and nonlinearity: Evidence from economic time series”, *Journal of Econometrics* 88, 251-282.
- [43] Lee, B., O. Rui, W. Wu (2008): “Market Segmentation and Stock Prices Discount in the Chinese Stock Market: Revisiting B-Share Discounts in the Chinese Stock Market,” *Asia-Pacific Journal of Financial Studies*, 37(1), 1-40.
- [44] Lunde, A. (1999): “A Generalized Gamma Autoregressive Conditional Duration Model,” Discussion paper, Aalborg University.
- [45] Maheu J. M. (2005): “Can GARCH Models Capture Long-Range Dependence?”, *Studies in Nonlinear Dynamics & Econometrics*, 9(4), Article 1.

- [46] Maheu J. M. and T. H. McCurdy (2011): "Do High-frequency Measures of Volatility Improve Forecasts of Return Distributions? ", *Journal of Econometrics*, 160(1), 69-76.
- [47] Muller, U., M. Dacorogna, R. Dave, R. Olsen, O. Pictet, and J. von Weizsacker (1997): "Volatilities of Different Time Resolutions — Analyzing the Dynamics of Market Components ", *Journal of Empirical Finance* 4(2-3), 213-39.
- [48] Oomen R. (2006): "Properties of Realized Variance Under Alternative Sampling Schemes," *Journal of Business and Economic Statistics*, 24(2), 219-237.
- [49] Tian, G, M. Guo (2007): "Interday and intraday volatility: Additional evidence from the Shanghai Stock Exchange," *Review of Quantitative Finance and Accounting* 28(3), 287-306.
- [50] Zhang, M., J. R. Russell and R. S. Tsay (2001): "A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data," *Journal of Econometrics*, 104(1), 179-207.
- [51] Zhou, H., J. Q. Zhu (2009): "An empirical examination of jump risk in asset pricing and volatility forecasting in China's equity and bond markets," *Pacific Basin Finance Journal*, Forthcoming.

Table 1: Data Filtering

	China			U.S.		
	Sinopec	XFPC	WKB	IBM	XOM	PFE
Total valid days	60	60	60	64	64	64
Total Obs.	79,229	29,925	22530	578,925	1,345,812	1,466,592
Filter 1	79,162	29,805	22410	576,765	1,341,833	1,462,852
	(99.92%)	(99.60%)	(99.47%)	(99.63%)	(99.70%)	(99.75%)
Filter 2	74,321	28,068	21339	399,453	669,419	608,026
	(93.81%)	(93.79%)	(94.71%)	(69.00%)	(49.74%)	(41.46%)
Obs./Hour	310	117	89	960	1609	1462
Filter 3	30,752	13,624	10032	209,474	333,932	269,670
	(38.81%)	(45.53%)	(44.53%)	(36.18%)	(24.81%)	(18.39%)
Obs./Hour	128	57	42	504	803	648
Range	54-193	13-142	9-97	371-770	633-1100	548-877
Startup Values	11,000	5,500	3,500	65,000	110,000	90,000
In-Sample	11,752	4,124	3,532	94,474	143,932	109,670
Out-of-Sample	8,000	4,000	3,000	50,000	80,000	70,000

This table reports the filtering process for transactions data from the 6 stocks: China Petroleum and Chemical Corporation (Sinopec), Xinfu Pharmaceutical Co. Ltd (XFPC), China Wanke Co. Ltd (WKB), IBM, Exxon Mobil Corporation (XOM) and Pfizer Inc. (PFE) in normal trading days from March 1, 2006 to May 31, 2006. The Chinese stock data is from 9:30am-11:30am and 1:00pm-3:00pm and U.S. data is between 9:30am and 4:00pm. The "Total Obs." row reports the total number of the observation in the data set. The "Filter 1" row reports the remaining number of observations by deleting the error records. The "Filter 2" row reports the remaining number after integrating all the transactions happening at the same time into a single transaction. The "Filter 3" row reports the total number of the observations after the records with no price changes are discarded. The "Obs./Hour" rows are the average number of observations per hour. The final panel reports how we divide total sample into in-sample and out-of-sample data.

Table 2: Summary statistics

		China			U.S.		
		Sinopec	XFPC	WKB	IBM	XOM	PFE
Duration	Mean	11.346	29.062	39.210	3.747	2.234	2.460
	Stdev	7.042	32.969	50.689	4.413	2.132	2.405
Thinned	Mean	11.486	29.748	41.560	3.911	2.270	2.578
Duration	Stdev	7.061	33.970	54.129	4.576	2.173	2.515
Return	Mean	2.33e-6	-1.01e-6	3.82e-6	-1.97e-6	-1.23e-6	-1.81e-6
	Stdev	6.97e-4	6.50e-4	7.49e-4	2.12e-4	1.69e-4	3.11e-4
Absolute Return	Mean	6.06e-4	4.58e-4	5.03e-4	1.53e-4	1.20e-4	2.46e-4
	Stdev	3.44e-4	4.63e-4	5.55e-4	1.46e-4	1.18e-4	1.91e-4

This table reports the statistics for the stocks in normal trading days from March 1, 2006 to May 31, 2006. The Chinese stock data is from 9:30am-11:30am and 1:00pm-3:00pm and IBM data is between 9:30am and 4:00pm.

Table 3: Estimation Results for ACD Model

		China			U.S.		
		Sinopec	XFPC	WKB	IBM	XOM	PFE
ω	0.0034	0.0176	0.0259	0.0114	0.0104	0.0507	
	(0.0004)	(0.0013)	(0.0021)	(0.0006)	(0.0006)	(0.0019)	
α	0.0152	0.0629	0.0773	0.0325	0.0286	0.0665	
	(0.0008)	(0.0029)	(0.0039)	(0.0008)	(0.0008)	(0.0013)	
β	0.9813	0.9195	0.8973	0.9562	0.9611	0.8829	
	(0.0010)	(0.0038)	(0.0053)	(0.0014)	(0.0014)	(0.0030)	
κ	2.4745	2.2470	2.0268	9.8942	6.4726	13.4348	
	(0.0125)	(0.0222)	(0.0224)	(0.7320)	(0.9004)	(1.4842)	
ϖ^2	0.4046	0.8474	0.8112	9.6993	4.2918	9.0024	
	(0.0089)	(0.0233)	(0.0258)	(0.7360)	(0.3993)	(1.3916)	

This table reports the posterior means and the standard deviations of the coefficients for all the stocks using Burr-ACD model for the full sample.

Table 4: Candidate Volatility Components

Components	Time	Number of ticks					
		China			U.S.		
		Sinopec	XFPC	WKB	IBM	XOM	FPE
Instantaneous	5 seconds				1**	1**	1**
	30 seconds	1**			5**	5**	5
S-T (1M $\leq H < 10M$)	1 minute	2**	1**		9**	13**	11
	2 minutes	4	2**	1**	15	25	22**
	5 minutes	10**	4	3**	40	70**	55
M-T (10M $\leq H \leq 1H$)	10 minutes	20	8**	6**	85**	135	110**
	15 minutes	30	13	9	125	200	160
	30 minutes	60	25**	20**	250	400	325**
	1 hour	125**	50	40	500**	800**	650
L-T (1H $< H \leq 1D$)	1.5 hours	180	75	60	750	1200	1000
	0.5 day	250	125	80	1750	2600	2100
	1 day	500	250	160	3200	5200	4200
V-L (1D $< H$)	1 week	2500	1250	800	16000	26000	21000
	1 month	11000	5500	3500	65000	111000	90000

This table reports the candidate volatility components (the possible HAR terms) that we consider. We classify all the volatility components into 5 groups: instantaneous components (Instantaneous: with the time horizon $H < 1$ minute), short-term components (S-T: 1 minute $\leq H \leq 10$ minutes), medium-term components (M-T: 10 minutes $< H \leq 1$ hour), long-term components (L-T: 1 hour $< H \leq 1$ day), and very-long term components (V-L: $H > 1$ day). Their approximate calendar times H are reported in the brackets. The “numbers of ticks” column reports the exact number of transactions that are used to construct the component in the conditional variance. Those cells with two stars are the optimally chosen components in the conditional variance.

Table 5: Comparison of the Models

	China			U.S.		
	Sinopec	XFPC	WKB	IBM	XOM	FPE
A: Log(BF)						
HAR-BACD vs HAR-EACD	3758	645	413	12026	12935	7847
HAR-BACD vs GARCH-BACD	80	48	77	198	552	279
B: Log(PBF)						
HAR-BACD vs HAR-EACD	2701	835	330	7436	10217	6175
HAR-BACD vs GARCH-BACD	44	71	48	121	672	331
C: RMSE						
GARCH-BACD	4.1133	7.4153	4.6521	21.9379	69.1231	5.8145
HAR-EACD	4.0842	7.4072	4.6410	21.8546	69.0271	5.7377
HAR-BACD	4.0842	7.4094	4.6418	21.8548	69.0271	5.7378
D: MAE						
GARCH-BACD	0.7463	2.5311	1.7713	2.5748	2.3442	1.6871
HAR-EACD	0.7459	2.5309	1.7714	2.5734	2.3424	1.6868
HAR-BACD	0.7438	2.5292	1.7731	2.5732	2.3420	1.6863

The table reports model comparison results for the GARCH-BACD and the HAR-EACD which have exponential duration innovations and the HAR-BACD model. Panel A reports the log-Bayes factors ($\log(\text{BF})$) for the HAR-BACD model against the other specifications using the in-sample data. Panel B reports the log-predictive Bayes factors ($\text{Log}(\text{PBF})$) for the out-of-sample data period. Panel C and Panel D report the out-of-sample root mean squared error (RMSE) and mean absolute error (MAE) as in equation (22). The sample periods are reported in Table 1.

Table 6: Estimation Results

		China			U.S.							
		Sinopac		WKB		IBM		XOM		FPE		
		Mean	Stdev	Mean	Stdev	Mean	Stdev	Mean	Stdev	Mean	Stdev	
A: Return Equation												
ρ	-0.5361*	(0.0138)	-0.3856*	(0.0272)	-0.3270*	(0.0311)	-0.2102*	(0.0114)	-0.3211*	(0.0078)	-0.2779*	(0.0081)
ϕ	0.0503*	(0.0167)	0.1257*	(0.0299)	0.0371	(0.0339)	-0.0068*	(0.0118)	0.1223*	(0.0083)	-0.0124	(0.0086)
ν	19.2073*	(2.8659)	7.9890*	(0.5137)	6.3176*	(0.4070)	7.0947*	(0.1210)	12.0941*	(0.2174)	8.2442*	(0.1822)
B: ACD Equation												
ω_a	0.0058*	(0.0016)	0.0218*	(0.0032)	0.0309*	(0.0050)	0.2687*	(0.0328)	0.0838*	(0.0158)	0.2720*	(0.0844)
α_a	0.0154*	(0.0021)	0.0627*	(0.0055)	0.0722*	(0.0068)	0.0849*	(0.0086)	0.0006*	(0.0002)	0.0066*	(0.0017)
β_a	0.9784*	(0.0033)	0.9094*	(0.0084)	0.8959*	(0.0108)	0.9142*	(0.0087)	0.9362*	(0.0120)	0.8641*	(0.0234)
κ	2.4986*	(0.0309)	2.4287*	(0.0445)	2.1209*	(0.0444)	4.6750*	(0.0431)	5.5241*	(0.3892)	5.8064*	(0.1543)
ϖ^2	0.3548*	(0.0203)	0.8579*	(0.0423)	0.9141*	(0.0501)	4.3615*	(0.0408)	4.7688*	(0.1973)	3.6702*	(0.1008)
C: Duration Effects												
γ_1	0.6816*	(0.0308)	12.7942*	(0.7362)	12.9265*	(0.7010)	0.6647*	(0.0072)	0.8559*	(0.0076)	0.8520*	(0.0071)
γ_2	0.0976*	(0.0220)	1.7614*	(0.4748)	1.0539*	(0.2758)	0.0146*	(0.0030)	-0.0030*	(0.0010)	-0.0086*	(0.0034)
γ_3	-0.4137*	(0.1105)	-6.9552*	(1.0640)	-5.0458*	(0.8600)	-0.8554*	(0.0931)	-1.8926*	(0.1077)	-1.3102*	(0.1377)
D: Volatility Components												
β_0	-0.1074	(0.1084)	-1.5764	(1.4786)	-2.3281*	(0.9074)	-0.0364*	(0.0220)	1.2688*	(0.0852)	0.5045*	(0.1494)
β_{I1}	0.0183	(0.0124)					0.0913*	(0.0042)	0.0835*	(0.0027)	0.0459*	(0.0023)
β_{I2}							0.0566*	(0.0061)	0.0397*	(0.0064)		
β_{S1}	0.0679*	(0.0184)	0.1198*	(0.0240)	0.1370*	(0.0265)	0.0352*	(0.0061)	0.0317*	(0.0072)	0.0277*	(0.0080)
β_{S2}	0.1611*	(0.0236)	0.1231*	(0.0300)	0.2105*	(0.0344)	0.0572*	(0.0068)	0.0403*	(0.0068)		
β_{M1}	0.1727*	(0.0378)	0.1594*	(0.0270)	0.0601	(0.0385)	0.0661*	(0.0078)	0.0690*	(0.0071)	0.0482*	(0.0135)
β_{M2}			0.1181*	(0.0216)	0.0986*	(0.0305)			0.0685*	(0.0071)	0.0685*	(0.0693)

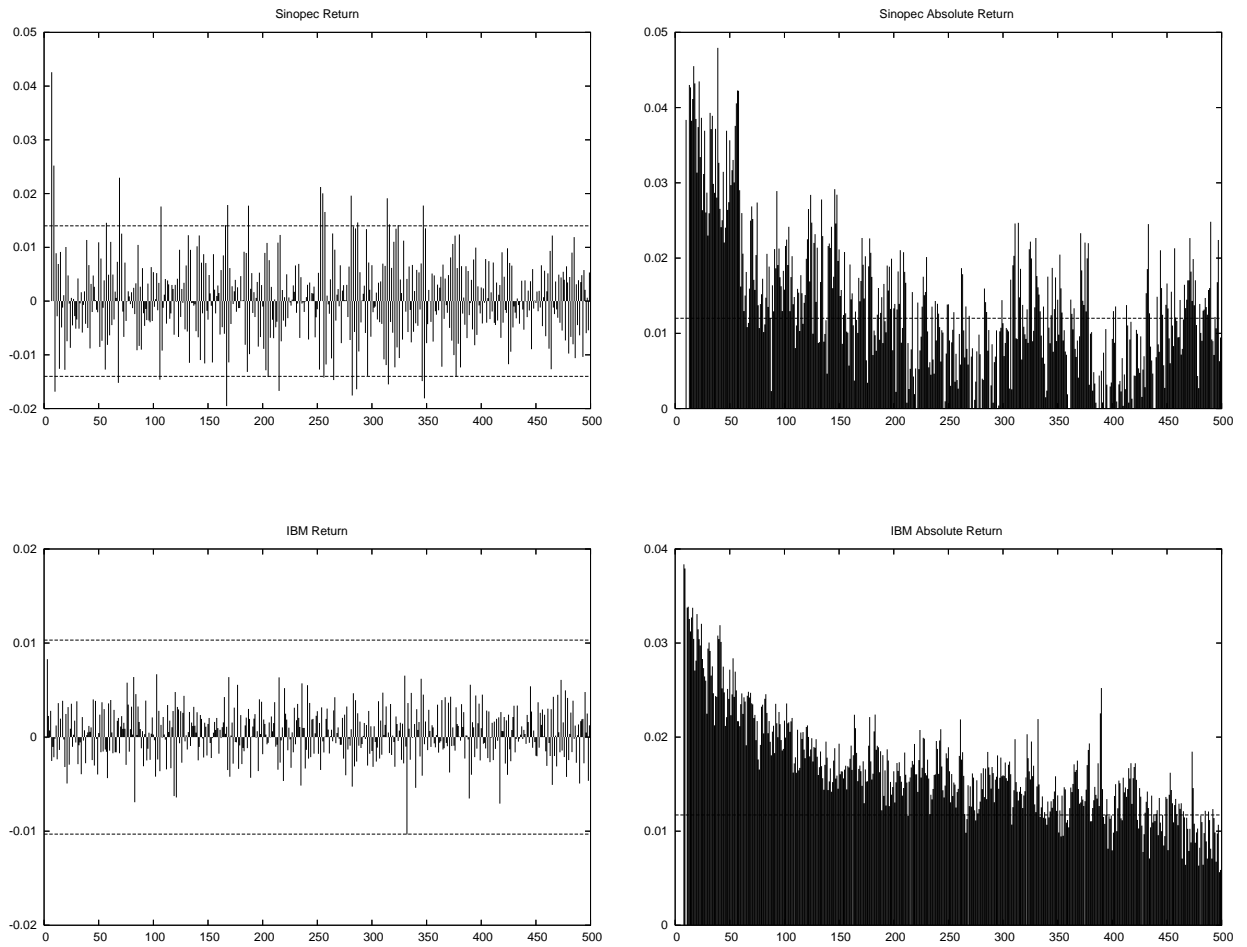
This table reports the posterior means and the standard deviations of the coefficients for all the stocks using HAR-BACD model. The cells with * have 0.95 density intervals (not reported) that exclude 0. The model is

$$\text{Return: } r_i = \rho r_{t-i} + u_i + \phi u_{i-1} \text{ with } u_i = \sqrt{q_i} \zeta_i, \zeta_i \stackrel{iid}{\sim} t_\nu(0, 1),$$

$$\text{Duration: } x_i = \psi_i z_i, z_i \sim \text{Burr}(1, \kappa, \varpi^2), \text{ with } \psi_i = \omega_a + \alpha_a x_{i-1} + \beta_a \psi_{i-1}$$

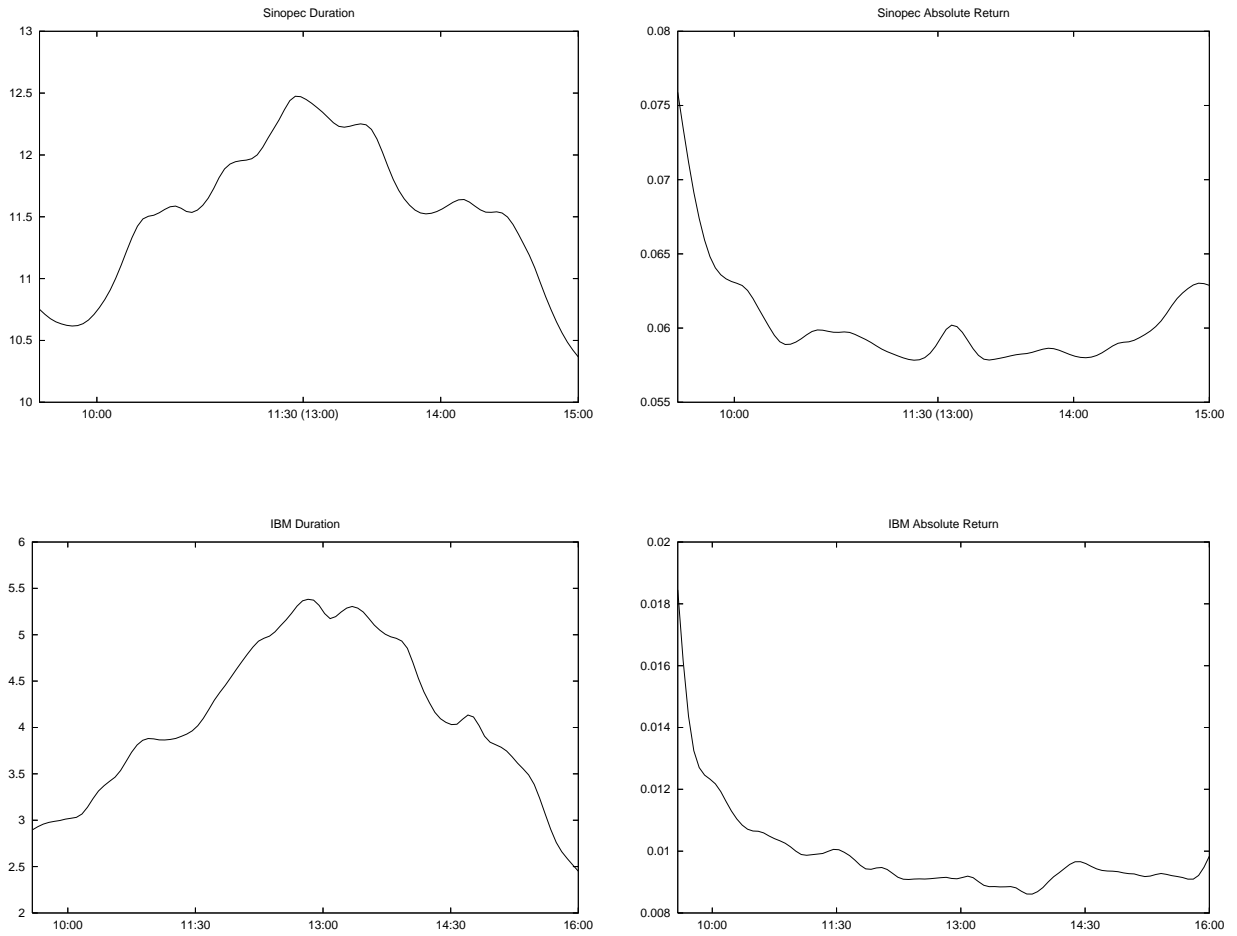
$$\text{Conditional Variance: } q_i = \beta_h + \sum_{m=1}^M \beta_m R V_{i-1, h_m} + \gamma_1 x_i^{-1} + \gamma_2 \frac{x_i}{\psi_i} + \gamma_3 \psi_i^{-1}$$

Figure 1: Autocorrelation Functions



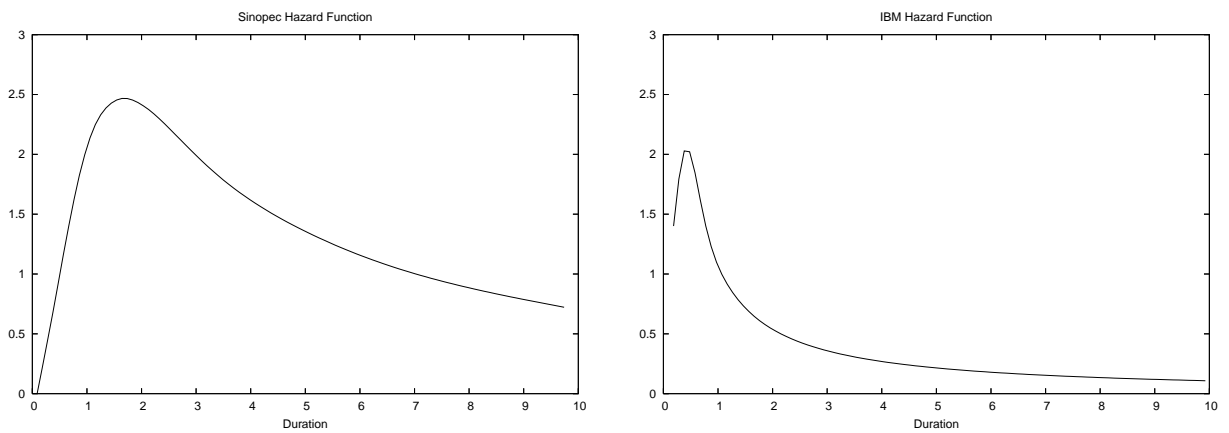
This figure reports the autocorrelation function of returns and absolute returns for Sinopec and IBM. The two dotted horizontal lines are the Bartlett standard error bands.

Figure 2: Diurnal Patterns



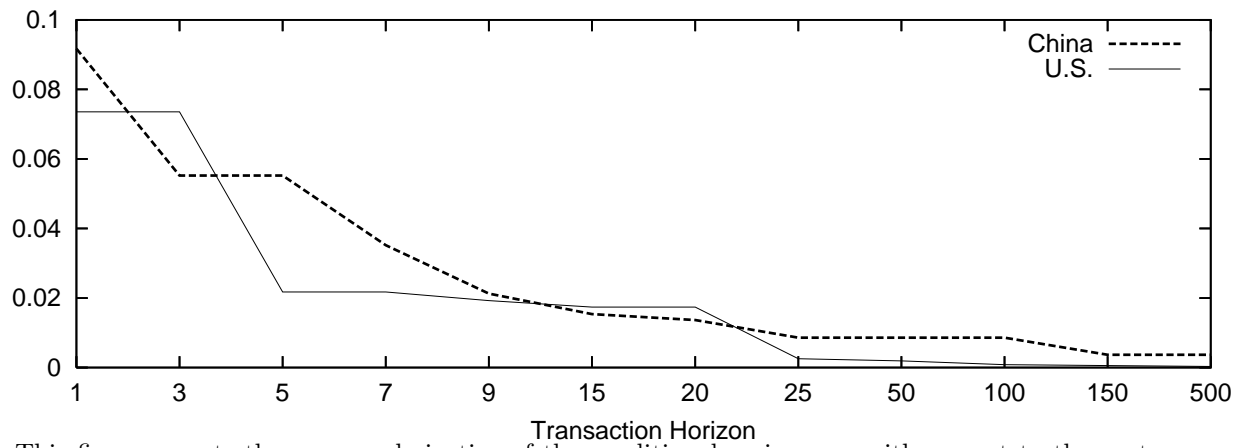
This Figure reports the daily spline estimates for duration and the absolute return. The top panel is for Sinopec data and bottom Panel for IBM. We use a cubic spline and the internal knots are set at 10:00, 11:00, 11:30 (1:00), 2:00 and 2:30 for Sinopec, and at 10:00, 11:00, 12:00, 1:00, 2:00, 3:00 and 3:30 for IBM.

Figure 3: Hazard Functions



This figure reports the conditional hazard functions for Sinopec and IBM.

Figure 4: Impact of a Volatility Shock on the Conditional Variance



This figure reports the average derivative of the conditional variance q_i with respect to the past squared innovation u_{i-h}^2 for Chinese stocks and U.S. stocks respectively.