

by

Oliver Linton
London School of Economics and Political Science

and

Yoon-Jae Whang
Korea University, Seoul

Contents:

Abstract

1. Introduction

2. Model and Null Hypothesis

3. Quantilogram

4. Numerical Results

5. Conclusions

6. Appendix

References

Figures 1 – 9

List of previous papers in this series

Discussion Paper
No.EM/03/463
November 2003

The Suntory Centre
Suntory and Toyota International Centres for
Economics and Related Disciplines
London School of Economics and Political Science
Houghton Street
London WC2A 2AE
Tel.: 020 7955 6698

* We would like to thank J Danielson, J M Dufour, M Halling, Y Hong, and R Koenker for interesting discussions. Oliver Linton's research was supported by the Economic and Social Research Council (ESRC) of the United Kingdom.

Abstract

In this note we propose a simple method of measuring directional predictability and testing for the hypothesis that a given time series has no directional predictability. The test is based on the correlogram of quantile hits. We provide the distribution theory needed to conduct inference, propose some model free upper bound critical values, and apply our methods to stock index return data. The empirical results suggest some directional predictability in returns, especially in mid-range quantiles like 5%-10%.

Keywords: Correlogram, dependence, efficient markets, quantiles.

JEL Nos.: C12, C13, C14, C22.

© by the authors. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without special permission provided that full credit, including © notice, is given to the source.

Contact addresses:

Professor Oliver Linton, Department of Economics, London School of Economics and Political Science, Houghton Street, London WC2A 2AE, UK. Email: o.linton@lse.ac.uk

Professor Yoon-Jae Whang, Department of Economics, Korea University, 5-1 Anamdong, Sungbukku, Seoul 136-701, Korea. Email: whang@korea.ac.kr

1 Introduction

In this note we propose a simple method of measuring directional predictability and testing for the hypothesis that a given time series has no directional predictability. There is a large literature in empirical finance attempting to find predictability in the direction of stock prices, based on the statistics of signs and ranks. Cowles and Jones (1937) proposed a statistic for testing market efficiency based on the frequencies of up movements relative to down movements. They computed the ‘C-J ratio’ for a number of stock price indexes, finding some evidence of predictability relative to what would be expected under a null hypothesis of i.i.d. mean zero. These and other results are reviewed in Campbell, Lo, and MacKinlay (1997, §2.2.2). They point out that when there is a non zero drift in stock returns the sign of returns has a non-zero mean and this could account for some of the earlier violations. They compute the distribution of the C-J ratio under the hypothesis of i.i.d. normal returns with drift and show how the limiting distribution can be corrected for drift.¹ Of course, this distribution and so the correction is heavily dependent on the normality assumption. Instead, one can just correct empirically the test statistic for the nonzero mean of signed returns. In that case, under the null hypothesis of i.i.d. returns there should be no predictability in the signed return series. Christoffersen and Diebold (2002) have recently investigated the predictability of signed returns under more general sampling schemes than were contemplated in this earlier work. For the most part they assumed that returns were conditionally normal but allowed for time varying volatility. They show that: (1) volatility dependence can induce sign dependence if expected returns are non-zero; (2) no mean dependence [market efficiency] is consistent with sign dependence and volatility. Thus a naive test of sign dependence is unlikely to reveal anything about market efficiency, unless we truly believe in a very simple null hypothesis.

Hong and Chung (2003) propose a new method for testing predictability of the direction of stock returns relative to a ‘fixed’ threshold. Their test is based on a generalized spectrum: it takes account of many lags and is consistent against a wide class of alternatives. They find evidence of predictability for a number of daily U.S. stock indexes.

We propose a simple diagnostic statistic for measuring the extent of directional predictability based on a sample correlation. In contrast to the fixed threshold of Hong and Chung (2003) we take our threshold to be an unconditional quantile. Our null hypothesis is thus that the chosen conditional quantile is not time varying. In the case of the median we are looking at the autocorrelation of returns signed relative to their unconditional median rather than the raw signs used in the C-J test and Diebold and Christoffersen (2002). We look at individual correlations but also aggregate into Box-Pierce type statistics that take account of a number of lags. In practice we must replace

¹They also show in a simple example that the C-J test will fail to pick up a simple 2 state markov process.

the population quantile by an estimate. This in general affects the limiting distribution, and our theory captures the leading effect of this estimation.² In some special cases that effect disappears. The advantage of our approach is: (a) conceptual - using the quantile in connection with counts is preferable from a statistical perspective to using a fixed threshold whose meaning is uncertain and depends on the time frame etc; (b) simplicity in computation and interpretation; (c) correct and simple asymptotic theory. Specifically, we give ‘model free’ upper bound critical values. We apply our test statistic to a sample of daily, weekly, and monthly returns on the S&P500. We find strong evidence of predictability in the high frequency data when a number of lags are take into account, and almost no evidence in monthly data. This seems to be contrary to some further predictions of Christoffersen and Diebold (2002): (3) sign dependence is not likely to be found via analysis of sign autocorrelations because the nature of sign dependence is nonlinear; (4) sign dependence is not likely to be found in high frequency data but more likely to be found in data with frequency of two or three months.

2 Model and Null Hypothesis

Suppose that random variables y_1, y_2, \dots are from a stationary process whose marginal distribution has quantiles μ_α for $0 < \alpha < 1$. Our null hypothesis is that some conditional quantiles are time invariant, which can be written more formally as: for some α

$$E[\psi_\alpha(y_t - \mu_\alpha)|\mathcal{F}_{t-1}] = 0 \text{ a.s., where } \psi_\alpha(x) = 1(x < 0) - \alpha \quad (1)$$

denotes the check function, while $\mathcal{F}_{t-1} = \sigma(y_{t-1}, y_{t-2}, \dots)$. One could call y_t a *quantilegale*, or in the special case where $\alpha = 1/2$, a *mediangale*. Under this hypothesis, if you are above the unconditional α -quantile today, the chance is no more than α that you will be above it tomorrow. In the absence of this property there is obviously some predictability in the process. We can distinguish between the cases where the hypothesis is about a particular quantile, about a set of quantiles, or when it is about all quantiles. The latter hypothesis is obviously much harder to satisfy, and is equivalent to y_t being i.i.d. We are just going to consider the single α case, although in the empirical work we look at a number of quantiles simultaneously.

Compare (1) with the usual weak form efficient markets hypothesis that for some μ ,

$$E[y_t - \mu|\mathcal{F}_{t-1}] = 0. \quad (2)$$

²Our approach is related to that taken in Engle and Manganelli (1999, §4) except that they take a more regression-based framework. Also, they do not present the full distribution theory for their test [Engle and Manganelli (1999, pp 25-26)].

It could be that the median is time invariant but the mean is time varying or vice versa. Under symmetry there is a one to one relationship between (2) and (1) with $\alpha = 1/2$, and in practice symmetry can be approximately true for many financial series. Bassett, Koenker, and Kordas (2003) show that a particular theory of decision making under uncertainty leads to a quantile regression. This gives some additional justification for looking at quantiles. To study the important concept of Value at Risk, Engle and Manganelli (2001) propose a class of models that makes the conditional quantiles time varying through past observations of y and past values of the conditional quantiles themselves.

The null hypothesis (1) is quite broad and includes many dependent processes. For example, suppose that

$$y_t = \mu_\alpha + \varepsilon_t \sigma_t, \quad (3)$$

where ε_t are i.i.d. with α -quantile zero for some single α , while σ_t^2 is some volatility process: stationary and measurable with respect to \mathcal{F}_{t-1} . In the case of symmetric ε_t distribution and $\alpha = 1/2$ this would include the standard strong GARCH process, and the process y_t is consistent with the usual efficient markets hypothesis. The process (3) satisfies (1) even when there is considerable dependence in the process through σ_t^2 . The process (3) is quite general, since we do not specify σ_t^2 . It is a more general straw man than the traditional i.i.d. assumption.³ If $\mu_\alpha = 0$ then no matter what value the mean takes the sign sequence is independent over time, which is contrary to finding (1) of Christoffersen and Diebold (2002).⁴

Compare our notion of predictability with that used in Hong and Chung (2003), which replaces μ_α by some fixed threshold c .⁵ Note that even if (3) is satisfied, then the process $1(y_t < c)$ will be predictable in the sense of Hong and Chung (2003) for any $c \neq \mu_\alpha$. Note that if (3) holds for some quantile α , then the conditional quantile at another α' is time varying, so we are subject to the same issues as Hong and Chung (2003).⁶

We compute an empirical test of the hypothesis (1) built around the quantilogram and establish its asymptotic properties. We present our tests graphically in the standard manner for time series analysis.

³Note that σ_t^2 may not be a conditional variance in this case. See Koenker and Zhao (1996) for discussion of estimation of quantiles in the presence of ARCH effects.

⁴For the most part they work with conditional normality, which makes the mean equal to the median. We note that it may be dangerous to work with normality as an assumption in this case where one statistical reason for looking at signs is related to their robustness with respect to moments.

⁵Actually, they scale the fixed c by an estimated standard deviation. However, they do not take account of this estimation in their distribution theory.

⁶Note that (1) allows even the semi-strong case where only the conditional α -quantile of ε_t is zero. In that case there is no implication about the behaviour of other quantiles other than on the magnitude.

3 Quantilogram

We first estimate μ_α by the quantile estimator $\hat{\mu}_\alpha$ which is defined by

$$\hat{\mu}_\alpha = \arg \min_{\mu \in \mathbb{R}} \sum_{t=1}^T \rho_\alpha(y_t - \mu), \text{ where } \rho_\alpha(x) = x [\alpha - 1(x < 0)].$$

Then let

$$\hat{\rho}_{\alpha k} = \frac{\frac{1}{T-k} \sum_{t=1}^{T-k} \psi_\alpha(y_t - \hat{\mu}_\alpha) \psi_\alpha(y_{t+k} - \hat{\mu}_\alpha)}{\sqrt{\frac{1}{T} \sum_{t=1}^T \psi_\alpha^2(y_t - \hat{\mu}_\alpha)} \sqrt{\frac{1}{T-k} \sum_{t=1}^{T-k} \psi_\alpha^2(y_{t+k} - \hat{\mu}_\alpha)}}, \quad k = 1, 2, \dots,$$

for any $\alpha \in [0, 1]$. Note that $-1 \leq \hat{\rho}_{\alpha k} \leq 1$ for any α, k because this is just a sample correlation based on data $\psi_\alpha(y_t - \hat{\mu}_\alpha)$. Compared with the correlogram of y_t itself this quantity is robust to the non-existence of moments. It also measures a different type of association from that given by the usual correlogram.

Under the null hypothesis (1) the population quantity

$$E[\psi_\alpha(y_t - \mu_\alpha) \psi_\alpha(y_{t+k} - \mu_\alpha)] = E[\psi_\alpha(y_t - \mu_\alpha)] E[\psi_\alpha(y_{t+k} - \mu_\alpha) | \mathcal{F}_{t+k-1}] = 0$$

for all k . Therefore, $\hat{\rho}_{\alpha k}$ should be approximately zero. Dufour, Hallin and Mizera (1998) establish various properties of the signogram (which corresponds to the case $\alpha = 1/2$) under independent sampling: when the median is known, they provide a test whose null distribution is known exactly; when the median is estimated, they show that this test is asymptotically distribution free under independent observations. In the dependent stochastic scaling process (3), this property no longer holds.

We now discuss the asymptotic properties of $\hat{\rho}_{\alpha k}$. We first assume:

ASSUMPTION 1. (a) $\{y_t : t = 1, \dots, T\}$ is a stationary and α -mixing sequence with mixing numbers satisfying $\sum_{m=1}^{\infty} \alpha(m)^{(p-2)/p} < \infty$ for some $p > 2$. (b) y_t has bounded unconditional density $f_y(\cdot)$ with respect to Lebesgue measure and has α -quantile zero and $f_y(\mu_\alpha) > 0$.

Under this assumption, we have the following Bahadur representation which is needed to discuss the asymptotic behavior of the quantilogram:

Lemma 1 *Suppose Assumption 1 holds. Then, we have*

$$\sqrt{T}(\hat{\mu}_\alpha - \mu_\alpha) = -\frac{1}{f_y(\mu_\alpha)} \frac{1}{\sqrt{T}} \sum_{t=1}^T \psi_\alpha(y_t - \mu_\alpha) + o_p(1).$$

Next we discuss the asymptotic property of $\hat{\rho}_{\alpha k}$ under the null hypothesis. Here, we focus on the scale process (3) and strengthen Assumption 1.

ASSUMPTION 2. (a) Suppose that y_t satisfies (3) and satisfies Assumption 1(a). (b) $\{\varepsilon_t : t = 1, \dots, T\}$ are i.i.d. with bounded density $f_\varepsilon(\cdot)$ with respect to Lebesgue measure and have α -quantile zero and $f_\varepsilon(0) > 0$. (c) σ_t^2 is stationary and measurable with respect to \mathcal{F}_{t-1} and $0 < E[\sigma_t^{-1}] < \infty$.

Theorem 2 Suppose Assumption 2 holds. Then, for $p = 1, 2, \dots$, we have

$$\sqrt{T}\widehat{\rho}_\alpha^{(p)} = \sqrt{T} \begin{bmatrix} \widehat{\rho}_{\alpha 1} \\ \vdots \\ \widehat{\rho}_{\alpha p} \end{bmatrix} \xrightarrow{d} N(0, V_\alpha^{(p)}), \text{ where } V_\alpha^{(p)} = (V_{\alpha jk})$$

$$V_{\alpha jj} = 1 + \left(\frac{E \left[\psi_\alpha(y_t - \mu_\alpha) \frac{1}{\sigma_{t+j}} \right]}{E[\psi_\alpha^2(y_t - \mu_\alpha)] E \left[\frac{1}{\sigma_t} \right]} \right)^2 \alpha(1 - \alpha)$$

$$V_{\alpha jk} = \frac{E \left[\psi_\alpha(\varepsilon_t) \frac{1}{\sigma_{t+k}} \right] E \left[\psi_\alpha(\varepsilon_t) \frac{1}{\sigma_{t+j}} \right]}{E[\psi_\alpha^2(y_t - \mu_\alpha)] E^2 \left[\frac{1}{\sigma_t} \right]}, \quad j \neq k.$$

The asymptotic variance does not explicitly depend on $f_y(\mu_\alpha)$ - there has been a cancellation of this quantity from estimation of the quantile to computing $\widehat{\rho}_{\alpha k}$. Nevertheless, the asymptotic variance is quite complicated since it generally depends on the process σ_t^2 . In some special cases, the correction factor in $V_{\alpha k}$ due to the estimation of μ_α , this is the term complicated term in parentheses, is zero. For example, when $\alpha = 1/2$: if ε_t is symmetric about zero and σ_{t+k}^2 is an even function of ε_t [as in GARCH processes], then

$$E \left[\psi_\alpha(\varepsilon_t) \frac{1}{\sigma_{t+k}} \right] = 0 \quad (4)$$

and $V_{\alpha kk} = 1$. In this case, the estimation of a model for σ_t^2 can be avoided.

This may be too restrictive a special case - since symmetric error/even σ_{t+k}^2 is often thought inappropriate for stock return data. In the more general case, given an estimated parametric model for σ_t^2 , one can estimate $V_{\alpha kk}$ by an obvious plug in approach. We next explore an approach that avoids this. Note that

$$\underline{V}_\alpha \equiv 1 \leq V_{\alpha kk} \leq 1 + \frac{[\max\{\alpha, 1 - \alpha\}]^2}{\alpha(1 - \alpha)} \equiv 1 + \bar{v}_\alpha \equiv \bar{V}_\alpha$$

because $|\psi_\alpha(\varepsilon_t)| \leq \max\{\alpha, 1 - \alpha\}$ and σ_t^{-1} is stationary. The upper bound is independent of k , like the usual Bartlett intervals for ordinary correlations. The upper bound increases as $\alpha \rightarrow 0, 1$, and so provides less information in such cases.

Under the null hypothesis (1) & (3), the quantilogram lies within $\pm z_{\gamma/2} \sqrt{\bar{V}_\alpha/T}$ with probability greater than $1 - \gamma$. If the additional conditions (4) are satisfied, the interval can be shrank to $\pm z_{\gamma/2} \sqrt{1/T}$. We call the smaller band liberal and the larger one conservative.

Note that

$$|V_{\alpha jk}| = \left| \frac{1}{\alpha(1-\alpha)} \frac{E \left[\psi_{\alpha}(\varepsilon_t) \frac{1}{\sigma_{t+k}} \right] E \left[\psi_{\alpha}(\varepsilon_t) \frac{1}{\sigma_{t+j}} \right]}{E^2 \left[\frac{1}{\sigma_t} \right]} \right| \leq \bar{v}_{\alpha},$$

although $V_{\alpha jk}$ itself can be positive or negative. The matrix $V_{\alpha}^{(p)}$ is dominated by a matrix $I + \bar{v}_{\alpha} ii^{\top}$ whose largest eigenvalue is $1 + p\bar{v}_{\alpha}$.

Consider the omnibus test statistic

$$Q_p = T \widehat{\rho}_{\alpha}^{(p)\top} \widehat{\rho}_{\alpha}^{(p)} \quad (5)$$

for any p . Then, $Q_p \leq [1 + p\bar{v}_{\alpha}] \times T \widehat{\rho}_{\alpha}^{(p)\top} [V_{\alpha}^{(p)}]^{-1} \widehat{\rho}_{\alpha}^{(p)}$. Let $\chi_{\gamma}^2(p)$ be the level γ critical value of a chi-squared(p) distribution. Under the null hypothesis, the rule:

$$\text{reject at level } \gamma \text{ if } Q_p > [1 + p\bar{v}_{\alpha}] \chi_{\gamma}^2(p),$$

has size less than or equal to γ . A liberal test can be constructed using the lower bound of one instead of $1 + p\bar{v}_{\alpha}$.

Compare our approach with the Engle and Manganelli (1999, pp11-12) dynamic conditional quantile test, which in our case would involve running the regression of $\psi_{\alpha}(y_t - \widehat{\mu}_{\alpha})$ on a constant and $\psi_{\alpha}(y_{t-1} - \widehat{\mu}_{\alpha}), \dots, \psi_{\alpha}(y_{t-p} - \widehat{\mu}_{\alpha})$ [and perhaps other variables] and then testing whether the coefficients in this regression are zero using the usual quadratic form.

A related test can be based on the partial quantilogram, which can be defined in the same way as in Brockwell and Davies (1991, p102). Specifically, define $\widehat{\phi}_{\alpha k}$ for each k as $\widehat{\phi}_{\alpha k} = \widehat{\phi}_{\alpha k k}$, where

$$\begin{bmatrix} \widehat{\phi}_{\alpha k1} \\ \widehat{\phi}_{\alpha k2} \\ \vdots \\ \widehat{\phi}_{\alpha k k} \end{bmatrix} = \begin{bmatrix} 1 & \widehat{\rho}_{\alpha 1} & \cdots & \widehat{\rho}_{\alpha, k-1} \\ \widehat{\rho}_{\alpha 1} & 1 & \widehat{\rho}_{\alpha 1} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \widehat{\rho}_{\alpha, k-1} & \widehat{\rho}_{\alpha, k-2} & \cdots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \widehat{\rho}_{\alpha 1} \\ \widehat{\rho}_{\alpha 2} \\ \vdots \\ \widehat{\rho}_{\alpha, k} \end{bmatrix}.$$

Theorem 3 *Suppose Assumption 2 holds. Then, for $p = 1, 2, \dots$, we have*

$$\sqrt{T} \widehat{\phi}_{\alpha}^{(p)} = \sqrt{T} \begin{bmatrix} \widehat{\phi}_{\alpha 1} \\ \vdots \\ \widehat{\phi}_{\alpha p} \end{bmatrix} \xrightarrow{d} N(0, V_{\alpha}^{(p)}).$$

Define the portmantau statistic

$$Q_p^* = T \widehat{\phi}_{\alpha}^{(p)\top} \widehat{\phi}_{\alpha}^{(p)}. \quad (6)$$

The same considerations apply as in the case of Q_p .

4 Numerical Results

We investigate samples of daily, weekly, and monthly returns on the S&P500 from 1955 to 2002, a total of 11,893, 2464, and 570 observations respectively. The daily data is quite heavy tailed.

In Figures 1,3,5 we give the quantilogram for quantiles in the range 0.01 – 0.99 and out to 100 lags. We also show the 95% confidence intervals (centered at 0) based on the lower and upper bound. There seems to be some evidence of predictability, but it depends on the data frequency and in some cases on which confidence interval you use. The evidence of predictability is strongest at the highest frequency, although this might be because of the better precision of estimation. For monthly data there are very few observations outside the liberal confidence bands, and none outside the conservative ones. The portmanteau tests give a clearer picture of the evidence of predictability, which is very pronounced for the daily data for all except the most extreme quantiles [where there is insufficient data]. This is consistent with the finding in Hong and Chung (2003) that the magnitude of predictability is small for any given lag but large when combined across many lags. It is interesting that the 0.05 quantile case has much more pronounced dependence than the 0.95 quantile case.

Note that the upper bound confidence interval/critical value becomes very large in the extreme quantile case and is perhaps too pessimistic.

*** FIGS 1-9 HERE ***

The conclusion is that there is evidence of directional predictability that is not consistent with the pure strong quantile volatility model, that is, for no quantile does (3) appear consistent with the data. The predictability could be coming through mean effects or time varying higher moments as discussed in Christoffersen and Diebold (2002).

We estimated on the daily data the following AR(2)/AGARCH(1,1) model

$$\begin{aligned}y_t &= \beta_0 + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \varepsilon_t \sigma_t \\ \sigma_t^2 &= \gamma_0 + \gamma_1 \sigma_{t-1}^2 + \gamma_2 u_t^2 + \gamma_3 u_t^2 1(u_t < 0),\end{aligned}$$

where $u_t = \varepsilon_t \sigma_t$ using the Gaussian qmle, which assumes that $\varepsilon_t \sim N(0, 1)$. We then examine the standardized residuals from this estimated model. In Figure 7,8 we show the quantilogram and portmanteau test statistic along with the lower bound critical values. Clearly, there is much less evidence of sign predictability left in the residuals, but there is still some on the downside.

5 Conclusions

We have proposed using a standard time series methodology for measuring linear dependence in quantile hits. We developed the distribution theory needed for the application of correlogram methods

to the quantile case. This methodology is used widely in econometrics for analyzing time series data and its computation is available in most standard packages. We think it is therefore likely to be a useful technique for analyzing directional predictability. The empirical results show that it is important to take account of the effects of many small contributions from different lags as is done in the Box-Pierce type statistics. We found very strong evidence of predictability in daily stock index returns at many different quantiles, and especially in the lower tails. This evidence remains, although it is much more muted, after fitting a time series model to the mean and variance of returns.

The ‘quantilogram’ can easily be extended to the vector case, where it is of interest to detect directional predictability from one series to another.

6 Appendix

PROOF OF LEMMA 1. The proof mimics Pollard (1991, Proof of Theorem 1). However, we need to generalize the latter result to allow for dependency of the errors and quantiles with $\alpha \neq 1/2$. Define $u_t = y_t - \mu_\alpha$ and note that $f_u(0) = f_y(\mu_\alpha)$.

For θ in \mathbb{R} , define

$$G_T(\theta) = \sum_{t=1}^T \left[\rho_\alpha \left(u_t - \frac{\theta}{\sqrt{T}} \right) - \rho_\alpha(u_t) \right].$$

This is a convex function minimized by

$$\hat{\theta}_T = \sqrt{T} (\hat{\mu}_\alpha - \mu_\alpha).$$

Assumption 1 ensures that the function

$$M(x) = E [\rho_\alpha(u_t - x) - \rho_\alpha(u_t)]$$

has a unique minimum at zero and

$$M(x) = \frac{1}{2}x^2 f_u(0) + o(x^2) \text{ for } x \text{ near zero} \quad (7)$$

via a Taylor expansion. Using (7), we have

$$\Gamma_T(\theta) \equiv EG_T(\theta) = \frac{1}{2}\theta^2 f_u(0) + o(1). \quad (8)$$

Let

$$D_t = 1(u_t < 0) - \alpha.$$

Note that $ED_t = 0$ since u_t has α -quantile zero. Define

$$\begin{aligned} R_{t,T}(\theta) &= \rho_\alpha\left(u_t - \frac{\theta}{\sqrt{T}}\right) - \rho_\alpha(u_t) - \frac{\theta}{\sqrt{T}}D_t \text{ and} \\ W_T &= \frac{1}{\sqrt{T}} \sum_{t=1}^T D_t. \end{aligned}$$

Then,

$$G_T(\theta) = \Gamma_T(\theta) + W_T\theta + \sum_{t=1}^T (R_{t,T}(\theta) - ER_{t,T}(\theta)). \quad (9)$$

We first establish that the sum of the centered terms $\xi_{t,T} = R_{t,T}(\theta) - ER_{t,T}(\theta)$ for fixed θ is $o_p(1)$. Observe that

$$|R_{t,T}(\theta)| \leq \frac{|\theta|}{\sqrt{T}} 1 \left(|u_t| \leq \frac{|\theta|}{\sqrt{T}} \right). \quad (10)$$

Write

$$E \left| \sum_{t=1}^T \xi_{t,T} \right|^2 = \sum_{t=1}^T E\xi_{t,T}^2 + 2 \sum_{t < s} E\xi_{t,T}\xi_{s,T}.$$

Then, we have

$$\sum_{t=1}^T E\xi_{t,T}^2 \leq \sum_{t=1}^T ER_{t,T}(\theta)^2 \leq \theta^2 E1 \left(|u_t| \leq \frac{|\theta|}{\sqrt{T}} \right) \leq C\theta^2 \frac{|\theta|}{\sqrt{T}} \rightarrow 0. \quad (11)$$

Also, for $p > 2$,

$$\begin{aligned} T^{1/p} \left| \sum_{t < s} E\xi_{t,T}\xi_{s,T} \right| &\leq 8T^{1/p} \sum_{t < s} \|R_{t,T}(\theta)\|_p \|R_{s,T}(\theta)\|_p \alpha(s-t)^{(p-2)/p} \\ &\leq CT^{-1} \sum_{t < s} \alpha(t-s)^{(p-2)/p} \\ &= C \sum_{m=1}^{T-1} \left(1 - \frac{m}{T}\right) \alpha(m)^{(p-2)/p} \\ &\rightarrow C \sum_{m=1}^{\infty} \alpha(m)^{(p-2)/p} < \infty, \end{aligned} \quad (12)$$

where $\|\cdot\|_p = (E|\cdot|^p)^{1/p}$ denotes the L^p -norm, the first inequality holds by the mixing inequality of Hall and Heyde (1980, Corollary A.2), the second inequality holds by using (10) and Assumption 1(b), and the last convergence follows from Toeplitz lemma. Now, (11) and (12) implies $\sum_{t=1}^T \xi_{t,T} = o_p(1)$ as desired. Therefore, this result and (8) imply that, for each fixed θ , we have

$$G_T(\theta) = \frac{1}{2}\theta^2 f_u(0) + W_T\theta + o_p(1). \quad (13)$$

The convexity lemma of Pollard (1991, p.187) strengthens the pointwise convergence result in (13) to uniform convergence on compact subsets of \mathbb{R} . That is, with $\eta_T = -W_T/f_u(0)$, we may write

$$G_T(\theta) = \frac{1}{2}f_u(0) |\theta - \eta_T|^2 - \frac{1}{2}f_u(0)\eta_T^2 + r_T(\theta),$$

where for each compact set K in R ,

$$\sup_{\theta \in K} |r_T(\theta)| = o_p(1).$$

Also, under Assumption 1, we have $W_T = O_p(1)$ by a CLT (see Hall and Heyde (1980, Corollary 5.1)). Finally, using an argument similar to Pollard (1991, Proof of Theorem 1) or Jureckova (1977, Proof of Lemma 5.2), we can show for each $\delta > 0$ that

$$\Pr \left[\left| \widehat{\theta}_T - \eta_T \right| > \delta \right] \rightarrow 0,$$

as desired. ■

PROOF OF THEOREM 2. Define

$$\begin{aligned} \widetilde{\sigma}_k(\mu) &= \frac{1}{T-k} \sum_{t=1}^{T-k} \psi_\alpha(y_t - \mu) \psi_\alpha(y_{t+k} - \mu), \\ \sigma_k(\mu) &= E[\psi_\alpha(y_t - \mu) \psi_\alpha(y_{t+k} - \mu)] \end{aligned}$$

and let $\widetilde{\sigma}_k = \widetilde{\sigma}_k(\mu_\alpha)$ and $\sigma_k = \sigma_k(\mu_\alpha)$. By rearranging terms and a Taylor expansion, we have

$$\sqrt{T}(\widetilde{\sigma}_k(\widehat{\mu}_\alpha) - \sigma_k) = \sqrt{T}(\widetilde{\sigma}_k(\widehat{\mu}_\alpha) - \sigma_k(\widehat{\mu}_\alpha)) + \left(\frac{\partial}{\partial \mu} [\sigma_k(\mu)] \right)_{\mu=\mu^*} \sqrt{T}(\widehat{\mu}_\alpha - \mu_\alpha), \quad (14)$$

where μ^* lies between $\widehat{\mu}_\alpha$ and μ_α . We first show that

$$\sqrt{T}(\widetilde{\sigma}_k(\widehat{\mu}_\alpha) - \sigma_k(\widehat{\mu}_\alpha)) = \sqrt{T}(\widetilde{\sigma}_k - \sigma_k) + o_p(1). \quad (15)$$

Consider the class of functions

$$\mathcal{F} = \{ \psi_\alpha(y_t - \mu) \psi_\alpha(y_{t+k} - \mu) : \mu \in \Theta \}.$$

For each $\delta > 0$ and $\mu \in \Theta$, we have

$$\begin{aligned} & E \sup_{\mu_1 \in \Theta: |\mu_1 - \mu| < \delta} [\psi_\alpha(y_t - \mu_1) \psi_\alpha(y_{t+k} - \mu_1) - \psi_\alpha(y_t - \mu) \psi_\alpha(y_{t+k} - \mu)]^2 \\ & \leq 16E \sup_{\mu_1 \in \Theta: |\mu_1 - \mu| < \delta} [\psi_\alpha(y_t - \mu_1) - \psi_\alpha(y_t - \mu)]^2 \\ & \leq 16E 1(\mu - \delta \leq y_t \leq \mu + \delta) \\ & = 16E \left[F_\varepsilon\left(\frac{\mu - \mu_\alpha + \delta}{\sigma_t}\right) - F_\varepsilon\left(\frac{\mu - \mu_\alpha - \delta}{\sigma_t}\right) \right] \\ & \leq C\delta, \end{aligned} \quad (16)$$

where the first inequality holds by stationarity of y_t and the result $|\psi_\alpha(\cdot)| \leq 2$ and the last inequality holds by Assumptions 2(b) and (c). (16) implies that the L^2 -bracketing number satisfies

$$N(\varepsilon, \mathcal{F}) \leq C(1/\varepsilon)^2 \forall \varepsilon > 0.$$

Thus, (15) follows from the consistency of $\hat{\mu}_\alpha$ for μ_α due to Lemma 1 and the stochastic equicontinuity result of Andrews and Pollard (1994, Theorem 2.2) by taking $Q = 2$ and $\gamma = p - 2$ in the latter paper.

Next, consider the second term on the rhs of (14). We have

$$\begin{aligned} \sigma_k(\mu) &= E[\psi_\alpha(y_t - \mu)\psi_\alpha(y_{t+k} - \mu)] \\ &= E\left[1\left(\varepsilon_t < \frac{\mu - \mu_\alpha}{\sigma_t}\right) - \alpha\right]\left[1\left(\varepsilon_{t+k} < \frac{\mu - \mu_\alpha}{\sigma_{t+k}}\right) - \alpha\right] \\ &= E\left\{\left[1\left(\varepsilon_t < \frac{\mu - \mu_\alpha}{\sigma_t}\right) - \alpha\right]\left[F_\varepsilon\left(\frac{\mu - \mu_\alpha}{\sigma_{t+k}}\right) - \alpha\right]\right\} \\ &= E\left[\int_{-\infty}^{\varepsilon_t \leq \frac{\mu - \mu_\alpha}{\sigma_t}} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{F_\varepsilon\left(\frac{\mu - \mu_\alpha}{\sigma_{t+k}}\right) - \alpha\right\} \prod_{j=0}^{k-1} f(\varepsilon_{t+j}) d\varepsilon_{t+j}\right] \\ &\quad - \alpha E\left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{F_\varepsilon\left(\frac{\mu - \mu_\alpha}{\sigma_{t+k}}\right) - \alpha\right\} \prod_{j=0}^{k-1} f(\varepsilon_{t+j}) d\varepsilon_{t+j}\right] \end{aligned}$$

by the law of iterated expectations. Therefore,

$$\begin{aligned} \left(\frac{\partial}{\partial \mu} \sigma_k(\mu)\right)_{\mu=\mu_\alpha} &= f_\varepsilon(0) E\left[\left(1(\varepsilon_t < 0) - \alpha\right) \frac{1}{\sigma_{t+k}}\right] \\ &\quad + f_\varepsilon(0) E\left[\frac{1}{\sigma_t} \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} \left\{F_\varepsilon(0) - \alpha\right\} \prod_{j=1}^{k-1} f(\varepsilon_{t+j}) d\varepsilon_{t+j}\right] \\ &= f_\varepsilon(0) E\left[\psi_\alpha(\varepsilon_t) \frac{1}{\sigma_{t+k}}\right] \end{aligned} \tag{17}$$

because $F_\varepsilon(0) = \alpha$. Furthermore, note that

$$f_u(0) = \left(\frac{\partial}{\partial x} E\left[F_\varepsilon\left(\frac{x}{\sigma_t}\right)\right]\right)_{x=0} = f_\varepsilon(0) E\left[\frac{1}{\sigma_t}\right]. \tag{18}$$

Therefore, by (14), (15), (17), (18) and consistency of $\hat{\mu}_\alpha$ for μ_α due to Lemma 1, we have

$$\sqrt{T}(\tilde{\sigma}_k(\hat{\mu}_\alpha) - \sigma_k) = \sqrt{T}(\tilde{\sigma}_k - \sigma_k) - \frac{E\left[\psi_\alpha(\varepsilon_t) \frac{1}{\sigma_{t+k}}\right]}{E\left[\frac{1}{\sigma_t}\right]} \frac{1}{\sqrt{T}} \sum_{t=1}^T \psi_\alpha(\varepsilon_t) + o_p(1). \tag{19}$$

On the other hand, for each $k = 0, 1, \dots$, and $\forall \varepsilon > 0$, we have

$$\begin{aligned}
& \Pr \left[\left| \frac{1}{T-k} \sum_{t=1}^{T-k} \psi_{\alpha}^2(y_{t+k} - \hat{\mu}_{\alpha}) - E\psi_{\alpha}^2(y_{t+k} - \mu_{\alpha}) \right| > \varepsilon \right] \\
& \leq \Pr \left[\sup_{\mu \in \Theta} \left| \frac{1}{T-k} \sum_{t=1}^{T-k} \{ \psi_{\alpha}^2(y_{t+k} - \mu) - E\psi_{\alpha}^2(y_{t+k} - \mu) \} \right| > \frac{\varepsilon}{2} \right] + \\
& \Pr \left[\left| (E\psi_{\alpha}^2(y_{t+k} - \mu))_{\mu=\hat{\mu}_{\alpha}} - E\psi_{\alpha}^2(y_{t+k} - \mu_{\alpha}) \right| > \frac{\varepsilon}{2} \right] + o(1) \\
& \rightarrow 0
\end{aligned} \tag{20}$$

where the inequality holds by triangle inequality and consistency of $\hat{\mu}_{\alpha}$ for μ_{α} and the convergence to zero holds by the following arguments: Consider the class of functions defined by $\mathcal{G} = \{ \psi_{\alpha}^2(y_{t+k} - \mu) : \mu \in \Theta \}$ for $k = 0, 1, \dots$. Using an argument analogous to (16), it is straightforward to see that the L^2 -bracketing number satisfies $N(\varepsilon, \mathcal{G}) < \infty \forall \varepsilon > 0$. Therefore, the first term on the rhs of (20) converges to zero by a uniform law of large numbers (LLN) using the pointwise WLLN result of Andrews (1988, Example 4, P.462) and an argument similar to Theorem 2.4.1 of van der Vaart and Wellner (1996, p.123). Next, the second term on the rhs of (20) is also $o(1)$ by Assumption 2, Lemma1 and a one term Taylor expansion, as desired.

Combining (19) and (20), we have

$$\sqrt{T}(\hat{\rho}_{\alpha k} - \rho_{\alpha k}) = \frac{1}{E[\psi_{\alpha}^2(y_t - \mu_{\alpha})]} \frac{1}{\sqrt{T}} \sum_{t=1}^T \psi_{\alpha}(\varepsilon_t) \psi_{\alpha}(\varepsilon_{t+k}) - \frac{E \left[\psi_{\alpha}(\varepsilon_t) \frac{1}{\sigma_{t+k}} \right]}{E[\psi_{\alpha}^2(y_t - \mu_{\alpha})] E \left[\frac{1}{\sigma_t} \right]} \frac{1}{\sqrt{T}} \sum_{t=1}^T \psi_{\alpha}(\varepsilon_t) + o_p(1).$$

Therefore, by a CLT for bounded rv's, see, e.g., Hall and Heyde (1980, Corollary 5.1, p.132), we have the desired asymptotic normality result of Theorem 2. The covariance between $\sqrt{T}(\hat{\rho}_{\alpha k} - \rho_{\alpha k})$ and $\sqrt{T}(\hat{\rho}_{\alpha j} - \rho_{\alpha j})$ for $j \neq k$ is determined by the second term, i.e.,

$$\text{acov}(\sqrt{T}(\hat{\rho}_{\alpha k} - \rho_{\alpha k}), \sqrt{T}(\hat{\rho}_{\alpha j} - \rho_{\alpha j})) = \frac{E \left[\psi_{\alpha}(\varepsilon_t) \frac{1}{\sigma_{t+k}} \right] E \left[\psi_{\alpha}(\varepsilon_t) \frac{1}{\sigma_{t+j}} \right]}{E[\psi_{\alpha}^2(y_t - \mu_{\alpha})] E^2 \left[\frac{1}{\sigma_t} \right]}.$$

■

References

- [1] ANDREWS, D. W. K. (1988): "Laws of large numbers for dependent non-identically distributed random variables," *Econometric Theory* 4, 458-467.

- [2] BASSETT, G.W., R. KOENKER, AND G. KORDA (2003): “Pessimistic portfolio allocation and Choquet expected utility,”
- [3] BLAKE, D., AND A. TIMMERMANN (2003): “Performance persistence in mutual funds: An independent assessment of the studies prepared by CRA for the IMA,” http://www.pensions-institute.org/reports/perormance_persistence.pdf
- [4] BROCKWELL, P.J., AND R.A. DAVIS (1991): *Time Series: Theory and Methods*. 2nd edition. Springer Verlag: Berlin.
- [5] CAMPBELL, J.Y., A.W. LO, AND A.C. MACKINLAY (1997): *The Econometrics of Financial Markets*, Princeton University Press, Princeton.
- [6] CHRISTOFFERSEN, P.F., AND F.X. DIEBOLD (2002): “Financial Asset Returns, market timing, and volatility dynamics,”
- [7] COWLES, A., AND H. JONES (1937): “Some A Posteriori Probabilities in Stock Market Action,” *Econometrica* 5, 280-294.
- [8] DUFOUR, J.M., M. HALLIN, AND I. MIZERA (1998): “Generalized Runs tests for heteroscedastic Time Series,” *Nonparametric Statistics* 9, 39-86.
- [9] DUFOUR, J.M., AND R. ROY (1986): “Generalized portmanteau statistics and tests of randomness,” *Communications in statistics: Theory and methods* 15, 2953-2972.
- [10] ENGLE, R.F. AND S. MANGANELLI (1999): CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles,” Manuscript.
- [11] HALL, P. AND C. C. HEYDE (1980): *Martingale Limit Theory and Its Applications*, Academic Press: New York.
- [12] HONG, Y., AND J. CHUNG (2003): “Are the directions of stock price changes predictable? Statistical Theory and Evidence,
- [13] JUREČKOVÁ, J. (1977): “Asymptotic relations of M-estimates and R-estimates in linear regression model,” *Annals of Statistics* 5, 464-472.
- [14] KOENKER, R., AND Q. ZHAO (1996). “Conditional quantile estimation and inference for ARCH models,” *Econometric Theory* 12, 793-814.
- [15] POLLARD, D. (1991): “Asymptotics for least absolute Deviation regression estimators,” *Econometric Theory* 7, 186-199.

- [16] TONKS, I. (2003): “Performance persistence of pension fund managers,”
- [17] VAN DER VAART, A. W. AND J. A. WELLNER (1996): *Weak Convergence and Empirical Processes*, Springer-Verlag: New York.

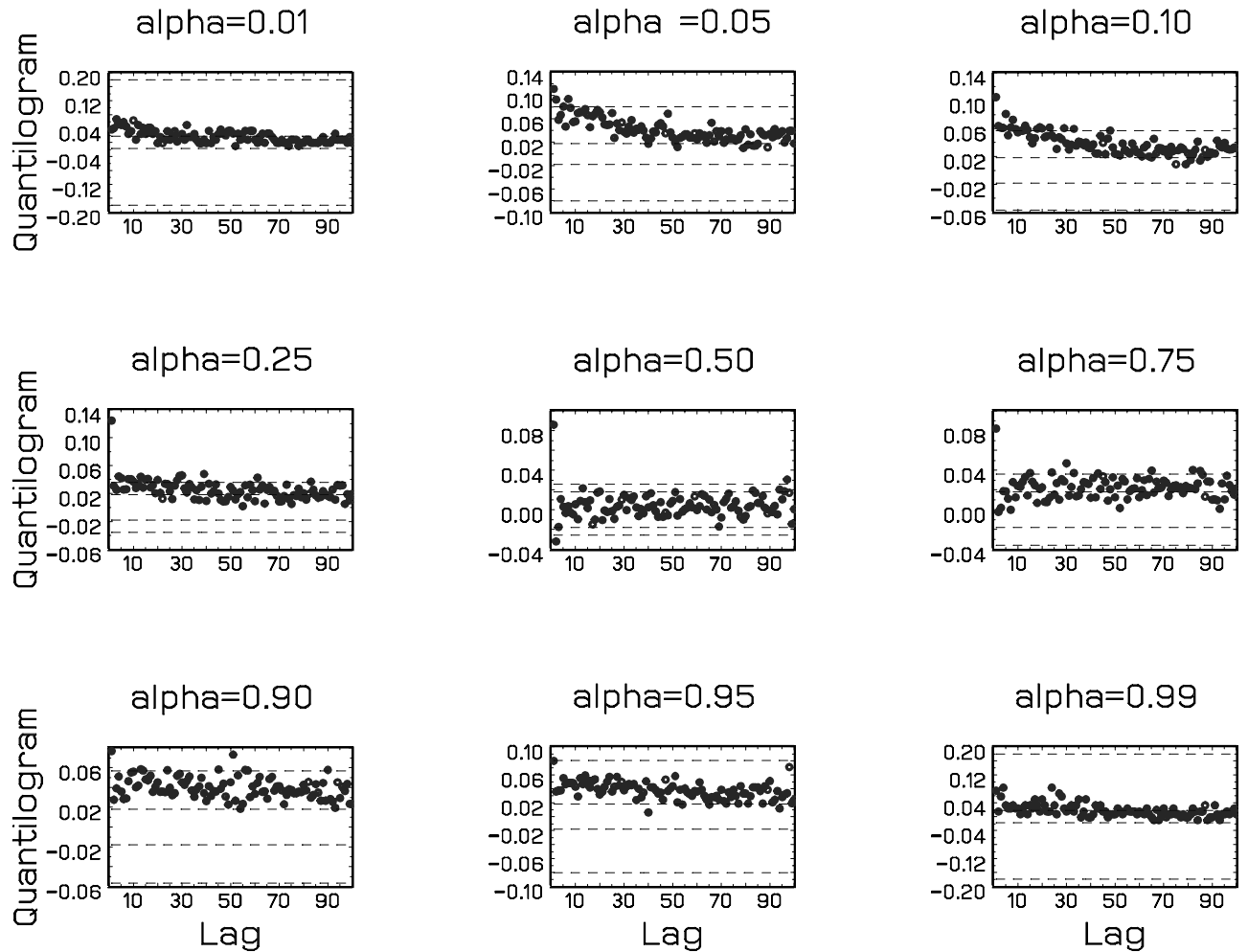


Figure 1. S&P500 Daily Data. Shown are the values of $\hat{\rho}_{\alpha k}$ along with the liberal and conservative 95% confidence intervals.

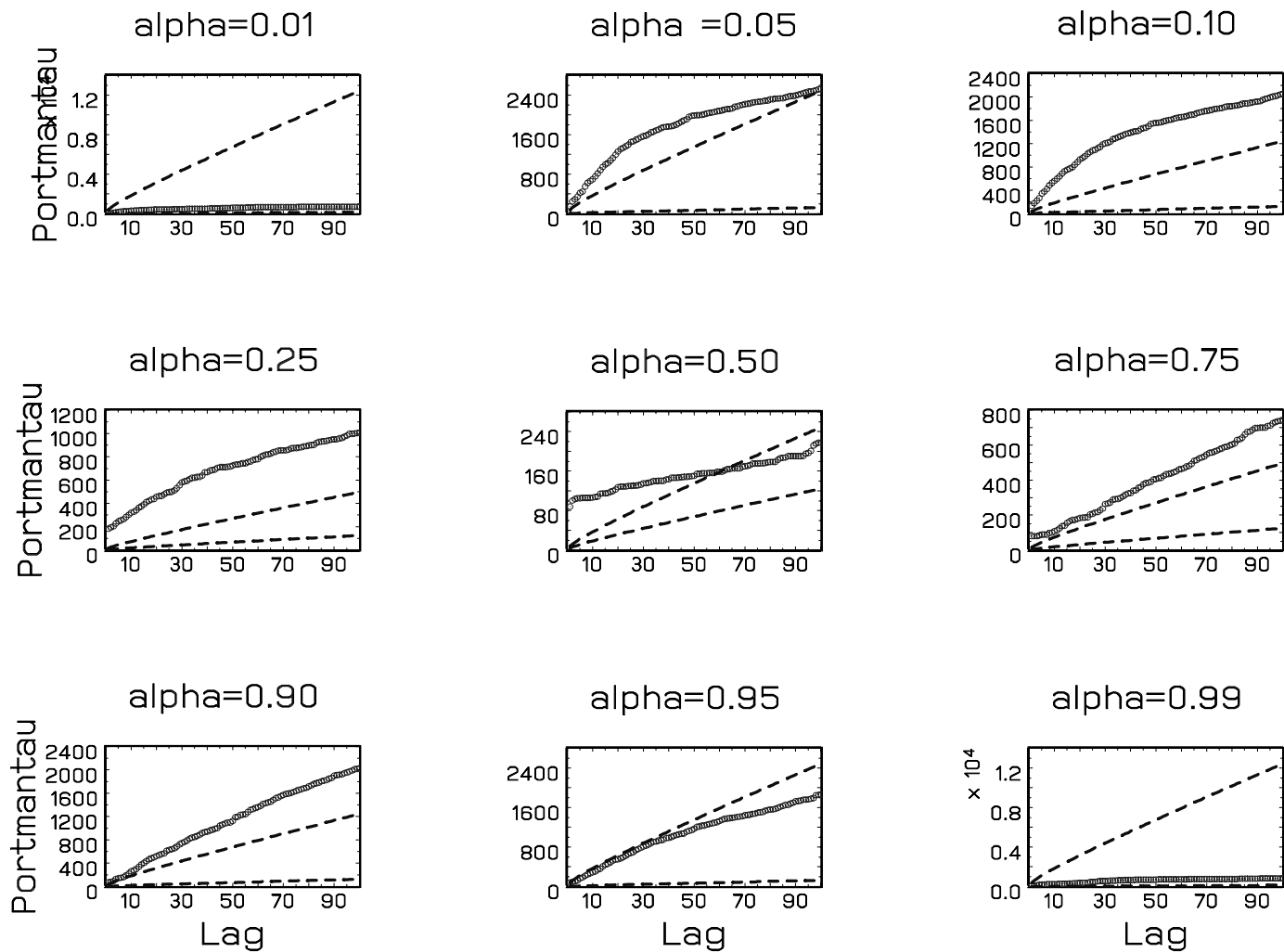


Figure 2. S&P500 Daily Data. Portmanteau Test statistic Q_p for each lag p and quantile α along with 95% liberal and conservative critical values

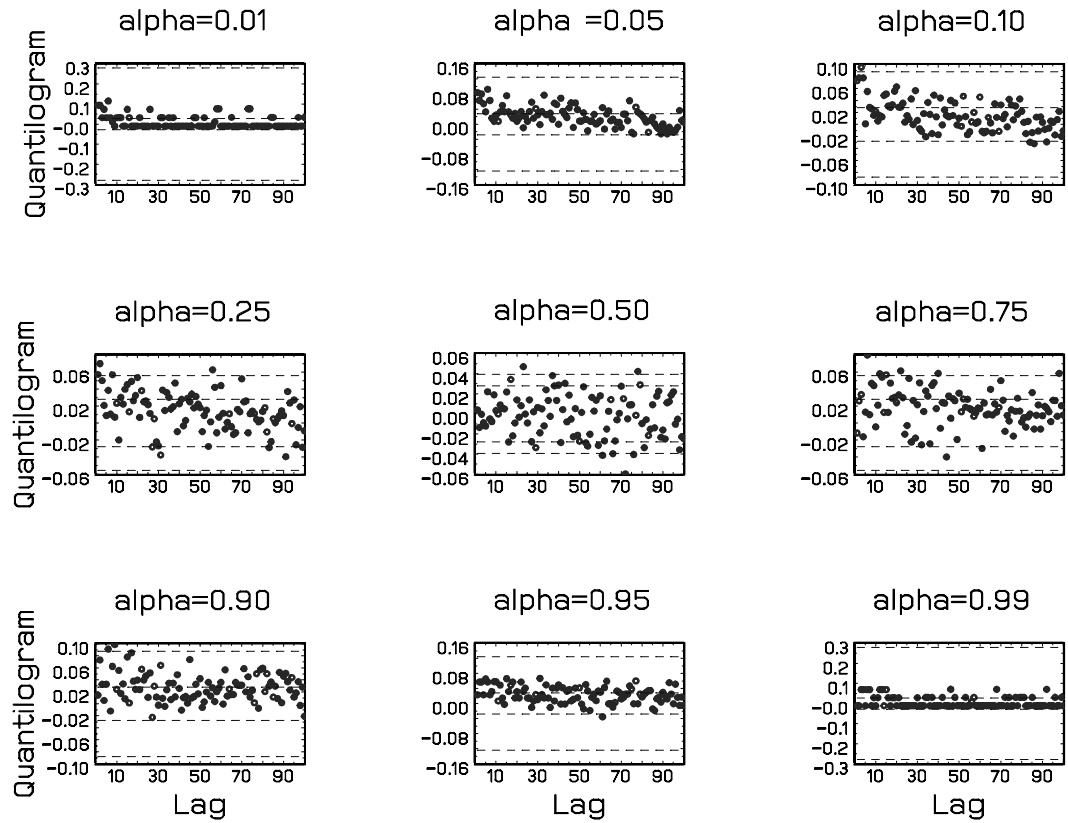


Figure 3. S&P500 Weekly Data; 95% confidence interval

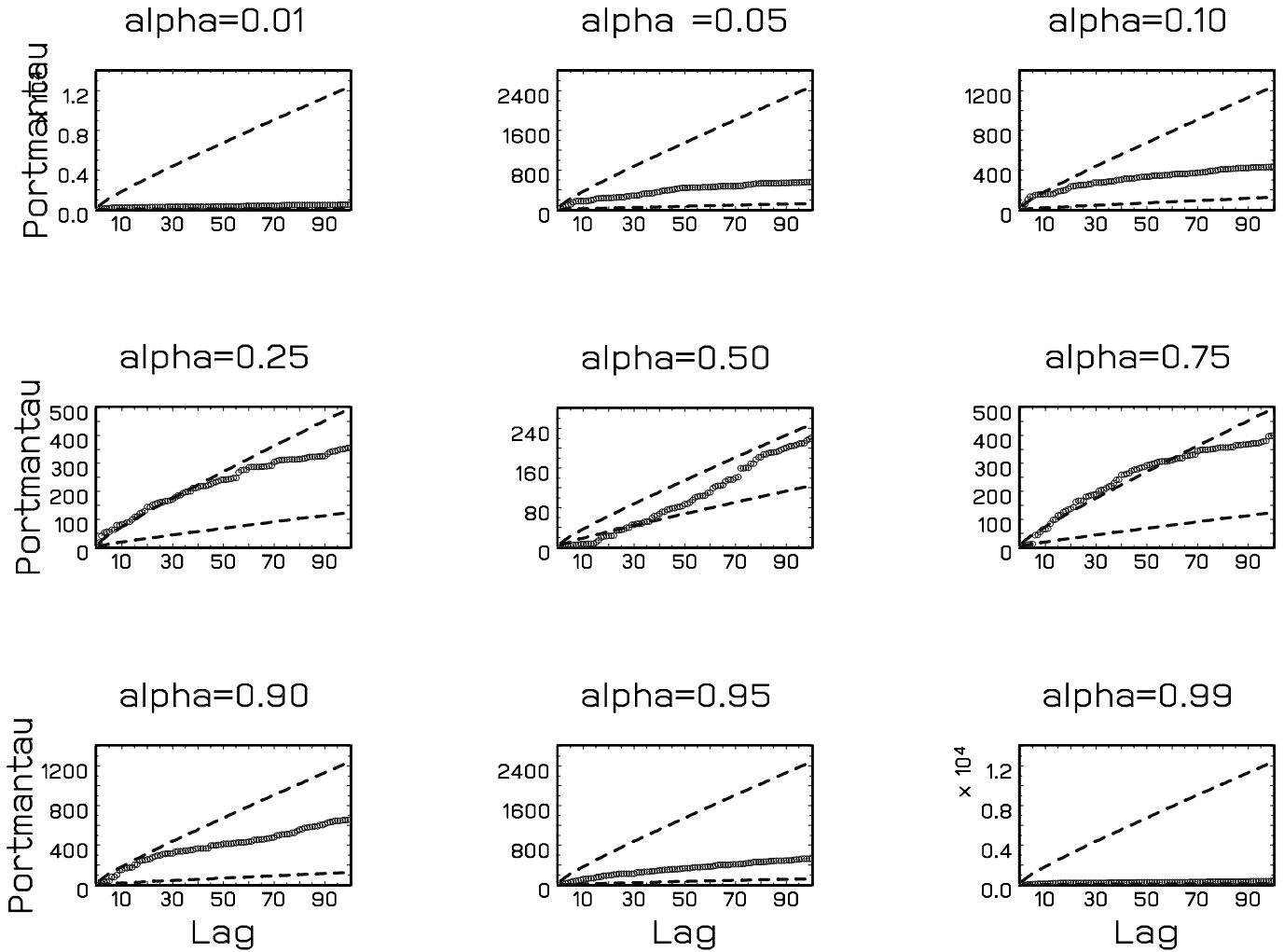


Figure 4. S&P500 Weekly Data; Portmanteau Test with 95% critical values

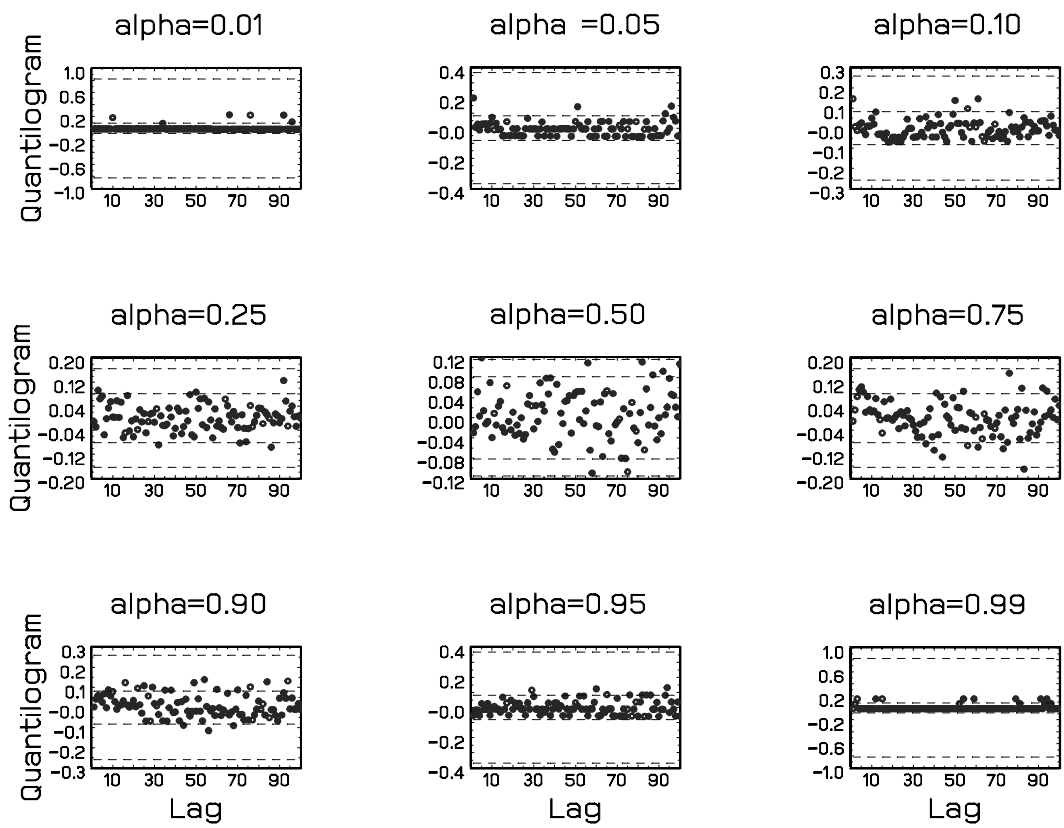


Figure 5. S&P500 Monthly Data; 95% confidence interval

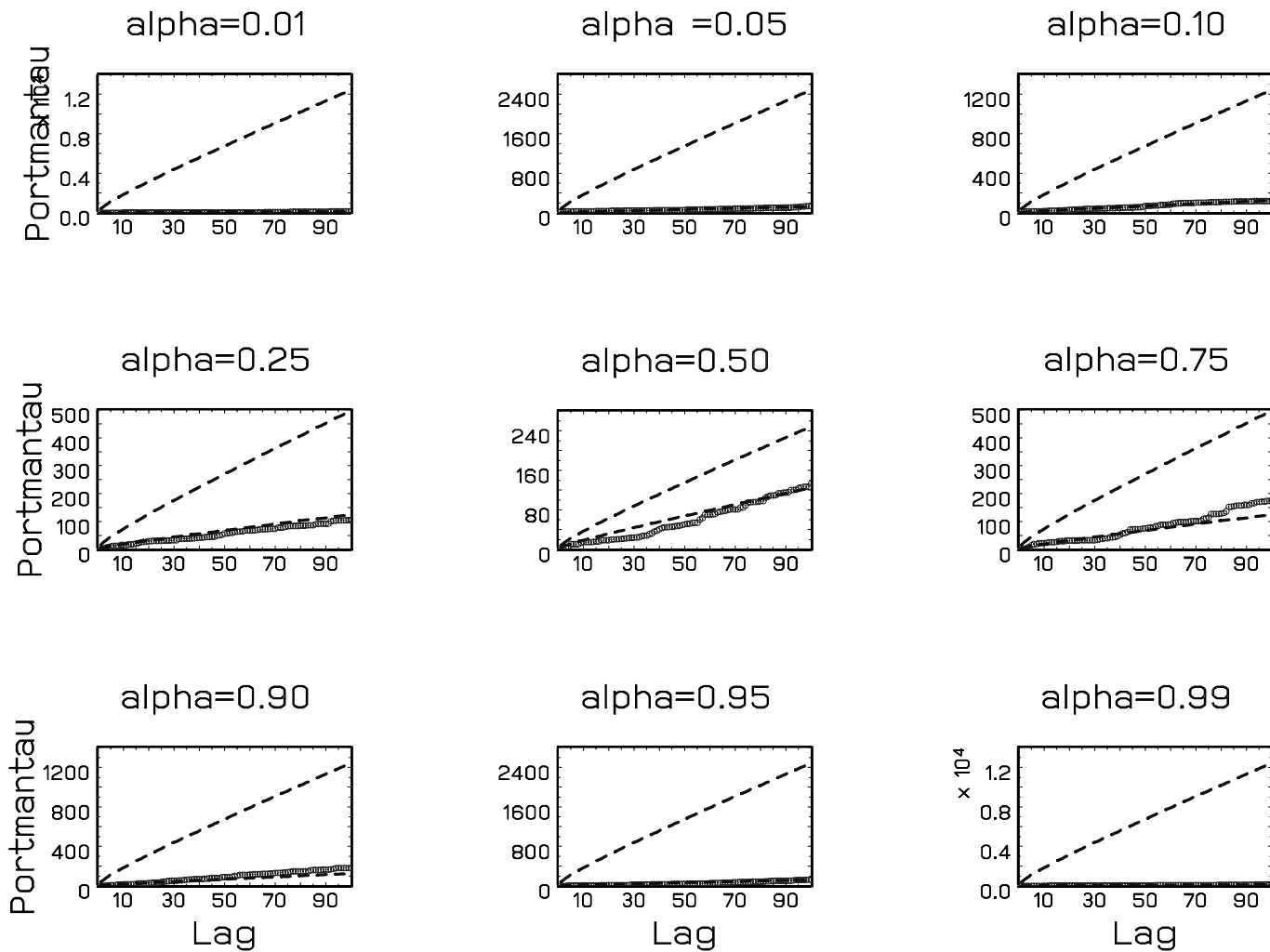


Figure 6. S&P500 Monthly Data; Portmanteau Test with 95% critical values

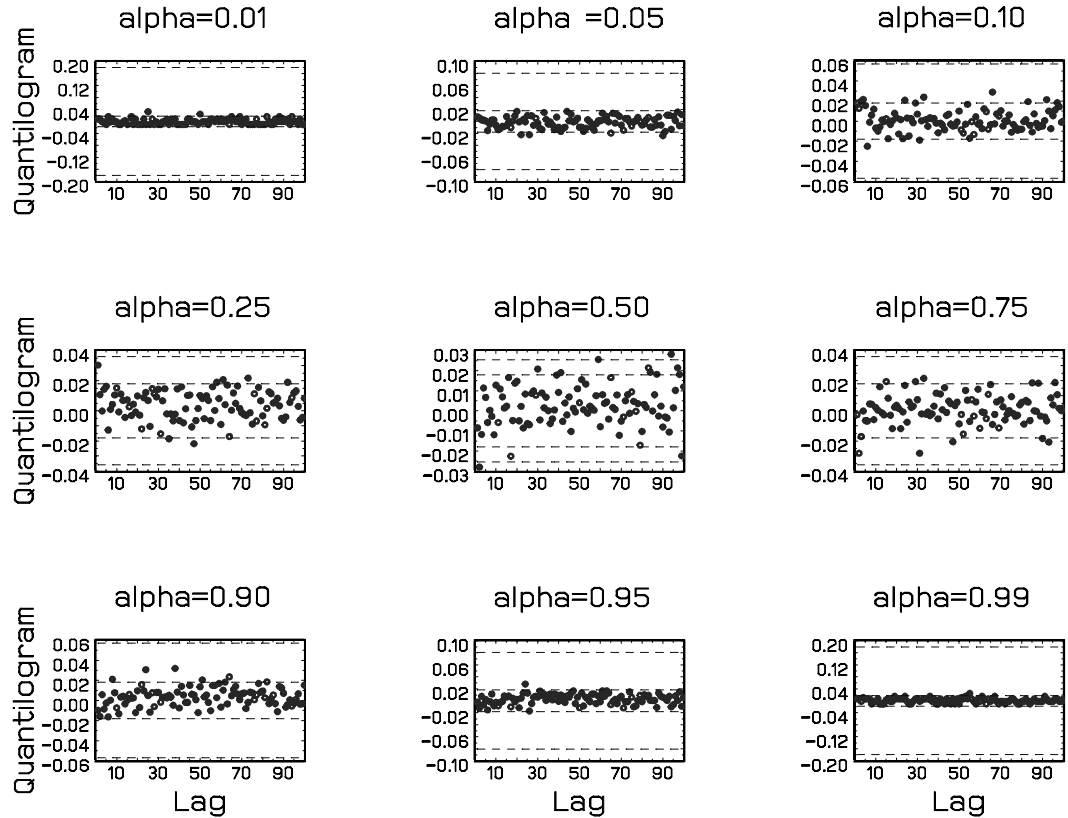


Figure 7. Standardized residuals from AR(2)/AGARCH(1,1) model fit on S&P500 Daily Data. Shown are the values of $\hat{\rho}_{\alpha k}$ along with the liberal 95% confidence intervals.

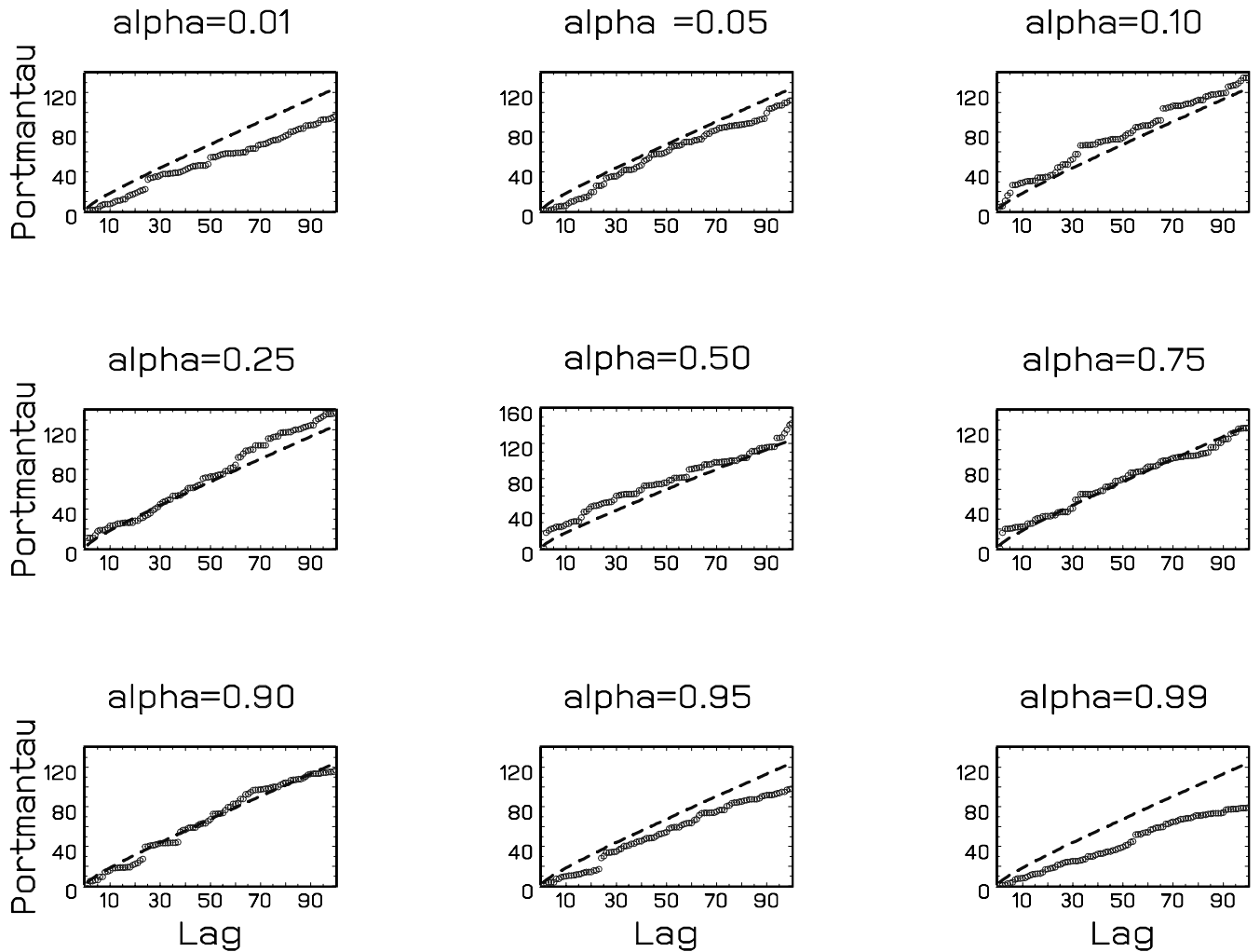


Figure 8. Standardized residuals from AR(2)/AGARCH(1,1) model fit on S&P500 Daily Data. Portmanteau Test statistic Q_p for each lag p and quantile α along with 95% liberal critical values

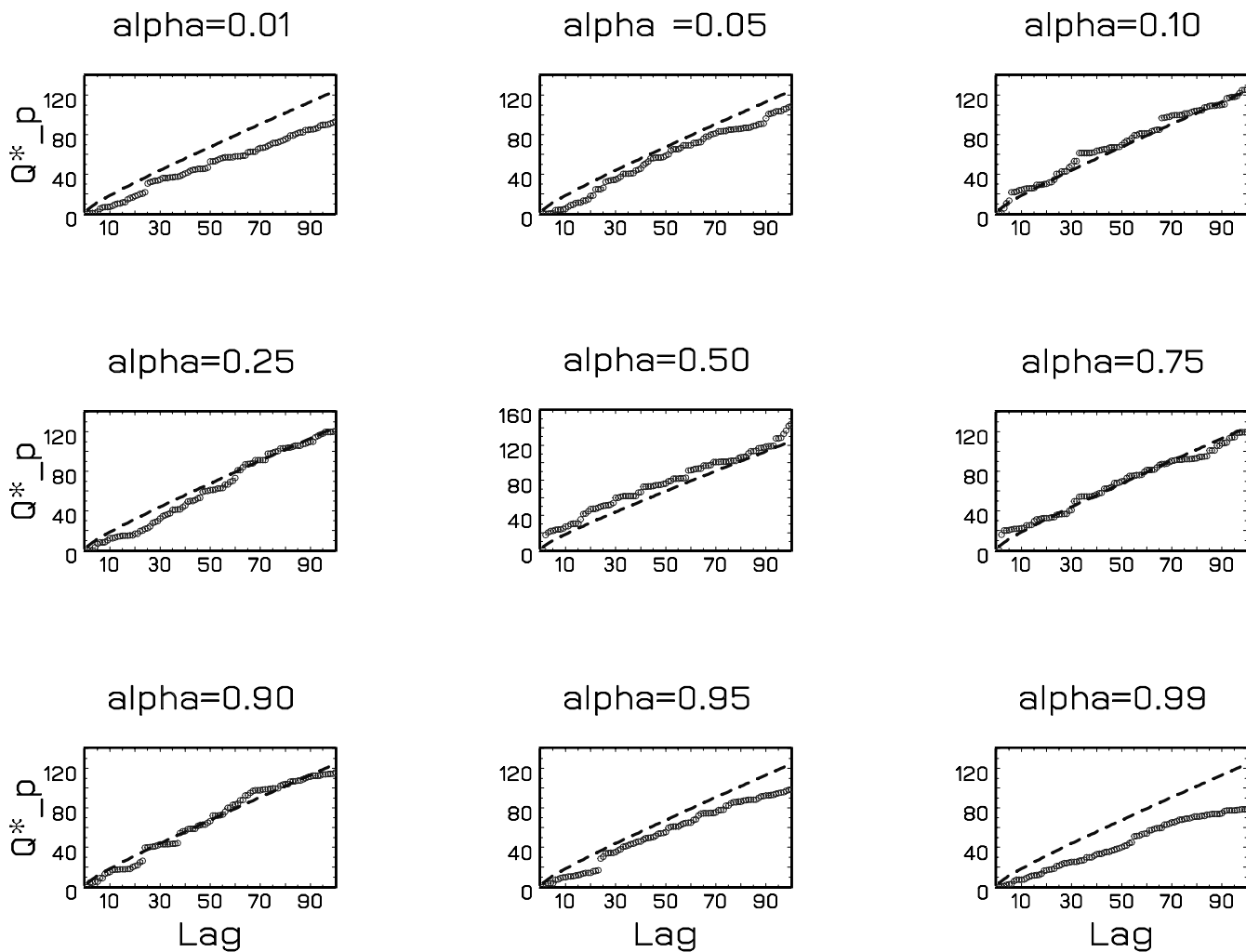


Figure 9. Standardized residuals from AR(2)/AGARCH(1,1) model fit on S&P500 Daily Data. Portmanteau Test statistic Q_p^* for each lag p and quantile α along with 95% liberal critical values