T E C H N I C A L
R E P O R T

0682

INTRADAILY SEASONALITY OF RETURNS DISTRIBUTION. A QUANTILE REGRESSION APPROACH AND INTRADAILY VaR ESTIMATION

CORONEO L. and D. VEREDAS

I A P S T A T I S T I C S
N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

Intradaily seasonality of returns distribution. A quantile regression approach and intradaily VaR estimation

Laura CORONEO[1] and David VEREDAS[2]

September 2006

Abstract

We investigate intradaily seasonal patterns on the distribution of high frequency financial returns. Using quantile regression we show the expansions and shrinks of the probability law through the day for three years of 15 minutes sampled stock returns. Returns are more dispersed and less concentrated around the median at the hours near the opening and closing. We provide intradaily value at risk assessments and we show how it adapts to changes of dispersion over the day.

Keywords: High frequency returns, Quantile Regression, Fourier series, Intradaily VaR.

JEL classification: C14, C22, C53, G10.

[1]ECARES, Universite Libre de Bruxelles; email: laura.coroneo@ulb.ac.be.

[2]ECARES, Universite Libre de Bruxelles and CORE, Universite catholique de Louvain; email: david.veredas@ulb.ac.be.

Corresponding address: David Veredas, ECARES, Universite Libre de Bruxelles. 50 Av F.D. Roosevelt CP 114, B1050 Brussels, Belgium. Phone: +32 2 6504218. Fax: +31 2 6504475.

We are grateful to Pierre Giot, Marc Hallin and Roberto Pascual for insightful remarks. Any remaining errors and inaccuracies are ours.

This text presents research results of the Belgian Program on Interuniversity Poles of Attraction initiated by the Belgian State, Prime Minister's Office, Science Policy Programming. The scientific responsibility is assumed by the authors.

1 Introduction and Motivation

We investigate intradaily seasonal patterns on the probability law of high frequency financial returns. The interest of this research rests on two facts. First, intradaily data has become a major pole of interest for researchers and financial agents that practice intradaily trading. For instance, high frequency hedge fund managers open and close positions within the day. For these managers intradaily risk evaluation is an important tool to follow the market. Within the day there are significant variations in asset prices, which imply different evaluations of the tails of the return's distribution through the day. And these variations are partly deterministic and due to the intradaily seasonality.

Second, the analysis of risk is intimately related with the analysis of probabilities and, therefore, the analysis of the conditional probability distribution. Asset returns are realizations of a random variable and hence their behavior is fully described by their conditional probability law. Any function, such as the density function, describing this law conveys information about the likelihood that the next realization will take a certain value. But within the day these odds depend partly on a deterministic seasonal component that makes the conditional probability density function to expand or shrink as a function of the time of the day. This effect is illustrated in the return kernel densities for different times of the day shown in Figure 1. Data are 15 minutes sampled standardized returns for three stocks (large, medium and small caps) traded at the Spanish stock exchange.[1] The kernel density estimates for returns at different times of the day vary significantly. Around lunch, the density is more peaked and the tails are thinner while it is more dispersed at the hours near the opening and closing. Intradaily value at risk evaluations therefore depend very much on the time of the day. If an intradaily trader does not take this seasonality into account in their risk estimations, she will underestimate the expected loss at the opening and closing and overestimate it at noon.

[FIGURE 1 ABOUT HERE]

Moreover, not only the dispersion presents seasonal movements, but also the kurtosis. Figure 2 shows Nadaraya-Watson estimators of the normalized returns to the power of 1 to 6 on the time of the day.[2] These estimates are proxies of the intradaily behaviour of different moments. There is no evidence of an intradaily seasonal pattern in the returns themselves. However, there is a very clear U-shaped pattern in the square of returns for all the stocks, as found in many former studies. The third panels suggest that none of the stocks presents a seasonal behavior for the cubes of returns, which can be interpreted as a lack of seasonality in the skewness. Instead, the Nadaraya-Watson estimates for r_t^4 have a significant seasonal behaviour. And a similar pattern is found for r_t^6.

[FIGURE 2 ABOUT HERE]

[1] All over the analysis we use standardized returns for comparison purposes. Otherwise the scale of the plots depends on the tick size, perturbing the interpretation.

[2] The estimator is

$$y_t(d_0) = \frac{\Sigma_{t=1}^{T} K(\frac{d_t - d_0}{h}) r_t^j}{\Sigma_{t=1}^{T} K(\frac{d_t - d_0}{h})}$$

for $j = 1, \ldots, 6$ and where t is the time index and d_t is the time of the day, $K(\cdot)$ is a quartic kernel function and h is its optimal bandwidth $2.78 s T^{-1/5}$, where s is the sample standard deviation and T is the number of observations. Note that, contrary to the standard time index, the time of the day is a variable with bounded support. It increases over the day, starting at closing and until closing of the market. And the same behaviour is repeated every day, creating a sort of forth and back scheme.

The standard approach to analyze the conditional distribution function of intradaily asset returns is fitting a model for the second moment as a function of two components. One for the dynamics and another for the seasonality. If prices are random walk and returns are Gaussian, the second moment provides information enough to describe the conditional probability law, as all the odd moments are zero and the even moments are functions of the second moment. Hence, in some sense, all the even moments are conditional moments, adapting the density to every update of the conditioning set. This property of the Gaussian distribution is very appealing since, for instance, kurtosis also changes and adapts to the time of the day, consistent with the features found in data.

However, it is now commonly accepted that asset returns are not normally distributed and more flexible distributions, such as the Student's t-distribution, are needed. The drawback of these laws is that moments beyond the second are either zero -e.g. the third moment- or functions of an invariant tail index -e.g. the fourth moment. For instance, a GARCH model with a Student's t-distribution has constant kurtosis given by a function of the estimated degrees of freedom, which is not consistent with the data features.

A possible solution to overcome this problem would be to fit models for different moments, similarly to Hansen (1994), but it is not clear the functional forms that these models should take and/or which regressors to use. Since our interest is the analysis of the seasonality of the conditional distribution, a natural alternative is to model directly the conditional probability. Among all the functions that characterize the conditional probability (density, cumulative, characteristic, Laplace, hazard, etc), the conditional quantiles are the better suited due to the existence of quantile regression, introduced by the seminal work of Koenker and Basset (1978).

Indeed quantile regression has a number of useful features. First, and as mentioned earlier, quantile regression is one of the possible ways to characterize the conditional probability law and, since there is a one to one relation with all the other possible characterizations, we can, indirectly, analyze the effect of the time of the day on the density function of asset returns. Second, quantile regression does not assume the existence of any moment. In fact, it does not assume anything about the moments. Often it happens that the tails of returns are so thick that some important moments do not exist. For instance, Table 1 shows the estimated parameters of a GARCH(1,1) with Student's t-distribution. The estimated degrees of freedom for the three stocks are very low. So low that, according to the model, kurtosis does not exist for any of them and variance does not even exist for one of them.[3]

[FIGURE 3 ABOUT HERE]

Third, quantile regression is robust in the sense that the estimated coefficients are not sensitive to outliers on the dependent variable. This is particularly useful in the analysis of high frequency financial returns since often we do find outliers or, al least, observations that are remarkably different to the rest of the process. For instance, Figure 3 shows the actual returns for the three stocks that we analyze. For all there is at least one observation that is unusually high. Fourth, quantile regression is a distribution free model. This is a very compelling feature. It does not rely on any distribution specification but, ironically, it is an estimate of the conditional probability distribution. As noted earlier, and shown in Table 1,

[3] We are here abusing a bit of the estimation results. According to the table, the variance does exist for the stock ANA, as the estimated degrees of freedom is higher than two, though very close to it. However, in the software we use -Eviews- the parameter is constrained to be greater than two. Therefore, in some sense, we can expect that the unbounded estimator for the degrees of freedom to be below two. Incidentally, the fact that the estimated degree of freedom reach it lower bound, produce nonsense estimates of the GARCH equation parameters.

assuming a parametric distribution for intradaily asset returns entails a series of problems that sometimes, e.g. infinite variance, are difficult to overcome.

[TABLE 1 ABOUT HERE]

The use quantile regression in asset returns is not new. One of the first to use it are Engle and Manganelli (1999) who introduce the CAViaR (Conditional Autoregressive Value at Risk). CAViaR extends the traditional linear quantile regression to a nonlinear framework and develop a new test of model adequacy, the Dynamic Quantile (DQ) test, utilizing the criterion that each period the probability of exceeding the VaR must be independent of all the past information. Recently Guorieroux and Jasiak (2005) introduce a new dynamic quantile model univariate series and panel data as well as the Quantile Factor Model. Less related, Bouye and Salmon (2003) use quantile regression in a copula context, that is they deduce the form of the non linear conditional quantile regression implied by the copula. As for intradaily VaR, Giot (2005) quantify intradaily VaR (15 and 30 minute returns) using normal GARCH, Student GARCH, RiskMetrics and high-frequency duration (Log-ACD) models. He shows that Student GARCH model performs best. And Dionnea et al. (2005) investigate the use of tick-by-tick data for market risk measurement and propose an intradaily Value at Risk at different horizons based on irregularly time-spaced high-frequency data by using an intradaily Monte Carlo simulation.

Using quote midpoints of three stocks traded at the Spanish stock exchange from January 2001 to December 2003, we show that indeed the conditional probability distribution depends on the time of the day. At the opening and closing the density flattens and the tails become thicker, while in the middle of the day returns concentrate around the median and the tails are thinner. Results are intuitive, in the sense that they confirm the general perception that in the opening and closing the probabilities of finding large price fluctuations are higher than at lunch. Results, in terms of quantiles, permit straightforward intradaily risk evaluations, such as value at risk. We show the intradaily variation of the maximum expected loss at 95% confidence level. The maximum expected loss is, as expected, maximal at the opening and closing and minimal at lunch time. Failure rates tests confirm that the model we use produces failure rates very close to the theoretical value 5%.

The structure of the paper is as follows. Section 2 introduces further more the data and the market. Section 3 briefs quantile regression, the model that is used for estimation and how to interpret results in term of density functions. Section 4 is the empirical application and the intradaily value at risk evaluation. Section 5 concludes.

2 Market and Data

Data come from the Spanish Stock Exchange (SSE), the 9th world largest stock exchange in terms of capitalization (the 3th among continental European markets), and the 7th in terms of total value of share trading (the 3th in continental Europe) according to the World Federation of Exchanges. The Spanish stock exchange interconnection system is the electronic platform that connects, since 1995, the four exchanges that compose the SSE (Barcelona, Bilbao, Madrid, and Valencia). This system holds all the Spanish stocks that achieve pre-determined minimum levels of trading frequency and liquidity. Every order submitted to the system is electronically routed to a centralized limit order book (LOB) to proceed with its immediate execution or storage. The matching of orders is, therefore, computerized. The LOB on the brokers' screens is updated each time there is a cancellation, execution, modification or new submission. The SSE is organized as an order-driven market with a daily continuous trading session from 9:00 a.m. to 5:30 p.m. and two call auctions that determine the opening and closing prices.

During the continuous trading session, a trade takes place if an only if an order hits the quotes. Pre-arranged trades are not allowed during the continuous session, and price-improvements are impossible. There are no market makers and there is no floor trading. The market is governed by a strict price-time priority rule, but an order may lose priority if modified. Stocks are quoted in euros. The minimum price variation (tick) equals 0.01 for prices below 50 and 0.05 for prices above 50. The minimum trade size is one share. There are three basic types of orders: market, limit, and market-to-limit. Market orders are executed against the best prices on the opposite side of the book. Any excess that cannot be executed at the best bid or ask quote is executed at less favorable prices by walking down (up) the book until the order is fulfilled. Market-to-limit orders do not specify a limit price but are limited to the best opposite-side price on the book at the time of entry. Any excess that cannot be executed is converted into a limit order at that price. Finally, limit orders are to be executed at the limit price or better. Any unexecuted part of the order is stored in front of the book at the limit price. By default, orders expire at the end of the session.

The official market index of the SSE is the IBEX-35, which includes the 35 most liquid and active stocks of the exchange, weighted by market capitalization. Its composition is regularly revised every semester. Our initial sample is formed by the 35 index constituents from Jannuary 2001 to December 2003. The data used in this study consists of 15 minutes sampled quote midpoints during 3 years, from Jannuary 2001 to December 2003, of the 35 companies listed in the IBEX-35. For simplicity, we will report the analysis only on 3 of the 35 companies but the results are valid for all of them and they are available upon request. Among the 35 companies of the IBEX-35, we choose to report the results for Telefonica (TEF), Endesa (ELE) and Aciona (ANA) that are, respectively, a big, medium and small company, weighting, approximately, 20%, 6% and 0.8% in the index.

3 Quantile Regression as Density Regression

The probability law of a random variable r_t can be characterized by means of different functions. Some, like the density or the cumulative functions, are common. Others, like the quantile function, the hazard function or the characteristic function are less used. Yet, any can be written as a function of the others and hence the knowledge of one implies the knowledge of the others. The quantile function is particularly compelling in the context of conditional distributions. This is due to the existence of a solid theory on quantile regression (see Koenker, 2005, for a survey). Let $Q_{r_t}(\tau)$ be the τ-th quantile of r_t. It is well known that

$$\begin{aligned} f(r_t) &= \frac{\partial}{\partial r_t} F(r_t) \quad \text{and} \\ Q_{r_t}(\tau) &= F^{-1}(\tau) = \inf\{r_t : F(r_t) \geq \tau\}, \quad \tau \in (0,1), \end{aligned}$$

where $f(r_t)$ is the probability density function, pdf hereafter, and $F(r_t)$ is the cumulative distribution function, cdf hereafter. We can easily pass from the pdf to the quantile function. Top row of Figure 4 shows this idea. The density is symmetric around the mean, which implies that the quantile is also symmetric around the median that equals the mean. The density is centered at zero, and hence the quantile function at the median, $Q_{r_t}(0.5)$, is zero.

One may question what happens with the pdf and the quantile function if there is a location-scale shift. Second to fourth rows of Figure 4 illustrate these cases. The second row shows a positive location shift in the density, which produces a parallel upward shift on the quantile function. Or, inversely, if the quantile function shifts, the density shifts the location. It is worth noticing that after the shift the quantile function at the median, $Q_{r_t}(0.5)$, is not zero anymore as the mean in the pdf is not zero anymore either.

[FIGURE 4 ABOUT HERE]

The third row shows the effect of a scale shift in the density. This shift produces a rotation of the quantile function around the median, or, inversely, a rotation of the quantile function around the median implies a scale shift in the density.[4] The scale shift shown in the figure is counter clockwise, which implies an increase in the dispersion, as it can clearly be seen in the pdf plot. This rotation in the quantile happens when we compare the probability law at, for instance, lunch and the closing, as already noted in Figure 1. By contrast, the dispersion of the observations decreases if we compare the probability law the opening and the closing which, in terms of quantiles, means a clockwise rotation.

Finally, last row illustrates a positive location shift and a scale shift in the density, implying an asymmetric shift -a mix of shift and rotation effects- in the quantile function. More complex shifs are possible. For instance a one-sided rotation in the quantile implies an increase of the dispersion in only one side of the density, creating skewness. Fat tails can be also created in the density if the quantile are stretched only at the highest and lowest values, say 1% and/or 99% quantiles. In sum, either a location shift or a scale shift, or both, in the pdf has a clear representation in terms of quantiles, as both functions contain the same information about the random variable of interest.

The understanding of the effect of these shifts and how the quantile and the density function are affected by them is important in a conditional context. In fact, the movements in the densities of Figure 1 are produced by the intradaily seasonality. It is therefore important to model and understand how the probability distribution evolves conditional to the time of the day. Quantile regression (QR henceforth), introduced by Koenker and Bassett (1978), is the appropriate tool. The problem of finding the τ-th unconditional quantile can be expressed as the solution of a simple linear optimization problem. Generalizing these results to the case in which the quantiles are linear functions of some explanatory variables leads to the QR method. The fundamental difference of QR with respect to mean regression is that the latter considers the effect of the regressor on the mean of the regressand while QR considers the effect of the regressor on the specific τ-th quantile of the regressand. Hence, for a sufficiently narrow grid of τ, the QR method can fully describe the quantile function. The basic QR model is

$$Q_{r_t}(\tau|x_t) = \omega(\tau) + \sum_{j=1}^{J} \beta_j(\tau) x_{t,j}, \quad \tau \in (0,1), \tag{1}$$

where the intercept, $\omega(\tau)$, and the slope parameters, $\beta_j(\tau)$, are functions of τ. While in the mean regression model there is a unique parameter β_j that describes the effect that $x_{t,j}$ has on the conditional mean of r_t, in QR for each $\tau \in (0,1)$ there is a parameter $\beta_j(\tau)$ that describes the effect of $x_{t,j}$ on the τ-th conditional quantile of r_t. In other words, QR measures the effect of the regressors on each quantile of the conditional distribution of the dependent variable. In this way it allows to analyze how a shock in the regressors affects the different quantiles and hence the pdf of returns.

[4]This type of shifts are of particular relevance for this article.

The set of regressors $x_{t,j}$ is divided in two parts. One accounts for the intradaily seasonality, the main object of interest, and the second controls for the dynamics. As for the seasonality, we model them using a Fourier series of order 3

$$seas_d(\tau) = \sum_{j=1}^{3} \alpha_j(\tau) \cos\left(2\pi j \frac{d}{D}\right) + \gamma_j(\tau) \sin\left(2\pi j \frac{d}{D}\right), \tag{2}$$

where d denotes the time of the day (i.e. the sequence 09:00, 09:15,..., 17:15 that is repeated every day) and D is the number of intradaily time intervals (34 for the 15 minutes sampled returns).[5] Fourier series are convenient expressions for seasonality as the combination of cosines and sines is flexible enough to capture virtually any seasonal pattern. A remark that will be useful in interpreting the results is that the cosine of zero and multiples of 2π is one, while the sine of zero and multiples of 2π is zero. This means that the cosines peak in the market opening and in the closure, while the sines peak in the middle of the day. We may therefore expect the cosine terms to capture most of the seasonal pattern.

To control for the dynamics, we follow Koenker (2006) choosing one lag of the absoute value of returns: $\beta(\tau)|r_{t-1}|$. More lags or other functions of r_t to capture the dynamics as, for instance, square returns are possible. However, in a robust setting, the choice of absolute values is more sensible.[6]

Putting all the elements together the model we estimate is

$$Q_{r_t}(\tau|d, |r_{t-1}|) = \omega(\tau) + \sum_{j=1}^{3} \alpha_j(\tau) \cos\left(2\pi j \frac{d}{D}\right) + \gamma_j(\tau) \sin\left(2\pi j \frac{d}{D}\right) + \beta(\tau)|r_{t-1}|. \tag{3}$$

Estimation has been implemented in GAUSS using a modified version of the library Qreg.[7] Parameters are estimated using the interior point method, as described by Portnoy and Koenker (1997). The chosen grid of quantiles is (0.05, 0.10,..., 0.95). The limiting covariance matrix has been computed in GAUSS using the procedure described in the Appendix.

4 Estimation Results and Intradaily VaR

Parameters in (3) depend on the quantile τ. There are as many parameters as quantiles times the number of explanatory variables plus those in the unconditional quantile $\omega(\tau)$. Because this number may become large, in our case is 168 per stock, we follow the literature, see for instance Koenker (2006), and we present all the results graphically. This presentation nicely dovetails with Figure 4 as the interpretation of density movements in terms of quantiles applies.

Figures 5, 6 and 7 show the estimated parameters for model (3) and for TEF, ELE and ANA respectively. Every point is an estimated parameter for a different quantile. We also plot the 5% confidence bands. Top left plots show $\hat{\omega}(\tau)$ and, for comparison purposes, the quantiles of a standard normal distribution (the dotted line). For all the stocks the unconditional quantiles are significatively different from zero and from the standard normal. Following the rationale of

[5]We tried higher orders of the Fourier series but not substantial change in the results are noticed.

[6]We have also tried with more lags of absolute returns and results, available upon request, don't change qualitatively.

[7]Qreg, a GAUSS library for computing quantile regression, D. Jacomy, J. Messines and T. Roncally (2000), Groupe de Recherche Operationelle, Credit Lyonnais, France: $http://gro.creditlyonnais.fr/content/rd/home_gauss.htm$

Figure 4, the unconditional dispersion of the returns for three stocks is way larger than that of a normal distribution, confirming the well document fact that the unconditional distribution of high frequency financial returns is far from being normal. This difference is particularly remarkable for ANA, which can be explained by the fact that it is a small company. Small caps are typically less liquid and shocks have a greater effect on price variations.

The coefficients for past absolute return, $\hat{\beta}(\tau)$, are in the top right plots. For all the stocks, the magnitude of the lagged value of the return is an important source of variation. But it affects differently the different quantiles of the distribution. For instance, the median is unaffected by a shock in $|r_{t-1}|$. Equivalently to Figure 4, the plot can be seen like a rotation from the zero vertical line, while there is no location shift and hence the median remains unchanged to any value of $|r_{t-1}|$. However, it does change the distribution for any quantile beyond and below 50%. For a given past absolute return, the effect on the extreme quantiles is larger than for the quantiles near the median. Exemplifying, if return at $t-2$ was zero, the density, conditional to the time of the day, remains unchanged. If at $t-1$ there is a large movement in returns, the density becomes more sparse around the median, that remains unchanged, increasing the probabilities of finding a large price variation the next period. If, by contrast, return at t is small, the density shrinks, decreasing the probabilities of finding large price variations.

The remaining six plots show the estimated values of the parameters for the Fourier series. The first three refer to the coefficients for the cosine terms (denoted by cos1, cos2 and cos3), while the other three are for the sinus terms (denoted by sin1, sin2 and sin3). The coefficients for the cosine terms are larger, for all stocks, to the sinus ones. This is due to the fact that the cosines peak at the opening and the closing. And these are the times of the day at which trading activity and return dispersion is more intense. The coefficients for the three cosine terms are virtually identical, in shape and magnitude, for all the stocks. The first one has the larger magnitude and is easily explained by its shape: it has two peaks -in the opening and the closing- and the minimum in the middle of the day. The other two cosine terms also have two maxima at the opening and at the closing, but their magnitude is smaller. On the other side, the coefficients of the three sinus terms are smaller and display differences among the stocks. The coefficients of the first sinus series have the same shape for all the stocks but they are significant for almost all quantiles for ANA, just for few for ELE and for even fewer quantiles for TEF. While the coefficients of the other two terms display some differences in the pattern among the three stocks, even if the estimated coefficients are small in magnitude and not significative for none of the stocks. None of the coefficients is different from zero for $\tau = 0.5$, meaning that the median quantile is not affected by past observation nor the time of the day. In other words, no profit strategies based on the time of the day are found. This is coherent with the results of the Nadaraya-Watson estimation in Figure 2. Consequently, since the estimated coefficient of $|r_{t-1}|$ for $\tau = 0.5$ is zero, the conditional median is equal to the relative unconditional median: $Q_{r_t}(0.5|t, |r_{t-1}|) = \hat{\omega}(0.5)$.

This discussion on the cosines and sines of the Fourier terms is rather technical but the conclusion is clear: the density of returns does depend on the time of the day. To grasp further insights, Figure 8 shows the estimated $seas_d(\tau)$ in (2). The plots read as follows. Each line is the seasonal component for different quantiles and for different times of the day. The combination of cosines and sines of the Fourier series produces different shapes of $seas_d(\tau)$ within the day. If $seas_d(\tau)$, for a given time of the day d, is upward, the density of returns at that time of the day expands, and viceversa if it is downwards. Some conclusions can be drawn. First, the shape and the magnitude of the seasonal component is fairly similar for all the stocks. In particular, there is no seasonal behaviour at the median. But there is beyond it and becomes more remarkable as we approach the extreme quantiles. Second, the seasonal component is clearly different at the opening and the closing, with a sharp rotation downwards with respect to the rest of the day.

To better see how the conditional distribution of returns moves though the day, Figure 9 plots model (3) conditional to a particular value of past absolute returns and for different times of the day:

$$Q_{r_t}(\tau|d, \overline{|r_{t-1}|}) = \omega(\tau) + \beta(\tau)\overline{|r_{t-1}|} + seas_d(\tau).$$

The choice of the conditioning value of $\overline{|r_{t-1}|}$ has a quantitative but not qualitative effect. For a given τ, $\beta(\tau)\overline{|r_{t-1}|}$ is constant. The only effect that the chosen value of $|r_{t-1}|$, its empirical median in Figure 9, has is to increase (if big) or decrease (if small) each τ-th quantile of the same amount during the different hours of the day. In other words, it rotates uniformly all the conditional quantiles by the same amount. The figure reads as follows: The closer the line is to the vertical zero line, the more concentrated is the density around the median. And the further it is the more dispersed it is. The time of the day at which we have the largest seasonal effect is at 17:00, the closure of the market, followed by the effect at 9:00, the opening. At these times the conditonal density becomes more dispersed. In the opposite direction, for all the companies, are the seasonal effects at 13:00 and 14:00. They decrease (in absolute value) the conditional quantiles, decreasing the dispersion. This effect can be associated to a reduced trading activity during the lunch break. Interestingly enough, for TEF and ELE we observe an increase of the dispersion around 15:00, which is due to the opening of the American markets, as both stocks have American depository receipts at NYSE and are present in South America and Caribbean countries.

4.1 Intradaily Value at Risk

As shown in Section 3, there is a one to one relation among the quantile and density functions. This is particularly appealing in the construction of risk measures, which are intimately related with the analysis of the tails of the density function. Using the results of Figure 9 and equation (5) in the Appendix, we compute the conditional density at different quantiles. Figure 10 shows the tails of the densities.[8] Each point of the conditional density is derived from its relative conditional quantile. As expected, the density mass at the extremes is way larger around the opening and closing than around lunch. This seasonal tail behaviour has to be taken into account in the computation of intradaily risk measures, such as VaR.

Value at Risk was developed to provide a single number that could summarize the information about the risk in a portfolio. Over the last ten years, this technique has been increasingly used by banks and regulators all over the world as a way to estimate possible losses related to the trading of financial assets, i.e. as a tool designed to quantify and forecast market risk. In particular, the goal of VaR is to assess the possible loss that can be incurred by a trader or bank, for a given portfolio of assets, over a given time period and for a certain confidence level. The time period and the confidence level are the two major parameters that should be chosen in a way appropriate to the overall goal of risk measurement. When the primary goal is to satisfy external regulatory requirements, such as bank capital requirements introduced by the Basel Capital Accord, the confidence level is typically very small (for example, 1%) and the time horizon is long (usually a 10 day period). However for an internal risk management model, used by a company to control the risk exposure, the typical confidence level is 5% and the time horizon shorter. In particular, for active market participants such as high frequency

[8] A full picture of the density is possible but not relevant as the financial interest lies on the tails and not around the median. And, moreover, it has been shown earlier that nothing interesting happens around the median.

traders, floor traders or market makers, the time horizon of their returns is really short and the corresponding trading risk must be assessed on such short time intervals. Therefore a VaR model that characterizes the market risk on an intradaily basis is useful for market participants (such as intradaily traders and market makers) involved in frequent intradaily trading. Further general information about VaR techniques can be found in Jorion (2000).

The VaR at level τ for the given sample is the loss at the α percent probability level, which can simply defined as the empirical quantile at τ:

$$Pr[r_t < -VaR_t|d, |r_{t-1}|] = \tau$$

or $Q_{r_t}(\tau|d, |r_{t-1}|)$. From an empirical point of view, the computation of the VaR of a portfolio of assets thus requires the computation of the empirical quantile at level τ of the distribution of the future returns of the portfolio. Engle and Manganelli (1999) introduced nonlinear QR as a method for computing VaR. The originality of our model relies on two points: The use of high frequency data to forecast VaR at intradaily time horizon and the use of the Fourier series to model the intradaily seasonality of returns. Our model defines the information set available up to time $t-1$ as including the lagged absolute value of returns and the three deterministic Fourier series.

In Figure 11 we displays the first 500 observations of the 15 minutes sampled returns for TEF, ELE and ANA with the relative VaR at a confidence level of 5%. The estimated VaR shows clearly the effects of the two components that we used to model the conditional quantiles. The seasonal component is responsible of the deterministic daily oscillations, while the dynamic one is amplifying or reducing the oscillation to take into account the dispersion clustering of the data. At first sight, it looks that the estimated VaR(5%) is close enough to the data, i.e. we are not overestimating the risk, and that the number of times that we are be below are not too big. As a check we computed the failure rate, that is the percentage of times that the observation are below the VaR(5%). If the VaR(5%) is well specified, then the empirical failure rate, denoted by $\hat{f}$, should be close to the confidence level 5%. These are 5.02% for TEF, 4.95% for ELE and 4.79% for ANA. All fairly close to the theoretical confidence level. As a second, and more quantitative, check we perform the Kupiec (1995) test to see if they are significatively different from the the theoretical value of 5%. The null hypothesis is $H_0 : f = 5\%$. The 5% confidence interval for f is given by $\hat{f} \pm 1.96\sqrt{\hat{f}(1-\hat{f})/T}$. For all the stocks, we accept the null hypothesis that the empirical failure rate is equal to the theoretical one.

5 Conclusions

We investigate intradaily seasonal patterns on the probability law of high frequency financial returns. Within the day there are significant variations in asset prices, which imply different evaluations of the tails of the return's distribution through the day. And these variations are partly deterministic and due to the intradaily seasonality. As returns are realizations of a random variable and as such their behavior is fully described by their conditional probability law. To analyze the intradaily behaviour of the probabiliy law we use quantile regression, where the regressors are Fourier series that capture the time of the day and past absolute returns.

Using quote midpoints of three stocks traded at the Spanish stock exchange from January 2001 to December 2003, we show that indeed the conditional probability distribution depends on the time of the day. At the opening and closing the density flattens and the tails become thicker, while in the middle of the day returns concentrate around the median and the tails are thinner. Results are intuitive, in the sense that they confirm the general perception that in the opening and closing the probabilities of finding large price fluctuations are higher than at lunch. Results, in terms of quantiles, permit straightforward intradaily risk evaluations, such as value at risk. We show the intradaily variation of the maximum expected loss at 95% confidence level. The maximum expected loss is, as expected, maximal at the opening and closing and minimal at lunch time.

Appendix

In this appendix we describe the estimation procedure that we followed for the estimation of the asymptotic covariance matrix of the QR estimates. We follow Koenker (2005). Consider the basic model presented in equation (1). The asymptotic distribution of the QR estimator in a non-iid setting

$$\sqrt(n)(\hat{\beta}(\tau) - \beta(\tau)) \rightsquigarrow N(0, \tau(1-\tau)H_n^{-1}J_nH_n^{-1})$$

where

$$J_n(\tau) = n^{-1}\sum_{i=1}^{n} x_i x_i'$$

and

$$H_n(\tau) = \lim_{n\to\infty} n^{-1}\sum_{i=1}^{n} x_i x_i' f_i(\xi_i(\tau)) \tag{4}$$

and $f_i(\xi_i(\tau))$ denotes the conditional density of the r_t evaluated at the τ-th percent conditional quantile. The asymptotic covariance among estimates at different quantiles has blocks

$$Cov(\sqrt(n)(\hat{\beta}(\tau_i) - \beta(\tau_i)), \sqrt(n)(\hat{\beta}(\tau_j) - \beta(\tau_j))) = [\tau_i \wedge \tau_j - \tau_i\tau_j]H_n(\tau_i)^{-1}J_nH_n(\tau_j)^{-1}$$

The conditional density $f_i(\xi_i(\tau))$ in (4) is estimated using the Hendricks and Koenker (1991) sandwich form. This estimation procedure requires at first to compute the optimal bandwidth for the τ's, h_n. To do it we used the optimal bandwidth suggested by Bofinger(1975)

$$h_n = T^{1/5}\left(\frac{4.5\phi^4(\tau)\Phi^{-1}(\tau)}{(2\Phi^{-1}(\tau)^2+1)^2}\right)^{1/5}$$

where T is the sample size and ϕ and Φ are the normal pdf and cdf respectively. Last, we re-perform the QR estimation for $\tau - h_n$ and $\tau + h_n$.

As we showed in figure 4, the cdf can be obtained inverting the quantile function and, once that we have the cdf, we can recover the density function differentiating. Hendricks and Bofinger suggest

$$\hat{f}_t = \max\{0, 2h_n/(x_t'\hat{\beta}(\tau+h_n) - x_t'\hat{\beta}(\tau-h_n) - \varepsilon_t)\} \tag{5}$$

where $\hat{\beta}(\tau+h_n)$ and $\hat{\beta}(\tau-h_n)$ are the estimated parameters at $\tau - h_n$ and $\tau + h_n$ and ε_t is a small tolerance parameter that we fixed to 0.01.

References

[1] Bouye, E. and Salmon, M. (2003), "Dynamic Copula Quantile Regressions and Tail Area Dynamic Dependence in Forex Markets". Mimeo.

[2] Dionnea, G., Duchesneb, P. and Pacurara, M. (2005), "Intraday Value at Risk (IVaR) Using Tick-by-Tick Data with Application to the Toronto Stock Exchange". Mimeo Dalhousie University

[3] Engle, R.F. and Manganelli, S. (1999), "CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles" University of California at San Diego, Economics Working Paper Series 99-20, Department of Economics, UC San Diego.

[4] Giot, P. (2005), "Market risk models for intraday data", *European Journal of Finance*, 11, 309-324.

[5] Gourieroux, C. and Jasiak, J. (2005), "Dynamic Quantile Models". Mimeo.

[6] Hansen, B. E. (1994), "Autoregressive Conditional Density Estimation", *International Economic Review*, 35(3), 705-730.

[7] Jorion, P. (2000), *Value-at-Risk*, McGraw-Hill. New York.

[8] Koenker, R. and Bassett G. (1978), "Regression quantiles", *Econometrica*, 46, 33-50.

[9] Portnoy, S. and Koenker R. (1997), "The Gaussian hare and the Laplacean tortoise: computability of squared-error versus absolute-error estimators, with discussion", *Statistical Science*, 12, 279-300.

[10] Koenker R. (2005), *Quantile Regression*, Cambridge University Press.

[11] Koenker R. and Xiao, Z. (2006), "Quantile Autoregression", *Journal of the American Statistical Association*, 101, 980-990.

[12] Kupiec, P. (1995), "Techniques for verifying the accuracy of risk measurament models", *Journal of Derivatives*, vol. 2, pp. 173-84.

Figure 1: Kernel estimates at different times of the day.

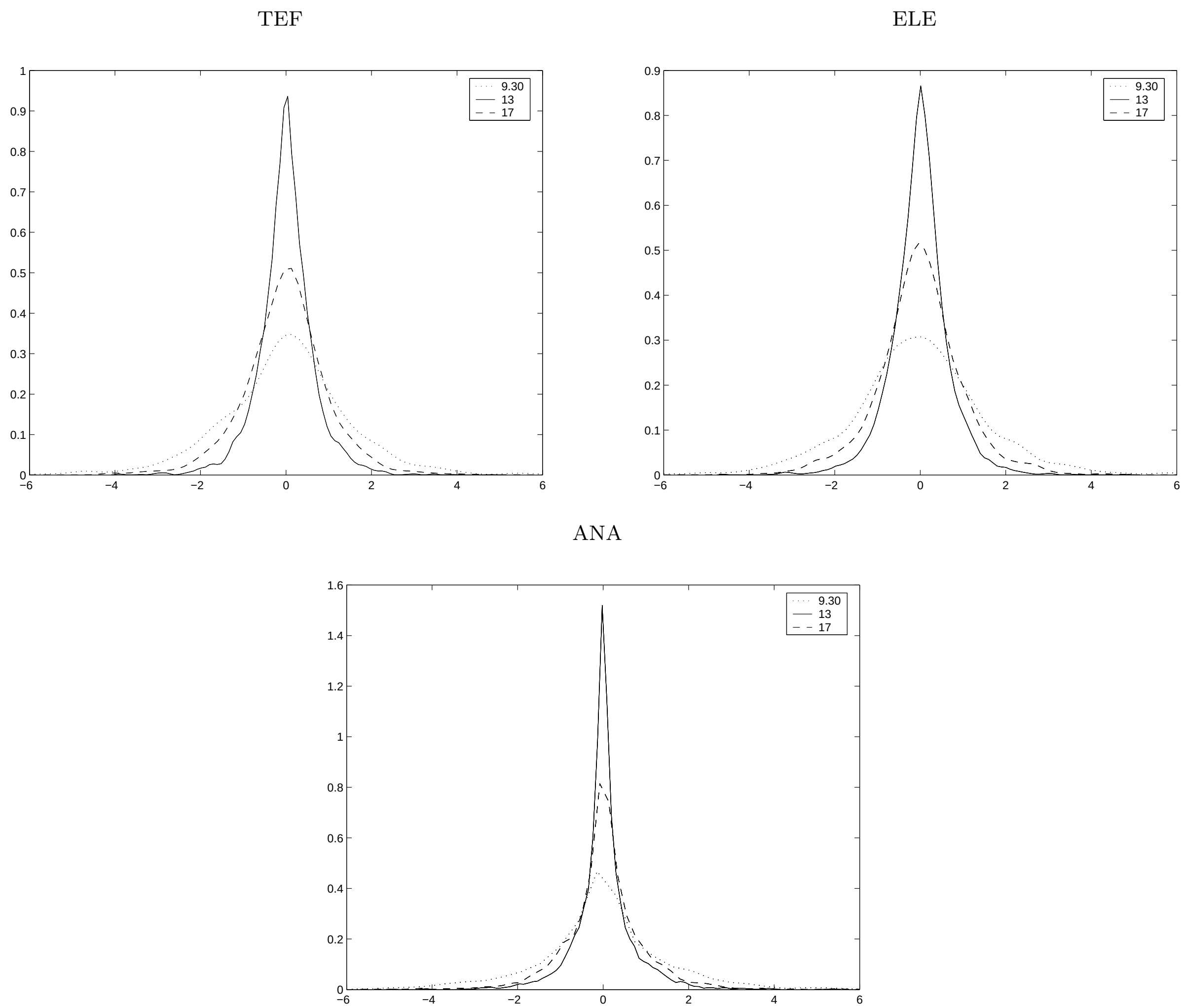
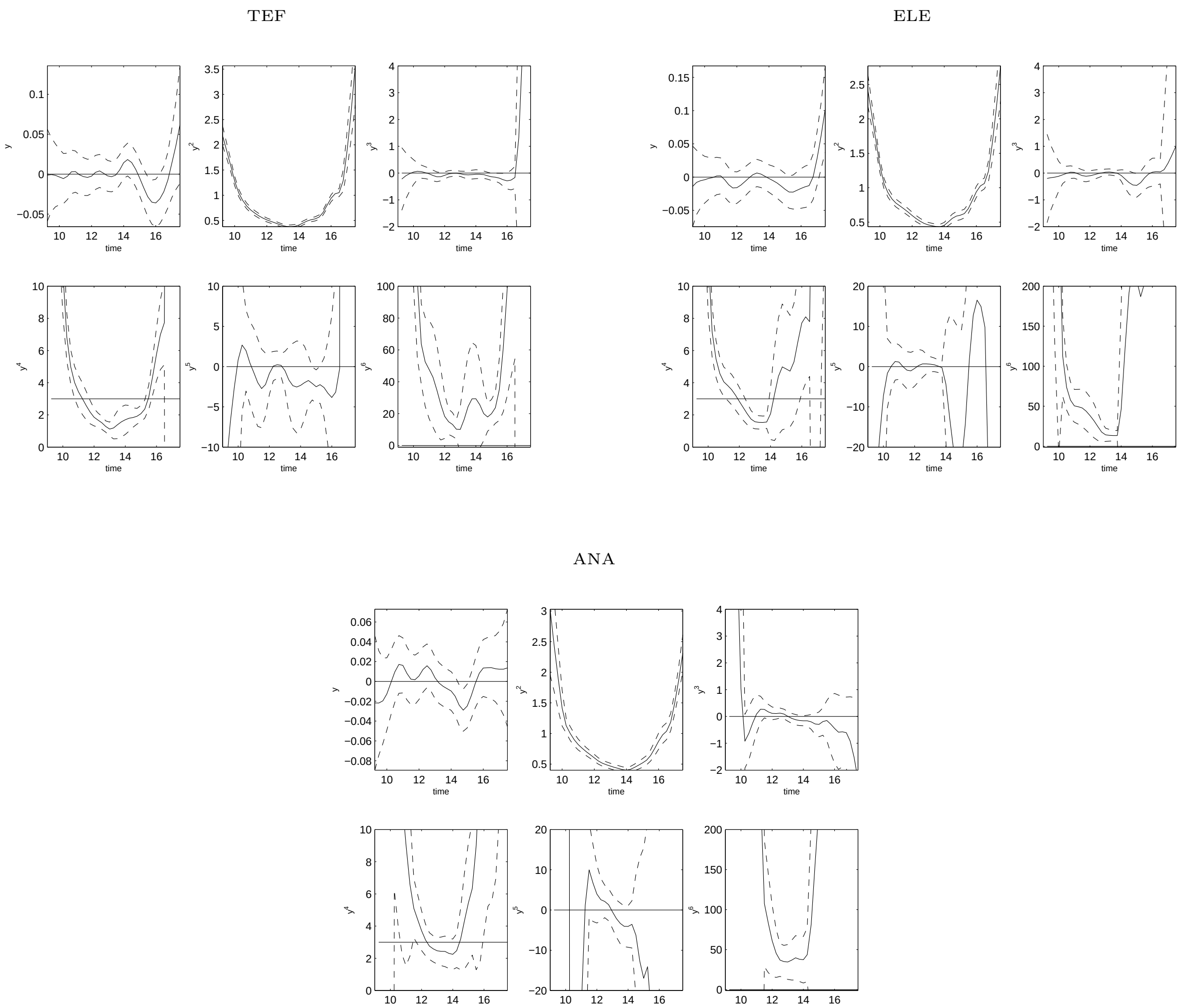


Table 1: GARCH(1,1) estimates with Students t-distribution

	ω	α	β	d.f.
ANA	3.1144	1.3777	0.5728	2.0001
ELE	0.0002	0.3666	0.6605	3.1211
TEF	0.0001	0.3507	0.7254	3.0105

Student's t GARCH(1,1) $h_t = \omega + \alpha r_i^2 + \beta h_{i-1}$ estimates. D.f. stands for degrees of freedom.

Figure 2: Nonparametric regressions of powered returns on the time of the day.



The estimates are obtained using equation (2) where t_i is the time of the day, h is the optimal bandwidth and K is the quartic density kernel. The first panel reports the results for returns, the second for squared returns and so on. The dashed lines are the 5% confidence bands and the straight line represents the theoretical value for a normal (0,1).

Figure 3: Plot of the 15 minutes returns

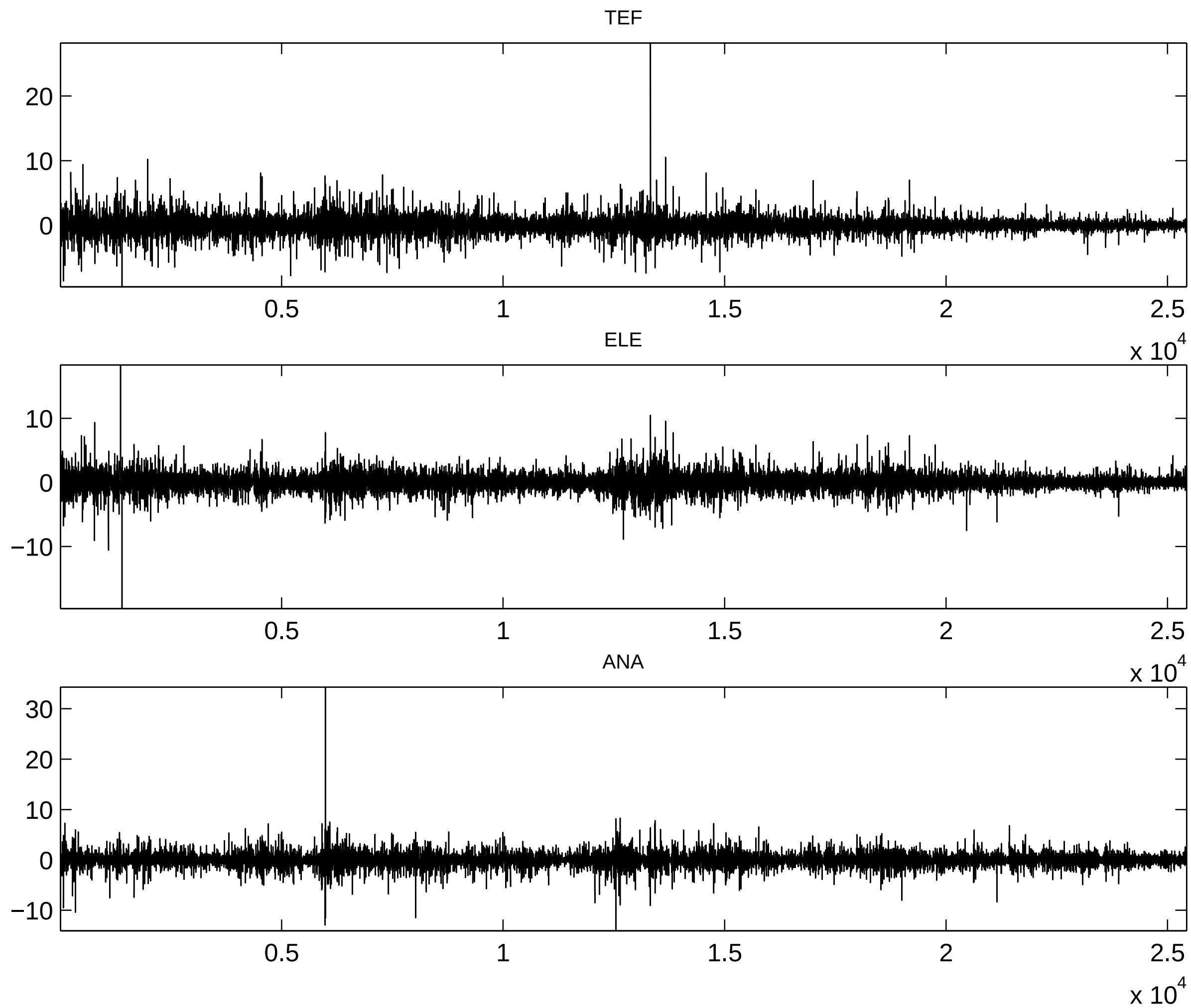
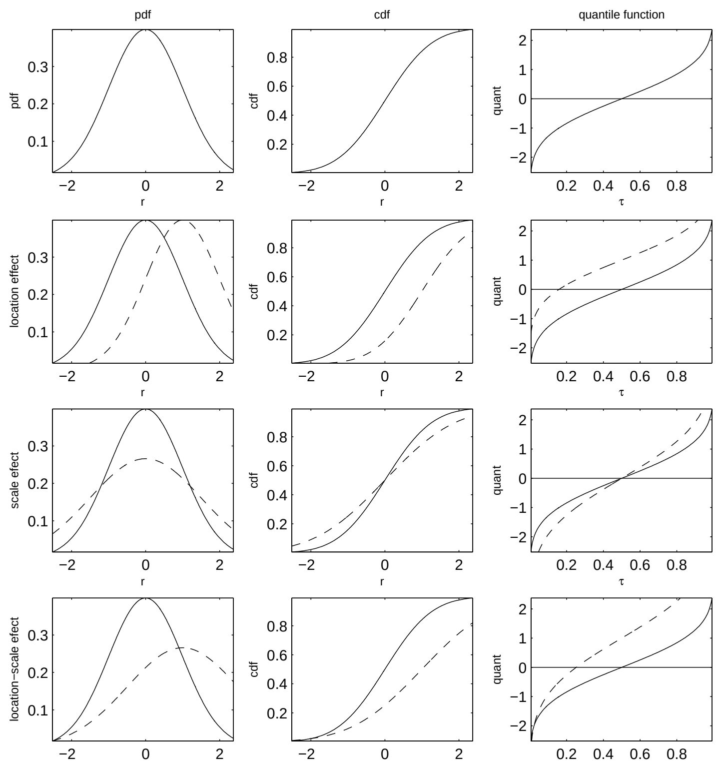
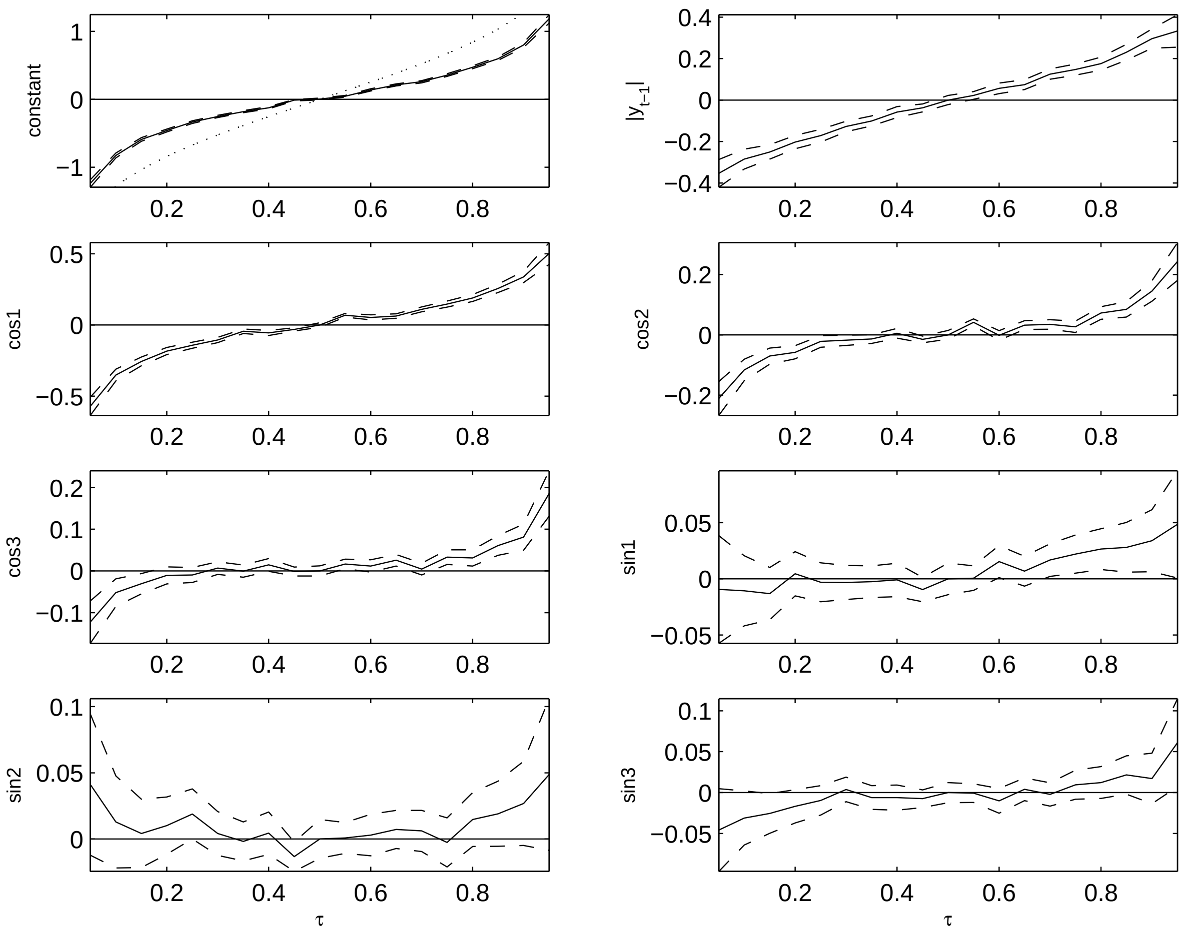


Figure 4: Changes in location and scale in the pdf through the quantile function.



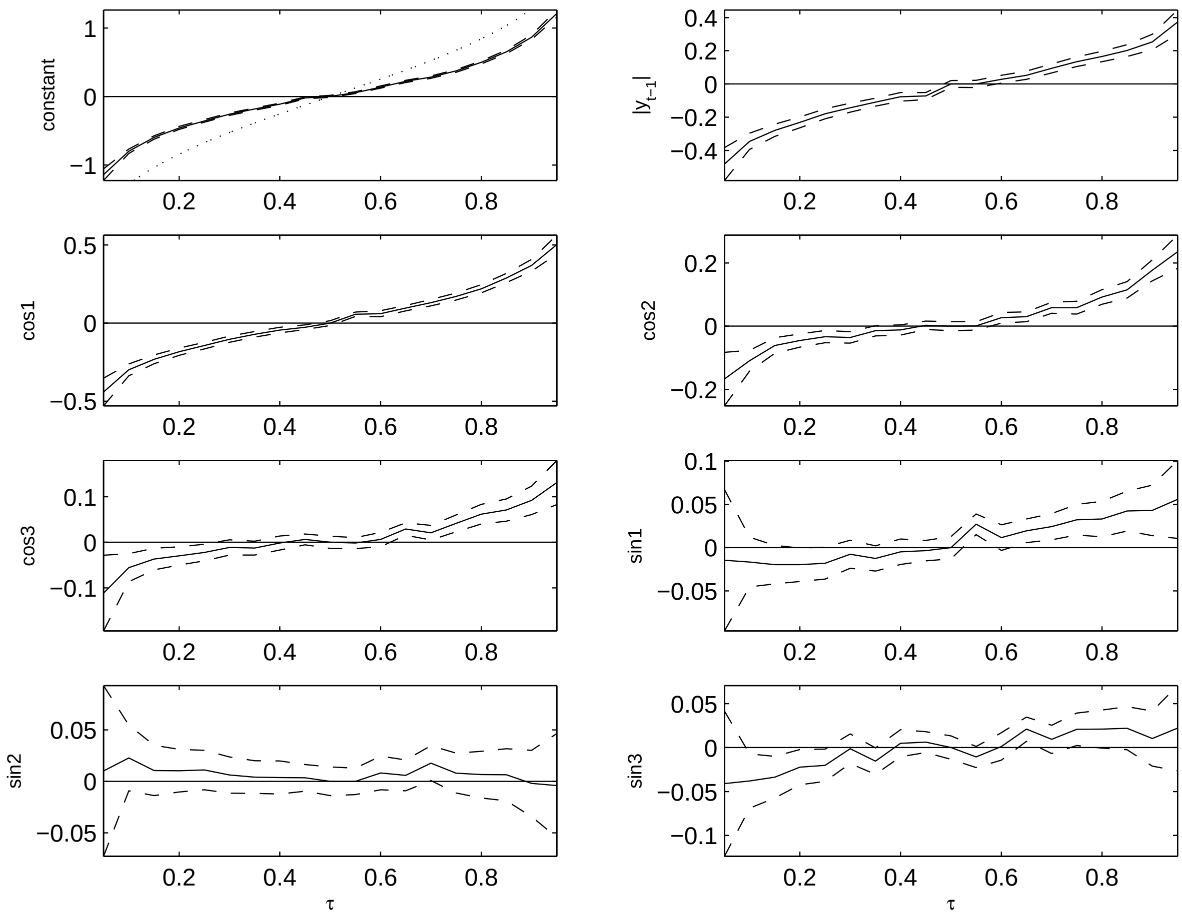
Top row shows the pdf, cdf and quantile function of a standardized normal. For the other three rows, the continuous line indicates the pdf, cdf and quantile function of the standardized normal. The dashed line in the second row refers to the pdf, cdf and quantile function of a normal with mean 1 and variance 1. In the third rown the dashed line indicates a normal with mean 0 and variance 1.5 and in the last row a normal with mean 1 and variance 1.5.

Figure 5: Estimated parameters for TEF



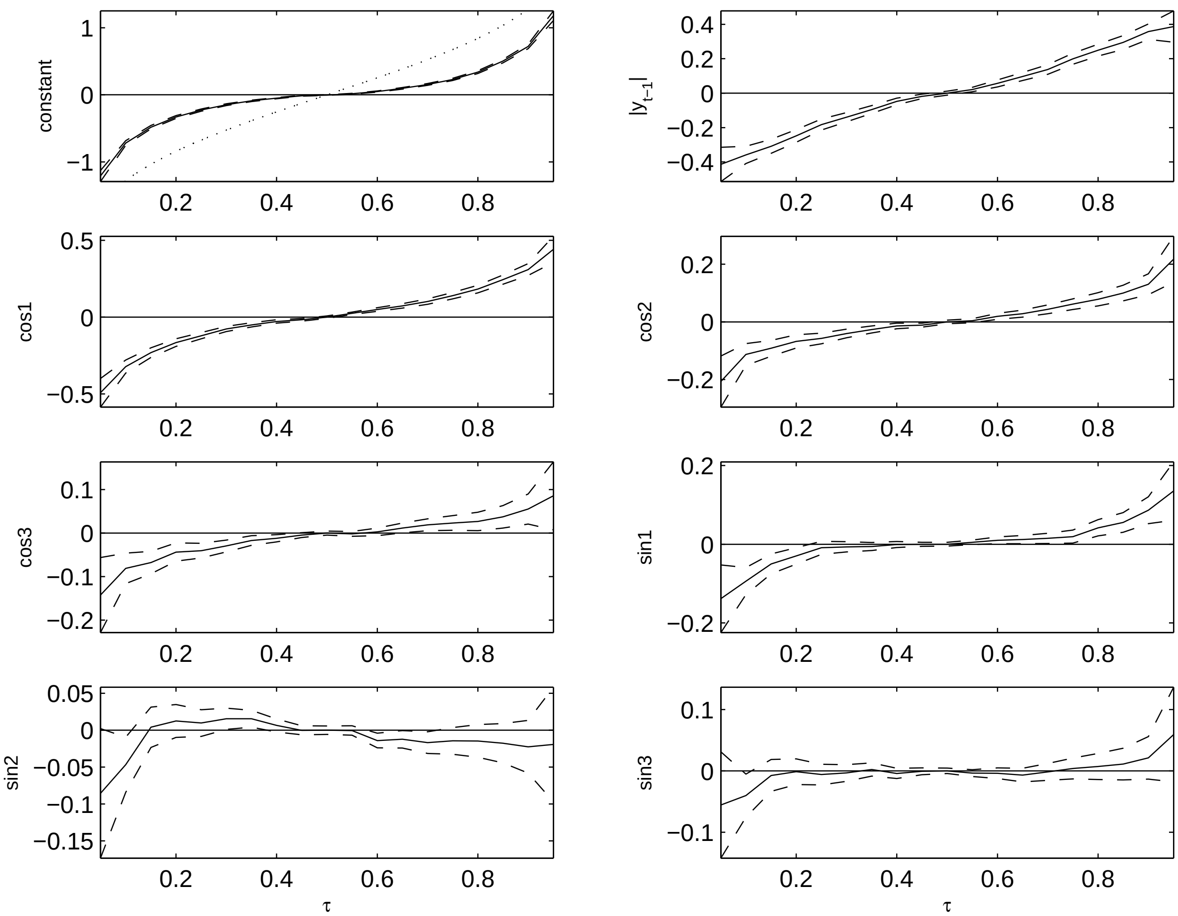
The figure displays the estimated parameters of (3). Continuous line indicates the estimated parameters for each τ quantile, while the dashed line refers to the 5% confidence bands. The top left plot also shows, in dotted line, the quantile of a standard normal distribution. Constant denotes $\hat{\omega}(\tau)$ and $\cos 1$, $\cos 2$, $\cos 3$, $\sin 1$, $\sin 2$, $\sin 3$ the estimated parameters for the cosine and sine Fourier terms respectively.

Figure 6: Estimated parameters for ELE



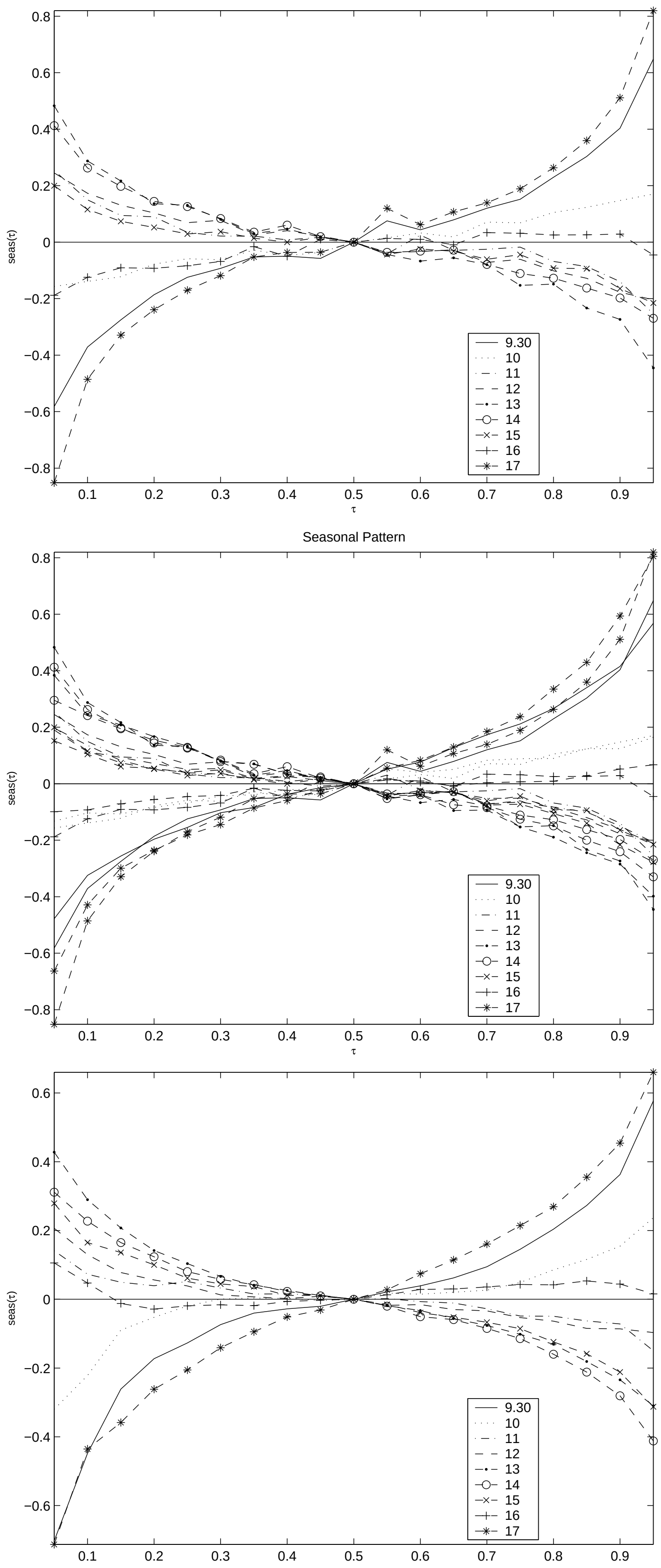
See legend Figure 5

Figure 7: Estimated parameters for ANA



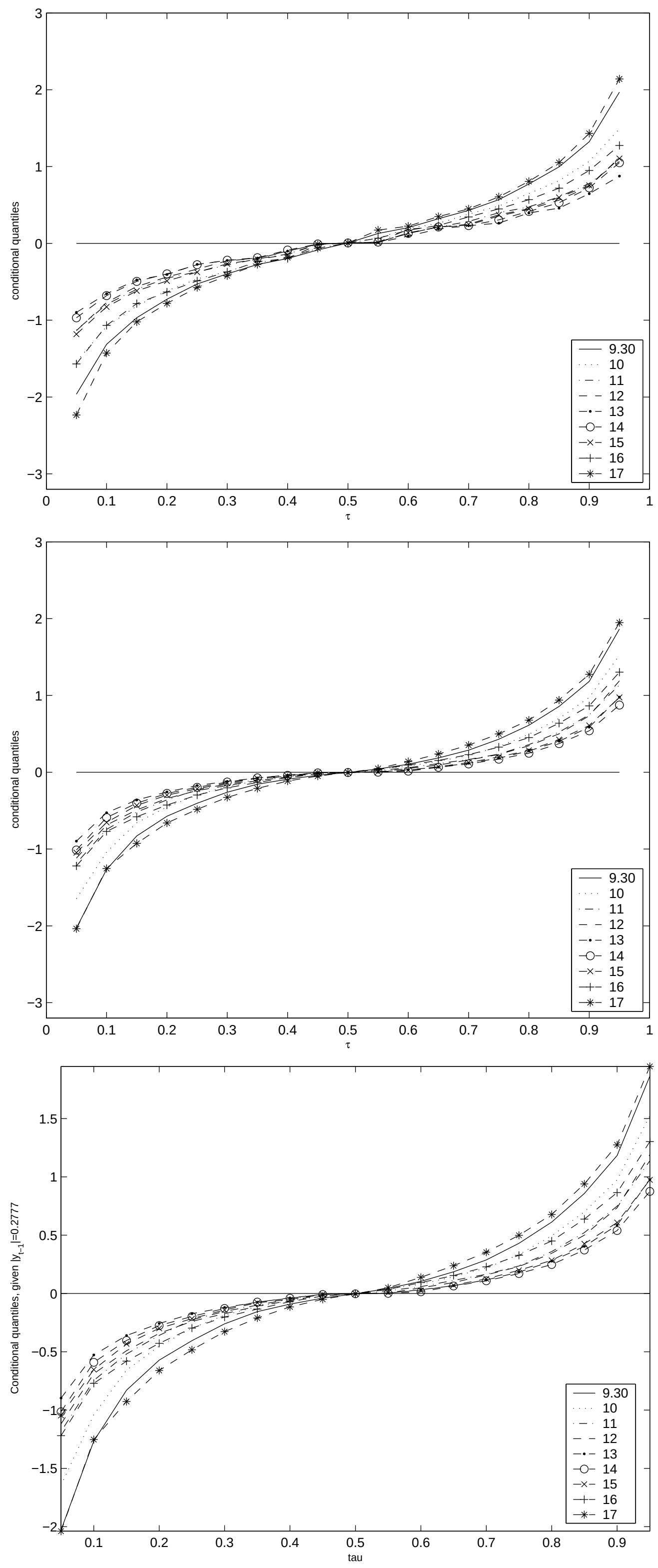
See legend Figure 5

Figure 8: Seasonal component



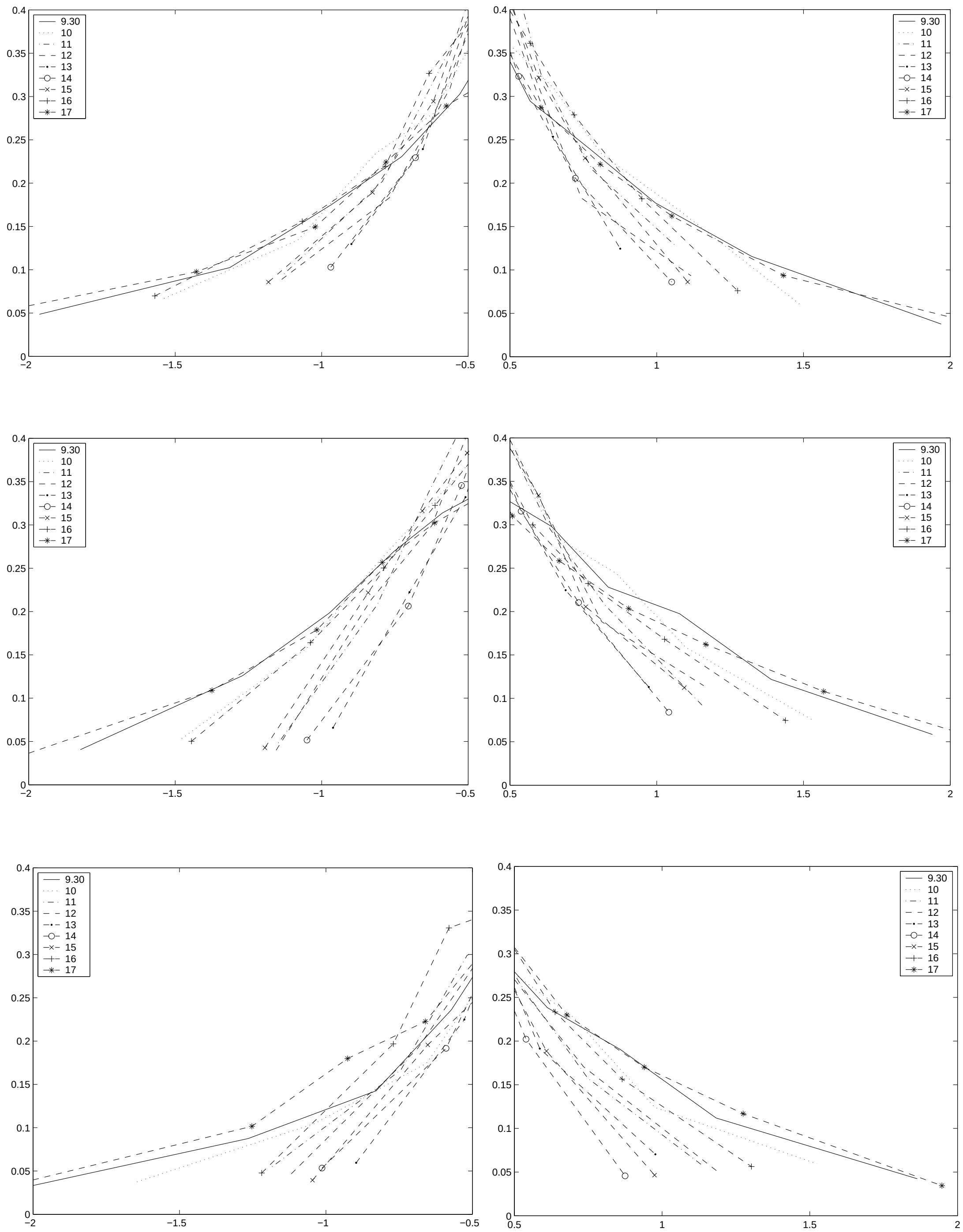
Estimated seasonal component $seas_t(\tau)$ for different times of the day. From top to bottom TEF, ELE and ANA

Figure 9: Seasonality in the quantiles



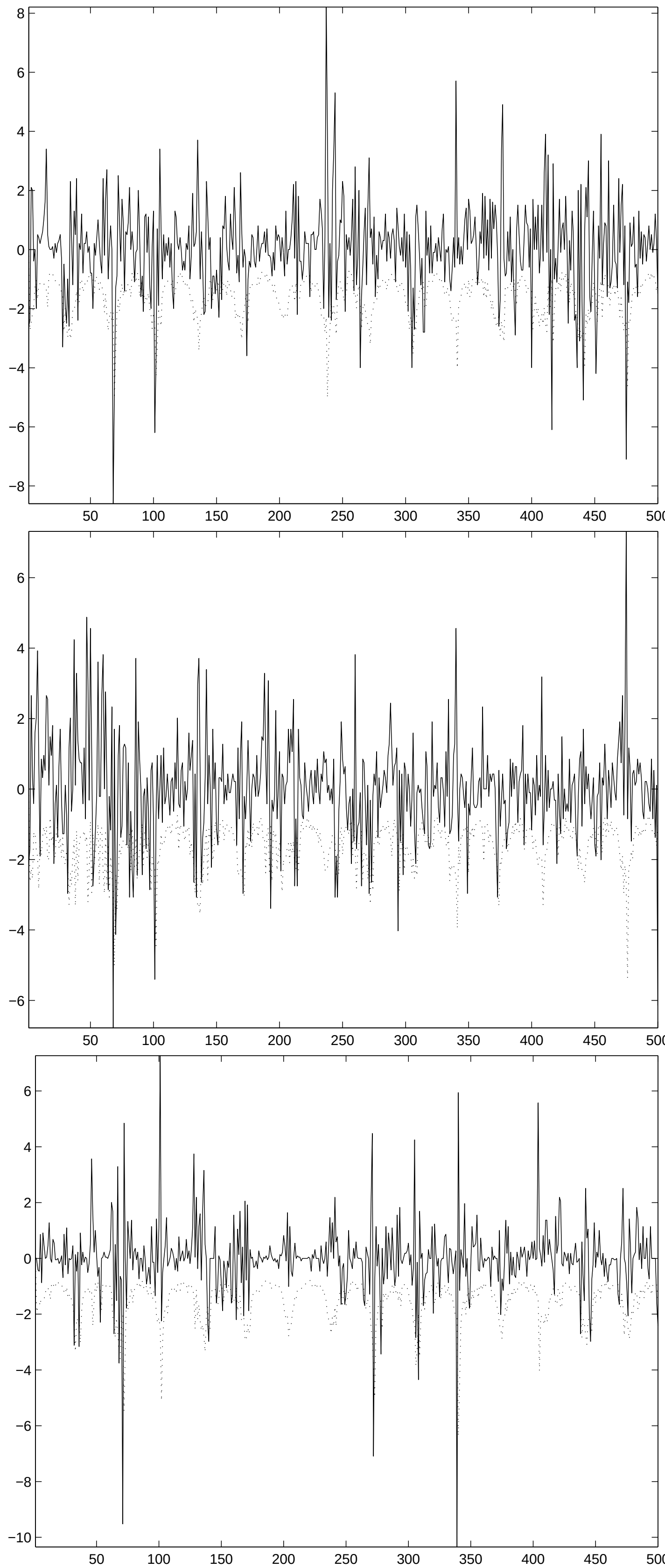
Conditional quantiles of r_t given $|r_{t-1}|$ equal to its 50 percent empirical quantile, $Q_{r_t}(\tau|t, |r_{t-1}| = Q_{|r_{t-1}|}(0.50))$, and for different times of the day . From top to bottom TEF, ELE and ANA.

Figure 10: Seasonality and the tails



Left and right tail of the conditional densities of r_t given $|r_{t-1}|$ equal to its 50 percent empirical quantile for different times of the day. The conditional density is computed using (5) in the Appendix. From top to bottom TEF, ELE and ANA.

Figure 11: 95% VaR.



First 500 observations and the 95% Value at Risk. From top to bottom TEF, ELE and ANA