PONTIFICIA UNIVERSIDAD CATÓLICA DEL PERÚ

ESCUELA DE POSGRADO

Duration Models and Value at Risk using High -Frequency Data for the Peruvian Stock Market

Tesis para optar el grado de Magíster en Economía con especialización en Finanzas y Mercados de Capitales que presentan

Giannio Enzo Téllez De Vettori
Ricardo Najarro Chuchón

Dirigido por
Gabriel Rodríguez

San Miguel, 2016

1 Introduction

Most empirical studies in finance use data on a daily basis which is obtained by retaining the last observation of the day and ignoring all intraday records. However, as a result of the increased automatization of financial markets and the evolution of computational trading systems, intraday data bases that record every transaction along with their characteristics have been stablished. These data sets prompted the development of a new area of research (finance with high frequency data), and in 1980 a literature based on the mechanisms of trading began (forms of trading, rules on securities trading, market structure, etc.), originating the Theory of Market Microstructure for the valuation of financial assets, whose models advocate that timing transmits information. Then the literature proposed an extension to risk management by calculating the implied volatility, which is estimated by the realized volatility on an intraday level, and its applications for a finer value at risk (VaR).

For this reason, Engle and Russell (1998) introduced the Autoregressive Conditional Duration (ACD) model to estimate the expected duration and justify time intervals between trades. The ACD model and all its extensions permits capturing the behavior of financial high frequency (intraday) series, driving other theoretical framework, such as trading strategies or risk measures. The importance of having finer measures of risk deals with the fact that the recent international crisis showed that the daily return risk measures used so far were insufficient to effectively reckon the risks of managed portfolios, especially when the market is very volatile, generating losses that could be monitored within the day. The problem with measures based on daily returns is that they neglect all intradays events[1], so even when the daily return might not be too volatile, at an intraday level this is not necessarily fulfilled, as the time dependence volatility is not suitable on a daily basis compared with an intraday frequency.

As the dynamics of intraday volatility of this type of data is complex, realized volatility models are estimated to capture those returns taking into account the period of financial transactions[2]. In the same way, the enormous growth of the trading activity of financial assets has led regulators to implement control measures (using quantitative techniques) to assess the potential loss that could be caused by these institutions. In that sense, the VaRs methodology has become the most wellknown technique as it provides an approach to measure market risk or expected financial losses at a time horizon with a given probability.

The objective of this research is to compute a better measure of intraday risk to infer policy recommendations to assess the regulators and the Portfolio Risk Managers using data from the portfolio of the General Index of the Lima Stock Exchange (GILSE). Regarding the revised literature, our results support the use of intraday data to calculate a better measure of risk as Fleming et. al. (2003), Giot and Laurent (2004), Louziset. et. al. (2011) and Khan (2011) did, who conducted similar research but in other sorts of markets. In the case of specific stocks, Dufour and Engle (2000) estimated that actively traded stocks at the New York Stock

[1] The intrinsic value of high frequency data is reflected in the events recorded. Moreover, these data are irregularly spaced in time.

[2] Note that in this type of data, there are patterns (and clustering) of volatility (on a daily basis) reflecting the daily cycle of activity in stock markets caused by the effects of macroeconomic announcements among others.

Exchange (NYSE), such as General Electric or Fannie Mae, are characterized by an average duration that ranges from 21 seconds to 1 minute, while more illiquid shares average duration ranges from about 2 minutes and 30 seconds (Potomac) to 4 minutes (Calfed). Similarly, Engle and Russell (2005) estimated that IBM, considered a highly liquid stock at NYSE, trades every 2 or 3 seconds, and a minor liquid stock such as Airgas trades every 3 minutes on average. Tse & Dong (2014) calculated that large-cap stocks are more frequently traded than small-cap stocks at the NYSE, for large-cap stocks the average duration per trade ranges from 2.27 seconds to 4.93 seconds, in contrast to small-cap stocks, that average duration per trade ranges from 7.68 seconds to 19.97 seconds. Regarding research in developing markets, Wong, Tan and Tian (2008) estimated that the average duration of liquid stocks at the Shanghai Stock Exchange (SHSE) oscillates between 21.8 and 34 seconds, whereas illiquid stocks average duration fluctuates between 85.4 and 160 seconds.

However, the literature of realized volatility estimation and VaR computing have not taken into account a very important element in the calculation of realized volatility, which is the justification of the duration, as this will depend on whether a high or low frequency exist between the returns. Furthermore, the duration directly affects the detection of jumps in the sample and influences the periodic effect (seasonal) in the series created to build the realized volatility. For example, in terms of risk management, we show that the intraday basis VaRs requires greater economic capital to offset any expected losses. Depending on the trading position, the best periodicity to compute the VaR are 2 - minutes interval for short positions and 5 - minutes intervals for long ones.

The rest of the paper is organized as follows. Section 2 reviews the theoretical framework associated with market microstructure, which generates the need for the ACD models at a high frequency framework, and the respective realized volatility and VaR applications. Section 3 provides a detailed description of the methodology to estimate the different models considered in our empirical study. Section 4 describes the data, which are sampled over seven different periods from 2008 to 2014. Also, major results obtained are described. Finally, Section 5 presents some conclusions and policy recommendations.

2 Theoretical Framework

The microstructure theoretical framework began with Demsetz (1968) developing a model which incorporates the immediacy with which the trades are executed, with the aim of analyzing the price formation, where agents with liquidity needs (impatient) and agents without monetary needs (patients) trade on the market at different stages of the negotiation. Given the above, Garman (1976) modeled a negotiation process with temporary imbalances at the trade flow (buy or sell), causing uncertainty in the estimated time of arrival of an order. These imbalances justify the presence of market maker who manages inventory stocks and cash, and solves the uncertainty problem in the arrival of an order. With respect to asymmetric information, Glosten and Milgrom (1985) modeled the problem generated by adverse selection cost of trading shares, in which agents with private information (insider traders) and operators that require liquidity (liquidity traders) trade in a perfectly competitive market and market maker revise their prices according to the information extracted from the order flow[3].

[3] Using a Bayesian learning process.

Glosten (1994) develops a theoretical model of the Limit Order Book (LOB)[4] with asymmetric information problems where investor behavior is modeled from an equation of price revision. This equation is defined as information to be conditional on public available information and private information of the trader[5]. If the marginal price valuation is greater or equal to the share price, the investor enters an order. Jong et. al. (1996) analyzed the effect on intra-day trading price and the components of bid-ask spread on the Paris Bourse from the model of Glosten (1994). They proposed that the spread is decomposed in two factors: order processing and asymmetric information. Finally, they concluded that the processing cost is higher and the adverse selection cost is lower in smaller operations.

Regarding the immediacy problem, Agudelo et. al. (2012) evaluated the probability of informed transactions and the effects of these on the daily and intraday markets shares of Peru, Argentina, Chile, Colombia, Brazil and Mexico yields, assuming a market with market makers, problems of asymmetric information and zero transaction costs. They used the model of Probability Informed Trading (PIN) to calculate the information asymmetry[6]. The conclusions are that persistent overnight was found in the arrival rate of uninformed agents and less persistence in the agents informed. Finally, the dynamic PIN variable with sign demonstrates the positive relationship between the level of asymmetry and expected asset returns.

From the LOB literature review, Chavez-Bedoya et. al. (2015) conducted an empirical test for the Lima Stock Exchange (LSE) and found that the balance point between immediacy and costs is higher on stocks with lower liquidity. In this sense, asymmetric information on the LSE is better used by specialist investors, which make profits at the expense of unskilled investors.

On the other hand, the use of high-frequency data within financial econometrics was introduced by the seminal paper of Hasbrouck (1991). This kind of data has important features such as irregular spacing between trades, discreteness of prices changes and intraday seasonalities, which are valuable sources of information for a better understanding of the transactions on the stock exchange. As a consequences modeling of high frequency data have captured much interest in researchers. Given the work of Hasbrouck (1991), Engle and Russell (1998) proposed the ACD model to estimate the expected duration and justify time intervals between trades.

Also, in this context, new econometric approaches emerge trying to improve the modeling of the high-frequency data in financial transactions. Thus, Zhang, Russell and Tsay (2001) proposed a more sophisticated model, which is called threshold autoregressive conditional duration (TACD) model, which permits calculating the nonlinear relation between the conditional expected duration and past information variables. They found that active and passive trading on the NYSE have a much differentiated dynamics and the durations have several structural changes, in line with real economy events.

[4] Up to this point, the theory of microstructure has focused on mechanisms in which there is a market maker. However, trading stock is not always promoted by this agent, although there is also the possibility that the LOB can provide liquidity. These models focus on the characteristics of the market structure or design,state of the LOB (full or empty) the problem of immediacy is analyzed with no asymmetric information or inventory costs.

[5] It is assumed that has no systematic behavior.

[6] In which a high PIN is interpreted as an increased probability of informed transaction, independent of management of such information.

From the VaR empirical review, Andersen et. al (2003) estimated a long-memory Gaussian VaR for the logarithmic realized volatilities of the Deutschemark/Dollar and Yen/Dollar spot exchange rates, and the results were more suitable than the conventional GARCH approaches and well-calibrated density forecast, but the realized volatility used in the paper do not discriminate between continuous price movements or jumps variability, so the forecasting model could be enhanced by taking into account the jump component. Also, Maheu, and McCurdy (2002) suggested that Markov switching models and non-linear components may improve the volatility forecasts.

On the other hand, Giot and Laurent (2004) calculated a daily VaR for two stock indexes (S&P500 and CAC40) using the realized volatility as an input. Although the VaR forecasts perform adequately, they also show that a daily return VaR delivers suitable results as well, so the realized volatility VaR does not improve much the forecast performance. It should be noted that their estimates do not consider jumps or seasonal periodicity effects, so the results could be improve.

Based on Barndorff-Nielsen and Shephard (2004b), Chun and Maheu (2005) investigated the benefits of volatility instruments, known as realized power variation and realized bipower variation in modeling and forecasting volatility, and conclude that realized power variation improves the volatility predictions for foreign exchange rates and equity markets because these measures are robust to jumps.

The Literature presented above has focused on forecasting realized volatility, regardless of the jumps in the sample, Clements and Liao (2013) took this into account examining how best to use the jump component of volatility for modeling and forecasting total volatility, considering the role of jump size and probability of jump occurrence. They concluded that jumps self-excite or cluster and the estimated intensity were used to forecast the volatility, founding that the jump intensity forecasting models were superior, and this improvement was robust to critical periods (jumps and high and variable volatility).

Khan (2011) proposed to compare the performance of a realized volatility support vector machines (SVM) ARCH type model versus a common daily ARCH model for the returns of the Nikkei 225 index and found that the SVM-ARCH type models are more suitable when 15 - minutes intraday returns are computed.

Finally, Louzis et. al (2011) evaluated the VaR prediction efficiency of six ARCH-type models, six realized volatility models and two GARCH models considering the realized volatility as a regressor for two S&P500 cash indexes, taking into account the 2007-2009 financial crisis period, and they found that the realized volatility and GARCH models with the Filtered Historical Simulation or the Extreme Value Theory quantile estimation methods result in superior VaR forecasts.

All the previosly research performs complex methodologies to analyze the intraday effects on the market and the respective empirical use of it to enhance risk management measured by the VaR, but obviates a very important element to calculate the realized volatility: does not justify the duration (taking randomly 5, 15, 30 minutes) as input for subsequent estimates, as this will depend on whether a high or low frequency exists between returns, so their results or findings may be biased depending on how irregular or periodic time intervals are. Furthermore, the duration affects directly the detection of jumps in the sample, and influences the periodic effect in the series created to build the realized volatility. In the current investigation, the calculation of the duration was justified to calculate the realized volatility adjusted by jumps and periodicity effects, estimating intraday basis VaRs and comparing them against a daily returns basis VaR.

3 Econometric Methodology

Realized volatility is the most common measure used, due to its easy calculation to generate a daily variable which allows us to measure intraday volatility. Furthermore, the realized volatility is very sensitive to the duration used as input, for this reason there are many possibilities of calculation depending on the length used, and likewise the results vary, ACD and TACD models are computed to estimate the duration to justify the time intervals between trades, given the fact that trade arrivals carry information on the state of the market.

It is important to note that, to standardize notations of econometric models, the following papers were taken into account: Tsay, R.S. (2002); Tsay, R.S. (2010); Zhang, Russell and Tsay (2001); Weisang (2008); and as a reference, the notes about ACD model of Tsay, R. S. (2007), situated in section 3.1. In section 3.2, we used as reference of the Realized Volatility theory and the framework of jumps and bipower variation the papers of Andersen and Benzoni (2008), Andersen and Teräsvirta (2009); and Boffelli (2013) and Boudt et al. (2009) for the section of jumps and intraday periodicity.

3.1 Autoregressive Conditional Duration (ACD) Models

Duration models in finance are involved with periods of time between trades, thus the behavior of durations contains information about trades in market activities. In that sense, Engle and Russell (1998) introduced the ACD model, which describes the evolution of time durations for stocks.

Let t_i be the ith transaction time, where $0 = t_0 < t_1 < ... < t_N$ and $x_i = t_i - t_{i-1}$ be the duration between trades. Define ψ_i as the conditional expectation of the adjusted duration between the $(i-1)th$ and the ith trades, that is:

$$\psi_i = E(x_i \mid x_{i-1}, ..., x_1) = E(x_i \mid F_{i-1}) \ , \tag{1}$$

where F_{i-1} represents the information set available at the $(i-1)th$ trade.

The ACD model in its simplest form is defined as:

$$x_i = \psi_i \epsilon_i, \tag{2}$$

where $\{\epsilon_i\}$ represents a sequence of independent and identically distributed nonnegative random variables with density $f(.)$ and $E(\epsilon_i) = 1$. Also note that ϵ_i is independent of F_{i-1}. Thus, from equation (1) a set of ACD model specifications can be defined by different distributions of ϵ_i and specifications of ψ_i. For example, ϵ_i follows a standard exponencial or weibull distribution, and ψ_i takes the following form[7]

$$\psi_i = \omega + \sum_{j=1}^{r} \gamma_j x_{i-j} + \sum_{l=1}^{s} \omega_l \psi_{i-l} \ , \tag{3}$$

[7] See, Engle and Russell (1998).

where r and s are nonnegative integers and the unknown parameters are $\theta = (\omega, \gamma_1, ..., \gamma_j, \omega_1, ..., \omega_l)$. The conditions $\omega > 0$, $\gamma_j \geq 0$ for $j \in \{1, ..., r\}$, $\omega_l \geq 0$ for $l \in \{1, ..., s\}$ and $\sum_{j=1}^{r} \gamma_j + \sum_{l=1}^{s} \omega_l < 1$ are required to ensure the non-negativity and weak stationarity of durations x_i. Since the durations are nonnegative variables, we use the Exponential, Weibull and Gamma distributions to model ACD structures [8]. When the distribution of ϵ_i is exponential, the model is called an EACD(r, s) model. Likewise, if ϵ_i follows a weibull and gamma distribution, the models are WACD(r, s) and GACD(r,s), respectively.
Similar to GARCH models, the process $\eta_i = x_i - \psi_i$ is a martingale difference sequence $[i.e., E(\eta_i \mid F_{i-1}) = 0]$, and the ACD model can be written as

$$x_i = \omega + \sum_{j=1}^{\max(r,s)} (\gamma_j + \omega_l) x_{i-j} - \sum_{l=1}^{s} \omega_l \eta_{i-l} + \eta_i, \tag{4}$$

which represents an ARMA process with non Gaussian innovations. Where $\gamma_j = 0$ for $j > r$ and $\omega_l = 0$ for $l > s$. This representation can be used to obtain the conditions for weak stationarity of the ACD model. For example, taking expectation on both sides of equation (4) and assuming weak stationarity, we have

$$E(x_i) = \frac{\omega}{1 - \sum_{j,l=1}^{\max(r,s)} (\gamma_j + \omega_l)}. \tag{5}$$

Therefore, we assume $\omega > 0$ and $1 > \sum_{jl} (\gamma_j + \omega_l)$ because the expected duration is positive.

3.1.1 ACD Estimation

For an ACD(r, s) model, let $i_0 = max(r, s)$ and $x_t = (x_1, ..., x_t)'$. The likelihood function of the durations $x_1, ..., x_T$ is

$$f(x_T \mid \theta) = \left[\prod_{i=i_0+1}^{T} f(x_T, F_{i-1} \mid \theta \right] \times f(x_{i_0} \mid \theta) \;, \tag{6}$$

where θ is the vector of model parameters and T is the sample size. The marginal probability density function $f(x_{i_0} \mid \theta)$ is rather complicated for a general ACD model. For this reason, the marginal density is ignored and the conditional likelihood method is used. For a WACD model, we use the probability density function and get the conditional log-likelihood function.

[8] See, Tsay, R.S. (2010).

$$\ell(x \mid \theta, x_{i_0}) = \sum_{i=i_0+1}^{T} \alpha \ln\left[\Gamma(1+\frac{1}{\alpha})\right] + \ln\left(\frac{\alpha}{x_i}\right) + \alpha\ln(\frac{x_i}{\psi_i}) - \left[\frac{\Gamma(1+\frac{1}{\alpha})x_i}{\psi_i}\right]^{\alpha} , \qquad (7)$$

where $\psi_i = \omega + \sum_{j=1}^{r}\gamma_j x_{i-j} + \sum_{l=1}^{s}\omega_l\psi_{i-l}$, $\theta = (\omega, \gamma_1, ..., \gamma_r, \omega_1, ..., \omega_s, \alpha)'$ and $x = (x_{i_0+1}, ..., x_T)'$. When $\alpha = 1$, the conditional log likelihood function is reduced to an EACD model. For a GACD model, the conditional log likelihood function is

$$\ell(x \mid \theta, x_{i_0}) = \sum_{i=i_0+1}^{T} \ln\left(\frac{\alpha}{\Gamma(\kappa)}\right) + (\kappa\alpha - 1)\ln(x_i) - \kappa\alpha\ln(\lambda\psi_i) - \left(\frac{x_i}{\lambda\psi_i}\right)^{\alpha} , \qquad (8)$$

where $\lambda = \Gamma(\kappa)/\Gamma(\kappa + \frac{1}{\alpha})$ and the parameter vector θ now also includes κ. As expected, when $\kappa = 1$, $\lambda = 1/\Gamma(1 + \frac{1}{\alpha})$ and the log likelihood function in equation (8) is reduced to a WACD model in equation (7).
In duration, the hazard function implied by a distribution function. For a random variable X, the survival function is defined as $S(x) = P(X > x) = 1 - P(X \leq x) = 1 - CDF(x)$, $x > 0$, which gives the probability that a subject, which follows the distribution of X, survives at the time x. The hazard function of X is then defined by $h(x) = \frac{f(x)}{S(x)}$, where $f(.)$ and $S(.)$ are the probability density function and survival function of X, respectively. In specific, for the weibull distribution, the hazard is a monotone function. If $\alpha > 1$ the hazard function is monotonously increasing and if $\alpha < 1$ the hazard function is monotonously decreasing[9]. For the generalized gamma distribution, the hazard function can show different patterns, including U shape or inverted U shape. If $\kappa\alpha < 1$ the hazard rate is decreasing for $\alpha \leq 1$, and U-shaped for $\alpha > 1$. Contrariwise, if $\kappa\alpha > 1$,the hazard rate is increasing for α ≥ 1, and inverted U-shaped for $\alpha < 1$. Lastly, if $\kappa\alpha = 1$,the hazard is decreasing for $\alpha < 1$, constant for $\alpha = 1$, and increasing for $\alpha < 1$[10].

3.1.2 Quasi maximum likelihood estimates

We know that the distribution of ϵ_i in an ACD model is unknown. Therefore, given the literature of these models, we use the conditional likelihood function of an EACD, WACD and GACD models to estimate the parameters. The results of the estimate are called the quasi maximum likelihood estimates. Thus, Engle and Russell (1998) proposed that, under regularity circumstances, quasi maximum likelihood estimates of an ACD model are consistent and asymptotically normal. Nevertheless, they are not efficient when ϵ_i does not follow an exponential distribution. Estimation procedures of the parameters have been done with the RATS package, as well as the BFGS algorithm has been used. Finally, the standard errors of the parameter estimates are the robust standard errors given by RATS.

[9] See, Tsay, R.S. (2010).
[10] See, Fernandes,M. and Gramming, J. (2005).

3.1.3 Threshold ACD Model

As mentioned in Zhang, Russell and Tsay (2001), the threshold autoregressive conditional duration (TACD) model is nearly associated to the TAR model and the more general threshold autoregressive moving average (TARMA) model. The TACD model is a generalization of the ACD model proposed by Engle and Russell (1998).
Based on notes about ACD model of Tsay, R. S. (2007), a simple TACD(r, s) model for x_i can be written as

$$x_i = \begin{cases} \psi_i \epsilon_{1i} & \text{, if } x_{t-d} \ \leq p, \\ \psi_i \epsilon_{2i} & \text{, if } x_{t-d} \ > p, \end{cases} \tag{9}$$

where d is a positive integer, x_{t-d} is the threshold variable, p is a threshold, and

$$\psi_i = \begin{cases} \omega_{10} + \sum_{j=1}^{r} \gamma_{1j} x_{i-j} + \sum_{l=1}^{s} \omega_{1l} \psi_{i-l} & if \ x_{t-d} \ \leq p, \\ \omega_{20} + \sum_{j=2}^{r} \gamma_{2j} x_{i-j} + \sum_{l=1}^{s} \omega_{2l} \psi_{i-l} & if \ x_{t-d} \ > p, \end{cases} \tag{10}$$

where $\omega_{l0} > 0$ and γ_{jj} and ω_{ll} satisfy the conditions of the standard ACD model for j and $l = 1$ and 2. Here j and l denote the regime. The innovations $\{\epsilon_{1i}\}$ and $\{\epsilon_{2i}\}$ are two independent $i.i.d.$ sequences. They can follow the exponential, weibull and gamma distribution and the resulting models can be TEACD, TWACD, and TGACD, respectively. The TACD model is a piecewise linear model in the space of x_{i-d}, and it is nonlinear when some of the parameters in the two regimes are different[11].

3.2 Realized Volatility

Based on Andersen and Benzoni (2008) notation, we assume a continuosly compounded return driven by a simple time-invariant Brownian motion:

$$ds(t) = \mu(t)dt + \sigma(t)dW(t), 0 \leq t \leq T, \tag{11}$$

where $s(t)$ is a logarithmic asset price at time t and the continuosly compounded returns over $[t-k, t]$ is given by $r(t, k) = s(t) - s(t-k)$, where $0 \leq t-k \leq t \leq T$ and $k = j/n$ for some positive integer j. W is a standard Brownian motion process, $\mu(t)$ and $\sigma(t)$ are functions with finite variation or strictly positive and square integrable, respectively, so that $E\left(\int_{t-k}^{t} \sigma_s^2 ds\right) < \infty$. Given the above, the processes $\mu(t)$ and $\sigma(t)$ stand for the conditional mean and volatility of the return. Therefore, the continuously compounded return over the time interval from $t-k$ to t, $0 < k \leq t$ is:

$$r(t, k) = s(t) - s(t-k) = \int_{t-k}^{t} \mu(\tau) d\tau + \int_{t-k}^{t} \sigma(\tau) dW(\tau), \tag{12}$$

and its quadratic variation $QV(t, k)$ is:

$$QV(t, k) = \int_{t-k}^{t} \sigma^2(\tau) dW(\tau). \tag{13}$$

[11] The model can be extended to have more than 2 regimes.

Equation (13) points out that innovations to $\mu(t)$ do not affect the sample path variation of the return[12], and the diffusive sample path variation over $[t-k,t]$ is also known as the integrated variance $IV(t,k)$,

$$IV(t,k) = \int_{t-k}^{t} \sigma^2(\tau) dW(\tau). \tag{14}$$

Equations (13) and (14) illustrate that the quadratic and integrated variation coincide[13], hence removing microstructure noise and measurent error, the quadratic return variation can be approximated by the cumulative squared return function.
Account a partion $\left\{t-k+\frac{j}{n}, j=1, ...n\times k\right\}$ of the $[t-k,t]$ interval, in this sense the realized volatility (RV) of the logarithmic price process is:

$$RV(t,k;n) = \sum_{j=1}^{n\times k} r(t-k+\frac{j}{n}, \frac{1}{n})^2. \tag{15}$$

As mentioned in Andersen and Benzoni (2008), the semimartingale theory guarantees that the realized volatility measure converges in probability to the return quadratic variation QV, previously defined in equation (13), when the sampling frequency n increases, so that:

$$RV(t,k;n) \rightarrow QV(t,k) \text{ as } n \rightarrow \infty. \tag{16}$$

The distributional result of the variance also generalizes directly, as we approach, for $n \rightarrow \infty$, to:

$$\sqrt{n\times k}\left(\frac{RV(t,k;n) - QV(t,k)}{\sqrt{2 \;\; IQ(t,k)}}\right) \rightarrow N(0,1), \tag{17}$$

where $IQ(t,k) = \int_{t-k}^{t} \sigma^4(\tau)d\tau$ is the integrated quarticity, independent from the limiting Gaussian distribution[14]. As explained in Andersen and Teräsvirta (2009), the equation (17) allows the ex-post inference of the realized return variation for a given period. Nevertheless, it should be noticed that the conclusion depends on the non appearance of jumps in the price process, which is subject that will be discussed in the next subsection.

3.2.1 Jumps and Bipower Variation

Equation (11) is quite restrictive for asset prices since under unexpected news and events, that hit the market, prices tend to evince irregular or abrupt discrete movements. Therefore, the return process should take into consideration a model that may exhibit jumps. Continuing with the notation of Andersen and Benzoni (2008), the presence of jumps in returns is defined by:

[12]Intuitively, this is because the mean is of lower order.

[13]This is no longer true for more general return process that will be discussed in the subsection 3.2.1 (stochastic volatility jump-diffussion model).

[14]This result was developed and brought into the realized volatility literature by Barndorff-Nielsen and Shephard (2002).

$$ds(t) = \mu(t)dt + \sigma(t)dW(t) + \xi(t)dq(t), \tag{18}$$

where $q(t)$ is a Poisson process uncorrelated with W and governed by the jump intensitity λ_t, i.e., $\Pr(dq_t = 1) = \lambda_t dt$, with λ_t positive and finite[15]. The adjustment factor $\xi(t)$ denotes the magnitude of the jump in the return process if a jump occurs at time t. In this respect, the quadratic return vaiation process over the interval from $t-k$ to t, $0 \leq k \leq t \leq T$, is the sum of the diffusive integrated variance and the cumulative squared jumps:

$$QV(t,k) = \int_{t-k}^{t} \sigma^2(s) + \sum_{t-k \leq s \leq t} J^2(s) = IV(t,k) + \sum_{t-k \leq s \leq t} J^2(s), \tag{19}$$

where $J(t) = \xi(t)dq(t)$ is non-zero if there is a jump at time t. As mentioned in Andersen and Benzoni (2008), the RV estimator persists as a consistent measure of the QV in the existence of jumps, generating that the result (16) continues[16]. However, since the volatility components have distinct persistence properties, we must acquire separate estimates of these factors in the decomposition of the equation (19).

With this objective the h-skip bipower variation, BV, introduced by Barndorff-Nielsen and Shephard (2004b) contributed to estimate a consistent IV component,

$$BV(t,k;h,n) = \frac{\pi}{2} \sum_{i=h+1}^{n \times k} \left| r(t-k+\frac{ik}{n}, \frac{1}{n}) \right| \left| r(t-k+\frac{(i-h)k}{n}, \frac{1}{n} \right|. \tag{20}$$

Given the notation of Andersen and Teräsvirta (2009), setting $h = 1$ in definition (20) yields the realized bipower variation $BV(t,k;n) = BV(t,k;1,n)$, so the bipower variation is robust to the presence of jumps and therefore, in combination with RV, it yields a consistent estimate of the cumulative squared jump component:

$$RV(t,k;n) - BV(t,k;n) \underset{n \to \infty}{\to} QV(t,k) - IV(t,k) = \sum_{t-k \leq s \leq t} J^2(s). \tag{21}$$

The results in equations (19), (20) and (21) along with the asymptotic distributions allow us to enhance the volatility forecast and design tests for jumps in the volatility.

3.2.2 Jumps and Intraday Periodicity

The intraday dynamics of financial time series volatility are complex due to the fact that exist intraday volatility patterns that reflect daily activity cycles, negotiation stages or irregular market events, for instance, regional holidays, weekend activity, unexpected political events, macroeconomic releases and standard volatility clustering. Given the above, if the time serie is affected by those events, it means that the variable exposes an intraday periodic structure. Given the notation of Boudt et. al. (2009), we dispose of T days of j equally-spaced and continuosly compounded intraday returns and that $r(t,k)$ is the $k-th$ return of day t, it is assumed that the log-price follows a Brownian semi martingale with finite activity jumps (BSMFAJ) diffusion process, and for small values of n the returns in a range without jumps

[15]This assumption implies that there can only be a finite number of jumps in the price path per time period.
[16]See Protter (1990) and the discussion in Andersen et al. (2004).

are normally distributed with mean zero and variance $\sigma^2_{t,k} = \int_{t-k}^{t} \sigma^2_s ds$. Due to daily activity cycles, negotiation stages or irregular market events, the high-frequency return variance $\sigma^2_{t,k}$ has a component $f^2_{t,k}$ which is the periodicity factor. Depending on the nature of the analysis, there exists a natural window lenght for which almost all variation in $\sigma^2_{t,k}$ during the window can be attributed to $f^2_{t,k}$, such that $s^2_{t,k} = \sigma^2_{t,k}/f^2_{t,k}$ is approximately constant over the local window. Henceforth, for its estimation we follow the suggestion of Andersen and Bollerslev (1997, 1998b), Andersen, Bollerslev, and Dobrev (2007) and Lee and Mykland (2008) to use local windows of one day. Furthermore, given the theoretical and econometric research, the estimated series and the periodicity factor must be stationary, so that the realized volatility will not be explosive.

Andersen and Bollerslev (1997, 1998b) suggested estimating $s_{t,k}$ by $\widehat{s}_t = \sqrt{\frac{1}{M} h_t}\ \forall\ k = 1, ..., M$, where h_t is the conditional variance of day t. Boudt et. al. (2009) concluded that under the BSM model, $\widehat{s}_t = \sqrt{\frac{1}{M} RV_t}$ is a more efficient estimator for $s_{t,k}$, and under the BSMJAJ model, $s_{t,k}$ is better approached by the normalized version of Barndorff-Nielsen and Shephard (2004b)'s realized bi-power variation $\widehat{s}_t = \sqrt{\frac{1}{M-1} BV_t}$, where BV is the bi-power variation computed on all the intraday returns of day t. Under this model, we have that when $n \to \infty$ and if $r(t,k)$ is not affected by jumps, the standarized high-frecuency return $\widehat{r}_t = r_{t,k}/s_{t,k}$ is normally distributed with mean zero and variance equal to the squared periodicity factor.

On the issue of intraday jumps, Lee and Mykland (2008) established that a return jump must be abnormally big and depends on the volatility condition prevailing at the time tested. In this sense, they proposed an statistic $J_{t,k}$ test whether a jump ocurred between intradaily time periods $r-1$ and r of day t. It is defined as the absolute return divided by an estimate of the local standards deviation $\widehat{\sigma}_{t,k}$, i.e. $J_{t,k} = |r_{t,k}| / \widehat{\sigma}_{t,k}$. Assuming that the return process follows a BSMJ model, a null hypothesis of no jump at the time, and a suitable choice of the window size for local volatility, they concluded that $r_{t,k}/\widehat{\sigma}_{t,k}$ asymptotically follows a standard normal distribution.

Finally, given the works of Brownlees and Gallo (2006) and Andersen, Bollerslev, and Dobrev (2007), Lee and Mykland (2008) proposed to overcome spurious jump detection by inferring jumps from a conservative critical value, which they obtained from the distribution of the statistic's maximum over the sample size. In this sense, if the statistic exceeds a plausible maximum, one rejects the null hypothesis of no jump.

3.3 Value at Risk (VaR)

3.3.1 VaR Measures

The VaR is defined as the maximum expected loss over a target time period for a given confidence interval, or with a error probability α[17]. Additionally, the VaR level α for a sample of returns is the empirical quantile at $\alpha\%$. For example, with probability $1-\alpha$, the losses will be smaller than the money amount given by the VaR. Also, given a certain confidence level α ϵ $(0, 1)$ the portfolio VaR confidence level α is given by the smallest number l such that the

[17] As proposed by the Basel Committee (1996), the perfect confidence interval is 99% (1% probability, -2.33 deviations).

probability that the loss L exceeds the value l is not greater than $(1-\alpha)$. The mathematical definition is:

$$VaR_\alpha = \inf\{l \,\epsilon\, \mathbb{R} : P(L > l) \leq 1-\alpha\} = \inf\{l \,\epsilon\, \mathbb{R} : F_L(l) \geq \alpha\}. \quad (22)$$

This definition considered an underlying probability distribution, which makes it true only for a parametric VaR. Risk managers often assume that a fraction of the losses will not be defined, either because markets are closed, illiquid, or there is an atypical event. Therefore, it is difficult to accept results based on the assumption of a probability distribution well defined, which is fundamental to the entire family of GARCH models[18] to model volatility.
The other two ways to calculate the VaR are based on simulations (Monte Carlo), and the non parametric methodology. In the first case, it is also inherent to know the distributions to originate the simulations, and is very sensitive to atypical market events. The non parametric VaR is based on the time series of returns[19], this investigation the realized returns adjusted with the realized volatility will be used as a finer input for calculating the VaR with this methodology.

3.3.2 VaR adequacy

The failure rate is a good and simple statistic to measure the effectiveness of a VaR model, and is defined as the number of times the returns exceed in absolute value the forecasted VaR. For example, when the VaR measure is correctly specified, the failure rate could be equal to the pre-specified VaR level. In the literature, the failure rate is also called the Kupiec (1995) LR test when the hypothesis is tested using a likelihood ratio test. The LR test is $LR = -2log\left(\frac{\beta^N(1-\beta)^{T-N}}{\widehat{f}^N(1-\widehat{f})^{T-N}}\right)$, where N is the number of VaR violations, T is the total number of observations and β is the theoretical failure rate. The LR test statistic is asymptotically distributed as $\chi^2_{(1)}$.
However, the VaR has some difficulties. According to Artzner, Delbaen, Eber, and Heath (1999), one of the diffiiculties is that it is not a coherent measure of risk. For this reason a measure of risk is the so-called expected shortfall (see Scaillet, 2000). Expected shortfall is a coherent measure of risk and it is defined as the expected value of the losses conditional on the loss being larger than the VaR.
In addition to the failure rate, another VaR measure should show a sequence of indicator functions that is not serially correlated. To analyze this, given the notation of Engle and Manganelli (1999) the variables are as follows $Hit_t(\alpha) = I\left[y_t < VaR_t(\alpha)\right] - \alpha$ and $Hit_t(1-\alpha) = I\left[y_t > VaR_t(1-\alpha)\right] - \alpha$, and suggests testing the following:

- $A1$: $E\left[Hit_t(\alpha)\right] = 0$ (respectively $E\left[Hit_t(1-\alpha)\right] = 0$) in the case of long trading positions,
- $A2$: $Hit_t(\alpha)$(or $Hit_t(1-\alpha)$) is uncorrelated with the variables included in the information set.

[18] Either Gaussian, Student, GED, Skewed-Student distributions.
[19] To characterize the models, we consider a collection of daily returns (in %), $y_t = 100\left[\log p_t - \log p_{t-1}\right]$, where $t = 1, ..., T$, and p_t is the price at time t.

Besides, testing $A1$ and $A2$ can be done using the artificial regression $Hit_t = X\lambda + \epsilon_t$, where X is a $T \times k$ matrix whose first column is a column of ones, the next p columns are $Hit_{t-1}, \ldots, Hit_{t-p}$ and the $k-p-1$ remaining columns are additional independent variables (including the VaR itself). Engle and Manganelli (1999) also showed that under the null $A1$ and $A2$, the Dynamic Quantile test statistic $\frac{\bar{\lambda}' X' X \bar{\lambda}}{\beta(1-\beta)} \rightarrow \chi^2(k)$ where $\bar{\lambda}$ is the OLS estimate of λ. A small sample version of this test (F-test) is readily obtained but the difference is negligible since the sample size is larger than 1,000 observations.

4 Empirical Results

4.1 Preliminaries

The study of high frequency series implies knowing the characteristics of the negotiation mechanism of a stock market, for that reason the most important features of the LSE will be explained below. The LSE has a mechanism that adds electronic trading orders or proposals for buying and selling in a LOB, that shows automatically low/high prices at their best time proposals in a process of continuous auction executed at different prices. There are no specialists or market makers[20], so liquidity is provided solely by the order book. There is a discriminatory pricing rule, which governs all phases of negotiation, which is the possibility that an order could be executed in parts at different prices. In this sense, the priority of the proposals, and their respective adjudications, are governed by: 1) the proposal which improves the price in the LOB and 2) a longer exposure in the market. For a proposal to improve the price in the LOB, it depends on whether you are buying or selling. Since the aim is to reduce the spread, a purchase proposal improve if the price is greater than the largest purchase proposal in the LOB. Conversely, a sell proposal improve if the price is less than the smallest selling proposal in the LOB. A longer market exposure references the fact that who initiates the proposal is being prioritized, in this regard who originates first is given priority. The formation of the initial price is given in a first phase call pre-opening, in which the auction system receives proposals without possibility of being canceled. At this stage, all market expectations are stored without being canceled. From this information, the system allocates a price in a variable period time[21], in order to pass the proposals to be implemented. The next phase is the continuous trading, in which traders enter pending proposals that fit automatically. These proposals are limit orders, which you can buy or sell specifying the quantity, price and an exposure period of the proposal. The limit orders can not enter prices that exceed the minimum limit of variation (tick) of 0.01, a maximum variation of 15% for domestic securities and 30% for foreign securities, being these variations from the last price adjudicated of the previous day. Thirty minutes before the next phase, the system performs an average of prices, and the last change of the actual stage is from that average, and the variation itself is not more/less than 2%.
The last phase is the closing, which works similarly to the pre-opening, determining a closing

[20] Although there is a regulatory promoter agent having market maker functions, in reality this figure is not done, see Loaiza (2013).

[21] This means that the first adjudication is in a time interval of +/- 2 minutes, with reference of the start time of the following stage.

price of the shares at an arbitrary time range similar to the previous phases. Once this price is determined, any proposal of purchase or sale is entered at that value. Thus, the formation of the price of each share, responds to the net aggregate demand (purchases minus sales) of each action that occurs in every moment of continuous trading. In this phase, dynamic prices can have asymmetric information problems between investors, and lack of immediacy in the execution of their transactions. These drawbacks create costs that are denominated in the literature of microstructure as frictions, because they impact on price formation.
Regarding the data implications, the GILSE series with an intraday basis from 2008 to 2014 were used. The data were obtained from a Bloomberg Terminal and the database of the LSE. The number of initial observations was 301,831, but only data from the continuous trading phase was considered[22], which were 264,874 observations. We divided the data into seven groups, which represent each year. For 2008, the number of observations was 55,165, for 2009 was 53,596, for 2010 was 43,798, for 2011 was 49,537, for 2012 was 29,929, for 2013 was 27,249 and for 2014 was 5,587. As explained in Tsay (2010), the intraday data depends seasonally on the negotiation phase. Therefore, the sample may not be representative and show bias. For that reason we only analyzed data from the continuous trading phase. The time duration between trades was obtained from each operation within the portfolio analyzed. Transactions do not occur at equally spaced time intervals. As such, the actual trading prices of an asset do not form equally spaced time series. For this reason, given the optimum durations of the ACD models, the data was transformed so the time intervals were equally spaced and thus realized volatility could be estimated. Finally, an intraday VaR was estimated with the duration criteria of ACD models, and a finer volatility calculation by estimating the realized volatility.

Table 1. Statistics of GILSE for the Duration characteristics from 2008 to 2014

Observations	264,874		
Sample Mean	123.135	Variance	36,632
Standard Error	191.396	SE of Sample Mean	0.372
t-Statistic (Mean=0)	331.107	Signif Level (Mean=0)	0.000
Skewness	8.031	Signif Level (Sk=0)	0.000
Kurtosis (excess)	163.950	Signif Level (Ku=0)	0.000
Jarque-Bera	299,502,866	Signif Level (JB=0)	0.000

4.2 Estimation Results

4.2.1 Lineal Models

We estimated the ACD (1,1) model, rather than the other ACD models of greater than one order, because the results obtained were favorable and consistent with the literature[23]. Therefore 3 different models were estimated: EACD(1,1), WACD(1,1) and GACD(1,1) to

[22] From 8:30 AM to 14:30 PM from the second Sunday of March to first Sunday of November, and 9:30 AM to 15:30 PM from the first Sunday of November to the second Sunday of March. Therefore, the data analyzed were 6 hours of the daily negotiation.

[23] The estimated parameters and expected duration are positive. Besides of Ljung–Box test and nonlinearity Tsay test (1989).

intraday series and for all periods analyzed. In apendix A, we present the results of the three models for each specific year, as well as a few statistics to analyze the models adequacy.
For instance, the estimated parameters of the three models in the period 2008 to 2014 are given in Table 2, the results of the three models show that the parameters are significant and similar. However the expected duration are differents, thus for EACD(1,1), WACD(1,1) and GACD(1,1) models the expected duration are 145, 188 and 175 seconds respectively, and the results of expected duration are above the sample mean (123 seconds). The average expected duration to ACD model from 2008-2014 period is 169 seconds (3 minutes), which is used as input to calculate the realized volatility. On the other hand, the three previous models described before, failed to pass the Ljung -Box test statistics at a 5% significance level. These results give us indications that there are some ACD effects in the estimated residuals. Another remarkable feature is the fact that the sum of the coefficients ($\gamma_1 + \omega_1$) are always extremely close to 1, which suggests a long memory effect. Furthermore, the proof of nonlinearity Tsay test (1989) advises rejecting the null hypothesis of linearity at a significance level of 1%. Finally, in this particular example, for the WACD(1,1) model, the estimated shape parameter (α) is greater than one, indicating that the conditional hazard function of duration is monotonously increasing, whereas for the generalized gamma distribution the hazard function is inverted U–shaped[24], these are consistent with the idea of volatility clustering generated by increased trade intensity.
In appendix A, we can find the estimation results for each year. It is important to mention that if the parameters in the model are not well-estimated, then the model is not adequate for describing the behavior of the data. For example the parameters estimated are significant for all the three models for each year. However we found mixed results in the Ljung-Box test statistics and some lags of Tsay test (1989), therefore, to try to correct this problem TACD models are taken into a count.
In appendix C.1, we show that when we use different initial values for the parameters of the ACD models[25], the results are not consistent with the literature. For instance, one of the expected duration is negative and some parameter estimates are negative. For that reason it is important that the initial values of the model parameters are specified correctly.

[24]The estimated shape parameters are α and κ.

[25]We consider as an example, estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2008-2014.

Table 2. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2008-2014.

Parameter		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		2.842	2.152	2.415
		(0.000)	(0.000)	(0.000)
γ_1		0.134	0.129	0.132
		(0.000)	(0.000)	(0.000)
ω_1		0.846	0.860	0.854
		(0.000)	(0.000)	(0.000)
α		-	1.017	0.525
		-	(0.000)	(0.000)
κ		-	-	3.593
		-	-	(0.000)
$\widehat{\psi}_i$		145	188	175
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	167.656	64.163	48.159
on $\widehat{\epsilon}_i$		(0.000)	(0.000)	(0.000)
	Q(20)	396.014	130.931	99.304
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	0.101	0.000	0.000
on $\widehat{\epsilon}_i^2$		(1.000)	(1.000)	(1.000)
	Q(20)	0.274	0.000	0.000
		(1.000)	(1.000)	(1.000)
NonlinearityTests	Tar-F(1)	28.346	28.672	25.252
		(0.000)	(0.000)	(0.000)
	Tar-F(2)	4.981	16.010	15.353
		(0.000)	(0.000)	(0.000)
	Tar-F(3)	4.093	8.947	9.076
		(0.001)	(0.000)	(0.000)
	Tar-F(4)	3.541	6.416	6.672
		(0.003)	(0.000)	(0.000)
L-likelihood function		-1507776.826	-1499551.012	-1491995.210

Notes: The sample size is 264874. The p-values of Ljung-Box statistic and nonlinearity Tsay test (1989) are in parentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized residual series and its squared process are reported.

4.2.2 Threshold models

The linear ACD models are applied in many situations, but in financial applications with high frequency series, the linearity assumption could be limited. To correct this problem, we have used the TACD model, for example, Zhang, Russell and Tsay (2001) proposed these models to improve the analysis of stock transaction durations. Therefore, we consider for each year nonlinear duration models to fit better the intraday series and show that they can improve the linear ACD models.

Table 3. Estimation results of TACD model[26]

	TEACD(1,1)		TEACD (1,2)		TWACD(1,1)		TWACD (2,2)		TGACD (1,1)	
Periods	R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
2008	167	153	-	-	240	82	201	145	412	189
2009	194	116	681	114	-	-	187	79	182	71
2010	138	117	-	-	-	-	-	-	164	118
2011	-	-	291	147	-	-	-	-	-	-
2012	-	-	-	-	156	185	-	-	-	-
2013	106	235	-	-	328	249	-	-	-	-
2014	405	312	-	-	-	-	-	-	-	-

Table 3, shows the expected durations for each year, we estimate 5 models with two regimes: TEDAD (1,1), TEACD (1,2), TWACD (1,1), TWACD (2,2) and TGACD (1,1). The thresholds were calibrated so the results enhance their fit. The threshold to TEACD and TWACD models was 145 seconds while the threshold for GACD was 200 seconds. The estimate for the period 2008-2014 was not considered, because the results obtained were not encouraging and inconsistent with the literature of the TACD models. However, in order to obtain the expected duration for the entire sample, we select the TACD models that best fit the series in each year. After obtaining the expected duration per year, we perform the simple average of the duration for the full period.

As shown in appendix B, the estimation results show that the TGACD models have a better fit, for example in the years 2008, 2009 and 2010 the models correct the problem of nonlinearity at a 1% significance level. But in the other years, there is still problems of nonlinearity in some lags of Tsay test (1989).

In appendix C.2, we can find that when we use different value of the thresholds (in this case 100)[27], the results are not consistent with the literature. For example, the expected duration of the first regime for all TACD models are negative, as well as some estimated parameters in the second regimen are negative too. For that reason it is important that thresholds are calibrated.

Notably, the estimated durations are an important source of information for the calculation of realized volatility, so the average expected duration are: 279 seconds (5 minutes) to the first regime from the 2008-2014 period, while in the second regime for the same period is 145 seconds (2 minutes).

4.2.3 Realized Volatility

Given the durations of time intervals between trades the realized volatility is calculated, but first the sample must be clean of jumps taking into account the effects of periodicity. To

[26] R1 means first regime, while R2 means second regime.

[27] We consider as an example, estimation results of TACD model for the intraday range of the GILSE from 2008.

detect them, the procedure of Lee and Mykland (2008) is estimated[28]. The results according to the optimum durations are:

Table 4. Estimation results of the Lee and Mykland test for jumps for the whole sample (2008-2014) depending on its duration

	2 minutes	3 minutes	5 minutes
Number of detected jumps	648	555	397
Number of periods with at least one significant jump	589	523	385
Proportion of periods with at least one significant jump	0.34	0.30	0.22
Critical value	5.94	5.90	5.85

The test used and average Bipower variation for local robust variance, and a WSD robust non parametric periodicity filter. The critical level of the test is 0.999.

As shown in Table 4, the number of jumps increases while time interval decreases, similarly the number of periods with at least one significant jump and the proportion as well grow if the time interval fell, being the jumps more sensitive while the duration is lower. In addition, in Appendix C, it can be seen that the periodicity factor tends to resemble as time interval reduces, and for the three cases, it tend to converge, on average, to 1.00. Also, the three durations indicate that there is a strong pattern effect at the beginning of the negotiation, which effect is rapidly diluted depending on the length selected. This fact is explained because in the first minutes of trading the stocks of dual-listed markets or American Depositary Receipt (ADR), in this case with the US stock market, arbitrate so the price is equated in both markets. After cleaning the series of jumps, the next step is to calculate the realized return (RR), realized variance (RV), and the bipower variation (BV). In Appendix D can be seen the graphs of the RR, RV, BV, and the returns standardized by RV and BV, depending on the durations. As can be seen, the RR are similar for all 3 durations but the volatility magnitude of the RV and BV depends on the duration. At 2 minutes, the RV is greater than the BV, but at 3 minutes is the other way round, and at 5 minutes both volatility measures tend to be the same.

In terms of capturing volatility, both measures detected a high degree of volatility since the fourth quarter of 2008 to first quarter of 2009 (initial period of the international financial crisis), third quarter of 2009 to the end of that year (pre eurozone crisis and agricultural prices decline), and the third quarter of 2011 to the secod quarter of 2012 (post electoral restlessness and the beginning of copper and gold prices decline). In this sense, both measures are successful in capturing and preceding high market volatility periods. Given the results, the returns must be standardized by the RV, because it is a more empirical/practical measure than the BV, and use this variable as input of the non parametric VaR.

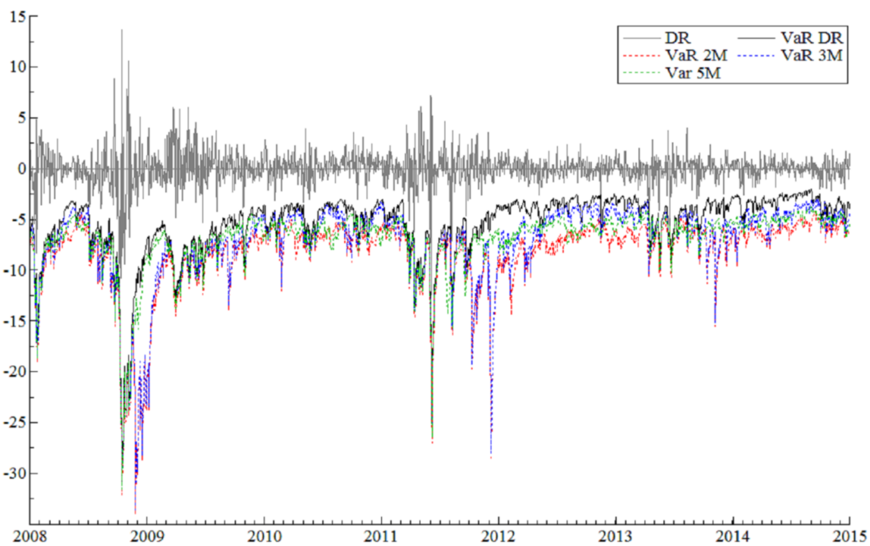
4.2.4 Value-at-Risk

Given the durations of time intervals between trades, the realized return and realized variance, the returns are standardized by the realized volatility to be inputs of the non parametric VaR. To analyze the contrast between risk calculations, a daily return (DR) basis VaR (at 99%)

[28] In Appendix C it can be seen the graphs of the jumps, and the intraday periodicity depending on the durations of time intervals between trades.

is calculated[29]. Also VaRs based on standardized returns by realized volatility, given the durations of 2 minutes (2M), 3 minutes (3M) and 5 minutes (5M), are estimated to determine the best measure of market portfolio risk.

Figure 1. Estimation results of the daily return, 2 minutes, 3 minutes, and 5 minutes basis VaR (at 99%) for the GILSE from 2008 to 2014



As shown in Figure 1, on a first instance, in terms of risk management the intraday basis VaRs requires greater economic capital to offset any expected losses. In any period the daily VaR reaches or exceeds the intraday basis VaRs. Apparently, these risk measures cover losses more adequately, particularly during periods of moderate and high volatility, this is due to the fact that the return adjusted with the realized volatility, which captures and even precede periods of high volatility returns, provides a better meterage to measure volatility by the intrinsic information inherent in the expected duration which depends on the market situation. For example, at the start of the international financial crisis the intraday basis VaRs require greater economic capital, given the volatility at that time, and the situation repeated again in 2011 by the political stress that lived Peru (presidential elections). In terms of VaR returns, the VaR at the time of the global financial crisis reported that the maximum expected loss, at 11/25/2008, was -12.1% for DR VaR, -34.0% for 2M VaR, - 33.6% for 3M VaR, and -15.5% for 5M VaR. Another example is the VaR generated by political uncertainty at election time, which reported that the maximum expected loss, at 12/09/2011, was -4.7% for DR VaR, -28.4% for 2M VaR, - 28.0% for 3M VaR, and -6.6% for 5M VaR. As can be seen, in any case the daily VaR exceed the intraday VaRs, also these show greater losses, mainly of 2 minutes and 3 minutes. Also, periods of crisis generate a nonlinear response of market risk, evidence found by comparing the values of nonlinear VaR. As can be seen, depending on the regimen,

[29] Based on the work by Rodríguez and Bedón (2015), a FIEGARCH (1,1) - Skewed Student model was estimated to capture the daily return volatility.

it requires a higher/lower level of economic capital. In this regard, a lower duration demands a higher capital requirement.
To test the adequacy of the VaR, the success/failure rate of the short/long trading positions are calculated, as well as the Kupiec test (1995), and the expected shortfalls (ESF1) and (ESF2).

Table 5. Estimation results of the VaR adequacy tests for Short positions

Quantile	Success rate	Kupiec LRT	p-value	ESF1	ESF2
		Short positions (DR)			
0.9000	0.897	0.127	0.721	2.562	1.517
0.9500	0.946	0.615	0.433	3.078	1.371
0.9750	0.971	1.129	0.288	3.518	1.273
0.9900	0.988	0.645	0.422	4.087	1.170
0.9950	0.995	0.071	0.790	4.226	1.118
0.9975	0.998	0.495	0.482	5.638	1.076
Average	0.966			3.852	1.254
		Short positions (2M)			
0.9000	0.893	0.822	0.365	3.421	1.374
0.9500	0.946	0.458	0.498	3.980	1.257
0.9750	0.975	0.000	0.985	4.537	1.205
0.9900	0.991	0.143	0.706	4.886	1.176
0.9950	0.995	0.071	0.790	5.613	1.170
0.9975	0.998	0.035	0.851	6.098	1.170
Average	0.966			4.756	1.225
		Short positions (3M)			
0.9000	0.888	2.827	0.093	2.327	1.485
0.9500	0.945	0.994	0.319	2.920	1.349
0.9750	0.976	0.018	0.893	3.571	1.289
0.9900	0.992	0.779	0.377	4.290	1.285
0.9950	0.996	0.388	0.533	4.386	1.275
0.9975	0.998	0.035	0.851	4.111	1.231
Average	0.966			3.601	1.319
		Short positions (5M)			
0.9000	0.891	1.483	0.223	3.184	1.393
0.9500	0.944	1.460	0.227	3.183	1.264
0.9750	0.974	0.104	0.747	4.241	1.219
0.9900	0.993	1.309	0.252	5.086	1.276
0.9950	0.995	0.006	0.940	5.773	1.228
0.9975	0.997	0.532	0.466	6.289	1.198
Average	0.966			4.626	1.263

Table 5 shows the estimation results of the VaR adequacy test for Short positions depending on the frequency used (daily, 2M, 3M and 5M). The first column represents the quantiles used for each VaR estimated, in the second column the succes rate of each quantile is observed, the third column shows the estimated Kupiec test with their respective p-value in the fourth column. Finally the expected shortfalls (ESF1) and (ESF2) are presented in the fifth and sixth column. The selection of the best VaR adequacy depends on the criterion shown chosen, because although exists some concordance between the measures, for example all p-values are greater than 0.05, so the null hypothesis of the Kupiec test (1995) is not rejected at a 5% significance level for any frequency, when quantiles increase the other measures begin to differ. If the 0.9000 quantile is selected, the greater success rate (0.897) is for a daily return VaR, but the ESF1 highest value (3.421) indicates that the 2M VaR is more appropriate, as ESF2 criterion lowest value (1.374) coincides with that frequency. If the 0.9900 quantile is selected, the greater success rate (0.993) is for a 5M return VaR, as ESF1 criterion highest value (5.086) matches that periodicity, but ESF2 lowest value (1.170) indicates daily return VaR is more suitable. Given above, the VaR frequency will depend on how risk adverse is the portfolio risk manager or regulator for choosing the desired quantile to select between the criteria (either Success rate, ESF1 or ESF2) for calculate its appropriate VaR. Broadly speaking, the average success rate is 0.966 for all periodicities, the largest ESF1 average (4.756) is the 2 minutes frequency VaR, as the lowest ESF2 average (1.225). So, if a short position is maintained is preferable, in average, to use a 2 minutes VaR.

Table 6. Estimation results of the VaR adequacy tests for Long positions

Quantile	Failure rate	Kupiec LRT	p-value	ESF1	ESF2
		Long positions (DR)			
0.1000	0.100	0.002	0.968	-2.673	1.517
0.0500	0.048	0.092	0.762	-3.487	1.371
0.0250	0.025	0.000	0.985	-3.835	1.273
0.0100	0.013	1.557	0.212	-4.636	1.170
0.0050	0.006	0.525	0.469	-5.889	1.118
0.0025	0.002	0.035	0.851	-5.154	1.076
Average	0.032			-4.279	1.254
		Long positions (2M)			
0.1000	0.097	0.129	0.719	-3.680	1.375
0.0500	0.047	0.275	0.599	-4.287	1.253
0.0250	0.023	0.197	0.657	-4.909	1.182
0.0100	0.006	3.883	0.049	-6.358	1.212
0.0050	0.003	0.993	0.319	-7.621	1.199
0.0025	0.002	0.035	0.851	-7.601	1.166
Average	0.030			-5.743	1.342
		Long positions (3M)			
0.1000	0.091	1.774	0.183	-2.743	1.594
0.0500	0.053	0.325	0.569	-3.434	1.414
0.0250	0.027	0.223	0.637	-4.056	1.346
0.0100	0.014	2.148	0.143	-4.843	1.230
0.0050	0.005	0.006	0.940	-5.963	1.251
0.0025	0.003	0.082	0.775	-6.788	1.214
Average	0.032			-4.638	1.340
		Long positions (5M)			
0.1000	0.099	0.014	0.905	-3.422	1.357
0.0500	0.044	1.183	0.277	-4.135	1.246
0.0250	0.019	2.468	0.116	-5.109	1.225
0.0100	0.007	1.994	0.158	-5.904	1.206
0.0050	0.003	0.993	0.319	-6.325	1.211
0.0025	0.002	0.035	0.851	-6.005	1.182
Average	0.029			-5.150	1.238

Table 6 shows the estimation results of the VaR adequacy test for Long positions depending on the frequency used (daily, 2M, 3M and 5M). The first column represents the quantiles used for each VaR estimated, in the second column the succes rate of each quantile is observed, the third column shows the estimated Kupiec test with their respective p-value in the fourth column. Finally the expected shortfalls (ESF1) and (ESF2) are presented in the fifth and sixth column. The selection of the best VaR adequacy depends on the criterion shown chosen, because although exists some concordance between the measures, for example all p-values are greater than 0.05, so the null hypothesis of the Kupiec test (1995) is not rejected at a 5% significance level for any frequency, when quantiles increase the other measures begin to differ. If the 0.1000 quantile is selected, the smaller failure rate (0.091) is for the 3M return VaR, but the ESF1 smallest value (-3.680) indicates that the 2M VaR is more appropriate, as ESF2 criterion lowest value (1.357) select the 5M VaR. If the 0.0100 quantile is chosen, the smaller failure rate (0.006) is for a 2M return VaR, as ESF1 criterion smallest value (-6.358) matches that periodicity, but ESF2 lowest value (1.170) indicates daily return VaR is more suitable. As in the short position case, the VaR frequency will depend on how risk adverse is the portfolio risk manager or regulator for choosing the desired quantile to select between the criteria (either Success rate, ESF1 or ESF2) for calculate its appropriate VaR. Broadly speaking, the minimum average failure rate is 0.029 for the 5M periodicity, the smallest ESF1 average (-5.473) is the 2 minutes frequency VaR, but the lowest ESF2 average (1.238) is at the 5 minutes periodicity. So, if a long position is maintained is preferable, in average, to use a 2 o 5 minutes VaR.

Now to test the VaR violations, that is not serially correlated, the test of Engle and Manganelli (1999) is performed. The results are detailed below:

Table 7. Estimation results of Engle and Manganelli test.

Short positions			Long positions		
Quantile	Stat.	p-value	Quantile	Stat.	p-value
Daily Returns (DR)					
0.9000	8.109	0.230	0.1000	8.501	0.204
0.9500	5.790	0.447	0.0500	13.564	0.305
0.9750	4.562	0.601	0.0250	9.974	0.126
0.9900	1.932	0.926	0.0100	6.917	0.329
0.9950	0.264	0.999	0.0050	13.776	0.032
0.9975	0.669	0.995	0.0025	0.084	0.999
2 Minutes (2M)					
0.9000	3.265	0.775	0.1000	3.735	0.712
0.9500	6.119	0.410	0.0500	3.148	0.790
0.9750	2.031	0.917	0.0250	16.701	0.010
0.9900	5.827	0.443	0.0100	21.895	0.001
0.9950	0.225	0.999	0.0050	49.765	0.000
0.9975	0.066	0.999	0.0025	0.084	0.999
3 Minutes (3M)					
0.9000	8.889	0.180	0.1000	11.447	0.076
0.9500	8.948	0.177	0.0500	16.311	0.012
0.9750	3.380	0.760	0.0250	47.220	0.000
0.9900	1.497	0.960	0.0100	28.052	0.000
0.9950	0.596	0.996	0.0050	0.245	0.999
0.9975	0.084	0.999	0.0025	0.148	0.999
5 Minutes (5M)					
0.9000	11.848	0.065	0.1000	5.068	0.535
0.9500	7.420	0.284	0.0500	7.921	0.244
0.9750	10.317	0.112	0.0250	12.501	0.052
0.9900	2.031	0.917	0.0100	13.360	0.038
0.9950	0.205	0.999	0.0050	1.393	0.966
0.9975	0.521	0.998	0.0025	0.084	0.999

In the Dynamic Quantile Regression, p=5.

Table 7 shows the estimation results of the Engle and Manganelli (1999) test violations test for Long and Short positions depending on the frequency used (daily, 2M, 3M and 5M). The first column, of any of the two subsections, represents the quantiles used for each periodicity estimated, in the second column the statistic is shown, and the third column shows the estimated p-value. For all short positions the null hypothesis of Engle and Manganelli (1999) test are not rejected at 5% significance level, so there are not VaR violations and the indicator functions are not serially correlated. But at long positions there are some drawbacks by the fact that not at all quantiles the null hypothesis is not rejected at 5% significance level. At a 10% significance level the efficiency improves, but there is still problems at 2 and 3 minutes periodicity. At 2 minutes there are problems at the quantiles 0.010 and 0.005, and at 3 minutes the 0.025 and 0.010 quantiles fail to non reject the null hypothesis of the Engle and Manganelli (1999) test.

Given the results of the Engle and Manganelli (1999) test and the VaR adequacy test previously shown, we can conclude that if a short position is maintained is preferable to use a 2 minutes VaR because, at average, it fulfill better the VaR adequacy tests (either Success rate, ESF1 or ESF2) and does not infringe any VaR violations. If a long position is maintained is preferable to use a 5 minutes VaR because, at average, it fulfill better the VaR adequacy tests (either Failure rate, ESF1 or ESF2) and does not infringe any VaR violations.

5 Conclusions and Policy Recommendations

Autoregressive conditional duration (ACD) models play an important role in financial modeling since they provide a better understanding of the timeout of each transaction into the stock market. The ACD models also show that multiple transactions can be grouped into high intensity (informed trading) and low intensity (non-informed trading) expected durations, which facilitates proper risk management and asset allocation. Above all these models are an important input to the calculation of realized volatility.

In this research, the expected duration is estimated through the ACD and TACD models to intraday series of GILSE. The average expected duration of the ACD models from 2008-2014 period is 169 seconds (3 minutes), while the average expected duration of the TACD models are 272 seconds (5 minutes) for the first regime, and 145 seconds (2 minutes) to the second one. Furthermore, we found that the TACD models are best suited to intraday series, specifically the TGACD models. However, even with these models we could not correct the problem of nonlinearity.

The high frequency analysis indicates that the number of jumps increases while time interval decreases, similarly the number of periods with at least one significant jump grow if the time interval diminishes, being the jumps more sensitive while the duration is lower. The periodicity factor tends to resemble as time interval reduces, and for the three duration cases, it tends to converge, on average, to 1. Also, the three durations indicate that there is a strong pattern effect at the beginning of the negotiation at the LSE, whose effect is rapidly diluted depending on the length selected.

In terms of risk management, the intraday basis VaRs requires greater economic capital to offset any expected losses. These risk measures cover losses more adequately, particularly during periods of low and high volatility. Periods of crisis generate a nonlinear response of market risk (evidence found by comparing the values of nonlinear VaR) this is due to the fact that the return adjusted with the realized volatility, which captures and even precedes periods of high volatility returns, provides a better meterage to measure volatility by the intrinsic information inherent in the expected duration which depends on the market situation. As can be seen, depending on the regimen, it requires a higher / lower level of economic capital. In this regard, a lower duration demands a higher capital requirement. Depending on the trading position, the best periodicity time interval to compute the VaR are, on average, 2 minutes for short positions, and 5 minutes for long positions.

According to the results obtained in this investigation and given the average expected duration of the ACD models it is much more efficient to compute an intraday VaR basis than a daily return VaR, since the first approach contains different types of measure depending on the position (short/long) of the portfolio. Additionally, the study allows to optimize the negotiation mechanism and improve the regulatory standards, while the calculation of the realized volatility generates a daily indicator, it is also possible to generate time intervals to measure the portfolio risks at a intraday level. Thus, given the duration, it is possible to generate recursive VaRs at certain time spaces.

References

[1] Agudelo, D., Giraldo, S. and Villarraga, E. (2012), “Asimetría en la información y su efecto en los rendimientos en los mercados accionarios latinoamericanos”, *Academia, Revista Latinoamericana de Administración* **50**, 100-117.

[2] Andersen, T. G, Benzoni, L. (2008), "Realized Volatility", *Federal Reserve Bank of Chicago Working Papers Series* No. 14.

[3] Andersen, T.G, Bollerslev, T. (1997), “Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns”, *Journal of Finance* **52**, 975-1005.

[4] Andersen, T. G, Bollerslev, T. (1998a), “Answering the skeptics: yes, standard volatility models do provide accurate forecast”, *International Economic Review* **39**, 885-905.

[5] Andersen, T. G, Bollerslev, T. (1998b), “Deusche Mark-Dollar volatility: intraday activity patterns, macroeconomic announcements, and longer run dependencies”, *Journal of Finance* **53**, 219-265.

[6] Andersen, T. G., Bollerslev, T., Diebold, F. X. and Labys, P. (2003), “Modeling and Forecasting Realized Volatility”, *Econometrica* **71**, 579–625.

[7] Andersen, T. G, Bollerslev, T. and Dobrev, D. (2007), “No-arbitrage semi-martingale restrictions for continuous-time volatility models subject to leverage effects, jumps and i.i.d noise: theory and testable distributional implications”, *Journal of Econometrics* **138**, 125-180.

[8] Andersen, T.G, Bollerslev, T. and Meddahi, N. (2004), “Analytic evaluation of volatility forecast”, *International Economic Review* **45**, 1079-1110.

[9] Andersen, T.G., Davis, R., KreiB, J. and Mikosch, T. (2009), “Handbook of Financial Time Series”, Springer, Berlin Heidelberg.

[10] Andersen, T. G, Teräsvirta, T. (2009), "Realized Volatility", *Handbook of Financial Time Series*, 555-575.

[11] Artzner, P., F. Delbaen, J.-M. Eber, and D. Heath (1999), “Coherent measures of risk”, *Mathematical Finance* **9**, 203–228.

[12] Barndorff-Nielsen, O.E.and Shephard, N. (2002), “Econometric analysis of realised volatility and its use in estimating stochastic volatility models”, *Journal of the Royal Statistical Society Series* **64**, 253-280.

[13] Barndorff-Nielsen, O.E.and Shephard, N. (2004a), “Econometric analysis of realised covariation: high-frequency covariance, regression and correlation in financial econometrics”, *Econometrica* **72**, 885-925.

[14] Barndorff-Nielsen, O.E.and Shephard, N. (2004b), “Power and bipower variation with stochastic volatility and jumps”, *Journal of Financial Econometrics* **2**, 1-37.

[15] Bedón, P. and Rodríguez G. (2015), "Univariate Autoregressive Conditional Heteroskedasticity Models: An Application to the Peruvian Stock Market Returns", Pontificia Universidad Católica del Perú, Economics Department Working Paper Series No. 400.

[16] Boffelli, S. (2013), "Comovements in European government bond spreads: jumps, cojumps, macrofactors and correlations", Thesis supported at the Universita degli studi di Bergamo. Available at <https://aisberg.unibg.it/retrieve/handle/10446/30383/15411/DT_Boffelli_Simona_2013.pdf>.

[17] Boudt, K., Croux, C., Laurent„ S. (2009), "Robust estimation of intraweek periodicity in volatility and jump detection", Université Lille 3, Faculty of Business and Economics Working paper. Available at <http://perso.univ-lille3.fr/~agautier/seminaire/Laurent.pdf>.

[18] Brownlees, C., and Gallo, G. (2006), "Financial Econometric Analysis at Ultra-high Frequency: Data Handling Concerns", *Computational Statistics & Data Analysis* **51**, 2232–2245.

[19] Chavez-Bedoya, L., Loaiza, C. and Téllez, G. (2015), "Precios de adjudicación, y componentes del spread en la Bolsa de Valores de Lima", *Cepal Review* **115**, 129-158.

[20] Chun, L. and Maheu, J. (2005). "Modeling and forecasting realized volatility: the role of power variation", University of Toronto Working Paper Series, <http://test.ccfr.org.cn/cicf2006/cicf2006paper/20060127131140.pdf.>

[21] Clements, A. and Liao, Y. (2013), "Modeling and forecasting realized volatility getting the most out of the jump component", NCER Working Paper Series, <http://www.ncer.edu.au/papers/documents/WP93.pdf.>

[22] Demsetz, H. (1968), "The Cost of Transacting", *Quarterly Journal of Economics* **82**, 33-53.

[23] Dufour, A. and Engle, R. (2000), "The ACD Model: Predictability of the Time Between Concecutive Trades", ICMA Centre Discussion Papers in Finance DP2000-05.

[24] Engle, R, and Russell, J. (1998), "Autoregressive conditional duration: a new model for irregularly spaced transaction data", *Econometrica* **66**, 1127-1162.

[25] Engle, R, and Russell, J. (2005), "A Discrete-State Continuous-Time Model of Financial Transactions Prices and Times: The Autoregressive Conditional Multinomial: Autoregressive Conditional Duration Model", Journal of Business & Economic Statistics **22**, 166-180.

[26] Engle, R., and S. Manganelli (1999), "CAViaR: conditional autoregressive Value at Risk by regression quantiles" Mimeo, San Diego, Department of Economics.

[27] Fernandes, M., and Gramming, J. (2005), "Nonparametric specification tests for conditional duration models", Journal of Econometrics **127(1)**, 35-68.

[28] Fleming, J., Kirby, C. and Ostdiek, B. (2003), "The economic value of volatility timing using "realized" volatility", *Journal of Financial Economics* **67**, 473-509.

[29] Garman, M. (1976), "Market Microstructure". *Journal of Financial Economics* **3**, 257-275.

[30] Giot, P. and Laurent, S. (2004), "Modelling daily Value-at-Risk using realized volatility and ARCH type models", *Journal of Empirical Finance* **11**, 379-398.

[31] Glosten, L. (1994), "Is the electronic open limit order book inevitable?", *Journal of Finance* **49**, 1127-1161.

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

[33] Hasbrouck, J. (1991a), "Measuring the information content of stock trades", *Journal of Finance* **46**, 179-207.

[34] Hasbrouck, J. (1991b), "The summary informativeness of stock trades: An econometric analysis", *Review of Financial Studies* **4**, 571-595.

[35] Hendricks, D. (1996), "Evaluation of Value-at-Risk Models using Historical Data", *Federal Reserve Bank of New York Economic Policy Review*, April 1996.

[36] Jong, F., Nijman, T., and Röell, A. (1996), "Price effects of trading and components of the bid-ask spread on the Paris Bourse", *Journal of empirical Finance* **3**, 193-213.

[37] Khan, M.A.I (2011), "Modelling daily value-at-risk using realized volatility, non-linear support vector machine and ARCH type models", *Journal of Economics and International Finance* **3**, 305-321.

[38] Kupiec, P. (1995), "Techniques for Verifying the Accuracy of Risk Measurement Models", *Journal of Derivatives* **2**, 173–84.

[39] Kuen Tse, Y., and Tao Yang, T. (2012), "Estimation of High-Frequency Volatility: An Autoregressive Conditional Duration Approach", School of Economics, Singapore Management University.

[40] Kuen Tse, Y., and Tao Yang, T. (2009), "Estimation of High-Frequency Volatility: Autoregressive Conditional Duration Models Versus Realized Volatility Methods", School of Economics, Singapore Management University.

[41] Lee, S., and Mykland P. A. (2008), "Jumps in Financial Markets: A New Nonparametric Test and Jump Dynamics", *Review of Financial Studies* **21**, 2535–2563.

[42] Loaiza, C. (2013), "Relación entre liquidez y la microestructura del mercado de renta variable al contado, en la Bolsa de Valores de Lima", Thesis supported at the Pontifical Catholic University of Peru (PUCP). Available at Thesis PUCP, <http://tesis.pucp.edu.pe/repositorio/handle/123456789/6>.

[43] Louzis, D., Refenes, A. and Xanthopoulos-Sisinis, S. (2011), “Are Realized Volatility Models Good Candidates for Alternative Value at Risk Prediction Strategies?”, available at SSRN: http://ssrn.com/abstract=1814171 or http://dx.doi.org/10.2139/ssrn.1814171.

[44] Maheu, J. and McCurdy, T. (2002), “Nonlinear Features of Realized FX Volatility”, *The Review of Economics and Statistics* **84**, 668-681.

[45] Meitz, M., and Teräsvirta, T. (2004), “Evaluating models of autoregressive conditional duration”, Working Paper Series in Economics and Finance No. 557.

[46] Protter, P. (1990), “Stochastic Integration and Differential Equations A New Approach”, Springer, Berlin Heidelberg.

[47] Scaillet, O. (2000), “Nonparametric Estimation and Sensitivity Analysis of Expected Shortfall”, Mimeo, Université Catholique de Louvain, IRES.

[48] Stoll, H. (2000), “Friction”, *Journal of Finance* **55**, 1479-1514.

[49] Tsay, R. S. (1989), “Testing and Modeling Threshold Autoregressive Processes”, *Journal of the American Statistical Association* **84**, 231-240.

[50] Tsay, R. S. (2002), “Analysis of Financial Time Series”, John Wiley & Sons, INC., Publication.

[51] Tsay, R. S. (2007), “Lecture 15: Autoregressive Conditional Duration Models”, Note based on a chapter that wrote for Handbook in Econometrics II by Palgrave Publishing Company, Bus 41910, Time Series Analysis, Mr. R. Tsay.

[52] Tsay, R. S. (2010), “Analysis of Financial Time Series”, John Wiley & Sons, INC., Publication, Third Edition.

[53] Tse, Y. and Dong, Y. (2014), "Intraday periodicity adjustments of transaction duration and their effects on high-frequency volatility estimation", Journal of Empirical Finance **28**, 352-361.

[54] Weisang, G. (2008), "ACD models: Models for data irregularly spaced in time", Department of Mathematical Sciences, Bentley College.

[55] Wong, W.,Tan, D. and Tian, Y. (2008), "Nonlinear ACD model and informed trading: Evidence from Shanghai Stock Exchange", Cardiff Economics Working Papers No. E2008/8.

[56] Zhang, Russell and Tsay (1999), “A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data”, *Graduate School of Business University of Chicago*,IL 60637.

[57] Zhang, Russell and Tsay (2001), “A Nonlinear Autoregressive Conditional Duration Model with Applications to Financial Transaction Data”, *Journal of Econometrics* **104**, 179–207.

APPENDIX A

Results of ACD model computed for each year.

Table A.1. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2008.

Parameters		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		5.202	3.686	3.135
		(0.000)	(0.000)	(0.000)
γ_1		0.198	0.180	0.179
		(0.000)	(0.000)	(0.000)
ω_1		0.752	0.791	0.798
		(0.000)	(0.000)	(0.000)
α		-	1.034	0.539
		-	(0.000)	(0.000)
κ		-	-	3.648
		-	-	(0.000)
$\widehat{\psi}_i$		104	127	136
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	10.336	6.038	5.177
on $\widehat{\epsilon}_i$		(0.412)	(0.812)	(0.879)
	Q(20)	22.007	10.321	10.296
		(0.340)	(0.962)	(0.962)
Ljung-Box	Q(10)	0.000	0.000	0.000
on $\widehat{\epsilon}_i^2$		(1.00)	(1.00)	(1.00)
	Q(20)	0.101	0.000	0.000
		(1.00)	(1.00)	(1.00)
NonlinearityTests	Tar-F(1)	2.840	6.418	6.159
		(0.014)	(0.000)	(0.000)
	Tar-F(2)	2.651	2.656	2.602
		(0.021)	(0.021)	(0.023)
	Tar-F(3)	5.511	1.116	0.987
		(0.000)	(0.349)	(0.424)
	Tar-F(4)	3.535	0.636	0.582
		(0.003)	(0.672)	(0.714)
L-likelihood function		-299272.143	-298097.885	-296176.165

Notes: The sample size is 55165. The p-values of Ljung-Box statistic and nonlinearity test of Tsayare inparentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized resudual series and its squared process.

Table A.2. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2009.

Parameters		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		5.127	3.533	3.597
		(0.000)	(0.000)	(0.000)
γ_1		0.173	0.164	0.162
		(0.000)	(0.000)	(0.000)
ω_1		0.776	0.808	0.808
		(0.000)	(0.000)	(0.000)
α		-	1.087	0.606
		-	(0.000)	(0.000)
κ		-	-	3.078
		-	-	(0.000)
$\widehat{\psi}_i$		102	127	122
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	28.206	32.327	33.867
on $\widehat{\epsilon}_i$		(0.002)	(0.000)	(0.000)
	Q(20)	41.475	42.460	44.231
		(0.003)	(0.002)	(0.001)
Ljung-Box	Q(10)	3.792	9.459	9.642
on $\widehat{\epsilon}_i^2$		(0.956)	(0.489)	(0.472)
	Q(20)	4.735	10.147	10.338
		(0.999)	(0.965)	(0.962)
NonlinearityTests	Tar-F(1)	10.172	6.431	6.627
		(0.000)	(0.000)	(0.000)
	Tar-F(2)	2.582	2.783	2.793
		(0.024)	(0.016)	(0.016)
	Tar-F(3)	2.032	4.169	4.006
		(0.071)	(0.000)	(0.001)
	Tar-F(4)	1.925	2.022	1.963
		(0.087)	(0.072)	(0.081)
L-likelihood function		-293029.602	-291321.323	-290011.971

Notes: The sample size is 53596. The p-values of Ljung-Box statistic and nonlinearity test of Tsay are inparentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized resudual series and its squared process.

Table A.3. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2010.

Parameters		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		10.352	8.037	6.669
		(0.000)	(0.000)	(0.000)
γ_1		0.197	0.184	0.178
		(0.000)	(0.000)	(0.000)
ω_1		0.718	0.756	0.774
		(0.000)	(0.000)	(0.000)
α		-	1.004	0.485
		-	(0.000)	(0.000)
κ		-	-	4.111
		-	-	(0.000)
$\widehat{\psi}_i$		121	134	139
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	13.210	45.276	39.609
on $\widehat{\epsilon}_i$		(0.212)	(0.000)	(0.000)
	Q(20)	26.655	56.022	52.143
		(0.145)	(0.000)	(0.000)
Ljung-Box	Q(10)	5.130	10.553	9.539
on $\widehat{\epsilon}_i^2$		(0.882)	(0.393)	(0.482)
	Q(20)	8.436	14.120	12.967
		(0.989)	(0.824)	(0.879)
NonlinearityTests	Tar-F(1)	1.605	22.053	21.253
		(0.155)	(0.000)	(0.000)
	Tar-F(2)	2.412	5.654	6.114
		(0.034)	(0.000)	(0.000)
	Tar-F(3)	2.758	3.936	4.046
		(0.017)	(0.001)	(0.001)
	Tar-F(4)	1.238	3.626	3.543
		(0.288)	(0.003)	(0.003)
L-likelihood function		-248002.396	-246735.454	-245283.198

Notes: The sample size is 43798. The p-values of Ljung-Box statistic and nonlinearity test of Tsay are inparentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized resudual series and its squared process.

Table A.4. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2011.

Parameter		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		0.633	0.600	0.968
		(0.000)	(0.000)	(0.000)
γ_1		0.066	0.070	0.085
		(0.000)	(0.000)	(0.000)
ω_1		0.929	0.925	0.908
		(0.000)	(0.000)	(0.000)
α		-	1.052	0.594
		-	(0.000)	(0.000)
κ		-	-	3.016
		-	-	(0.000)
$\widehat{\psi}_i$		131	130	129
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	20.830	3.471	3.118
on $\widehat{\epsilon}_i$		(0.022)	(0.969)	(0.978)
	Q(20)	42.368	7.468	6.748
		(0.002)	(0.995)	(0.997)
Ljung-Box	Q(10)	0.000	0.000	0.000
on $\widehat{\epsilon}_i^2$		(1.000)	(1.000)	(1.000)
	Q(20)	0.000	0.000	0.000
		(1.000)	(1.000)	(1.000)
NonlinearityTests	Tar-F(1)	4.816	4.402	2.077
		(0.000)	(0.000)	(0.065)
	Tar-F(2)	4.566	5.840	5.737
		(0.000)	(0.000)	(0.000)
	Tar-F(3)	1.142	1.430	1.469
		(0.336)	(0.210)	(0.196)
	Tar-F(4)	2.946	2.797	3.924
		(0.012)	(0.016)	(0.001)
L-likelihood function		-275817.769	-274620.313	-273427.007

Notes: The sample size is 49537. The p-values of Ljung-Box statistic and nonlinearity test of Tsay are inparentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized resudual series and its squared process.

Table A.5. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2012.

Parameter		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		6.033	5.196	5.666
		(0.000)	(0.000)	(0.000)
γ_1		0.070	0.070	0.071
		(0.000)	(0.000)	(0.000)
ω_1		0.896	0.900	0.897
		(0.000)	(0.000)	(0.000)
α		-	0.997	0.507
		-	(0.000)	(0.000)
κ		-	-	3.558
		-	-	(0.000)
$\widehat{\psi}_i$		179	174	176
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	19.025	126.490	113.291
on $\widehat{\epsilon}_i$		(0.040)	(0.000)	(0.000)
	Q(20)	29.674	152.678	140.905
		(0.075)	(0.000)	(0.000)
Ljung-Box	Q(10)	57.843	1780.216	1516.878
on $\widehat{\epsilon}_i^2$		(0.000)	(0.000)	(0.000)
	Q(20)	70.025	1802.800	1539.038
		(0.000)	(0.000)	(0.000)
NonlinearityTests	Tar-F(1)	2.191	1.976	1.793
		(0.052)	(0.079)	(0.110)
	Tar-F(2)	4.155	3.304	3.353
		(0.000)	(0.006)	(0.005)
	Tar-F(3)	3.753	2.469	2.573
		(0.002)	(0.030)	(0.025)
	Tar-F(4)	6.969	5.649	5.697
		(0.000)	(0.000)	(0.000)
L-likelihood function		-183566.076	-182252.427	-181585.228

Notes: The sample size is 29929. The p-values of Ljung-Box statistic and nonlinearity test of Tsay are inparentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized resudual series and its squared process.

Table A.6. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2013.

Parameter		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		5.017	4.349	5.911
		(0.000)	(0.000)	(0.000)
γ_1		0.069	0.069	0.076
		(0.000)	(0.000)	(0.000)
ω_1		0.906	0.907	0.893
		(0.000)	(0.000)	(0.000)
α		-	0.959	0.506
		-	(0.000)	(0.000)
κ		-	-	3.286
		-	-	(0.000)
$\widehat{\psi}_i$		198	188	191
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	19.014	12.270	7.253
on $\widehat{\epsilon}_i$		(0.040)	(0.267)	(0.701)
	Q(20)	55.863	36.234	25.160
		(0.000)	(0.014)	(0.195)
Ljung-Box	Q(10)	4.277	3.902	3.882
on $\widehat{\epsilon}_i^2$		(0.934)	(0.952)	(0.953)
	Q(20)	14.392	12.805	12.162
		(0.810)	(0.886)	(0.910)
NonlinearityTests	Tar-F(1)	3.480	2.851	2.257
		(0.003)	(0.014)	(0.046)
	Tar-F(2)	6.322	5.051	4.789
		(0.000)	(0.000)	(0.000)
	Tar-F(3)	2.209	1.351	1.538
		(0.051)	(0.240)	(0.174)
	Tar-F(4)	2.895	1.653	1.743
		(0.013)	(0.142)	(0.121)
L-likelihood function		-169484.496	-168089.051	-167560.092

Notes: The sample size is 27249. The p-values of Ljung-Box statistic and nonlinearity test of Tsay are inparentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized resudual series and its squared process.

Table A.7. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2014.

Parameter		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		16.308	13.078	17.859
		(0.000)	(0.000)	(0.000)
γ_1		0.087	0.084	0.092
		(0.000)	(0.000)	(0.000)
ω_1		0.852	0.868	0.843
		(0.000)	(0.000)	(0.000)
α		-	0.917	0.467
		-	(0.000)	(0.000)
κ		-	-	3.460
		-	-	(0.000)
$\widehat{\psi}_i$		265	271	274
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	5.224	4.933	3.414
on $\widehat{\epsilon}_i$		(0.876	(0.896)	(0.970)
	Q(20)	13.775	11.010	9.546
		(0.842)	(0.946)	(0.976)
Ljung-Box	Q(10)	3.314	2.590	2.201
on $\widehat{\epsilon}_i^2$		(0.969)	(0.990)	(0.995)
	Q(20)	9.728	6.716	6.172
		(0.973)	(0.998)	(0.999)
NonlinearityTests	Tar-F(1)	1.107	1.480	1.289
		(0.354)	(0.192)	(0.266)
	Tar-F(2)	1.218	1.076	1.075
		(0.298)	(0.371)	(0.372)
	Tar-F(3)	1.310	1.489	1.384
		(0.257)	(0.190)	(0.227)
	Tar-F(4)	0.608	0.551	0.557
		(0.694)	(0.738)	(0.733)
L-likelihood function		-36481.424	-36058.818	-35955.310

Notes: The sample size is 5587. The p-values of Ljung-Box statistic and nonlinearity test of Tsay are inparentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized resudual series and its squared process.

APPENDIX B

Results of threshold ACD model computed for each year.

Table B.1. Estimation results of TACD model for the intraday range of the GILSE from 2008.

Parameter		TEACD(1,1)		TWACD (1,1)		TWACD(2,2)		TGACD (1,1)	
		Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
ω		3.592	15.610	3.283	3.514	6.689	11.429	3.440	19.569
		(0.000)	(0.000)	(0.000)	(0.057)	(0.000)	(0.000)	(0.000)	(0.000)
γ_1		0.229	0.147	0.207	0.180	0.263	0.164	0.251	0.139
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
γ_2		-	-	-	-	0.098	0.099		
		-	-	-	-	(0.000)	(0.000)		
ω_1		0.757	0.751	0.779	0.777	0.235	0.537	0.741	0.758
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ω_2		-	-	-	-	0.371	0.121		
		-	-	-	-	(0.000)	(0.182)		
α		-	-	1.055	0.990	1.059	0.997	0.581	0.429
		-	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
κ		-	-	-	-	-	-	3.212	4.940
		-	-	-	-	-	-	(0.000)	(0.000)
$\widehat{\psi}_i$		267	153	240	82	201	145	412	189
		(0.000)	(0.000)	(0.000)	(0.007)	(0.000)	(0.000)	(0.013)	(0.000)
NonlinearityTests	Tar-F(1)	1.977		1.102		1.983		1.811	
		(0.079)		(0.357)		(0.078)		(0.107)	
	Tar-F(2)	1.146		0.380		2.923		1.109	
		(0.333)		(0.863)		(0.012)		(0.353)	
	Tar-F(3)	2.451		1.807		3.858		2.119	
		(0.031)		(0.108)		(0.002)		(0.060)	
	Tar-F(4)	0.831		0.735		2.594		0.705	
		(0.527)		(0.597)		(0.024)		(0.619)	
L-likelihood function		-299209.010		-297726.903		-297726.079		-296106.124	

Notes: The sample size is 55165. The p-values and nonlinearity test of Tsay are inparentheses

Table B.2. Estimation results of TACD model for the intraday range of the GILSE from 2009.

Parameter		TEACD(1,1)		TEACD (1,2)		TWACD(2,2)		TGACD (1,1)	
		Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
ω		3.411	13.552	2.085	15.860	4.172	6.549	3.708	6.753
		(0.000)	(0.000)	(0.000)	(0.057)	(0.000)	(0.009)	(0.000)	(0.000)
γ_1		0.195	0.149	0.197	0.145	0.205	0.193	0.206	0.180
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
γ_2		-	-	-	-	0.042	0.128	-	-
		-	-	-	-	(0.000)	(0.000)	-	-
ω_1		0.787	0.734	0.609	0.607	0.351	0.147	0.774	0.726
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.215)	(0.000)	(0.000)
ω_2		-	-	0.192	0.109	0.379	0.450	-	-
		-	-	(0.000)	(0.000)	(0.000)	0.182	-	-
α		-	-	-	-	1.111	1.017	0.644	0.392
		-	-	-	-	(0.000)	(0.000)	(0.000)	(0.000)
κ		-	-	-	-	-	-	2.784	6.017
		-	-	-	-	-	-	(0.000)	(0.000)
$\widehat{\psi}_i$		194	116	681	114	187	79	182	71
		(0.000)	(0.000)	(0.541)	(0.000)	(0.000)	(0.000)	(0.000)	(0.004)
NonlinearityTests	Tar-F(1)	7.826		3.413		2.304		1.927	
		(0.000)		(0.0047)		(0.042)		(0.086)	
	Tar-F(2)	1.935		3.533		1.959		0.470	
		(0.085)		(0.003)		(0.081)		(0.799)	
	Tar-F(3)	1.345		0.677		3.145		2.281	
		(0.242)		(0.641)		(0.008)		(0.044)	
	Tar-F(4)	1.011		0.337		0.722		1.602	
		(0.409)		(0.891)		(0.606)		(0.156)	
L-likelihood function		-292989.464		-214596.872		-291116.381		-289950.497	

Notes: The sample size is 53596. The p-values and nonlinearity test of Tsay are inparentheses

Table B.3. Estimation results of TACD model for the intraday range of the GILSE from 2010.

Parameter		TEACD(1,1)		TGACD (1,1)	
		Regime 1	Regime 2	Regime 1	Regime 2
ω		9.599	11.091	6.653	14.618
		(0.000)	(0.000)	(0.000)	(0.000)
γ_1		0.212	0.195	0.210	0.194
		(0.000)	(0.000)	(0.000)	(0.000)
ω_1		0.719	0.710	0.750	0.681
		(0.000)	(0.000)	(0.000)	(0.000)
α		-	-	0.535	0.382
		-	-	(0.000)	(0.000)
κ		-	-	3.495	5.849
		-	-	(0.000)	(0.000)
$\widehat{\psi}_i$		138	117	164	118
		(0.000)	(0.000)	(0.000)	(0.004)
NonlinearityTests	Tar-F(1)	1.970		1.656	
		(0.080)		(0.142)	
	Tar-F(2)	2.602		3.012	
		(0.023)		(0.010)	
	Tar-F(3)	2.538		2.834	
		(0.026)		(0.015)	
	Tar-F(4)	1.025		1.805	
		(0.401)		(0.109)	
L-likelihood function		-247999.973		-245147.159	

Notes: The sample size is 43798. The p-values and nonlinearity test of Tsay are inparentheses

Table B.4. Estimation results of TACD model for the intraday range of the GILSE from 2011.

Parameter		TEACD (1,2)	
		Regime 1	Regime 2
ω		1.567	19.215
		(0.000)	(0.057)
γ_1		0.125	0.053
		(0.000)	(0.000)
γ_2			-
		-	-
ω_1		0.444	0.761
		(0.000)	(0.000)
ω_2		0.425	0.056
		(0.000)	(0.68)
α		-	-
		-	-
κ		-	-
		-	-
$\widehat{\psi}_i$		291	147
		(0.415)	(0.000)
NonlinearityTests	Tar-F(1)	0.600	
		(0.699)	
	Tar-F(2)	5.197	
		(0.000)	
	Tar-F(3)	1.894	
		(0.092)	
	Tar-F(4)	0.580	
		(0.715)	
L-likelihood function		-90737.794	

Notes: The sample size is 49537. The p-values and nonlinearity test of Tsay are inparentheses

Table B.5. Estimation results of TACD model for the intraday range of the GILSE from 2012.

Parameter		TWACD(1,1)	
		Regime 1	Regime 2
ω		3.242	17.434
		(0.000)	(0.000)
γ_1		0.077	0.070
		(0.000)	(0.000)
ω_1		0.902	0.836
		(0.000)	(0.000)
α		0.999	1.016
		(0.000)	(0.000)
κ		-	-
		-	-
$\widehat{\psi}_i$		156	185
		(0.000)	(0.000)
NonlinearityTests	Tar-F(1)	6.694	
		(0.000)	
	Tar-F(2)	2.858	
		(0.014)	
	Tar-F(3)	2.007	
		(0.074)	
	Tar-F(4)	3.758	
		(0.002)	
L-likelihood function		-182146.205	

Notes: The sample size is 29929. The p-values and nonlinearity test of Tsay are inparentheses

Table B.6. Estimation results of TACD model for the intraday range of the GILSE from 2013.

Parameter		TEACD(1,1)		TWACD (1,1)	
		Regime 1	Regime 2	Regime 1	Regime 2
ω		4.807	11.956	57.252	43.456
		(0.000)	(0.000)	(0.000)	(0.000)
γ_1		0.042	0.065	0.396	0.125
		(0.000)	(0.000)	(0.000)	(0.000)
ω_1		0.912	0.884	0.429	0.701
		(0.000)	(0.000)	(0.000)	(0.000)
α		-	-	0.932	0.956
		-	-	(0.000)	(0.000)
κ		-	-	-	-
		-	-	-	-
$\widehat{\psi}_i$		106	235	328	249
		(0.000)	(0.000)	(0.000)	(0.000)
NonlinearityTests	Tar-F(1)	3.292		9.492	
		(0.006)		(0.000)	
	Tar-F(2)	6.223		3.144	
		(0.023)		(0.008)	
	Tar-F(3)	2.093		3.942	
		(0.063)		(0.001)	
	Tar-F(4)	3.00		3.286	
		(0.010)		(0.006)	
L-likelihood function		-169472.691		-168580.304	

Notes: The sample size is 27249. The p-values and nonlinearity test of Tsay are inparentheses

Table B.7. Estimation results of TACD model for the intraday range of the GILSE from 2014

Parameter		TEACD(1,1)	
		Regime 1	Regime 2
ω		10.900	35.518
		(0.000)	(0.000)
γ_1		0.124	0.079
		(0.000)	(0.000)
ω_1		0.849	0.807
		-	-
α		-	-
		-	-
κ		-	-
		-	-
$\widehat{\psi}_i$		405	312
		(0.000)	(0.000)
NonlinearityTests	Tar-F(1)	0.723	
		(0.606)	
	Tar-F(2)	1.051	
		(0.386)	
	Tar-F(3)	1.129	
		(0.342)	
	Tar-F(4)	0.519	
		(0.762)	
L-likelihood function		-36475.711	

Notes: The sample size is 5587. The p-values and nonlinearity test of Tsay are inparentheses

APPENDIX C

Results of ACD models that are not consistent with the literature (sensitivity to different initial value for the parameters)

Table C.1. Estimation results of EACD (1,1), WACD(1,1) and GACD(1,1) models for the intraday range of the GILSE from 2008-2014.

Parameters		EACD(1,1)	WACD(1,1)	GACD(1,1)
ω		0.538	6.153	123.135
		(0.000)	(0.000)	(0.000)
γ_1		0.122	0.177	-0.000
		(0.000)	(0.000)	(0.000)
ω_1		0.880	0.781	-0.000
		(0.000)	(0.000)	(0.000)
α		-	0.986	9.000
		-	(0.000)	(0.000)
κ		-	-	-6.999
		-	-	(0.000)
$\widehat{\psi}_i$		-237	145	123
		(0.000)	(0.000)	(0.000)
Ljung-Box	Q(10)	363.862	27.698	37554.211
on $\widehat{\epsilon}_i$		(0.00)	(0.002)	(0.000)
	Q(20)	986.549	58.229	49346.327
		(0.00)	(0.000)	(0.000)
Ljung-Box	Q(10)	0.000	0.000	0.106
on $\widehat{\epsilon}_i^2$		(0.00)	(1.00)	(1.00)
	Q(20)	0.228	0.000	0.123
		(1.00)	(1.00)	(1.00)
NonlinearityTests	Tar-F(1)	70.983	140.732	1208.474
		(0.000)	(0.000)	(0.000)
	Tar-F(2)	18.516	87.183	770.415
		(0.000)	(0.000)	(0.023)
	Tar-F(3)	7.487	29.157	577.510
		(0.000)	(0.000)	(0.424)
	Tar-F(4)	4.129	14.089	531.757
		(0.000)	(0.000)	(0.714)
L-likelihood function		-1508647.698	-1504935.724	104806275.334

Notes: The sample size is 55165. The p-values of Ljung-Box statistic and nonlinearity test of Tsayare inparentheses. The Ljung-Box statistic with Q(10) and Q(20) for standardized resudual series and its squared process.

Table C.2. Estimation results of TACD model for the intraday range of the GILSE from 2008.

Parameter		TEACD(1,1)		TWACD (1,1)		TWACD(2,2)		TGACD (1,1)	
		Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2	Regime 1	Regime 2
ω		3.153	8.900	2.863	2.523	5.004	6.572	1.912	9.924
		(0.000)	(0.000)	(0.000)	(0.013)	(0.000)	(0.000)	(0.000)	(0.000)
γ_1		0.265	0.157	0.237	0.178	0.295	0.168	0.283	0.147
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
γ_2		-	-	-	-	0.063	0.028		
		-	-	-	-	(0.000)	(0.386)		
ω_1		0.748	0.777	0.772	0.797	0.354	0.845	0.756	0.784
		(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
ω_2		-	-	-	-	0.293	-0.081		
		-	-	-	-	(0.000)	(0.474)		
α		-	-	1.047	1.039	1.052	1.044	0.586	3.144
		-	-	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
κ		-	-	-	-	-	-	0.471	4.711
		-	-	-	-	-	-	(0.000)	(0.000)
$\widehat{\psi}_i$		-246	135	-314	101	-901	161	-49.159	144
		(0.077)	(0.000)	(0.167)	(0.000)	(0.541)	(0.000)	(0.013)	(0.000)
NonlinearityTests	Tar-F(1)	2.257		1.939		2.091		3.614	
		(0.046)		(0.084)		(0.063)		(0.003)	
	Tar-F(2)	1.184		0.547		1.619		2.135	
		(0.314)		(0.741)		(0.151)		(0.058)	
	Tar-F(3)	2.215		1.566		2.920		1.680	
		(0.050)		(0.166)		(0.012)		(0.135)	
	Tar-F(4)	0.710		0.607		1.772		0.320	
		(0.616)		(0.695)		(0.115)		(0.901)	
L-likelihood function		-299193.455		-297743.243		-297695.731		1142745966.490	

Notes: The sample size is 55165. The p-values and nonlinearity test of Tsay are inparentheses

APPENDIX D

Lee and Mykland (2008) estimation of intraday jumps, and Andersen and Bollerslev (1997, 1998b) periodicity factor.

Figure D.1. Estimation of intraday jumps, and periodicity factor for 2 minutes

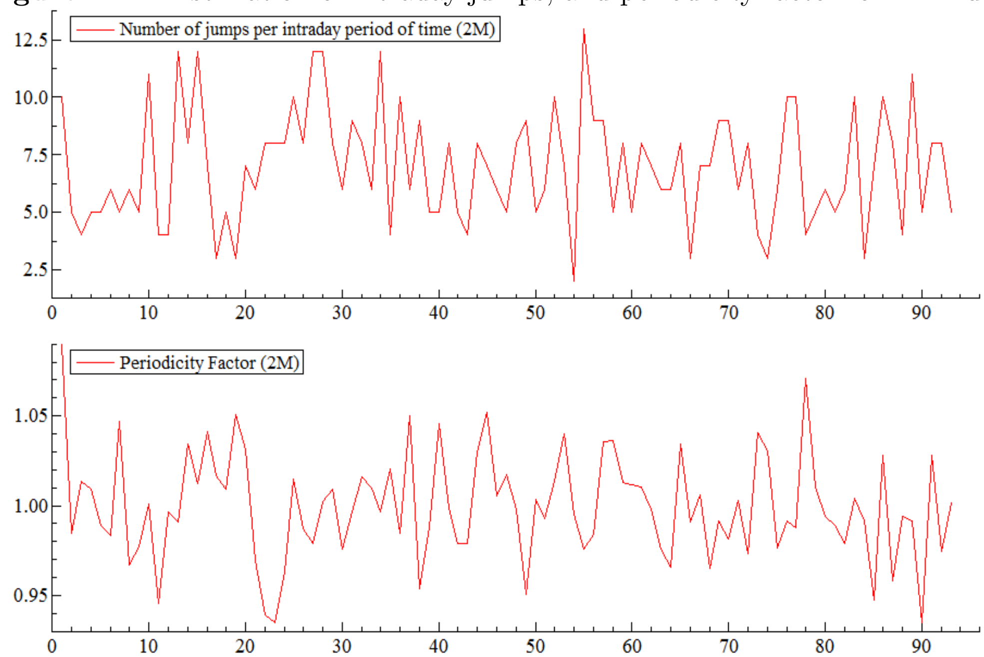


Figure D.2. Estimation of intraday jumps, and periodicity factor for 3 minutes

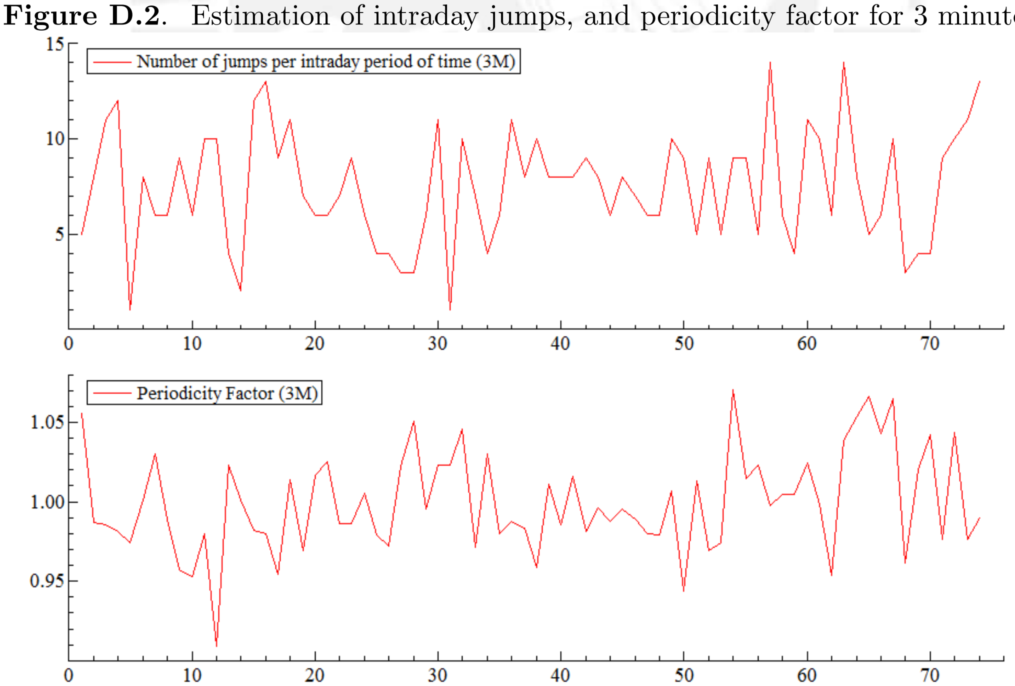
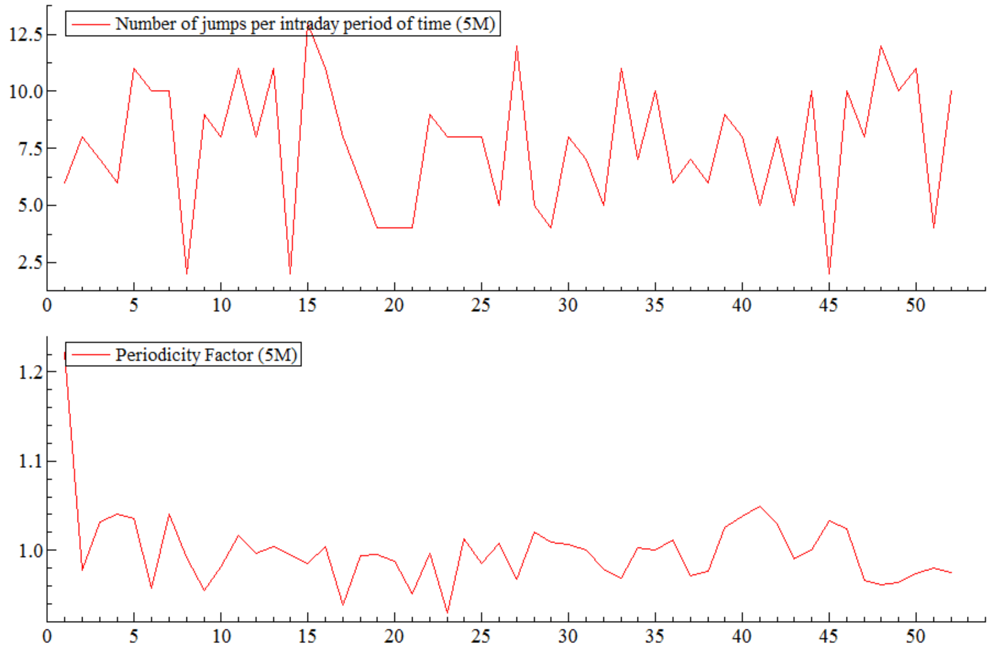


Figure D.3. Estimation of intraday jumps, and periodicity factor for 5 minutes



APPENDIX E

Realized Return (RR), Realiced Varience (RV), Bipower Variation (BV), Returns Standardized (RS) by Realized Variance and Bipower Variation.

Figure E.1. RR, RV, BV, RS by RV and BV for 2 minutes periodicity.

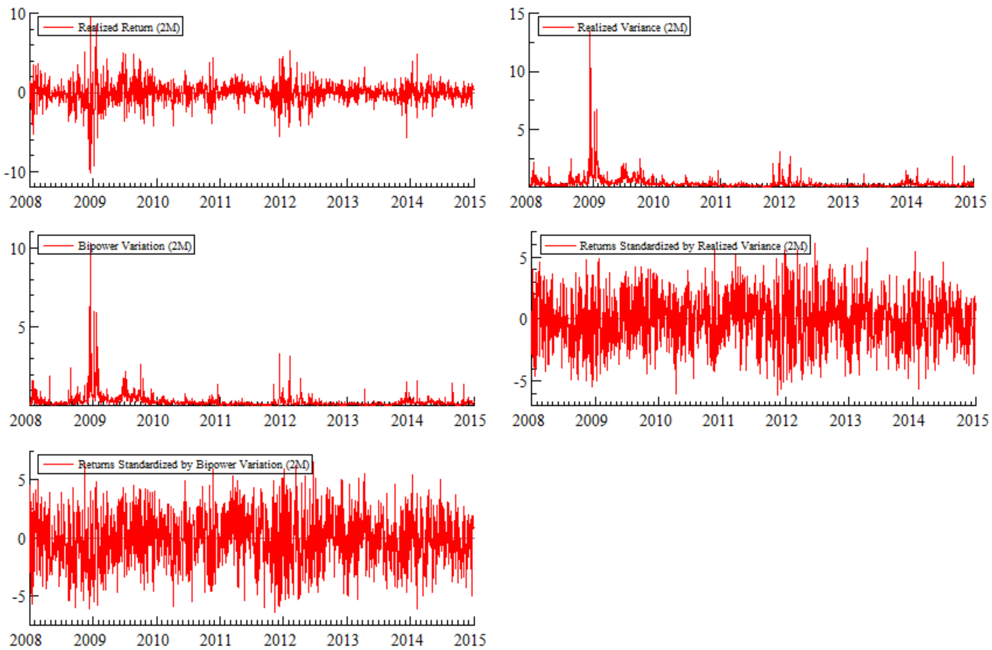


Figure E.2. RR, RV, BV, RS by RV and BV for 3 minutes periodicity.

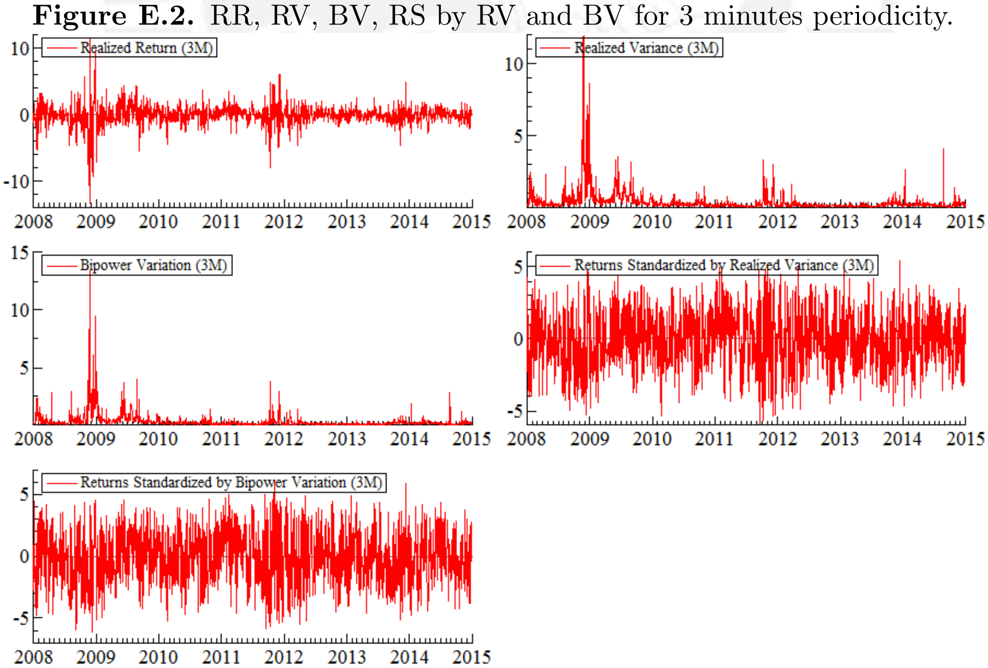


Figure E.3. RR, RV, BV, RS by RV and BV for 5 minutes periodicity.