UNIVERSITY OF CALIFORNIA
RIVERSIDE

Essays on Multivariate Modeling in Financial Econometrics

A Dissertation submitted in partial satisfaction
of the requirements for the degree of

Doctor of Philosophy

in

Economics

by

Emre Yoldas

December 2008

Dissertation Committee:
Prof. Gloria González-Rivera, Chairperson
Prof. Marcelle Chauvet
Prof. Tae-Hwy Lee
Prof. Aman Ullah

The Dissertation of Emre Yoldas is approved:

_____

_____

_____

_____
Committee Chairperson

University of California, Riverside

ACKNOWLEDGEMENTS

First of all, I am grateful to my dissertation advisor, Prof. Gloria González-Rivera, for her excellent guidance, constant support, and invaluable suggestions. Her open-minded approach and appreciation for critical thinking always made me feel comfortable to express my own views and led to very fruitful academic discussions. I would not be able to become an independent academic researcher if it were not for her.

I am also indebted to the other members of my dissertation committee, Prof. Tae-Hwy Lee, Prof. Marcelle Chauvet, and Prof. Aman Ullah, for helping me at various stages of my studies and spending their valuable time with me whenever I needed. I am especially thankful to Prof. Lee for showing me the importance of hard-work and dedication, to Prof. Chauvet for always encouraging me to do my best and inviting me to a conference to present a paper, which was a great experience, and to Prof. Ullah for always being a source of wisdom and inspiration as an extraordinary academic.

I regard Prof. Jang-Ting Guo as the honorary member of my dissertation committee and I am grateful to him for being a great mentor and a close friend. He has also been a role model for me as a dedicated academic.

Special thanks go to my dear wife, Zeynep. She has been not only my loving and supporting wife, but also a great friend, office-mate, and co-author. She was always there for me, believed in me, and encouraged me to do my best. Without her I would not be able to achieve what I did.

Even though they have been far away most of the time, I have always felt the positive energy of my mom and sister with me. Talking to them on the phone over the weekends was one of the best ways to get over the endless stress caused by the graduate school. I cannot thank them enough. I would also like to thank my mother, father, and brother in-law for their support and always being there for Zeynep and me.

I would also like to thank Deniz Baglan, Daniel Farhad, Alan Krause, Charles Mutsalklisana, Ernesto Rattia-Lima, and Heather L.R. Tierney for their friendship and support in many ways. Heather especially inspired me with her uniquely positive attitude towards life and other people.

Finally, I would like to thank the staff members of UCR Economics Department and the Graduate Division for their help during my studies. Among them, Gary Kuzas deserves special thanks for always going the extra mile to be able to help everybody around him.

Thank you all for making the graduate school a great experience.

To my mom, dad, and sister

Mubeccel, Metin and Esra Yoldas

ABSTRACT OF THE DISSERTATION

Essays on Multivariate Modeling in Financial Econometrics

by

Emre Yoldas

Doctor of Philosophy, Graduate Program in Economics
University of California, Riverside, December 2008
Dr. Gloria González-Rivera, Chairperson

The main theme of this dissertation is multivariate modeling in financial econometrics. The first chapter uses the fundamental properties of the multivariate distributions of independent random variables to develop a new specification testing methodology for dynamic models. The second chapter generalizes this methodology to tests of distributional assumptions and dynamic specification in multivariate models. In the third chapter we focus on testing and modeling asymmetries in the second moments of multiple equity returns.

The methodological advances in nonlinear time series models with non-normal density functions and in density forecasting have emphasized the need for developing dynamic specification tests for the joint hypothesis of i.i.d.-ness and density functional form. In Chapter I, we propose a new battery of tests that rely on the fundamental properties of independent random variables with identical distributions and we introduce a graphical device -the autocontour-that helps to visualize the modeling problems. Based

on the theoretical probability coverage of the autocontours, we construct a battery of asymptotic t-tests and chi-squared tests, which have standard convergence rates. The tests are very powerful against violations of both hypotheses. They do not require either a transformation of the original data or an assessment of goodness-of-fit à-la Kolmogorov and explicitly account for parameter uncertainty. Monte Carlo simulations show that their finite sample performance is very good even in relatively small samples. We illustrate the usefulness of this methodology within the context of GARCH and ACD models using returns and duration data from the US equity markets.

In Chapter II, we generalize the testing methodology developed in Chapter I to time series models with multivariate GARCH disturbances. The tests are applied to the vector of generalized errors that must be i.i.d. with a certain parametric multivariate probability density function under the null hypothesis of correct specification. We develop t-tests based on a single autocontour and also more powerful chi-squared tests based on multiple autocontours. In the spirit of goodness-of-fit tests, we also propose an additional test that focuses on the multivariate density functional form of the vector of innovations. We perform Monte-Carlo simulations to investigate the size and power properties of the test statistics in finite samples. We apply our tests to multivariate GARCH models fitted to excess returns on portfolios sorted according to market capitalization.

In Chapter III we test and model asymmetries in time-varying means, volatilities, correlations, and betas of equity returns in a multivariate threshold framework. We consider alternative specifications in which the threshold variable is based on market

excess return, the Fama-French size and value factors, realized volatility of the market portfolio, and variables reflecting economic fundamentals. We find strong threshold effects with respect to market excess return, value premium, and term spread. Our results indicate that the threshold model based on market excess return provides a flexible and computationally inexpensive specification for modeling asymmetries, especially when dimensionality is high. We find that small caps, value stocks, and the Durables industry exhibit the strongest expected return asymmetries. Correlations of small caps, value firms, and defensive industries with the market tend to be significantly larger during market downturns. Correlation asymmetry usually translates into asymmetry in beta. Regime dependent volatility is a common characteristic of all portfolio groups. We evaluate performance of the proposed threshold models in an out-of-sample setup and find that there can be substantial economic gains in incorporating asymmetries in portfolio decisions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

AUTOCONTOURS: DYNAMIC SPECIFICATION TESTING

1 Introduction

When we specify a time series model, our purpose is to capture the time dependence of the variable of interest as parsimoniously as possible. If this task is accomplished, the innovation of the model will be an independent random variable. In this chapter, we focus on models in which all the dependence is contained in the first and second moments such that for a process $\{y_t\}$ we write $y_t = \mu_t(\theta_{01}, \Im_{t-1}) + \sigma_t(\theta_{02}, \Im_{t-1})\varepsilon_t$ for $t = 1,2,\ldots$ where $\mu_t(.)$ is the conditional mean and $\sigma_t^2(.)$ is the conditional variance, both functions of an information set $\Im_{t-1}$, $\theta_0 = (\theta_1', \theta_{02}')'$ is a parameter vector, and ε_t is an innovation that is independent and identically distributed. The innovation, $\varepsilon_t \equiv (y_t - \mu_t)/\sigma_t$, will be characterized by a parametric probability density function (p.d.f.), say $D(0,1)$. Within this context, it is easy to understand the importance of correctly specifying the conditional mean and conditional variance of $\{y_t\}$, as these two moments are the filters that contain the dependence of the data. However, what is the importance of correctly specifying the p.d.f. when robust estimation procedures are available? The correct specification of the p.d.f. has important implications for estimation, testing, and prediction, particularly in those instances in which the modeling of the conditional variance is the object of interest. Given that in many cases a quasi-maximum likelihood

(QML) estimator is available, the implementation of a maximum likelihood (ML) estimator, which requires the specification of the true density, must be desirable when the efficiency losses are large. In González-Rivera and Drost (1999), the efficiency losses of QML versus ML estimators are calculated both analytically and numerically for a variety of time series models. For some commonly used densities the losses are larger for the variance estimators than for the mean estimators. For instance, when the true density is a Student-t with 5 degrees of freedom, the efficiency loss for the variance parameters is about 150% compared to 25% for the parameter estimators in the mean. For a chi-square with 10 degrees of freedom, a similar comparison brings a loss of 167% versus 67%. These results underscore the importance of testing not only for independence in the innovation of the model but also for the correct density function.

There is an extensive literature on testing for density functional form. The pioneering work of Kolmogorov (1933) and Smirnov (1939) (K-S) on goodness-of-fit represents the basis of numerous parametric and non-parametric tests. The original work of K-S assumes that the observations are i.i.d. Lillefors (1967) shows that when parameter estimation is involved the asymptotic distribution of the K-S test depends on the true c.d.f. and population parameters, and critical values need to be tabulated. Andrews (1997) and Stinchcombe and White (1998) extend the K-S framework to test for the conditional density implied by a regression model within the context of i.i.d. observations. Nonparametric kernel based methods for testing conditional and unconditional densities of i.i.d. random variables are also proposed in Fan (1994) and Zheng (2000).

Within the context of the time series literature, testing for uncorrelatedness in the innovations of a time series model is a primary diagnostic check for model specification. During the 1970's with the advent of the Box-Jenkins methodology, portmanteau statistics as proposed by Box and Pierce (1970) and Ljung and Box (1978) were designed to test for autocorrelation in the residuals of ARMA and ARMA-X specifications. Since the prevalent distributional assumption in these studies was the normal density function, checking for the absence of autocorrelation implied an i.i.d. innovation process. During the 1980's and 1990's the econometric developments in nonlinear time series modeling and in particular the introduction of ARCH models by Engle (1982) brought further diagnostic checks to test for i.i.d-ness. The classical Box-Pierce-Ljung tests were not sufficient to test for the adequacy of non-linear specifications. Tests on squared-residual autocorrelations are found in McLeod and Li (1983), Li and Mak (1994), and, more recently, in Chen (2007). A different set of diagnostic tests to check for independence can be found in Brock et al. (1996), whose BDS test is based on the correlation dimension, and in Hong and Lee (2003), who propose a statistic based on the generalized spectral density. Additionally, time series modeling of the conditional variance underscored the importance of specifying the correct density function, mainly when unconditional and conditional normality has been almost universally rejected in the econometric analysis of financial time series. Thus, nonlinear modeling in conjunction with non-normal densities emphasizes the need for development of dynamic specification tests for the joint hypothesis of i.i.d.-ness and density functional form.

With the revival of density forecasting towards the end of the 1990's, testing for density functional form has stirred a greater interest both in academia and among practitioners. Most of the density forecast evaluation methods are based on Rosenblatt's probability integral transform. Diebold et al. (1998) show that if the proposed density forecast, say $f_t(.)$, is correct, then the transformed random variables $u_t = F_t(y_t \mid \Im_{t-1}; \theta_0)$ will be i.i.d. uniformly distributed. The testing of i.i.d. uniformly distributed random variables is carried out in an informal way by means of a histogram of $\{u_t\}$ and autocorrelograms of different powers of it. Berkowitz (2001) proposes a further transformation to normality of the process $\{u_t\}$. If $\Phi(.)$ is the normal cumulative distribution function then $z_t \equiv \Phi^{-1}(u_t)$ must be an i.i.d. $N(0,1)$ random variable. Since Berkowitz's test is a likelihood ratio test, it has power only against fixed alternatives. Chen and Fan (2004) generalize Berkowitz's approach by using copula functions. None of the three aforementioned articles deals with parameter uncertainty since they treat $\{u_t\}$ as the primitive process. Bai (2003) studies this type of tests further and proposes a conditional Kolmogorov test to test for the properties of $\{u_t\}$. He deals with parameter uncertainty using a martingale transformation and obtains an asymptotically distribution free test statistic. Corradi and Swanson (2006) note that Bai's test does not have power against violations of independence and they propose a Kolmogorov type test that, like the non-parametric test by Hong (2001), is robust to dependence. Corradi and Swanson rely on bootstrap techniques to deal with the effects of parameter uncertainty on the limiting distribution of the statistic. An alternative route to the construction of robust tests against

dependence is to search for tests with power against violations of both dependence and density functional form. In this line, Hong and Li (2005), again within the context of density forecast, emphasize the importance of jointly checking for i.i.d.-ness and uniformity of $\{u_t\}$. They propose a non-parametric test statistic based on the squared distance between the joint density of u_t and u_{t-j} under the null and its nonparametric kernel based estimator. The asymptotic distribution of their test is a standard normal and it is immune to parameter estimation due to the fact that parameter estimators converge at the standard rate $\sqrt{T}$ while the nonparametric kernel estimators that form the test statistic converge at slower rates. As in any nonparametric procedure, the choice of the optimal bandwidth is an issue.

In a different vein Bontemps and Meddahi (2005, 2006) adopted the GMM approach for testing distributional assumptions based on moment conditions for the Pearson's family of distributions. They describe moment conditions that are robust against parameter uncertainty.

In this chapter we propose a battery of tests for the joint hypothesis of i.i.d.-ness and density functional form that are very powerful against violations of both. The tests have standard convergence rates and standard limiting distributions. They do not require either a transformation of the original data or an assessment of goodness-of-fit *à-la* Kolmogorov, and explicitly account for parameter uncertainty. The proposed tests focus on fundamental properties of independent random variables with identical distributions. Let the process under consideration be $\{\varepsilon_t\}$ with density $f(.)$. The random variables in this process are independent if and only if their multivariate distribution is equal to the

product of their marginal distribution functions, in which case the null hypothesis simply boils down to $f(\varepsilon_{t-k_1},\varepsilon_{t-k_2},...,\varepsilon_{t-k_m}) = f(\varepsilon_{t-k_1})f(\varepsilon_{t-k_2})\cdots f(\varepsilon_{t-k_m}), \{k_j\}_{j=1}^{m} \in \mathrm{N}$. The specification tests we propose are based on a new concept that we term *autocontour*. Under the null, we horizontally slice the joint density at different levels and project the resulting segments down to the hyperplane $(\varepsilon_{t-k_1},\varepsilon_{t-k_2},...,\varepsilon_{t-k_m})$. The projection is the autocontour containing a known percentage of the observations. Based on the sample estimates of these percentages we construct a battery of *t*-statistics and chi-squared statistics, which have standard asymptotic distributions. Our tests can be applied to primitive series and to residuals series, in which case we need to address the parameter uncertainty problem. While it is possible to obtain analytical results for special cases for the effect of parameter uncertainty on the asymptotic distribution of the tests, we show that a general bootstrap procedure to obtain their asymptotic variance delivers standard asymptotic tests with the right size.

The structure of the chapter is as follows. In Section 2, we formalize the notion of autocontour and present the general framework of our testing methodology by introducing the resulting *t* and chi-squared test statistics. We also illustrate the application of our methodology to commonly used distributions: normal, Student-t, and exponential. In Section 3 we explicitly deal with the parameter uncertainty problem by deriving the asymptotic distributions of our test statistics and using a bootstrap procedure to make them feasible. In Section 4, we provide extensive Monte Carlo evidence regarding the size and power properties of the proposed tests. In Section 5, we show several empirical

applications within the context of GARCH and ACD modeling; and in Section 6 we conclude.

2 The Joint Test of Density and Independence

The class of dynamic models that we are interested in are of the following form,

$$y_t = \mu_t(\theta_{01}, \Im_{t-1}) + \sigma_t(\theta_{02}, \Im_{t-1})\varepsilon_t, \quad t = 1,...,T, \tag{1}$$

where $\Im_{t-1}$ denotes the information set available at time $t-1$, $\mu_t(.)$ and $\sigma_t(.)$ are fully parameterized by $\theta_0 = (\theta_{01}', \theta_{02}')'$ and measurable with respect to $\Im_{t-1}$, and $\{\varepsilon_t\}_{t=1}^T$ is a series of i.i.d. innovations having a particular density function, $f(.)$. Usually, ε_t is assumed to have zero mean and unit variance, but for nonnegative data it naturally has a nonzero mean. For the moment we assume that ε_t is observable, i.e. θ_0 is known. Later on we will relax this assumption to account for the effects of estimation on distribution of our test statistics.

2.1 Autocontour

Under correct dynamic specification the null hypothesis in its most general form is stated as,

$$H_0 : \varepsilon_t \text{ is i.i.d. with density } f(.)$$
$$H_1 : \text{negation of } H_0$$

Under this null hypothesis the multivariate density function for an *m*-dimensional vector $(\varepsilon_{t-k_1},\ldots,\varepsilon_{t-k_m})$ is written as $f(\varepsilon_{t-k_1},\varepsilon_{t-k_2},...,\varepsilon_{t-k_m}) = f(\varepsilon_{t-k_1})f(\varepsilon_{t-k_2})\cdots f(\varepsilon_{t-k_m})$. We define the (α, m)-autocontour, ACR_α^m, as the set of points in the hyperplane

$(\varepsilon_{t-k_1},\ldots,\varepsilon_{t-k_m})$ that results from horizontally slicing the multivariate density function at a certain value to guarantee that the resulting set contains $\alpha\%$ of observations, that is,

$$ACR_a^m := \left\{ B(\varepsilon_{t-k_1},\ldots,\varepsilon_{t-k_m}) \subset \Re^m \left| \int_{l_1}^{u_1} \cdots \int_{l_m}^{u_m} f(\varepsilon_{t-k_1}) \ldots f(\varepsilon_{t-k_m}) d\varepsilon_{t-k_1} \ldots d\varepsilon_{t-k_m} \le \alpha \right. \right\} \quad 2)$$

where B is a set in $\Re^m$, $u_m \equiv u_m(\varepsilon_{t-k_1},\ldots,\varepsilon_{t-k_{m-1}})$, $l_m \equiv l_m(\varepsilon_{t-k_1},\ldots,\varepsilon_{t-k_{m-1}})$, and $\{k_j\}_{j=1}^m \in \mathrm{N}$. Let us focus on the bivariate case to show the construction of the autocontour for some specific distributions commonly encountered in financial econometrics: standard normal, Student-t, and exponential. In the bivariate case, the null hypothesis reduces to $H_0 : f(\varepsilon_t, \varepsilon_{t-k}) = f(\varepsilon_t) f(\varepsilon_{t-k})$ for $k = 1,\ldots,K$ and K is some positive integer. The implied bivariate autocontour is given by

$$ACR_{\alpha,k} := \left\{ B(\varepsilon_t, \varepsilon_{t-k}) \subset \Re^2 \left| \quad \int_{l_0}^{u_0} \int_{l_k(\varepsilon_t)}^{u_k(\varepsilon_t)} f(\varepsilon_t) f(\varepsilon_{t-k}) d\varepsilon_t d\varepsilon_{t-k} \le \alpha \right. \right\} \qquad (3)$$

where B is a set on the plane $(\varepsilon_t, \varepsilon_{t-k})$ and the limits of integration are such that the contour shape of the hypothesized density is preserved.

2.1.1 Standard Normal Distribution

Suppose $\varepsilon_t \sim \text{i.i.d.}\, N(0,1)$, then the joint density of interest is given by

$$f(\varepsilon_t, \varepsilon_{t-k}) = \frac{1}{2\pi} \exp\left(-\frac{1}{2}(\varepsilon_t^2 + \varepsilon_{t-k}^2) \right) \qquad (4)$$

For a fixed value of this density, say $\bar{f}$, we have $\varepsilon_t^2 + \varepsilon_{t-k}^2 = a^2$ where $a = \sqrt{-2\ln(2\pi\bar{f})}$. Thus, autocontours are circles with radius a. The value of a that describes the

autocountour with $\alpha\%$ coverage can be computed by numerical integration using the following equation,

$$\int_{-a}^{a}\int_{-g(\varepsilon_t)}^{g(\varepsilon_t)} f(\varepsilon_t, \varepsilon_{t-k}) d\varepsilon_t d\varepsilon_{t-k} = \alpha \tag{5}$$

where $g(\varepsilon_t) = \sqrt{a^2 - \varepsilon_t^2}$. Alternatively one can get the *a* values based on the c.d.f. of a chi-squared random variable with two degrees of freedom due to the normality assumption, e.g. for 90% autocontour $a = \sqrt{4.61}$.

2.1.2 Student-t Distribution

Now let us consider the case where $\varepsilon_t \sim$ i.i.d. Student-t(v). In this case, the joint density under the null is given by,

$$f(\varepsilon_t, \varepsilon_{t-k}) = A\left(1 + \frac{\varepsilon_t^2}{v}\right)^{\frac{-(v+1)}{2}}\left(1 + \frac{\varepsilon_{t-k}^2}{v}\right)^{\frac{-(v+1)}{2}} \tag{6}$$

where $A = [\Gamma((v+1)/2)]^2 \big/ \pi v[\Gamma(v/2)]^2$. The equation pertaining to autocontours of this joint density is given by $a = 1 + (\varepsilon_t^2 + \varepsilon_{t-k}^2)/v + \varepsilon_t^2\varepsilon_{t-k}^2/v^2$ where $a = (\bar{f}/A)^{-2/(v+1)}$.[1] We rely on numerical integration to find the value of *a* that describes the autocountour with $\alpha\%$ coverage

$$\int_{-\sqrt{v(a-1)}}^{+\sqrt{v(a-1)}}\int_{-g(\varepsilon_t)}^{g(\varepsilon_t)} f(\varepsilon_t, \varepsilon_{t-k}) d\varepsilon_t d\varepsilon_{t-k} = \alpha \tag{7}$$

where $g(\varepsilon_t) = \sqrt{(a - 1 - \varepsilon_t^2/v)/(1/v + \varepsilon_t^2/v^2)}$.

[1] Note that the joint distribution given by the product of two marginal Student-t densities is not bivariate Student-t as opposed to the case of normal distribution.

2.1.3 Exponential Distribution

Finally we consider the case where $\varepsilon_t \sim$ i.i.d. exp(β). The bivariate exponential density under the null is given by,

$$f(\varepsilon_t, \varepsilon_{t-k}) = \frac{1}{\beta^2} \exp\left(-\frac{1}{\beta}(\varepsilon_t + \varepsilon_{t-k}) \right) \tag{8}$$

For some fixed value of this density, the equation for the autocontours is given by $a = \varepsilon_t + \varepsilon_{t-k}$ where $a = -\beta \ln(\beta^2 \bar{f})$. Since we restrict our attention to the positive quadrant we obtain contours that are isosceles triangles. For a given α, the following equation can be iterated on a, until the desired probability is obtained.

$$\int_0^a \int_0^{a-\varepsilon_t} f(\varepsilon_t, \varepsilon_{t-k}) d\varepsilon_t d\varepsilon_{t-k} = \alpha \tag{9}$$

In Figure 1.1, we show the graphical contours corresponding to three aforementioned bivariate density functions.

2.2 Test Statistics and Asymptotic Distributions

For a given autocontour $ACR_{\alpha_i,k}$, we define a binary variable as follows

$$I_t^{k,i} = \begin{cases} 1 \text{ if } (\varepsilon_t, \varepsilon_{t-k}) \notin ACR_{\alpha_i,k} \\ 0 \text{ if } (\varepsilon_t, \varepsilon_{t-k}) \in ACR_{\alpha_i,k} \end{cases} \quad t = k+1, ..., T, \tag{10}$$

where $k = 1, \ldots, K$ and $i = 1, \ldots, C$, i.e. K is the number of lags and C is the number of autocontours. Hence, this Bernoulli random variable takes on value 1 if an observation falls outside the autocontour and 0 otherwise. Since $ACR_{\alpha_i,k}$ contains α_i% of observations, we expect to have $(1-\alpha_i)$% outside the autocontour. Let $p_i \equiv 1 - \alpha_i$. Under

the null we have $E[I_t^{k,i}] = p_i$ and $Var(I_t^{k,i}) = p_i(1-p_i)$. Furthermore, the indicator is a linearly dependent process with a MA structure. Its autocovariance function is given by

$$\gamma_h^i \equiv Cov(I_t^{k,i}, I_{t-h}^{k,i}) = \begin{cases} P(I_t^{k,i}=1, I_{t-h}^{k,i}=1) - p_i^2 & \text{if } h = k \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

Our first test statistic is a *t*-statistic based on this indicator series.

Proposition 1 Let $\hat{p}_i^k = (T-k)^{-1}\sum_{t=1}^{T-k} I_t^{k,i}$. Under the null hypothesis,

$t_{k,i} = \sqrt{T-k}(\hat{p}_i^k - p_i) \xrightarrow{d} N(0, \sigma_{k,i})$ where $\sigma_{k,i}^2 = p_i(1-p_i) + 2\gamma_k^i$.

Proof. Please see Appendix A for the proofs of Propositions 1-4.

Note that since p_i is given under the null, to make this test operational we need to replace only the autocovariance, γ_k^i, with its consistent estimate given by

$$\hat{\gamma}_k^i = \frac{1}{T-2k}\sum_{t=1}^{T-2k} I_{t+k}^{k,i} I_t^{k,i} - \left(\frac{1}{T-k}\sum_{t=1}^{T-k} I_t^{k,i}\right)^2.$$

For a given autocontour *i*, we can examine the lag structure of $t_{k,i}$ for $k = 1,\ldots,K$ and collect those *t*-statistics in a graph, which we call *autocontourgram.* As an illustration, consider the daily returns of NYSE index from June 1, 1995 to December 31, 2004. In Figure 1.2-a, we plot the standardized return and the bivariate normal autocontours for $k = 1$. Figure 1.2-b shows the autocontourgram that summarizes the values of $t_{k,i}$ for $k = 1,\ldots,25$ based on three autocontours ($\alpha \in \{10, 50, 99\%\}$).

As expected, we clearly reject the null hypothesis. It can be seen from the autocontourgram that the rejection in the largest autocontour $\alpha = 99\%$ is due to the larger than expected theoretical number of observations outside of the autocontour, thus the tails are thicker than those of a normal. On the contrary, the rejection in the central autocontours $\alpha = 10\%$ and $\alpha = 50\%$ is due to the smaller than expected number of observations outside of the autocontours, thus the central empirical autocontours are more dense than the theoretical ones.

In the spirit of Box-Pierce-Ljung statistics, the information contained in the individual $t_{k,i}$ statistics can be pooled either across *K* lags or across *C* contours. The following two test statistics consider the joint distribution of asymptotically normal random variables associated with different lags or with different contours.

Proposition 2 For a given autocontour *i*, consider all lags up to *K*. Let $q_{k,i} = \sqrt{T-k}(\hat{p}_i^k - p_i)$, $k = 1,\ldots,K$ and stack them in a vector $\mathbf{q}_i = (q_{1,i},...,q_{K,i})'$. Under the null we have $\mathbf{q}_i \xrightarrow{\text{d}} N(\mathbf{0},\mathbf{\Omega}_i)$ where any element $\omega_{i,kl}$ in $\mathbf{\Omega}_i$ is given by

$$\omega_{i,kl} = \begin{cases} Cov(I_t^{k,i}, I_t^{l,i}) + Cov(I_t^{k,i}, I_{t-k}^{l,i}) + Cov(I_t^{l,i}, I_{t-l}^{k,i}) + Cov(I_t^{l,i}, I_{t-l+k}^{k,i}) , & l > k \\ p_i(1-p_i) + 2\gamma_k^i , & l = k \end{cases}$$

It directly follows that $Q_i^K \equiv \mathbf{q}_i'\mathbf{\Omega}_i^{-1}\mathbf{q}_i \xrightarrow{d} \chi_K^2$.

Proposition 3 For a given lag k, consider multiple contours. Let $z_{i,k} = \sqrt{T-k}(\hat{p}_i^k - p_i)$, $i = 1,\ldots,C$, and stack them in a vector $\mathbf{z}_k = (z_{1,k},...,z_{C,k})'$. Under the null we have $\mathbf{z}_k \xrightarrow{\text{d}} N(\mathbf{0},\Xi_k)$ where any element $\xi_{k,ij}$ in Ξ_k is given by $\xi_{k,ij} = \min(p_i, p_j) - p_i p_j + Cov(I_t^{k,i}, I_{t-k}^{k,j}) + Cov(I_{t-k}^{k,i}, I_t^{k,j})$, $\forall i, j$. Then, it directly follows that $J_k^C \equiv \mathbf{z}_k' \Xi_k^{-1} \mathbf{z}_k \xrightarrow{d} \chi_C^2$.

To make these test statistics operational we replace the covariance terms with their corresponding sample counterparts. This has no impact on the asymptotic distributions.

3 Parameter Uncertainty

Even though the tests we propose can be applied to raw data, they will be most useful as a diagnostic tool for model specification. Thus, in practice we will be analyzing residuals, $\hat{\varepsilon}_t(\hat{\theta})$, which depend on parameter estimates, instead of the true error $\varepsilon_t(\theta_0)$. Our tests are subject to the uncertainty created by parameter estimation. The following discussion will be centered around the *t*-statistics, $t_{k,i}$, considered in Proposition 1, but the same conclusions will apply to the Q_i^K and J_k^C tests given in Propositions 2 and 3 respectively.

To understand how parameter estimation affects the tests, let us consider the following mean value expansion[2]

[2] For notational parsimony we use T instead of $(T-k)$ since they are asymptotically equivalent.

$$\sqrt{T}\left(\hat{p}_i^k(\hat{\theta}) - p_i\right) = \sqrt{T}\left(\hat{p}_i^k(\theta_0) - p_i\right) + \sqrt{T}\left(\hat{\theta} - \theta_0\right)'\left(\frac{\partial \hat{p}_i^k(\theta)}{\partial \theta}\Big|_{\theta=\theta^*} \right), \quad (12)$$

where θ^* is between θ_0 and $\hat{\theta}$. Applying Slutsky's Theorem yields

$$\sqrt{T}\left(\hat{p}_i^k(\hat{\theta}) - p_i\right) = \sqrt{T}\left(\hat{p}_i^k(\theta_0) - p_i\right) + \sqrt{T}\left(\hat{\theta} - \theta_0\right)' \lim_{T\to\infty} E\left(\frac{\partial \hat{p}_i^k(\theta)}{\partial \theta}\Big|_{\theta=\theta_0} \right) + o_p(1). \quad (13)$$

We make the following assumptions to obtain the asymptotic distribution under parameter uncertainty:

A1. $\sqrt{T}\left(\hat{\theta} - \theta_0\right) \xrightarrow{d} N(0, A^{-1}BA^{-1})$ where $A \equiv E\left[-H(\theta_0)\right]$, $B \equiv E\left[S(\theta_0)S(\theta_0)'\right]$, and $H(\theta_0) = T^{-1}\sum_{t=1}^{T} H_t(\theta_0)$ and $S(\theta_0) = T^{-1/2}\sum_{t=1}^{T} s_t(\theta_0)$ are the Hessian matrix and the score vector corresponding to QML estimation.

A2. Let $D \equiv \lim_{T\to\infty} E\left(\frac{\partial \hat{p}_i^k(\theta)}{\partial \theta}\Big|_{\theta=\theta_0} \right)$. Assume $D_j < \infty$ for $j = 1, \ldots, q$ where q is the dimensionality of the parameter space.

A3. $Cov(I_t^{k,i}, s_t(\theta_0)) < \infty$ and $Cov(I_t^{k,i}, s_{t-k}(\theta_0)) < \infty$.

A1 is based on standard QMLE arguments. A2 guarantees boundedness of the gradient vector. A3 is a weak assumption that is required to have a well defined asymptotic variance. A2 and A3 can be analytically verified for commonly used models (please see Proposition 5).

Proposition 4 Under A1-A3 we have $\sqrt{T}\left(\hat{p}_i^k(\hat{\theta}) - p_i\right) \xrightarrow{d} N(0, \tau_{k,i}^2)$ where

$\tau_{k,i}^2 = \sigma_{k,i}^2 + D'(A^{-1}BA^{-1})D + 2E\left[\sqrt{T}(\hat{p}_i - p_i)S(\theta_0)'\right]A^{-1}D$.

Proposition 5 For a Gaussian location-scale model, $y_t = \mu_0 + \sigma_0\varepsilon_t$, $\varepsilon_t \sim$ i.i.d. N(0,1), the gradient vector *D* is equal to $D = (0, -a_i f_{z_t}(a_i)/\sigma_0^2)'$ where f_{z_t} is the p.d.f. of z_t, a chi-squared random variable with 2 degrees of freedom; and the covariance term is

$E[\sqrt{T}(\hat{p}_i - p_i)S(\theta_0)'] = (0,\ (E[I_t^{k,i}\varepsilon_t^2] - p_i)/2\sigma_0^2)'$.

Proof. See Appendix B.

This proposition says that the estimation of the parameters in the mean does not affect the asymptotic distribution of the test ($D_1 = 0$).[3] In general it would be difficult to obtain an empirical counterpart of the gradient vector *D*. In addition, for some models the covariance terms given in A3 may be difficult to estimate, e.g. simulation based methods reviewed in Gouriéroux and Monfort (1996). Therefore, we propose to estimate the asymptotic variance $\tau_{k,i}^2$ using a bootstrap procedure. This is a commonly used approach in the literature to overcome the difficulties associated with asymptotic variance estimation in various contexts (see Efron (1979), Buchinsky (1995) and Ledoit et al. (2003) among others). The bootstrap estimator of the variance in Proposition 4 is given by

[3] This result also holds for a model with Student-t innovations. We conjecture that with homoscedastic innovations, the symmetry of the density function is a sufficient condition for the estimation of the mean parameters not to have any effect in the asymptotic distribution of the test.

$$\hat{\tau}_{k,i}^2 = \frac{T}{B-1}\sum_{b=1}^{B}\left(\tilde{p}_i^k(\tilde{\theta}_b) - \frac{1}{B}\sum_{b=1}^{B}\tilde{p}_i^k(\tilde{\theta}_b)\right)^2, \tag{14}$$

where B is the total number of bootstrap samples, $\tilde{\theta}_b$ is the parameter estimator from the b^{th} bootstrap sample.[4] We prefer a parametric bootstrap since the null hypothesis fully specifies the parametric DGP.[5] In particular, bootstrap samples are obtained from Equation (1) by replacing θ_0 with $\hat{\theta}$ and generating ε_t from the specified parametric distribution. Under suitable regularity conditions, this estimator should be consistent as proven within the linear regression context for i.i.d observations (Liu and Singh, 1992) and for dependent data (Goncalves and White, 2005).

Alternatively, one can bootstrap the full distribution of the test statistics. However, our test statistics are not asymptotically pivotal under parameter uncertainty and this implies that the bootstrap distribution does not necessarily provide a superior approximation to the finite sample distributions of test statistics, see Horowitz (2001). Monte Carlo results (to be presented in the following section) indicate that bootstrapping the asymptotic covariance matrices and using standard asymptotic critical values delivers remarkable results in terms of size and power of the tests.

[4] For the chi-squared statistics, the covariance matrix estimators are defined analogously.
[5] See Horowitz (2001) for a detailed discussion on the practical implementation of bootstrap techniques.

4 Monte Carlo Simulations

In this section we investigate size and power properties of our test statistics in finite samples by Monte Carlo (MC) simulations. We first present results for the case of observable data followed by simulation evidence on model residuals.

4.1 The Case of Observable Data

For size simulations we consider the following three cases: (i) $\varepsilon_t \sim$ i.i.d. $N(0,1)$, (ii) $\varepsilon_t \sim$ i.i.d. Student-t(5), and (iii) $\varepsilon_t \sim$ i.i.d. exp(1). The Gauss 7.0 random number generator is used to generate pseudo random numbers from these three distributions. The number of Monte Carlo replications is equal to 10,000. We consider 13 autocontours $(C = 13)$ with coverage levels (%): 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 99, spanning the entire density function. We start with a sample size of 250 and consider increments of 250 up to 2,000. The figures show results for all sample sizes whereas the tables report results for sample sizes 250, 500, 1,000 and 2,000. In all cases the nominal size is 5%.[6]

In Figure 1.3 we show the simulated size results associated with $t_{1,i}$, $i = 1,\ldots,13$ for normal, Student-t and exponential distributions respectively. For all three cases, we observe that that our *t*-statistics do not suffer from any systematic size distortions with the exception of 1% and 99% autocontours when $T = 250$. This result is not surprising because for small samples there is not enough variation in the indicator series at the extreme autocontours (1% or 99% coverage). In order to check the size robustness of the *t*-statistics for different lags, we exclusively focus on the 50% autocontour $(i = 7)$ and

[6] We also have simulation results for 1% and 10% nominal size levels, which are in line with the results for the 5% level. They are all available from the authors upon request.

consider $k = 1,\ldots,5$. We report these results in Table 1.1. Overall the simulated size values are around 5% indicating good size for all distributions.[7]

Simulated size values for the chi-squared statistic across lags for the 50% autocontour, i.e. Q_7^K, $K = 1,\ldots,5$, are reported in Table 1.2. In general, this test statistic is oversized. The distortions in size are larger for small samples and especially for large values of *K*. This result stems from the difficulty associated with the estimation of off-diagonal terms of the covariance matrix given in Proposition 2. For large values of *K* the number of terms to be estimated in the covariance matrix increases at a higher rate than *K*. Therefore, one needs more data to achieve a reasonable size. We have run an experiment with 5,000 observations and we have observed that even for $K = 5$ there is almost no distortion. From a practical point of view, one may want to choose a small number of lags if the sample size is relatively small.

Simulated size values for the chi-squared statistic across autocontours for $k = 1$ (J_1^{13} and J_1^7) are reported in Table 1.3. Although this test has a better size than the *Q*-statistic, it is still oversized in small samples, especially when we consider the full set of 13 contours. As we mentioned before, in the extreme contours (i.e., 1 and 99%) the indicator series may not exhibit enough variation in small samples and it may affect the finite sample performance of the *J*-statistic. We have dropped the first three (1, 5, 10%) and the last three (90, 95, 99%) autocontours to simulate the empirical size of J_1^7. There is a significant improvement in the size for all three distributions and for all sample sizes.

[7] These results are robust across all the autocontours and also available from the authors upon request.

The size is acceptable for the most reasonable sample sizes encountered in financial data sets. From a practical point of view, one may not want to use autocontours too small or too large when the sample size is small.

To analyze the power properties of our test statistics we consider three alternative data generating processes:

(i) $\varepsilon_t = \varphi\varepsilon_{t-1} + u_t$ where $u_t \sim$ i.i.d. $N(0,1-\varphi^2)$,

(ii) $\varepsilon_t = u_t\sqrt{(v-2)/v}$ where $u_t \sim$ i.i.d. Student-t(v),

(iii) $\varepsilon_t = \sqrt{h_t}u_t$ where $u_t \sim$ i.i.d. $N(0,1)$ and $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$

For all three cases the null hypothesis is $\varepsilon_t \sim$ i.i.d. $N(0,1)$. In case (i) we investigate departures from the independence hypothesis by considering different values of $\varphi \in \{0.3, 0.5, 0.7, 0.9\}$. In case (ii), we maintain the independence hypothesis and investigate departures from the hypothesized density functional form by generating i.i.d. data from Student-t distribution for three different values of the shape parameter $v \in \{5, 10, 15\}$.[8] Finally, in case (iii) we analyze departures from both dependence and functional form by generating data from a GARCH (1,1) model with $\alpha \in \{0.05, 0.1, 0.15\}$, $\beta \in \{0.9, 0.85, 0.8\}$. We set $\omega = 1-\alpha-\beta$ to normalize the unconditional variance to one.

The results of power simulations for $t_{1,i}, i = 1,\ldots,13$ are presented in Figures 1.4-1.6. Figure 1.4 exhibits the results for case (i) for different degrees of dependence. For high values of the autoregressive parameter, i.e. $\varphi = 0.7$ and 0.9, and for autocontours

[8] The variance is normalized to unity to control for the effects of different moment structures of normal and Student-t distributions on the simulation results.

ranging from 5 to 50% coverage, the power of the *t*-statistics approaches 1 rather quickly. There is a substantial drop in power for autocontours with 70-90% coverage even with a high degree of dependence. There is not a general explanation for this phenomenon; we understand it as a result due to the specific DGP and the density under the null. In the following power simulations we also observe similar behavior of the power surfaces around the same contours. From a practical point of view, one may want to focus on the most informative autocontours, which are those covering the center and the tails of the distribution, i.e. $\leq 50\%$ or $\geq 95\%$.

The results for case (ii) are reported in Figure 1.5. The range of autocontours that yield the highest power is wider for this case ranging from 5 to 70%. The power surface exhibits the lowest values at the 90% contour. In the 99% autocontour, the power is as high as in the central autocontours for all sample sizes. This is expected because of the leptokurtosis of the Student-t density. As anticipated, rejection rates decrease for larger values of *v*, for which the null and the alternative become less distinguishable.

The results for case (iii) are presented in Figure 1.6. The data generated under this DGP is uncorrelated but nonlinearly dependent, and there is excess kurtosis relative to the normal distribution. Persistence is the same across three alternative parameterizations $(\alpha + \beta = 0.95)$, but kurtosis is increasing with α. The power of the test is the highest for the largest levels of kurtosis indicating that the *t*-tests are stronger at detecting departures from the correct density functional form than departures from the independence hypothesis. When $\alpha = 0.05$, the excess kurtosis in the data is only 0.16, so deviations from the null are mainly due to the dependence in higher moments.

In Table 1.4 we report the power simulation results for J_k^{13}, $k = 1,\ldots,5$ for the three DGP's mentioned above.[9] A common characteristic across all the three DGP's is that, for a given sample size, the power of the *J*-statistics is roughly the same as the maximum power of the *t*-statistics. The pattern on rejection frequencies is the same as that of the individual *t*-statistics. In case (i), the rejection is stronger when there is high dependence even for small samples. We note that there is a decrease in power as *k* increases, which is more evident when the autoregressive parameter is small. This is somehow expected because the DGP is only an AR(1) process. In case (ii), the test is more powerful for small degrees of freedom but we should mention that even for $v = 15$ the power of the test is still around 43% for 2000 observations. In case (iii), we have stronger rejections rates when the data exhibits higher kurtosis.

4.2 Model Residuals

In this section, we analyze the size and power properties of our tests when they are applied to the residuals generated by model estimation. For the size experiments we use the same set of distributions as in section 4.1 and consider location-scale models with the following specifications:

(i) $y_t = \mu + \sigma\varepsilon_t$, $\varepsilon_t \sim$ i.i.d. $N(0,1)$, $\mu = 1.25$ and $\sigma = 2$;

(ii) $y_t = \mu + \sigma\varepsilon_t\sqrt{(v-2)/v}$, $\varepsilon_t \sim$ i.i.d. Student $-$ t(v), $\mu = 1.25$, $\sigma = 2$, and $v = 5$;

(iii) $y_t = \beta\varepsilon_t$, $\varepsilon_t \sim$ i.i.d. exp(1), and $\beta = 1.25$.

[9] Power simulations for *Q*-statistics are deferred to the next section on "model residuals".

For each case, we generate the data y_t and proceed to estimate μ and σ. We retrieve the residuals and apply the tests to the properly standardized residuals. Because of computational considerations, we reduce the number of Monte Carlo replications to 1,000. The number of bootstrap replications is 500.

In Table 1.5 we report the results for $t_{1,i}, i = 1,...,13$. There are not systematic deviations from the nominal size either across autocontours or across distributions. These results follow very closely those for observable data reported in Figure 1.3 and Table 1.1.

In Table 1.6, we report the size for 50% autocontour for lags 1 to 5 and we observe that the size of the *t*-statistics is very robust to the choice of lags.

In Tables 1.7 and 1.8 we report the size results for the *Q*-statistic and the *J*-statistic respectively. We note that bootstrapping the variance helps tremendously to correct the size of the *Q*-statistic, which consistently over-rejects in the case of observable data reported in Table 1.2. For the *J*-statistic, when we consider all 13 autocontours, there is a tendency for the test to over-reject but when we focus on the middle autocontours by removing the first and the last three, the empirical size improves substantially, approaching the nominal size even for small samples.

To investigate the power of our test statistics under parameter uncertainty, we consider the same DGP's as in section 4.1, and apply the tests to the standardized residuals generated by the estimation of the following location-scale models:

(i) $y_t = \mu + \sigma\varepsilon_t,\ \varepsilon_t = \varphi\varepsilon_{t-1} + u_t$ where $u_t \sim$ i.i.d. $N(0,1-\varphi^2)$,

(ii) $y_t = \mu + \sigma\varepsilon_t\sqrt{(v-2)/v}$ where $\varepsilon_t \sim$ i.i.d. Student $-t(v)$,

(iii) $y_t = \mu + \sigma\varepsilon_t,\ \varepsilon_t = \sqrt{h_t}u_t$ where $u_t \sim \text{i.i.d.}N(0,1),\ \ h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$

In Figures 1.7, 1.8 and 1.9 we present the power surfaces for $t_{1,i}, i = 1,\ldots,13$ for samples sizes ranging from 250 to 2000 observations and for the 13 autocontours. These figures are similar to Figures 1.4, 1.5, and 1.6 for the case of observable data. The common characteristic to these three figures is that bootstrapping the variance of the test lifts the power surfaces up. In case (i) this increase in power is more evident for the largest three autocontours (90, 95, 99%). We also observe higher power at lower dependence levels in all autocontours with the exception of 80%. In this case the drop in power is similar to that we observe in Figure 1.4. For case (ii), power increases for all autocontours especially for larger values of the shape parameter. Finally, in case (iii) we observe that using bootstrap standard errors boosts the power to around 40% in those cases when the null and the alternative are almost undistinguishable.[10]

In Table 1.9, we report the power of the *Q*-statistics. In each case we present power results for a different autocontour to offer a comprehensive analysis. In all three cases the rejection rates are high and behave in the right direction. The power is close to 1 for samples of size 1,000 and larger when there is high dependence, or a large departure from normality, or strong ARCH effects. In case (i) when dependence is high we observe an increase in rejection rates as *K* increases. This is expected given the sensitivity of the test to linear dependence. On the other hand, we do not observe a similar pattern in the GARCH (1,1) where the dependence comes through higher moments.

[10] We also investigated the power of the *t*-statistics for different lags. The results are equally robust to parameter uncertainty.

In Table 1.10, we report the results for the *J*-statistic. Overall the power results under parameter uncertainty are similar to those for observable data. However, when detecting departures from normality, case (ii), bootstrapping increases the power of the test considerably.

Finally, we apply our tests to the standardized residuals of a more complicated model to assess the robustness of our results. We estimate a GARCH (1,1) model: $y_t = \sqrt{h_t}\varepsilon_t$, $\varepsilon_t \sim$ i.i.d. $N(0,1)$ and $h_t = \omega + \alpha y_{t-1}^2 + \beta h_{t-1}$, with $\omega = 0.1$, $\alpha = 0.15$ and $\beta = 0.8$. In Table 1.11, we summarize the size results for all test statistics. We observe that the results are very similar to those obtained when a location-scale model is estimated. Thus, provided that estimators are $\sqrt{T}$ consistent, the finite sample properties of our test statistics are robust to alternative parameterizations.

5 Empirical Applications

To illustrate the application of the proposed autocontour tests, we provide two examples dealing with financial data. The first entails the estimation of a GARCH (1,1) model for stock returns with conditional normal and Student-t distributions. The second deals with the estimation of an ACD model for duration data with a conditional exponential density.

5.1 GARCH Models

We consider the daily returns to the NYSE Composite Index from June 1, 1995 to December 31, 2004, with a total of 2,411 observations. In Figure 1.2, we have already provided the normal autocontours and the autocontourgram for the standardized data and

concluded that departures from normality come not only from the tail behavior but also from the central body of the distribution. We now proceed to estimate a GARCH (1,1) with a constant mean specification

$$\begin{aligned} y_t &= \mu + \sqrt{h_t}\,\varepsilon_t \\ h_t &= \omega + \alpha(y_{t-1} - \mu)^2 + \beta h_t, \end{aligned} \tag{15}$$

where y_t denotes the continuously compounded return on the NYSE Index. We retrieve the residuals and apply our tests to their standardized version. First, we consider the case where $\varepsilon_t \sim$ i.i.d. $N(0,1)$. In Figure 1.10, we present the results for the *t*-statistics up to 25 lags pertaining to the 10, 50, and 99% autocontours,[11] and the J-statistics for all 13 autocontours up to 25 lags.

Even though the modeling of the dependence in the data clearly helps to bring down the *t*-statistics reported in Figure 1.2, we still reject the i.i.d. normality hypothesis for the innovations of GARCH (1,1) model. The rejection is stronger in the 10 and 50% autocontours than in the tail area (99% autocontour). Thus, the empirical distribution has more probability mass in the central autocontours than the hypothesized density. This finding suggests that we should aim towards a different distributional assumption, which is expected given the stylized facts of financial returns, see Bollerslev et al. (1994), and Engle and Patton (2001) among others. Summarizing the information across all the 13 autocontours, the *J*-statistics overwhelmingly reject the null. In Table 1.12, we present the results for the *Q*-statistics for all 13 autocontours and for $K \in \{5, 10, 15, 20, 25\}$. From the 5% to 80% autocontours, the tests reject the null for all values of *K*. It is only in

[11] We select these three contours to represent the behavior of both the center and the tail of the distribution.

the 90% autocontour that the test fails to reject the null. However, this must be a reflection of the lack of power that we have observed in the Monte Carlo simulation results.

Next, we consider the same model with a conditional Student-t distribution. We assume $\varepsilon_t = \sqrt{(v-2)/v}\, u_t$ where $u_t \sim$ i.i.d. Student-t(v). Following Bai (2003) we rely on the QMLE results of Lee and Hansen (1994) and set $\hat{u}_t = \hat{\varepsilon}_t \sqrt{v/(v-2)}$ where we obtain $\hat{\varepsilon}_t$ from a Gaussian QML estimator. We estimate the value of v by GMM using the fourth moment condition, $E[\varepsilon_t^4] = 3(v-2)/(v-4)$, as in Bontemps and Meddahi (2006).[12] The GMM estimate is 7.48, thus we set $v = 7$ under the null. The results for $v = 8$ are very similar. In Figure 1.11, the *t*-statistics fail to reject the null at the 5% level except for the 50% autocontour at the first lag. The *J*-statistics also deliver the same message.

The *Q*-statistics, reported in Table 1.13, also fail to reject the null with the exception of the 30% and 80% autocontours. However, the values of the *Q*-statistics are considerably much lower than those in Table 1.12. Overall, the Student-t is a remarkable improvement although some of the autocontours indicate the need for further investigation.

In Figure 1.12, we describe graphically how the autocontours can guide the modeling of the NYSE returns. In Panel a, the standardized returns are superimposed on the normal and Student-t(7) autocontours. The rejection of normality is very clear and is mainly due to the many outliers spread all over the four quadrants. A similar picture with

[12] This moment condition is used since the first two moments are normalized to zero and one respectively.

autocontours corresponding to the Student-t(7) helps on picking up some of the outliers but yet there is a rejection of the joint hypothesis of the i.i.d.-ness and density function. In Panel b, we proceed to model the dependence with a symmetric GARCH(1,1) but maintain the Student-t(7) hypothesis. The symmetric GARCH model is successful enough but yet there is a cluster of outliers in the south-west quadrant that alerts about the possibility of asymmetries in the model. A GARCH(1,1) with a leverage term in the variance will take care of this asymmetric behavior.

5.2 ACD Models

We estimate an Autoregressive Conditional Duration (ACD) model for the durations (time intervals between transactions) of the Airgas common stock from March 1 to December 31, 2001 with a total of 32,366 intra daily observations.[13] Let $t_1, t_2, ..., t_i, ...$ denote a sequence of transaction times. Durations are defined as $x_i = t_i - t_{i-1}$. Engle and Russell's model is specified as

$$\begin{aligned} x_i &= \psi_i \varepsilon_i \\ \psi_i &= \omega + \sum_{j=1}^{p} \alpha_j x_{i-j} + \sum_{j=1}^{q} \beta_j \psi_{i-j} \end{aligned} \tag{16}$$

where $\psi_i = E[x_i \mid x_{i-1}, x_{i-2}, ..., x_1]$ and ε_i is i.i.d. with density $f(.)$. Following Engle and Russell (2004) we concentrate on the case where $p = 3, q = 2$ and $f(.)$ is the exponential density with $\beta = 1$. The results of the *t*-tests and the *J*-statistics are presented in Figure 1.13. Both tests reject the null hypothesis of i.i.d exponential innovations very strongly at all lags for all autocontours.

[13] We are grateful to Jeff Russell for providing us with the data set.

The Q-statistics, reported in Table 1.14, deliver extremely large values indicating strong rejection of the null with the exception of the 70% autocontour. These results are in line with those of Engle and Russell (1998) who also reject the exponential distribution assumption using more conventional methods. In Figure 1.14 we offer a comparison of the standardized durations (durations divided by the sample mean) and the ACD (3,2) residuals. The dependence modeled by the ACD specification helps tremendously to reduce the magnitude of the standardized durations, however this is not enough to make it consistent with the hypothesized exponential distribution.

6 Conclusion

The methodological advances in two fronts of time series analysis -nonlinear models with non-normal density functions, and density forecasting- have emphasized the need for developing dynamic specification tests for the joint hypothesis of i.i.d.-ness and density functional form. In this chapter we have proposed a new battery of tests that rely on the fundamental properties of independent random variables with identical distributions and we have introduced a graphical device -the autocontour- that helps to visualize the modeling problems. On reviewing the most relevant tests in the literature, from the pioneering work of Kolmogorov (1933) to the most recent insights, we believe that our tests bring considerable advantages. Among these, our tests are very powerful against violations of both hypotheses, i.i.d.-ness and density function. They have standard convergence rates and standard limiting distributions. They do not require either a

transformation of the original data or an assessment of goodness-of-fit *à-la* Kolmogorov and explicitly account for parameter uncertainty.

We have introduced our methodology within the context of pair-wise independence but it can be extended to higher dimensions. On going further than the bivariate case we will be losing the graphical representation of the autocontour, which is helpful for the understanding of the modeling problem; however once the analytical functional form of the autocontour is obtained, the indicator variable is trivial to construct and the *t*-tests and chi-squared tests that we propose will follow naturally. Implementation of this methodology in the context of multivariate densities is presented in Chapter II.

References

Anderson, T.W. (1971), *The Statistical Analysis of Time Series*, New York: Wiley.

Andrews, D.W.K. (1997), A Conditional Kolmogorov Test, *Econometrica*, 65, 1097 1128.

Arfken, G.B. (2005), *Mathematical Methods for Physicists*, Amsterdam: Elsevier.

Bai J. (2003) Testing Parametric Conditional Distributions of Dynamic Models, *Review of Economics and Statistics*, 85(3), 532-549.

Box, G.E.P. and D.A. Pierce (1970), Distribution of Residual Autocorrelations in Autoregressive-Integrated Moving Average Time Series Models, *Journal of the American Statistical Association*, 65(332), 1509-1526.

Berkowitz, J. (2001), Testing Density Forecasts with Applications to Risk Management, *Journal of Business and Economic Statistics*, 19(4), 465-474.

Bollerslev, T., R.F. Engle and D.B. Nelson (1994), ARCH Models, in R.F. Engle and D. McFadden (eds.), *Handbook of Econometrics*, Volume IV, 2959-3038, Amsterdam: North-Holland.

Bontemps, C. and N. Meddahi (2006), Testing Distributional Assumptions: A GMM Approach, Working Paper.

Bontemps, C. and N. Meddahi (2005), Testing Normality: A GMM Approach, *Journal of Econometrics*, 124, 149-186.

Brock, W., D. Dechert, J. Scheinkman and B. LeBaron (1996), A Test for Independence based on the Correlation Dimension, *Econometric Reviews*, 15, 197-235.

Brock, W., D. Hsieh and B. LeBaron (1991), *Nonlinear Dynamics, Chaos, and Instability: Statistical Theory and Economic Evidence*, Cambridge: MIT Press.

Buchinsky, M. (1995), Estimating the Asymptotic Covariance Matrix for Quantile Regression Models: A Monte Carlo Study, *Journal of Econometrics*, 68, 303-338.

Chen, Y.T. (2007), A Unified Approach to Standardized Residuals-Based Correlation Tests for GARCH Type Models, *Journal of Applied Econometrics*, forthcoming.

Chen, X. and Y.Q. Fan (2004), Evaluating Density Forecasts via the Copula Approach, *Finance Research Letters*, 1, 74-84.

Christoffersen, P. (1998), Evaluating Interval Forecasts, *International Economic Review*, 39, 841-862.

Corradi, V. and N. Swanson (2006), Predictive Density Evaluation, in G. Elliot, C.W.J. Granger and A. Timmerman (eds.), *Handbook of Forecasting*, Volume I, 197-284. Elsevier Science.

Diebold, F.X., T. Gunther, and A. Tay (1998), Evaluating Density Forecasts with Applications to Financial Risk Management, *International Economic Review,* 39, 863-88.

Efron, B. (1979), Bootstrap Methods: Another Look at the Jackknife, *The Annals of Statistics*, 7(1), 1-26.

Engle R. F. and A. Patton (2001), What Good is a Volatility Model?, *Quantitative Finance*, 1, 237-245.

Engle, R.F. and J. R. Russell (2004), Analysis of High Financial Frequency Data, in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*, Amsterdam: North Holland, forthcoming.

Engle, R.F. and J. R. Russell (1998), Autoregressive Conditional Duration: A New Model for Irregularly Spaced Transaction Data, *Econometrica*, 66(5), 1127-1162.

Fan, Y.Q. (1994), Testing the Goodness-of-Fit Test of a Parametric Density Function, *Econometric Theory*, 10, 316–356.

Goncalves, S. and H. White (2005), Bootstrap Standard Error Estimates for Linear Regression, *Journal of the American Statistical Association*, 100(471), 970-978.

González-Rivera, G. and F. Drost (1999), Efficiency Comparisons of Maximum Likelihood based Estimators in GARCH Models, *Journal of Econometrics*, 93, 93-111.

Gouriéroux, C. and A. Monfort (1996), *Simulation Based Econometric Methods*, CORE Lectures, Oxford.

Hong, Y. and H. Li (2005), Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates, *Review of Financial Studies*, 18, 37-84.

Hong, Y. and T.-H. Lee (2003), Diagnostic Checking for the Adequacy of Nonlinear Time Series Models, *Econometric Theory*, 19, 1065-1121.

Hong, Y. (2001), Evaluation of Out-of-Sample Density Forecasts with Applications to Stock Prices, Working Paper, Cornell University.

Horowitz, J.L. (2001), The Bootstrap in Econometrics, in J.J. Heckman and E.E. Leamer (eds.), *Handbook of Econometrics*, Volume V, 3159-3228. Elsevier Science.

Kolmogorov, A.N. (1933), Sulla Determinoziane Empiricadi una Legge di Distribuzione, Giornalle dell'Instutitio Italiano degli Attuari, 4, 83-91.

Ledoit, O.P., P. Santa-Clara and M. Wolf (2003), Flexible Multivariate GARCH Modeling with an Application to International Stock Markets, *The Review of Economics and Statistics*, 85(3), 735-747.

Lee, S.W. and B.E. Hansen (1994), Asymptotic Theory for the GARCH (1,1) Quasi-Maximum Likelihood Estimator, *Econometric Theory*, 10, 29-52.

Li, W.K. and T.K. Mak (1994), On the Squared Residual Autocorrelations in Nonlinear Time Series with Conditional Heteroskedasticity, *Journal of Time Series Analysis*, 15, 627-636.

Lilliefors, H.W. (1967), On the Kolmogorov-Smirnov Test for Normality with Mean and Variance Unknown, *Journal of the American Statistical Association*, 62(318), 399-402.

Ljung, G.M. and G.E.P. Box (1978), On a Measure of Lack of Fit in Time Series Models, *Biometrika*, 65(2), 297-303.

MacKinnon, J. (2006), Bootstrap Methods in Econometrics, *Queen's University Working Paper No.* 1028.

McLeod, A.I. and W.K. Li (1983), Diagnostic Checking ARMA Time Series Models Using Squared Residual Autocorrelations, *Journal of Time Series Analysis*, 4, 269-273.

Phillips, P.C.B., (1991), A Shortcut to LAD Estimator Asymptotics, *Econometric Theory*, 7, 451-463.

Randles, R. H. (1982), On the Asymptotic Normality of Statistics with Estimated Parameters, *The Annals of Statistics*, 10(2), 462-474.

Stinchcombe, M.B. and H. White (1998), Consistent Specification Testing with Nuisance Parameters Present Only Under the Alternative, *Econometric Theory*, 14, 295-325.

Smirnov, N.V. (1939), On the Estimation of the Discrepancy between Empirical Curves of Distribution for two Independent Samples, *Bul. Math. De l'Univ. de Moscou*, 2, 3-14.

Zheng, J.X. (2000), A Consistent Test of Conditional Parametric Distribution, *Econometric Theory*, 16, 667–691.

Appendix A: Proofs of Propositions 1-4

Proof of Proposition 1:

The result directly follows from the Central Limit Theorem for covariance stationary processes presented in Anderson (1971). ■

Proof of Proposition 2:

Asymptotic Normality

Without loss of generality, let us consider the joint distribution of $q_{k,i}$ and $q_{l,i}$. Let $x \equiv \lambda_1 q_{k,i} + \lambda_2 q_{l,i}$ where λ_1 and λ_2 are arbitrary constants and assume $k < l$. Then

$$x = \lambda_1 \sqrt{T-k}(\hat{p}_i^k - p_i) + \lambda_2 \sqrt{T-l}(\hat{p}_i^l - p_i).$$

For simplicity we drop first $(l-k)$ observations on $I_t^{k,i}$. Since all results are asymptotic it is convenient to use the same scaling factor

$$x = \frac{1}{\sqrt{T-l}} \sum_{t=1}^{T-l} \{\lambda_1 (I_t^{k,i} - p_i) + \lambda_2 (I_t^{l,i} - p_i)\}.$$

Let $e_t \equiv \lambda_1 (I_t^{k,i} - p_i) + \lambda_2 (I_t^{l,i} - p_i)$, then $x = (T-l)^{-1/2} \sum_{t=1}^{T-l} e_t$. The first two moments of e_t and its autocovariance function are given by

$$E[e_t] = 0,$$

$$Var(e_t) = (\lambda_1^2 + \lambda_2^2) p_i (1-p_i) + 2\lambda_1 \lambda_2 Cov(I_t^{l,i}, I_t^{k,i}) < \infty,$$

$$Cov(e_t, e_{t-h}) = \begin{cases} \lambda_1^2 \gamma_k^i + \lambda_1 \lambda_2 Cov(I_t^{k,i} I_{t-k}^{l,i}) & \text{if } h = k \\ \lambda_1^2 \gamma_l^i + \lambda_1 \lambda_2 Cov(I_t^{l.i}, I_{t-l}^{k.i}) & \text{if } h = l \\ \lambda_1 \lambda_2 Cov(I_t^{l.i}, I_{t-l+k}^{k.i}) & \text{if } h = l-k \\ 0 \quad \text{otherwise} \end{cases}$$

Therefore, $Cov(e_t, e_{t-h}) < \infty, \forall h$. From the CLT for covariance stationary processes, $x \xrightarrow{d} N(0, \sigma_x^2)$ where $\sigma_x^2 = Var(e_t) + 2Cov(e_t, e_{t-k}) + 2Cov(e_t, e_{t-l}) + 2Cov(e_t, e_{t-l+k})$. Thus, we have shown that any linear combination of $q_{k,i}$ and $q_{l,i}$ is asymptotically normal establishing their joint asymptotic normality. ■

Asymptotic Covariance Matrix

Given that $E[q_{k,i}] = E[q_{l,i}] = 0$, we have $Cov(q_{k,i}, q_{l,i}) = E[q_{k,i} q_{l,i}]$ where

$$E[q_{k,i} q_{l,i}] = (T-l) E\left[\frac{1}{T-l} \sum_{t=1}^{T-l} (I_t^{k,i} - p_i) \frac{1}{T-l} \sum_{t=1}^{T-l} (I_t^{l,i} - p_i) \right]$$

$$= \frac{1}{T-l} \left\{ \sum_{t=1}^{T-l} E[(I_t^{k,i} - p_i)(I_t^{l,i} - p_i)] + \sum_{t=1}^{T-l-1} \sum_{t'=t+1}^{T-l} E[(I_t^{k,i} - p_i)(I_{t'}^{l,i} - p_i)] + \sum_{t=1}^{T-l-1} \sum_{t'=t+1}^{T-l} E[(I_t^{l,i} - p_i)(I_{t'}^{k,i} - p_i)] \right\}$$

$$= Cov(I_t^{k,i}, I_t^{l,i}) + Cov(I_t^{k,i}, I_{t-k}^{l,i}) + Cov(I_t^{l,i}, I_{t-l}^{k,i}) + Cov(I_t^{l,i}, I_{t-l+k}^{k,i})$$

as $T \to \infty$ ■

Proof of Proposition 3:

Asymptotic Normality

Without loss of generality, let us consider the joint distribution of $z_{i,k}$ and $z_{j,k}$. Let $x \equiv \lambda_1 z_{i,k} + \lambda_2 z_{j,k}$. Then we have

$$x = \lambda_1 \sqrt{T-k}(\hat{p}_i^k - p_i) + \lambda_2 \sqrt{T-k}(\hat{p}_j^k - p_j)$$

$$= \frac{1}{\sqrt{T-k}} \sum_{t=1}^{T-k} \{\lambda_1 (I_t^{k,i} - p_i) + \lambda_2 (I_t^{k,j} - p_j)\}$$

Defining $e_t \equiv a(I_t^{k,i} - p_i) + b(I_t^{l,j} - p_j)$ yields $x = (T-k)^{-1/2} \sum_{t=1}^{T-k} e_t$. The first two moments of e_t are given by

$$E[e_t] = 0$$

$$Var(e_t) = \lambda_1^2 p_i(1-p_i) + \lambda_2^2 p_j(1-p_j) + 2\lambda_1\lambda_2\{\min(p_i, p_j) - p_i p_j\} < \infty,$$

since $E[I_t^{k,i} I_t^{k,j}] = \min(p_i, p_j)$. Furthermore,

$$Cov(e_t, e_{t-h}) = \begin{cases} \lambda_1^2 \gamma_k^i + \lambda_2^2 \gamma_k^j + \lambda_1\lambda_2(\gamma_k^{i,j} + \gamma_k^{j,i}) < \infty & \text{if } h = k \\ 0 \qquad \text{otherwise} \end{cases}$$

where $\gamma_k^{i,j} = Cov(I_t^{i,k}, I_{t-k}^{j,k})$ and $\gamma_k^{j,i} = Cov(I_t^{j,k}, I_{t-k}^{i,k})$. In general, $\gamma_k^{i,j} \neq \gamma_k^{j,i}$ since they are cross autocovariances. Thus, $Cov(e_t, e_{t-h}) < \infty, \forall h$. It directly follows, from the CLT, that $x \xrightarrow{d} N(0, \sigma_x^2)$ where $\sigma_x^2 = Var(e_t) + 2Cov(e_t, e_{t-k})$. We have shown that any linear combination of $z_{i,k}$ and $z_{j,k}$ is asymptotically normal establishing their joint asymptotic normality. ■

Asymptotic Covariance Matrix

Given that $E[z_{i,k}] = E[z_{j,k}] = 0$ we have $Cov(z_{i,k}, z_{j,k}) = E[z_{i,k} z_{j,k}]$ where

$$E[z_{i,k} z_{j,k}] = (T-k)E\left[\frac{1}{T-k}\sum_{t=1}^{T-k}(I_t^{k,i} - p_i)\frac{1}{T-k}\sum_{t=1}^{T-k}(I_t^{k,j} - p_j)\right]$$

$$= \frac{1}{T-k}\left\{\sum_{t=1}^{T-k}E[(I_t^{k,i} - p_i)(I_t^{k,j} - p_j)] + \sum_{t=1}^{T-k-1}\sum_{t'=t+1}^{T-k}E[(I_t^{k,i} - p_i)(I_{t'}^{k,j} - p_j)] + \sum_{t=1}^{T-k-1}\sum_{t'=t+1}^{T-k}E[(I_t^{k,j} - p_j)(I_{t'}^{k,i} - p_i)]\right\}$$

$$= \min(p_i, p_j) - p_i p_j + Cov(I_t^{k,i}, I_{t-k}^{k,j}) + Cov(I_{t-k}^{k,i}, I_t^{k,j}) \quad \text{as } T \to \infty \qquad \blacksquare$$

Proof of Proposition 4:

From assumption A1 we have,

$$\sqrt{T}(\hat{\theta} - \theta_0) = A^{-1}\frac{1}{\sqrt{T}}\sum_{t=1}^{T} s_t(\theta_0) + o_p(1).$$

Defining $x \equiv \lambda_1\sqrt{T}(\hat{p}_i^k(\theta_0) - p_i) + \lambda_2\sqrt{T}(\hat{\theta} - \theta_0)'D$, and suppressing the argument of the score we obtain,

$$x = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\{\lambda_1(I_t^{k,i} - p_i) + \lambda_2 s_t' A^{-1} D\} + o_p(1).$$

Now, let $e_t \equiv \lambda_1(I_t^{k,i} - p_i) + \lambda_2 s_t' A^{-1} D$, then $x = T^{-1/2}\sum_{t=1}^{T} e_t + o_p(1)$. The first two moments of e_t are given by

$$E[e_t] = 0$$

$$Var(e_t) = \lambda_1^2 p_i(1-p_i) + \lambda_2^2 D' A^{-1} B A^{-1} D + 2\lambda_1\lambda_2 Cov(I_t^{k,i}, s_t') A^{-1} D < \infty$$

Furthermore,

$$Cov(e_t, e_{t-h}) = \begin{cases} \lambda_1^2 \gamma_k^i + \lambda_1\lambda_2 Cov(I_t^{k,i}, s_{t-k}') A^{-1} D < \infty & \text{if } h = k \\ 0 \qquad \text{otherwise} \end{cases}$$

Note that finiteness of the variance and covariance expressions follows from A2-A3. As a result applying the CLT, we have $x \xrightarrow{d} N(0, \sigma_x^2)$ where $\sigma_x^2 = Var(e_t) + 2Cov(e_t, e_{t-k})$. The variables in the right hand side of (13) are both asymptotically normal due to Proposition 1 and A1. We showed that their linear combinations are also asymptotically normal, which completes the proof. ■

Appendix B: Covariance and Gradient Terms of Gaussian Location-Scale Model

The model is given by

$$y_t = \mu_0 + \sigma_0 \varepsilon_t, \quad \varepsilon_t \sim \text{i.i.d. N(0,1)} \tag{B.1}$$

Let $\theta_0 = (\mu_0, \sigma_0^2)'$. In this case the indicator series is constructed as follows

$$I_t^{k,i}(z_t(\theta_0)) = \begin{cases} 1 & \text{if } \; z_t(\theta_0) - a_i > 0 \\ 0 & \text{otherwise} \end{cases} \tag{B.2}$$

where $z_t(\theta_0) = \varepsilon_t^2(\theta_0) + \varepsilon_{t-k}^2(\theta_0)$. From the mean value expansion in the text we need the following:

$$V_{12} \equiv \lim_{T\to\infty} E[\sqrt{T}(\hat{p}_i^k(\theta_0) - p_i)\sqrt{T}(\hat{\theta} - \theta_0)] \tag{B.3}$$

$$D \equiv \lim_{T\to\infty} E\left(\left. \frac{\partial \hat{p}_i^k(\theta)}{\partial \theta} \right|_{\theta=\theta_0} \right) \tag{B.4}$$

Let us start with the first term given in (B.3). From the properties of the QML estimator

$$\begin{aligned} V_{12} &\equiv \lim_{T\to\infty} E\left[\sqrt{T}(\hat{p}_i^k(\theta_0) - p_i) \frac{1}{\sqrt{T}} \sum_{t=1}^{T} A^{-1} s_t(\theta_0) \right] \\ &= A^{-1} \lim_{T\to\infty} E\left[\frac{1}{T} \sum_{t=1}^{T} (I_t^{k,i}(z_t(\theta_0)) - p_i) \sum_{t=1}^{T} s_t(\theta_0) \right] \\ &= A^{-1}[Cov(I_t^{k,i}(z_t(\theta_0)), s_t(\theta_0)) + Cov(I_t^{k,i}(z_t(\theta_0)), s_{t-k}(\theta_0))] + o(1) \end{aligned}$$

For the model under consideration, we have $s_t = (\varepsilon_t / \sigma_0,\ (\varepsilon_t^2 - 1)/2\sigma_0^2)'$.[14] Hence, we need to obtain (i) $E[I_t^{k,i}\varepsilon_t]$, (ii) $E[I_t^{k,i}\varepsilon_{t-k}]$, (iii) $E[I_t^{k,i}\varepsilon_t^2]$ and (iv) $E[I_t^{k,i}\varepsilon_{t-k}^2]$ to get the covariance term.

$$\begin{aligned} \text{(i)}: \quad E[I_t^{k,i}\varepsilon_t] &= P(\varepsilon_{t-k}^2 > a_i) E[I_t^{k,i}\varepsilon_t \mid \varepsilon_{t-k}^2 > a_i] + P(\varepsilon_{t-k}^2 \le a_i) E[I_t^{k,i}\varepsilon_t \mid \varepsilon_{t-k}^2 \le a_i] \\ &= 0 + P(\varepsilon_{t-k}^2 \le a_i) E\left[\int_{-\infty}^{-g(\varepsilon_{t-k})} \varepsilon_t f(\varepsilon_t) d\varepsilon_t + \int_{g(\varepsilon_{t-k})}^{\infty} \varepsilon_t f(\varepsilon_t) d\varepsilon_t \mid \varepsilon_{t-k}^2 \le a_i \right] \\ &= 0 \end{aligned}$$

where $g(\varepsilon_{t-k}) = \sqrt{a_i - \varepsilon_{t-k}^2}$ and the third line follows from the fact that ε_t has a symmetric distribution with zero mean.

(ii): $E[I_t^{k,i}\varepsilon_{t-k}] = 0$, using analogous conditioning arguments as in (i).

$$\begin{aligned} \text{(iii)}: \quad E[I_t^{k,i}\varepsilon_t^2] &= P(\varepsilon_{t-k}^2 > a_i) E[I_t^{k,i}\varepsilon_t^2 \mid \varepsilon_{t-k}^2 > a_i] + P(\varepsilon_{t-k}^2 \le a_i) E[I_t^{k,i}\varepsilon_t^2 \mid \varepsilon_{t-k}^2 \le a_i] \\ &= P(\varepsilon_{t-k}^2 > a_i) + P(\varepsilon_{t-k}^2 \le a_i) E\left[\int_{-\infty}^{-g(\varepsilon_{t-k})} \varepsilon_t^2 f(\varepsilon_t) d\varepsilon_t + \int_{g(\varepsilon_{t-k})}^{\infty} \varepsilon_t^2 f(\varepsilon_t) d\varepsilon_t \mid \varepsilon_{t-k}^2 \le a_i \right] \\ &= P(\varepsilon_{t-k}^2 > a_i) + P(\varepsilon_{t-k}^2 \le a_i) E\left[1 - \int_{-g(\varepsilon_{t-k})}^{g(\varepsilon_{t-k})} \varepsilon_t^2 f(\varepsilon_t) d\varepsilon_t \mid \varepsilon_{t-k}^2 \le a_i \right] \end{aligned}$$

[14] We suppress the arguments of the score and the indicator to simplify notation.

$$= 1 - \int_{-\sqrt{a_i}}^{\sqrt{a_i}} \int_{-g(\varepsilon_{t-k})}^{g(\varepsilon_{t-k})} \varepsilon_t^2 f(\varepsilon_t) f(\varepsilon_{t-k}) d\varepsilon_t d\varepsilon_{t-k} \, .$$

The last line follows from the properties of truncated distributions, i.e. $f(x \mid x^2 < a) = f(x) / P(-\sqrt{a} < x < -\sqrt{a})$. The resulting integral is obtained by numerical methods.

(iv) : $E[I_t^{k,i} \varepsilon_{t-k}^2] = 1 - \int_{-\sqrt{a_i}}^{\sqrt{a_i}} \int_{-g(\varepsilon_t)}^{g(\varepsilon_t)} \varepsilon_{t-k}^2 f(\varepsilon_{t-k}) f(\varepsilon_t) d\varepsilon_{t-k} d\varepsilon_t$, using similar conditioning arguments as in (iii). Numerically, this integral is identical to that in (iii).

Therefore, we conclude that

$$V_{12} = 2A^{-1}(0, \ (E[I_t^{k,i} \varepsilon_t^2] - p_i) / 2\sigma_0^2)' \tag{B.5}$$

Now let us concentrate on D. Let $D = (D_1, \ D_2)'$ then we have

$$D_1 = (-2/\sigma_0) \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[\delta(z_t - a_i)(\varepsilon_t + \varepsilon_{t-k})] \tag{B.6}$$

$$D_2 = (-1/\sigma_0^2) \lim_{T \to \infty} \frac{1}{T} \sum_{t=1}^{T} E[\delta(z_t - a_i) z_t] \tag{B.7}$$

where $\delta(x)$ is the Dirac delta function[15]. To simplify D_1 and D_2, we use the following properties of the Dirac delta function:

$$\delta(x - a) = 0 \ \text{ for } \ x \neq a \tag{B.8}$$

$$\int_{-\infty}^{\infty} h(x) \delta(x - a) dx = h(a) \tag{B.9}$$

$$\delta(x^2 - a^2) = \frac{1}{2a} [\delta(x - a) + \delta(x + a)] \tag{B.10}$$

[15] See Phillips (1991) for a similar application of the Dirac delta function in econometrics. See Arfken (2005) for further details on Dirac delta function.

where a is a finite constant and h is any continuous function. Let us start with D_1:

$$E[\delta(z_t - a_i)(\varepsilon_t + \varepsilon_{t-k})] = P(\varepsilon_{t-k}^2 > a_i)E[\delta(z_t - a_i)(\varepsilon_t + \varepsilon_{t-k}) \mid \varepsilon_{t-k}^2 > a_i] + P(\varepsilon_{t-k}^2 \le a_i)E[\delta(z_t - a_i)(\varepsilon_t + \varepsilon_{t-k}) \mid \varepsilon_{t-k}^2 \le a_i]$$

$$= P(\varepsilon_{t-k}^2 \le a_i)E[\delta(z_t - a_i)(\varepsilon_t + \varepsilon_{t-k}) \mid \varepsilon_{t-k}^2 \le a_i]$$

since the first term vanishes due to (B.8). Rearranging yields

$$E[\delta(z_t - a_i)(\varepsilon_t + \varepsilon_{t-k}) \mid \varepsilon_{t-k}^2 \le a_i] = E[\delta(\varepsilon_t^2 - c)(\varepsilon_t + \varepsilon_{t-k}) \mid \varepsilon_{t-k}^2 \le a_i]$$

where $c \equiv a_i - \varepsilon_{t-k}^2$. Using (B.10) we get

$$E[\delta(\varepsilon_t^2 - c)(\varepsilon_t + \varepsilon_{t-k}) \mid \varepsilon_{t-k}^2 \le a_i] = \frac{1}{2\sqrt{c}} E[\{\delta(\varepsilon_t - \sqrt{c}) + \delta(\varepsilon_t + \sqrt{c})\}\varepsilon_t \mid \varepsilon_{t-k}^2 \le a_i] + \frac{1}{2\sqrt{c}} E[\{\delta(\varepsilon_t - \sqrt{c}) + \delta(\varepsilon_t + \sqrt{c})\}\varepsilon_{t-k} \mid \varepsilon_{t-k}^2 \le a_i].$$

Using (B.9) directly yields

$$E[\{\delta(\varepsilon_t - \sqrt{c}) + \delta(\varepsilon_t + \sqrt{c})\}\varepsilon_t \mid \varepsilon_{t-k}^2 \le a_i] = \frac{1}{2\sqrt{c}} E[\sqrt{c}f(\sqrt{c}) - \sqrt{c}f(-\sqrt{c}) \mid \varepsilon_{t-k}^2 \le a_i]$$

$$= 0, \text{ and}$$

$$E[\{\delta(\varepsilon_t - \sqrt{c}) + \delta(\varepsilon_t + \sqrt{c})\}\varepsilon_{t-k} \mid \varepsilon_{t-k}^2 \le a_i] = \frac{1}{2\sqrt{c}} E[\{f(\sqrt{c}) - f(-\sqrt{c})\}\varepsilon_{t-k} \mid \varepsilon_{t-k}^2 \le a_i]$$

$$= \frac{\exp(-0.5a_i^2)}{\sqrt{2\pi}P(\varepsilon_{t-k}^2 \le a_i)} \int_{-\sqrt{a_i}}^{\sqrt{a_i}} \frac{\varepsilon_{t-k}}{\sqrt{a_i - \varepsilon_{t-k}^2}} d\varepsilon_{t-k}$$

$$E[\{\delta(\varepsilon_t - \sqrt{c}) + \delta(\varepsilon_t + \sqrt{c})\}\varepsilon_{t-k} \mid \varepsilon_{t-k}^2 \le a_i] = \frac{\exp(-0.5a_i^2)}{\sqrt{2\pi}P(\varepsilon_{t-k}^2 \le a_i)} \{\sqrt{a_i - a_i} - \sqrt{a_i - a_i}\}$$

$$= 0$$

This establishes that $D_1 = 0$. Finally, applying (B.9) to (B.7) yields $D_2 = -a_i f_{z_t}(a_i)/\sigma_0^2$ where f_{z_t} is the p.d.f. of z_t, a chi-squared random variable with 2 degrees of freedom. Thus, we conclude that

$$D = (0, -a_i f_{z_t}(a_i)/\sigma_0^2)'. \tag{B.11}$$

■

The table below provides the size of *t*-statistics under parameter uncertainty for the model given in (B.1) where the covariance and the gradient terms are obtained from (B.5) and (B.11) respectively. These results are directly comparable to those presented in Table 5, Panel-a, which is based on bootstrapping the asymptotic variance.

T	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	$t_{1,4}$	$t_{1,5}$	$t_{1,6}$	$t_{1,7}$	$t_{1,8}$	$t_{1,9}$	$t_{1,10}$	$t_{1,11}$	$t_{1,12}$	$t_{1,13}$
250	2.2	5.3	5.6	5.2	5.0	4.6	5.3	6.5	7.0	6.2	5.6	9.1	16.4
500	4.4	5.3	6.7	5.2	5.4	4.4	5.1	6.6	6.2	6.3	6.0	7.8	7.6
1000	5.1	5.4	5.0	4.7	4.0	4.9	4.6	5.6	5.3	6.3	5.6	5.4	5.7
2000	4.9	4.4	5.2	4.5	5.0	4.2	4.6	5.4	5.4	5.5	5.3	5.1	5.2

Tables and Figures

Table 1.1: Size of the *t*-Statistics

T	$t_{1,7}$	$t_{2,7}$	$t_{3,7}$	$t_{4,7}$	$t_{5,7}$
Panel a: Normal					
250	5.32	5.43	5.28	5.29	5.37
500	4.98	4.77	4.86	5.01	4.89
1000	4.98	4.84	4.72	5.01	5.08
2000	5.21	5.10	5.64	4.90	5.10
Panel b: Student-t					
250	5.05	5.11	4.96	4.93	4.89
500	4.83	4.61	4.51	5.08	4.79
1000	5.31	5.07	5.35	5.09	5.19
2000	5.18	5.08	5.04	5.14	5.26
Panel c: Exponential					
250	5.42	5.28	5.55	5.53	5.30
500	5.07	4.75	5.17	4.91	5.02
1000	5.37	4.79	5.26	5.00	4.99
2000	4.95	4.79	4.86	4.93	5.04

Table 1.2: Size of the *Q*-Statistics

T	Q_7^2	Q_7^3	Q_7^4	Q_7^5
Panel a: Normal				
250	8.19	12.08	15.40	17.31
500	6.86	9.12	13.39	17.43
1000	5.92	6.91	9.45	12.10
2000	6.01	6.02	7.15	8.30
Panel b: Student-t				
250	8.10	11.15	14.38	16.32
500	6.58	8.91	13.60	18.16
1000	6.23	6.93	9.45	12.41
2000	5.43	5.84	7.03	8.12
Panel c: Exponential				
250	7.93	11.31	15.27	18.45
500	5.96	8.01	11.29	15.26
1000	5.63	6.11	7.68	9.88
2000	5.08	5.54	6.28	7.06

Table 1.3: Size of the *J*-Statistics

T	Normal	Student-t	Exponential
Panel a: J_1^{13}			
250	9.34	9.17	10.38
500	6.59	6.78	7.16
1000	5.47	6.23	5.90
2000	5.35	6.14	5.27
Panel b: J_1^7			
250	6.40	6.73	7.43
500	5.58	5.90	6.14
1000	5.44	5.63	5.47
2000	5.53	5.21	5.38

Notes: Simulated size (%) of the test statistics under three DGPs: (i) $\varepsilon_t \sim$ i.i.d.$N(0,1)$, (ii) $\varepsilon_t \sim$ i.i.d. Student-t(5), (iii) $\varepsilon_t \sim$ i.i.d. exp(1). Table 1.1 presents the size of the *t*-statistic based on the 50% autocontour for $k = 1,...,5$. Table 1.2 presents the size of the *Q*-statistic based on the 50% autocontour for $K = 2,...,5$. Table 1.3 presents the size of the *J*-statistic based on all contours (Panel a) and 7 contours, excluding the first and the last three, (Panel b) for $k = 1$. Number of MC replications is 10,000 and nominal size is 5% in all cases.

Table 1.4: Power of the *J*-Statistics

Panel a: AR(1)

	T	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
$\varphi = 0.9$	250	76.23	67.12	62.03	58.79	57.53
	500	95.69	84.85	73.61	66.98	61.99
	1000	99.96	97.74	90.89	82.30	73.66
	2000	100.00	100.00	99.28	95.99	88.64
$\varphi = 0.7$	250	36.88	23.41	20.92	19.99	19.55
	500	56.91	24.97	18.00	16.80	16.09
	1000	88.90	36.46	19.22	15.72	15.49
	2000	99.83	60.48	25.85	17.21	15.47
$\varphi = 0.5$	250	17.28	12.88	11.86	12.89	11.62
	500	18.75	9.54	9.58	9.11	8.88
	1000	31.64	8.79	7.93	7.93	8.00
	2000	61.41	9.90	7.54	8.01	7.84
$\varphi = 0.3$	250	10.59	9.88	10.26	10.33	10.02
	500	8.57	6.71	7.54	7.19	7.48
	1000	8.13	6.62	6.30	6.62	6.67
	2000	10.67	6.06	6.30	5.99	6.34

Panel b: i.i.d. Student-t

	T	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
$\nu = 5$	250	58.31	58.19	58.73	58.06	57.82
	500	88.85	88.84	88.65	88.54	88.55
	1000	99.76	99.78	99.78	99.85	99.77
	2000	100.00	100.00	100.00	100.00	100.00
$\nu = 10$	250	18.44	18.65	17.96	18.23	17.88
	500	26.09	25.94	26.22	25.18	26.13
	1000	48.73	48.58	49.26	48.79	48.70
	2000	84.68	84.44	84.36	84.20	84.22
$\nu = 15$	250	12.50	12.88	13.01	13.81	13.11
	500	13.34	13.41	13.55	13.61	13.47
	1000	20.87	20.87	21.14	21.62	21.37
	2000	43.29	43.61	43.29	43.10	42.22

Table 1.4 (Continued)
Panel c: GARCH(1,1)

	T	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
$\alpha = 0.15$	250	58.74	58.00	58.97	58.66	57.93
	500	77.54	76.86	76.67	75.64	75.60
$\beta = 0.8$	1000	95.54	94.90	94.76	94.20	93.98
	2000	99.96	99.92	99.91	99.84	99.80
	250	38.14	38.46	38.35	38.55	38.18
$\alpha = 0.1$	500	47.17	46.93	46.35	46.14	45.69
$\beta = 0.85$	1000	67.15	66.87	65.73	65.22	64.45
	2000	90.61	90.83	89.63	88.76	88.37
	250	20.53	20.05	19.88	19.98	20.23
$\alpha = 0.05$	500	18.69	18.90	18.59	18.92	19.10
$\beta = 0.9$	1000	21.61	20.97	21.10	20.90	20.33
	2000	29.17	29.21	28.40	28.13	27.04

Notes: Simulated power (%) of the *J*-statistic for $k = 1,\ldots,5$ under the following DGPs: (i) $\varepsilon_t = \varphi\varepsilon_{t-1} + u_t$ where $u_t \sim$ i.i.d. $N(0,1-\varphi^2)$ (Panel a), (ii) $\varepsilon_t = u_t\sqrt{(v-2)/v}$ where $u_t \sim$ i.i.d. Student-t(v) (Panel b), (iii) $\varepsilon_t = \sqrt{h_t}u_t$ where $u_t \sim$ i.i.d. $N(0,1)$, and $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$ (Panel c). The null hypothesis is $\varepsilon_t \sim$ i.i.d. $N(0,1)$. The test statistic is based on all 13 autocontours. Number of MC replications is 10,000 and nominal size is 5%.

Table 1.5: Size of the t-Statistics under Parameter Uncertainty (All Contours)

T	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	$t_{1,4}$	$t_{1,5}$	$t_{1,6}$	$t_{1,7}$	$t_{1,8}$	$t_{1,9}$	$t_{1,10}$	$t_{1,11}$	$t_{1,12}$	$t_{1,13}$
Panel a: Normal													
250	5.3	5.9	7.0	5.8	5.1	5.3	5.6	4.9	4.8	4.8	5.8	4.4	2.4
500	4.0	4.6	4.0	5.1	4.4	5.5	5.1	5.4	5.8	4.6	4.3	4.1	3.9
1000	4.7	5.7	6.8	5.5	4.8	4.9	4.4	5.3	5.7	5.7	4.6	4.5	6.0
2000	4.5	5.2	5.9	4.0	4.8	4.0	4.9	5.6	7.2	6.0	4.5	5.9	4.6
Panel b: Student-t													
250	3.6	4.4	4.7	4.3	4.1	4.2	4.3	5.6	4.6	5.3	5.2	4.5	2.0
500	4.2	5.5	5.5	5.3	5.5	5.1	5.2	5.2	5.3	4.9	4.7	4.3	5.7
1000	3.9	3.8	3.9	4.4	3.7	3.6	4.1	3.8	4.3	3.5	3.4	3.6	5.8
2000	5.4	5.0	4.0	4.5	4.4	4.0	5.1	5.1	4.6	4.7	5.4	3.2	4.0
Panel c: Exponential													
250	4.9	4.7	4.8	4.4	5.2	5.6	5.0	4.1	5.3	4.4	5.4	4.8	4.2
500	4.7	4.8	5.0	4.8	4.8	4.2	6.2	5.6	6.4	4.6	5.4	4.9	5.6
1000	4.0	5.4	4.8	5.1	5.1	5.6	5.4	5.3	4.7	6.4	4.3	4.1	5.0
2000	3.7	5.5	4.5	4.0	4.4	4.9	4.4	4.7	5.4	5.0	3.9	5.6	5.5

Notes: Simulated size (%) of t-statistics ($k=1$) applied to standardized residuals under three DGPs: (i) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t \sim$ i.i.d. $N(0,1)$, (ii) $y_t = 1.25 + 2\varepsilon_t\sqrt{3/5}$, $\varepsilon_t \sim$ i.i.d. Student-t(5) (iii) $y_t = 1.25\varepsilon_t$, $\varepsilon_t \sim$ i.i.d. exp(1). Number of MC replications:1,000; bootstrap replications:500; nominal size 5%.

Table 1.6: Size of the *t*-Statistics under Parameter Uncertainty

T	$t_{1,7}$	$t_{2,7}$	$t_{3,7}$	$t_{4,7}$	$t_{5,7}$
Panel a: Normal					
250	5.0	6.2	4.9	6.4	5.8
500	5.0	5.1	6.6	5.0	4.8
1000	4.7	5.4	5.5	5.1	4.7
2000	5.0	4.6	4.8	5.3	4.6
Panel b: Student-t					
250	4.5	4.6	4.1	4.6	4.6
500	5.3	5.0	5.5	4.9	4.3
1000	4.2	4.4	4.1	4.4	4.0
2000	5.0	5.4	4.7	5.5	5.2
Panel c: Exponential					
250	5.3	5.3	5.6	4.6	5.0
500	7.0	6.2	5.3	4.4	5.4
1000	5.8	4.5	5.0	4.7	5.6
2000	4.5	4.5	5.5	5.3	6.0

Table 1.7: Size of the *Q*-Statistics under Parameter Uncertainty

T	Q_7^2	Q_7^3	Q_7^4	Q_7^5
Panel a: Normal				
250	5.3	6.7	6.4	6.5
500	6.0	6.5	6.7	6.6
1000	5.5	6.6	6.4	6.5
2000	4.4	5.1	4.6	4.5
Panel b: Student-t				
250	4.4	4.3	4.4	5.1
500	6.1	6.9	7.1	6.5
1000	5.0	6.5	6.3	6.0
2000	5.5	5.6	5.4	6.5
Panel c: Exponential				
250	5.9	5.5	5.3	4.7
500	7.3	6.0	5.6	5.5
1000	5.1	5.2	6.0	6.1
2000	4.5	4.4	5.2	5.2

Table 1.8: Size of the *J*-Statistics under Parameter Uncertainty

T	Normal	Student-t	Exponential
Panel a: J_1^{13}			
250	6.6	6.0	7.1
500	6.2	7.4	6.0
1000	6.6	6.1	6.9
2000	5.8	6.9	5.2
Panel b: J_1^7			
250	5.7	4.8	5.5
500	6.0	6.2	5.4
1000	5.4	5.4	5.7
2000	5.4	6.3	5.2

Notes: Simulated size (%) of the test statistics for three DGPs: (i) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t \sim$ i.i.d. $N(0,1)$, (ii) $y_t = 1.25 + 2\varepsilon_t\sqrt{3/5}$, $\varepsilon_t \sim$ i.i.d. Student-t(5) (iii) $y_t = 1.25\varepsilon_t$, $\varepsilon_t \sim$ i.i.d. exp(1). Table 1.6 presents the size of the *t*-statistic based on the 50% autocontour for $k = 1,...,5$. Table 1.7 presents the size of the *Q*-statistic based on the 50% autocontour for $K = 2,...,5$. Table 1.8 presents the size of the *J*-statistic based on all autocontours (Panel a) and 7 autocontours, excluding the first and the last three, (Panel b) for $k = 1$. All test statistics are based on standardized residuals. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5% in all cases.

Table 1.9: Power of the Q-Statistics under Parameter Uncertainty

Panel a: AR(1)

	T	Q_4^2	Q_4^3	Q_4^4	Q_4^5
$\varphi = 0.9$	250	83.9	87.5	91.8	93.5
	500	99.5	99.7	100.0	100.0
	1000	100.0	100.0	100.0	100.0
	2000	100.0	100.0	100.0	100.0
$\varphi = 0.7$	250	40.1	42.6	45.9	48.0
	500	70.7	74.7	79.7	80.7
	1000	96.3	98.4	99.3	99.5
	2000	99.9	100.0	100.0	100.0
$\varphi = 0.5$	250	11.9	12.6	11.9	11.6
	500	19.2	18.2	18.1	17.7
	1000	40.0	41.7	42.3	43.9
	2000	67.6	72.3	75.7	75.0
$\varphi = 0.3$	250	6.1	6.8	5.3	5.7
	500	4.5	5.4	6.4	6.1
	1000	8.5	8.7	9.5	9.8
	2000	14.0	13.6	11.9	13.3

Panel b: i.i.d. Student-t

	T	Q_8^2	Q_8^3	Q_8^4	Q_8^5
$v = 5$	250	73.8	71.0	69.2	67.3
	500	94.8	93.5	92.9	92.3
	1000	99.8	99.7	99.8	99.8
	2000	100.0	100.0	100.0	100.0
$v = 10$	250	27.2	25.5	23.5	21.7
	500	48.5	42.9	39.8	37.5
	1000	71.2	66.7	64.9	62.2
	2000	95.8	95.2	94.5	93.9
$v = 15$	250	15.8	14.1	12.3	11.3
	500	21.5	19.3	17.5	16.1
	1000	38.7	36.0	34.8	33.4
	2000	67.7	62.6	59.7	57.9

Table 1.9 (Continued)
Panel c: GARCH(1,1)

	T	Q_6^2	Q_6^3	Q_6^4	Q_6^5
$\alpha = 0.15$	250	41.3	38.4	36.8	35.6
	500	69.5	66.4	64.6	63
$\beta = 0.9$	1000	94.1	92.7	91.5	90.6
	2000	99.9	99.8	99.8	99.7
$\alpha = 0.1$	250	19.7	18	15.8	15.8
	500	40.2	36.1	35	32.4
$\beta = 0.85$	1000	67.2	63.3	60.7	58.5
	2000	91.9	90.8	88.5	86.2
$\alpha = 0.05$	250	5.2	5.1	5.9	5.7
	500	10.5	10.1	9.6	8.9
$\beta = 0.8$	1000	16.7	14.4	14.2	13.1
	2000	29.8	27.5	24.5	22.4

Notes: Simulated power (%) of the Q-statistic for $K = 2,\ldots,5$ under the following DGPs: (i) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t = \varphi\varepsilon_{t-1} + u_t$, $u_t \sim$ i.i.d. $N(0,1-\varphi^2)$ (Panel a), (ii) $y_t = 1.25 + 2\varepsilon_t\sqrt{(v-2)/v}$, $\varepsilon_t \sim$ i.i.d. Student-t(v) (Panel b), (iii) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t = \sqrt{h_t}u_t$, $u_t \sim$ i.i.d. $N(0,1)$, and $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$ (Panel c). The null hypothesis is $\varepsilon_t \sim$ i.i.d. $N(0,1)$. All test statistics are based on standardized residuals. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%.

Table 1.10: Power of the *J*-Statistics under Parameter Uncertainty

Panel a: AR(1)

	T	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
$\varphi = 0.9$	250	82.8	59.4	43.6	35.2	29.6
	500	98.0	84.8	69.4	53.2	44.9
	1000	100.0	99.0	90.6	78.8	65.7
	2000	100.0	100.0	99.5	94.9	85.5
$\varphi = 0.7$	250	43.1	17.9	10.7	9.1	10.1
	500	69.2	27.2	16.1	9.1	9.0
	1000	94.1	41.8	19.5	12.4	10.1
	2000	99.8	66.5	25.1	13.0	8.2
$\varphi = 0.5$	250	17.1	8.4	8.3	7.0	7.1
	500	27.8	9.2	7.2	7.7	7.3
	1000	46.4	11.2	7.5	7.5	6.9
	2000	75.0	12.8	8.6	4.4	6.3
$\varphi = 0.3$	250	8.9	5.9	7.4	7.1	7.0
	500	9.0	6.6	7.8	7.4	7.4
	1000	12.5	7.5	5.8	7.0	7.9
	2000	17.4	6.5	5.7	6.8	6.1

Panel b: i.i.d. Student-t

	T	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
$\nu = 5$	250	83.4	81.9	81.8	81.1	81.0
	500	97.3	97.5	96.9	97.3	95.7
	1000	100.0	99.9	100.0	100.0	100.0
	2000	100.0	100.0	100.0	100.0	100.0
$\nu = 10$	250	38.3	37.1	37.3	35.8	35.6
	500	54.6	55.4	54.7	53.7	52.7
	1000	79.2	78.7	76.0	77.5	75.8
	2000	97.2	97.4	97.4	97.6	97.5
$\nu = 15$	250	21.9	22.4	20.2	20.5	19.7
	500	27.3	28.6	27.7	27.3	28.2
	1000	47.7	47.4	47.0	45.4	48.3
	2000	69.3	69.1	69.7	69.0	68.9

Table 1.10 (Continued)
Panel c: GARCH(1,1)

	T	J_1^{13}	J_2^{13}	J_3^{13}	J_4^{13}	J_5^{13}
	250	46.8	45.2	45.9	41.7	40.4
$\alpha = 0.15$	500	72.8	70.8	70.9	70.3	68.1
$\beta = 0.8$	1000	94.1	93.4	93.4	92.5	92.2
	2000	100.0	99.6	99.7	99.4	99.6
	250	24.9	25.6	23.2	22.8	21.6
$\alpha = 0.1$	500	44.4	44.6	44.0	41.9	39.0
$\beta = 0.85$	1000	71.1	67.7	69.7	64.4	64.9
	2000	94.1	92.4	92.4	91.2	90.8
	250	12.1	10.0	11.3	9.9	10.1
$\alpha = 0.05$	500	15.9	13.8	14.7	14.2	14.6
$\beta = 0.9$	1000	20.7	21.2	19.3	19.5	18.7
	2000	36.1	34.0	33.1	31.1	29.3

Notes: Simulated power (%) of the *J*-statistic for $K = 1,\ldots,5$ for the following DGPs: (i) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t = \varphi\varepsilon_{t-1} + u_t$, $u_t \sim$ i.i.d. $N(0,1-\varphi^2)$ (Panel a), (ii) $y_t = 1.25 + 2\varepsilon_t\sqrt{(v-2)/v}$, $\varepsilon_t \sim$ i.i.d. Student-t(v) (Panel b), (iii) $y_t = 1.25 + 2\varepsilon_t$, $\varepsilon_t = \sqrt{h_t}u_t$, $u_t \sim$ i.i.d. $N(0,1)$, and $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$ (Panel c). The null hypothesis is $\varepsilon_t \sim$ i.i.d. $N(0,1)$. All test statistics are based on standardized residuals. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%.

Table 1.11: Size of the Tests applied to GARCH Residuals under Normal Distribution

T	$t_{1,1}$	$t_{1,2}$	$t_{1,3}$	$t_{1,4}$	$t_{1,5}$	$t_{1,6}$	$t_{1,7}$	$t_{1,8}$	$t_{1,9}$	$t_{1,10}$	$t_{1,11}$	$t_{1,12}$	$t_{1,13}$
500	4.6	3.9	4.5	5.3	5.8	5.3	6.0	6.5	6.5	4.7	4.1	4.0	4.2
1000	5.1	5.7	6.9	5.0	5.6	5.2	5.2	6.3	6.5	4.0	6.7	5.1	5.7
2000	4.3	4.6	6.3	4.4	4.4	4.2	4.6	4.8	6.9	6.4	3.2	4.9	4.9
T	$t_{1,4}$	$t_{2,4}$	$t_{3,4}$	$t_{4,4}$	$t_{5,4}$		Q_7^2	Q_7^3	Q_7^4	Q_7^5			J_1^{13}
500	6.0	6.6	5.4	6.1	6.2		6.8	7.0	6.9	6.1			7.0
1000	5.6	5.6	5.7	4.6	4.9		5.5	6.5	6.8	6.3			5.9
2000	4.9	5.4	4.6	5.6	5.1		5.2	5.2	5.8	5.2			6.0

Notes: Simulated size (%) of t, Q and J-statistics for the following DGP: $y_t = \sqrt{h_t}\varepsilon_t$ where $h_t = 0.1 + 0.15y_{t-1}^2 + 0.8h_{t-1}$, and $\varepsilon_t \sim \text{i.i.d.}N(0,1)$. The null hypothesis is $\varepsilon_t \sim \text{i.i.d.}N(0,1)$. All test statistics are based on standardized residuals. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%.

Table 1.12: Q-Statistics for GARCH (1,1) under Normal Distribution

	$K=5$	$K=10$	$K=15$	$K=20$	$K=25$
Q_1^K	4.96	8.55	10.40	13.58	15.76
Q_2^K	**20.39**	**31.29**	**32.42**	**43.41**	**45.99**
Q_3^K	**31.51**	**40.82**	**43.76**	**45.39**	**47.13**
Q_4^K	**51.54**	**68.02**	**70.88**	**83.26**	**96.83**
Q_5^K	**44.58**	**58.35**	**62.14**	**64.33**	**79.00**
Q_6^K	**47.09**	**50.78**	**58.11**	**60.27**	**68.73**
Q_7^K	**41.95**	**53.74**	**56.76**	**58.54**	**59.64**
Q_8^K	**31.05**	**32.25**	**33.60**	**38.36**	**46.20**
Q_9^K	**18.66**	**20.84**	**27.83**	**32.24**	**34.19**
Q_{10}^K	**32.10**	**37.84**	**44.30**	**50.78**	**58.86**
Q_{11}^K	5.01	8.42	13.86	18.73	21.68
Q_{12}^K	10.58	**19.11**	22.54	26.39	**37.38**
Q_{13}^K	**11.69**	**16.12**	19.47	24.54	29.31

Notes: Q-statistics for all 13 autocontours applied to the standardized residuals of the GARCH(1,1) model for daily NYSE returns: $y_t = \mu + \sqrt{h_t}\varepsilon_t$, $h_t = \omega + \alpha(y_{t-1} - \mu)^2 + \beta h_{t-1}$, and $\varepsilon_t \sim \text{i.i.d.} N(0,1)$. The null hypothesis is $\varepsilon_t \sim \text{i.i.d.} N(0,1)$. Covariance matrices are estimated by the parametric bootstrap procedure described in the text with 500 replications. Bold numbers indicate significance at 5% level.

Table 1.13: Q-Statistics for GARCH (1,1) under Student-t Distribution

	$K=5$	$K=10$	$K=15$	$K=20$	$K=25$
Q_1^K	3.61	4.34	6.23	8.81	11.65
Q_2^K	5.51	9.28	11.34	17.33	21.14
Q_3^K	7.02	13.25	18.11	20.70	22.58
Q_4^K	9.78	16.79	21.33	23.44	32.13
Q_5^K	**17.34**	**27.51**	**31.67**	**34.34**	**51.42**
Q_6^K	8.46	12.11	20.14	23.93	35.65
Q_7^K	**14.38**	**19.89**	**26.01**	30.09	35.00
Q_8^K	10.47	11.97	13.03	17.51	22.57
Q_9^K	4.15	6.73	10.66	14.77	18.04
Q_{10}^K	**12.93**	**29.41**	**37.82**	**43.68**	**55.37**
Q_{11}^K	5.44	7.16	11.62	20.33	30.50
Q_{12}^K	6.47	16.18	17.30	22.01	34.15
Q_{13}^K	1.71	5.86	8.89	13.74	16.58

Notes: Q-statistics for all 13 autocontours and five lag values ($K \in \{5,10,15,20,25\}$) for the standardized residuals of the following GARCH(1,1) model fitted to daily NYSE returns: $y_t = \mu + \sqrt{h_t(v-2)/v}\varepsilon_t$ where $h_t = \omega + \alpha(y_{t-1} - \mu)^2 + \beta h_t$, $\varepsilon_t \sim$ i.i.d. Student-t(v), and $v = 7$. The null hypothesis is $\varepsilon_t \sim$ i.i.d. Student-t(7). Covariance matrices are estimated by the parametric bootstrap procedure described in the text with 500 replications. Bold numbers indicate significance at 5% level.

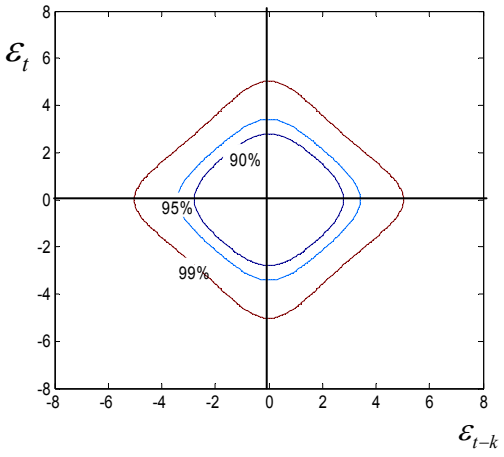
Table 1.14: Q-Statistics for ACD(3,2) under Exponential Distribution

	$K=5$	$K=10$	$K=15$	$K=20$	$K=25$
Q_1^K	**15,707.94**	**18,063.22**	**18,257.23**	**19,077.22**	**19,613.71**
Q_2^K	**8,053.69**	**8,648.05**	**9,079.96**	**9,149.96**	**9,357.13**
Q_3^K	**5,387.60**	**5,721.03**	**5,840.05**	**5,989.33**	**6,128.28**
Q_4^K	**2,825.55**	**2,941.91**	**2,992.49**	**3,054.07**	**3,098.97**
Q_5^K	**1,619.36**	**1,646.27**	**1,697.05**	**1,738.21**	**1,777.32**
Q_6^K	**899.63**	**922.26**	**943.10**	**952.72**	**958.68**
Q_7^K	**384.22**	**389.80**	**398.84**	**400.46**	**402.91**
Q_8^K	**105.79**	**120.92**	**129.72**	**131.98**	**134.26**
Q_9^K	**11.58**	16.11	22.97	25.70	30.01
Q_{10}^K	**187.85**	**188.81**	**191.41**	**193.92**	**197.34**
Q_{11}^K	**679.17**	**739.94**	**744.11**	**758.21**	**762.26**
Q_{12}^K	**1,166.99**	**1,223.89**	**1,238.99**	**1,271.75**	**1,287.17**
Q_{13}^K	**1,901.58**	**1,987.34**	**2,123.62**	**2,132.91**	**2,196.79**

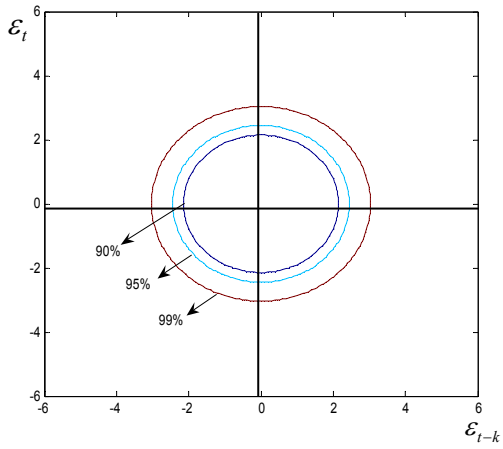
Notes: Q-statistics for all 13 autocontours and five lag values $(K \in \{5,10,15,20,25\})$ for the standardized residuals of the following ACD(3,2) model fitted to Airgas intra-day transaction data: $x_i = \psi_i \varepsilon_i$ where $\psi_i = \omega + \sum_{j=1}^{3} \alpha_j x_{i-j} + \sum_{j=1}^{2} \beta_j \psi_{i-j}$, $\varepsilon_i \sim$ i.i.d. exp(1). The null hypothesis is $\varepsilon_i \sim$ i.i.d. exp(1). Covariance matrices are estimated by the parametric bootstrap procedure described in the text with 500 replications. Bold numbers indicate significance at 5% level.

Figure 1.1 Sample Autocontours of Bivariate Distributions under Independence

Panel a: Standard Normal



Panel b: Student-t (ν = 5)



Panel c: Exponential (β = 1)

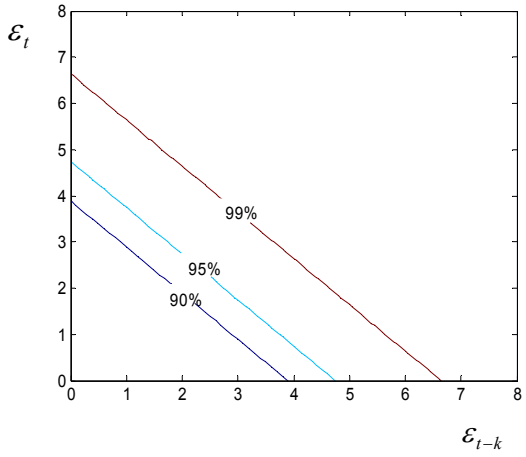
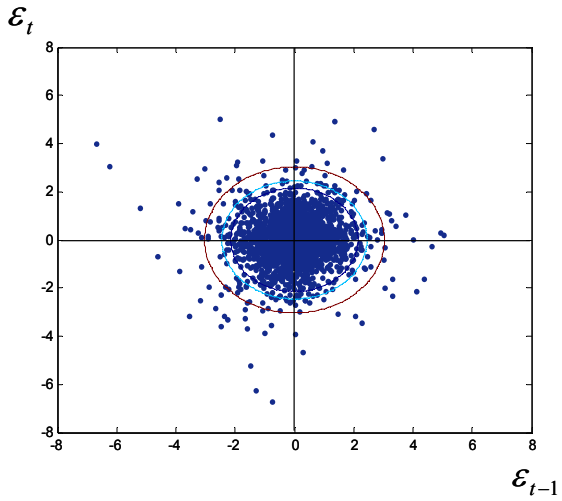


Figure 1.2
Panel a: 90, 95, 99% Autocontours under Normal Distribution and Standardized NYSE Returns



Panel b: Autocontourgrams of NYSE Returns under Normal Distribution

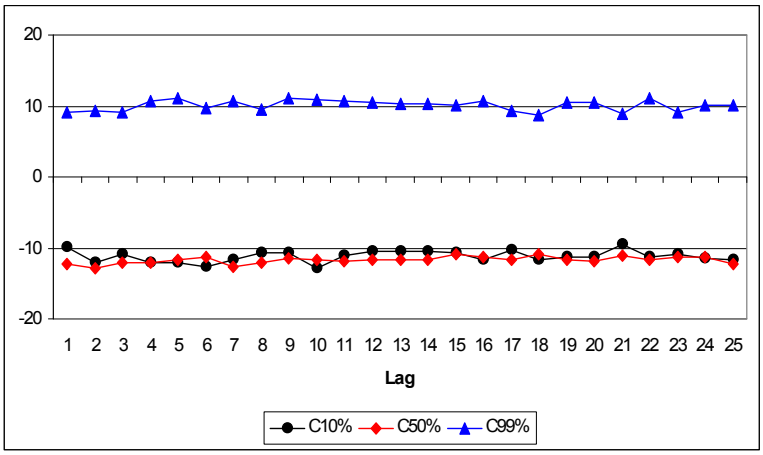
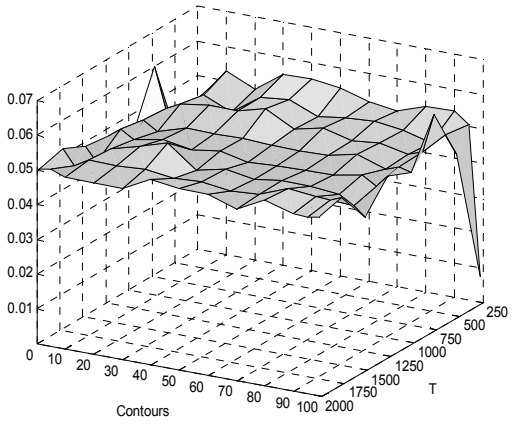
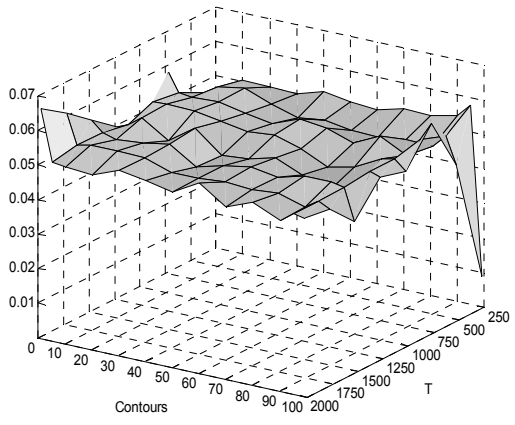


Figure 1.3: Size of the *t*-Statistics

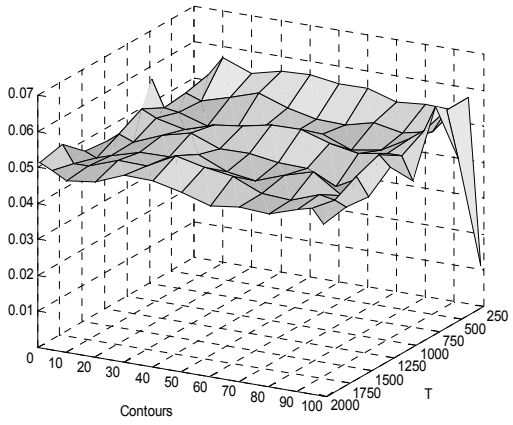
Panel a: $\varepsilon_t \sim$ i.i.d.$N(0,1)$



Panel b: $\varepsilon_t \sim$ i.i.d. Student-t(5)

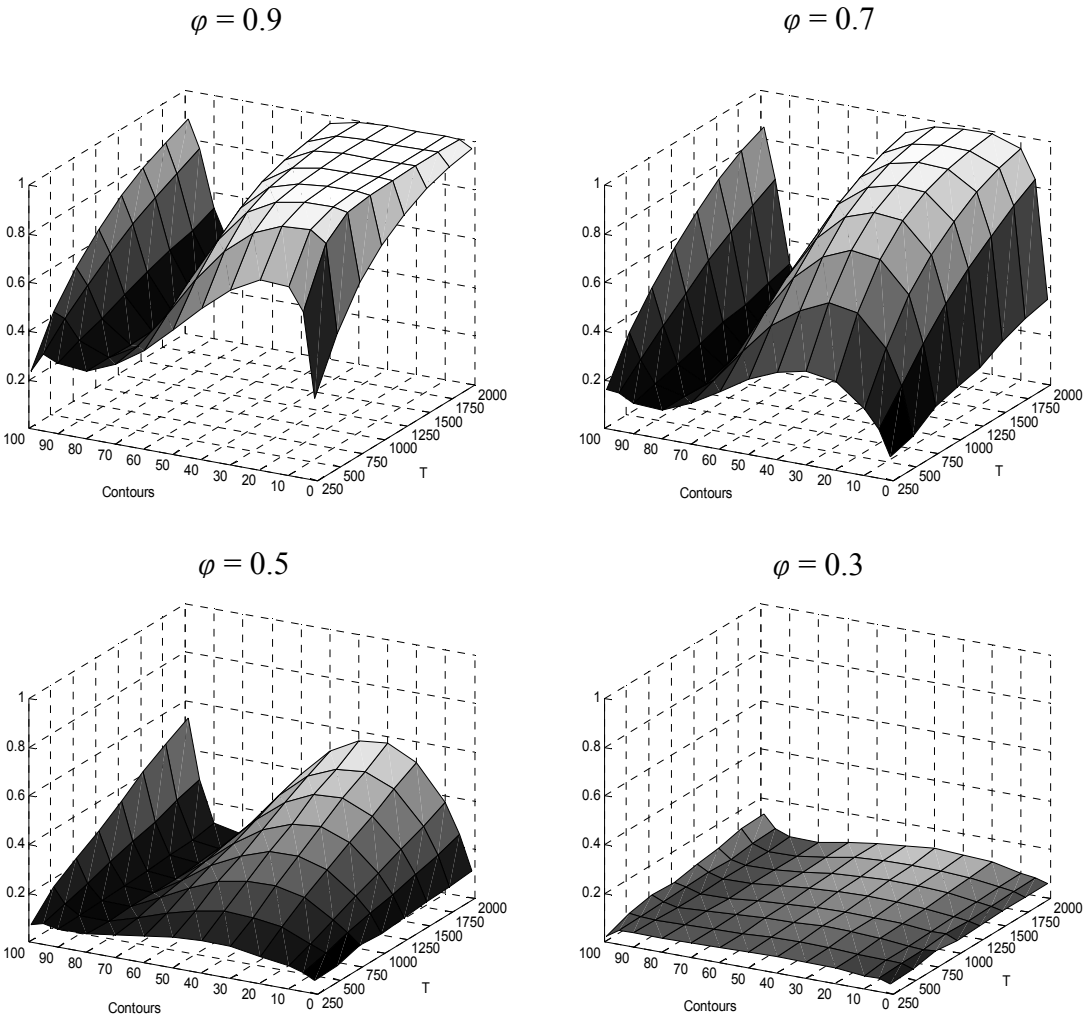


Panel c: $\varepsilon_t \sim$ i.i.d. exp(1)



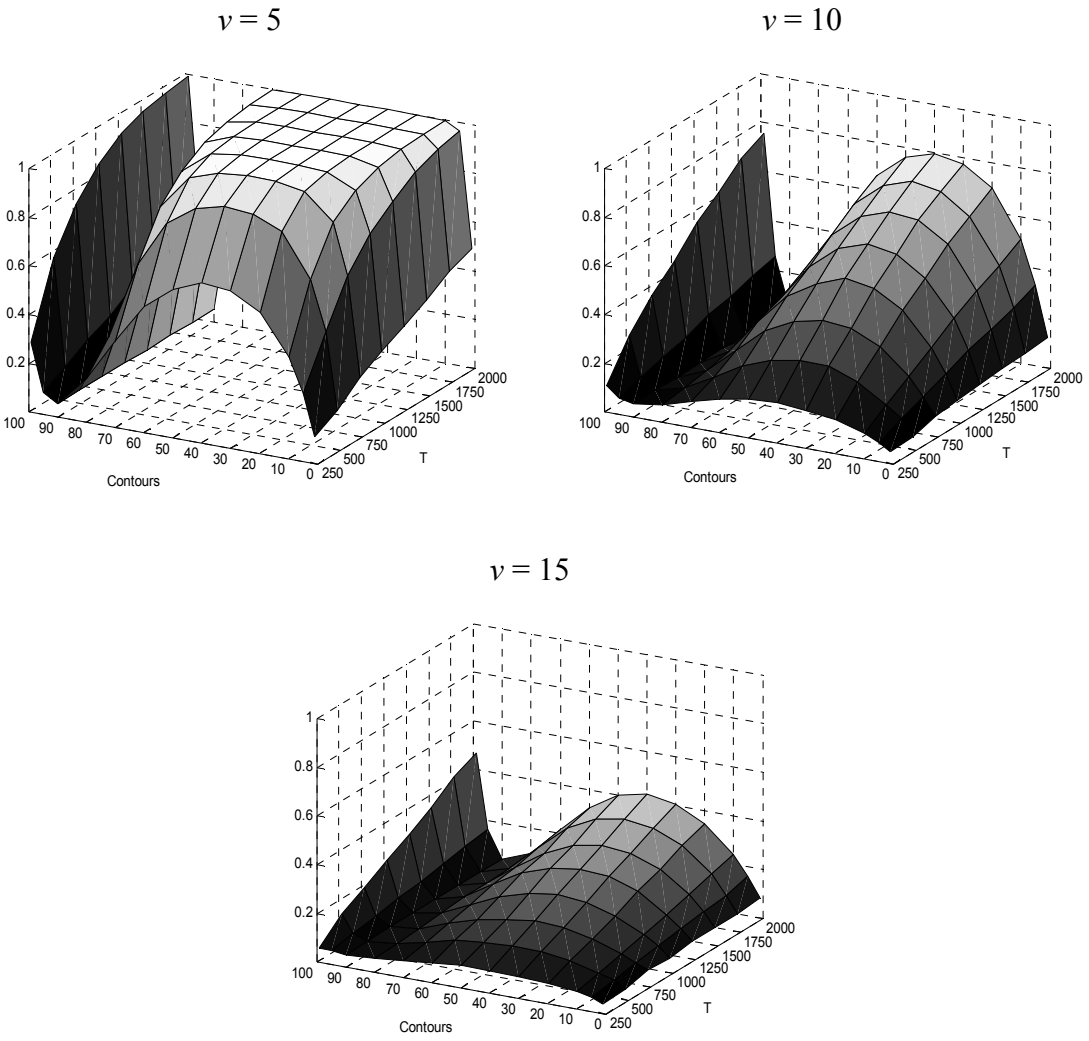
Notes: Simulated size of the t-statistic for all 13 autocontours and the first lag $(k = 1)$ under three alternative DGPs: (i) $\varepsilon_t \sim$ i.i.d.$N(0,1)$ (ii) $\varepsilon_t \sim$ i.i.d. Student-t(5), (iii) $\varepsilon_t \sim$ i.i.d. exp(1) . Number of MC replications is 10,000 and nominal size is 0.05.

Figure 1.4: Power of the *t*-Statistics under AR(1) DGP



Notes: Simulated power of the *t*-statistic for all 13 autocontours and for $k = 1$ under the following DGP: $\varepsilon_t = \varphi\varepsilon_{t-1} + u_t$ where $u_t \sim$ i.i.d. $N(0,1-\varphi^2)$. Number of MC replications is 10,000 and nominal size is 5%.

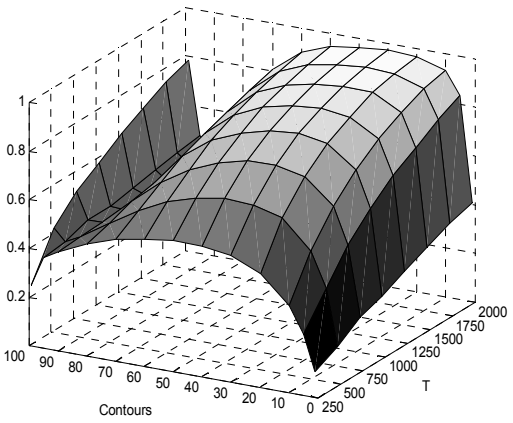
Figure 1.5: Power of the *t*-Statistics under i.i.d. Student-t DGP



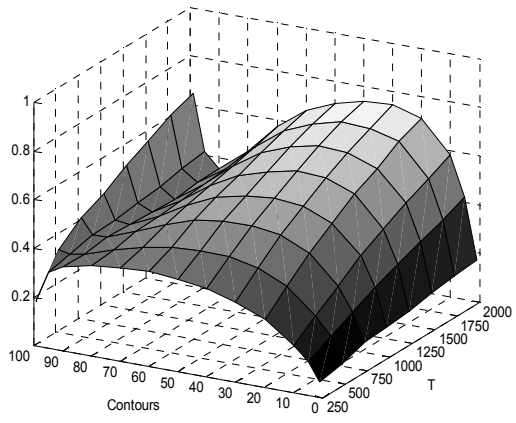
Notes: Simulated power of the *t*-statistic for all 13 autocontours and the first lag ($k = 1$) under the following DGP: $\varepsilon_t = u_t\sqrt{(v-2)/v}$ where $u_t \sim$ i.i.d. Student-t(v). Number of MC replications is 10,000 and nominal size is 5%.

Figure 1.6: Power of the *t*-Statistics under GARCH (1,1) DGP

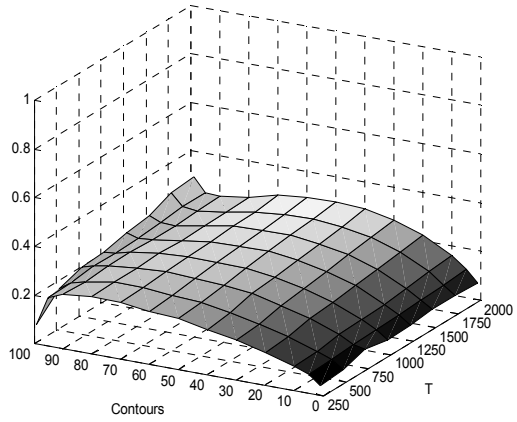
$\alpha = 0.15$
$\beta = 0.8$



$\alpha = 0.1$
$\beta = 0.85$

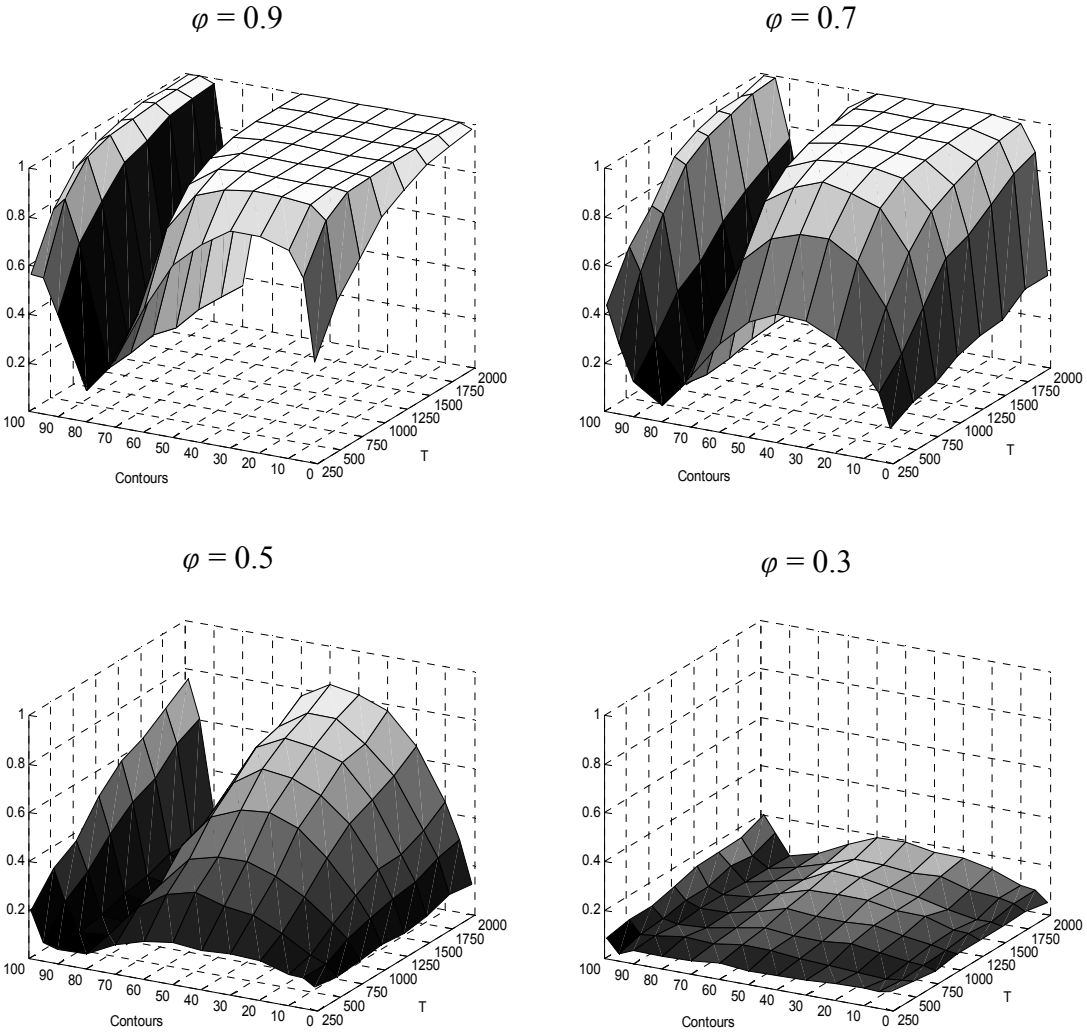


$\alpha = 0.05$
$\beta = 0.9$



Notes: Simulated power of the *t*-statistic for all 13 autocontours and the first lag $(k = 1)$ under the following DGP: $\varepsilon_t = \sqrt{h_t}u_t$ where $u_t \sim$ i.i.d. $N(0,1)$ and $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$. Number of MC replications is 10,000 and nominal size is 5%.

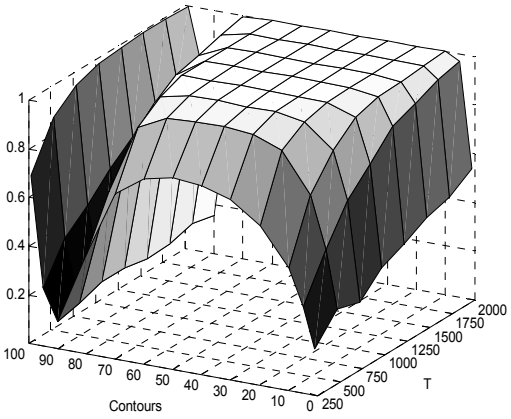
Figure 1.7: Power of the *t*-Statistics under AR (1) DGP and Parameter Uncertainty



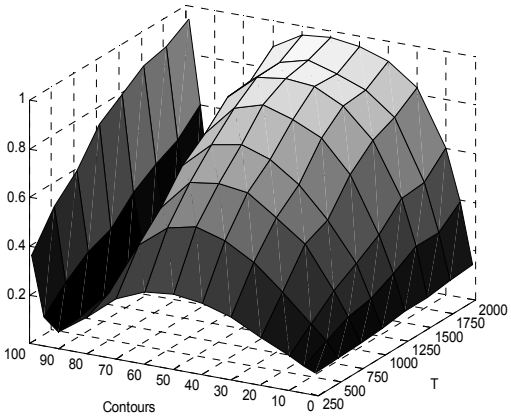
Notes: Simulated power of the *t*-statistic for all 13 autocontours and the first lag $(k = 1)$ under the following DGP: $y_t = 1.25 + 2\varepsilon_t$ where $\varepsilon_t = \varphi\varepsilon_{t-1} + u_t$, and $u_t \sim$ i.i.d. $N(0,1-\varphi^2)$. The null hypothesis is $\varepsilon_t \sim$ i.i.d.$N(0,1)$. All test statistics are based on standardized residuals. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%.

Figure 1.8: Power of the *t*-Statistics under i.i.d. Student-t DGP and Parameter Uncertainty

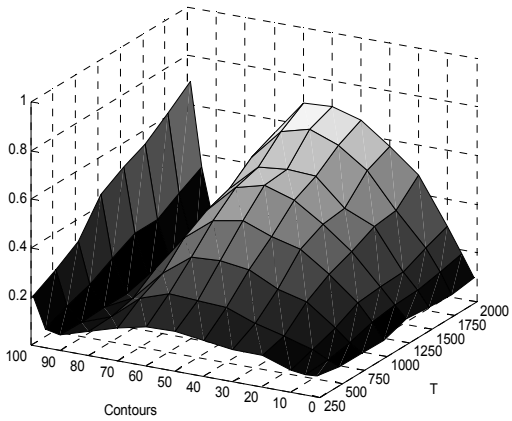
$v = 5$



$v = 10$



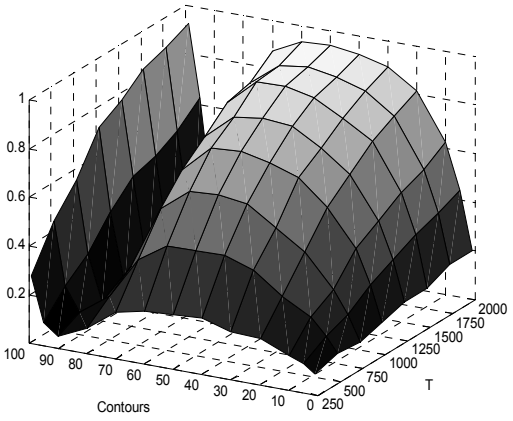
$v = 5$



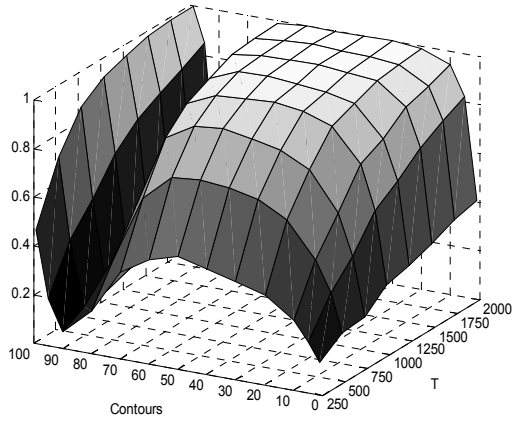
Notes: Simulated power of the *t*-statistic for all 13 autocontours and the first lag ($k = 1$) under the following DGP: $y_t = 1.25 + 2\varepsilon_t\sqrt{(v-2)/v}$ where $\varepsilon_t \sim$ i.i.d. Student-t(v). The null hypothesis is $\varepsilon_t \sim \text{i.i.d.}N(0,1)$. All test statistics are based on standardized residuals. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%.

Figure 1.9: Power of the *t*-Statistics under GARCH (1,1) DGP and Parameter Uncertainty

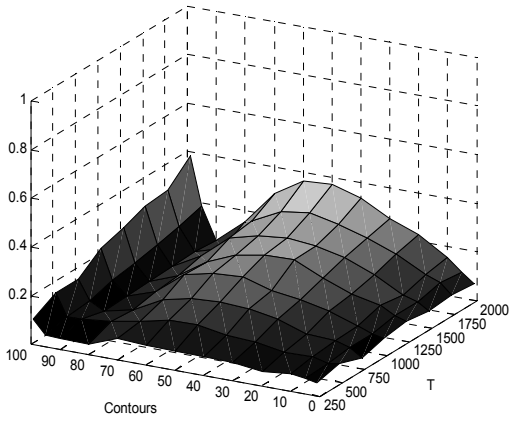
$\alpha = 0.15$
$\beta = 0.8$



$\alpha = 0.1$
$\beta = 0.85$

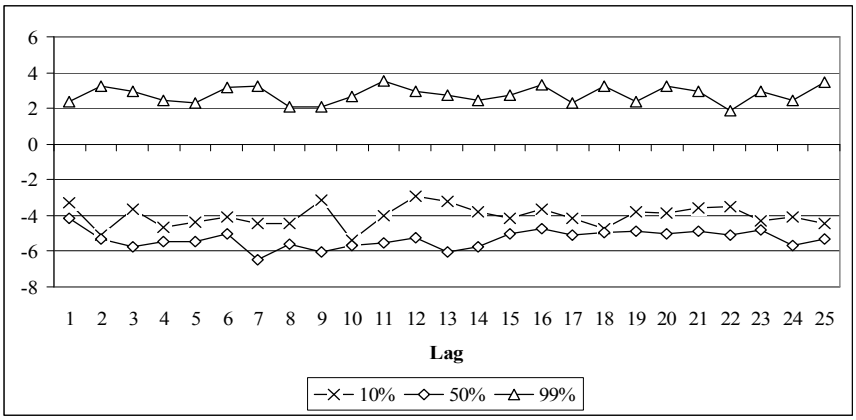


$\alpha = 0.05$
$\beta = 0.9$

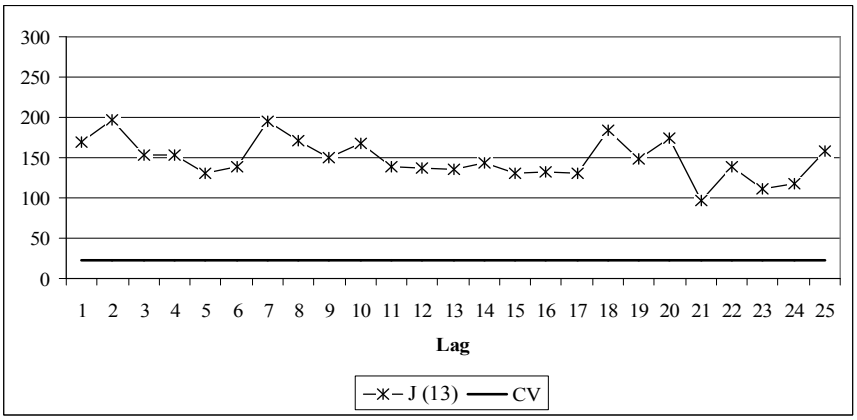


Notes: Simulated power of the *t*-statistic for all 13 autocontours and the first lag ($k = 1$) under the following DGP: $y_t = 1.25 + 2\varepsilon_t$ where $\varepsilon_t = \sqrt{h_t}u_t$, $u_t \sim \text{i.i.d.}N(0,1)$, $h_t = \omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}$, and $\omega \equiv 1 - \alpha - \beta$. The null hypothesis is $\varepsilon_t \sim \text{i.i.d.}N(0,1)$. All test statistics are based on standardized residuals. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%**.**

Figure 1.10
Panel a: *t*-Statistics for GARCH (1,1) Model of NYSE Returns under Normal Distribution

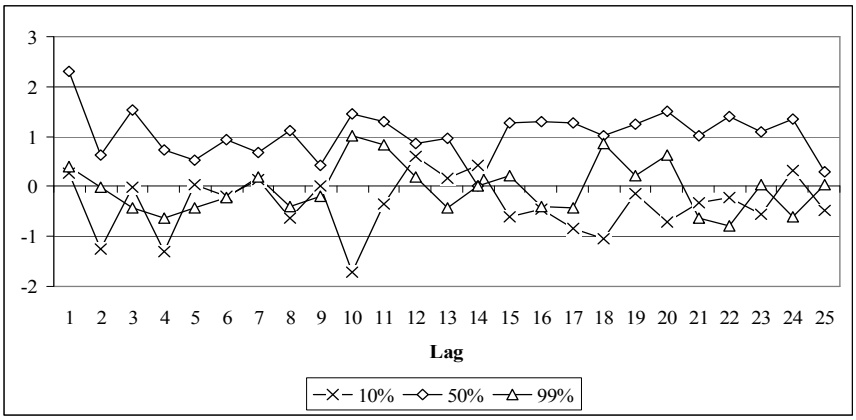


Panel b: *J*-Statistics for GARCH (1,1) Model of NYSE Returns under Normal Distribution

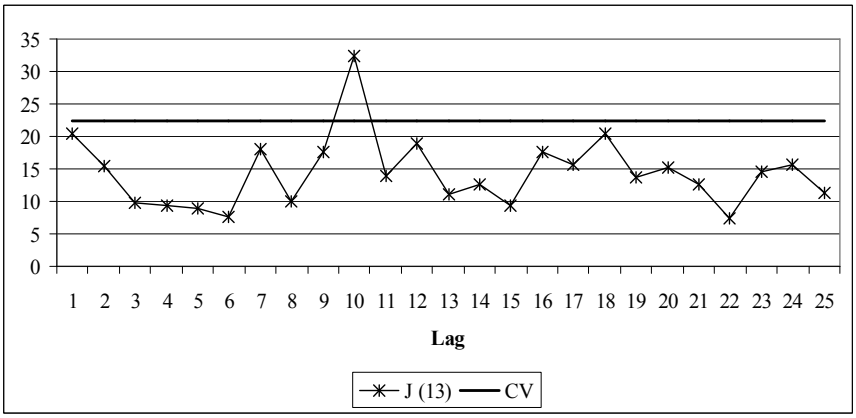


Notes: *t*-statistic for three autocontours (10, 50, 99%) up to 25 lags for the residuals of the normal GARCH(1,1) model fitted to daily NYSE returns (Panel a). *J*-statistic based on all 13 autocontours up to 25 lags for the residuals of the normal GARCH(1,1) model (Panel b). CV denotes 5% critical value.

Figure 1.11
Panel a: *t*-statistics for GARCH (1,1) Model of NYSE Returns
under Student-t Distribution



Panel b: *J*-statistics for GARCH (1,1) Model of NYSE Returns
under Student-t Distribution

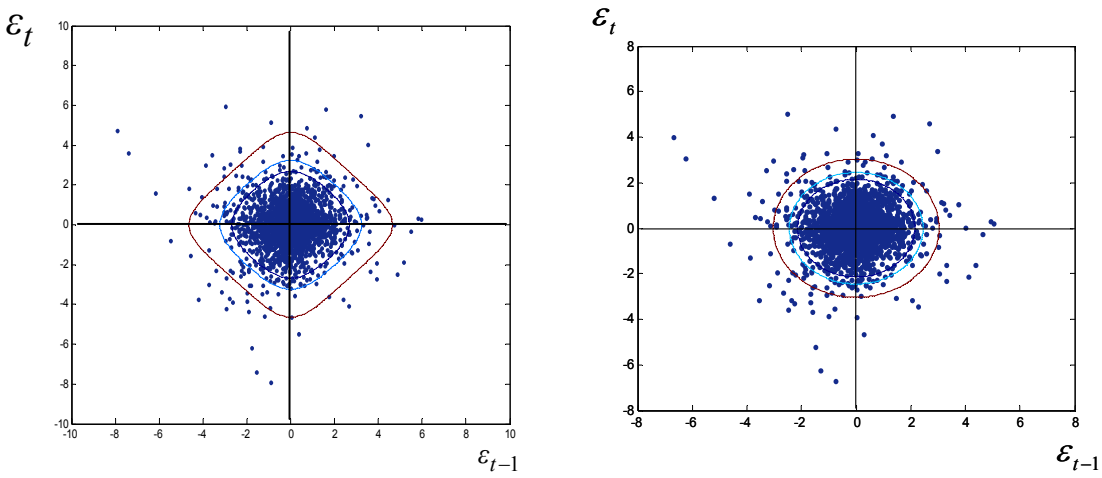


Notes: *t*-statistic for three autocontours (10, 50, 99%) up to 25 lags for the residuals of the Student-t GARCH(1,1) model fitted to daily NYSE returns (Panel a). *J*-statistic based on all 13 autocontours up to 25 lags for the residuals of the Student-t GARCH(1,1) model (Panel b). CV denotes 5% critical value.

Figure 1.12

Panel a: 90, 95, 99% Autocontours and Standardized NYSE Returns

under Normal distribution under Student-t(7) Distribution



Panel b: 90, 95, 99% Autocontours under Student-t(7) Distribution and Standardized GARCH(1,1) Residuals

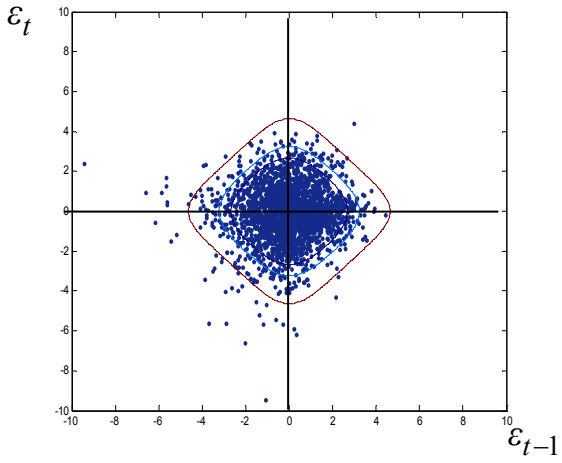
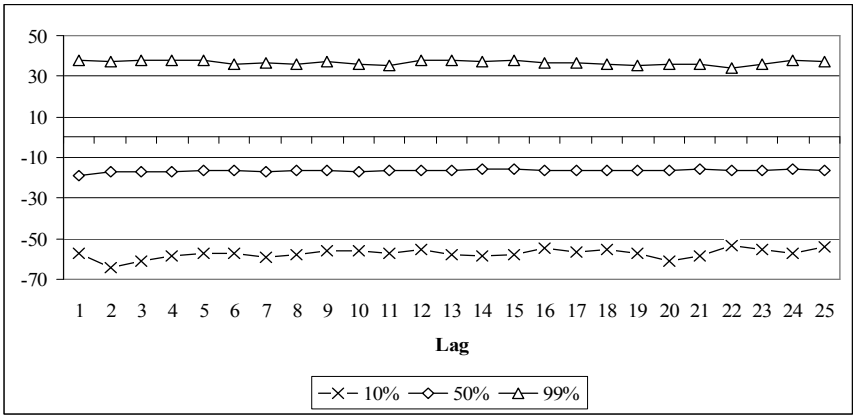
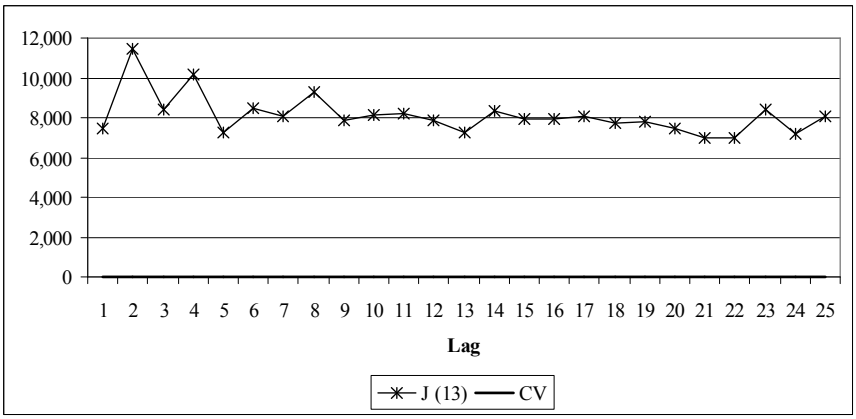


Figure 1.13
Panel a: *t*-Statistics for ACD (3,2) Model of Airgas Transaction Durations under Exponential Distribution



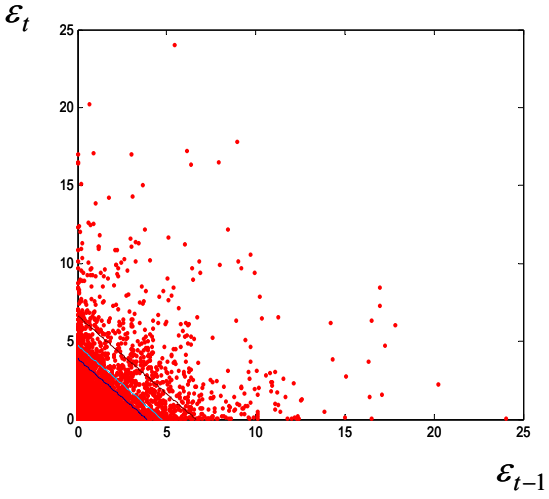
Panel b: *J*-Statistics for ACD (3,2) Model of Airgas Transaction Durations under Exponential Distribution



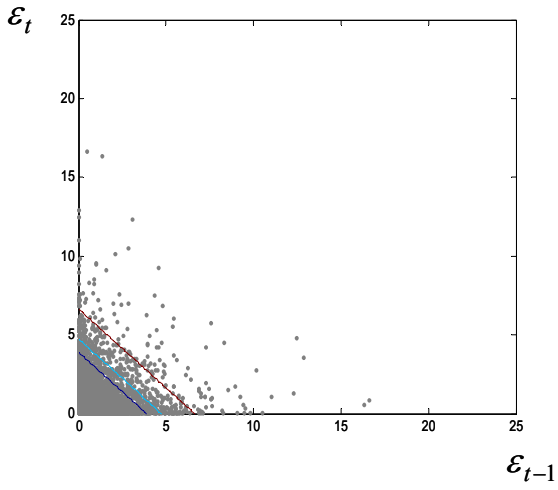
Notes: *t*-statistic for three autocontours (10, 50, 99%) up to 25 lags for the residuals of the exponential ACD(3,2) model fitted to Airgas intra-day transaction durations (Panel a). *J*-statistic based on all 13 autocontours up to 25 lags for the residuals of the exponential ACD(3,2) model (Panel b). CV denotes 5% critical value.

Figure 1.14: 90, 95, 99% Autocontours under Exponential Distribution and Standardized Airgas Durations

Panel a: Standardized Durations



Panel b: Standardized ACD Residuals



CHAPTER II

MULTIVARIATE AUTOCONTOURS FOR SPECIFICATION TESTING IN MULTIVARIATE GARCH MODELS

1 Introduction

Even though there is an extensive literature on specification tests for univariate time series models, the development of new tests for multivariate models has been very slow. As an example, in the ARCH literature we have numerous univariate specifications for which we routinely scrutinize the standardized residuals for possible neglected dependence and deviation from the assumed conditional density. However, for multivariate GARCH models we rarely test for the assumed multivariate density and for cross-dependence in the residuals. Given the inherent difficulty of estimating multivariate GARCH models, the issue of dynamic misspecification at the system level -as important as it may be- seems to be secondary. Though univariate specification tests can be performed in each equation of the system, these tests are not independent from each other, and an evaluation of the system will demand adjustments in the size of any joint test that combines the results of the equation-by-equation univariate tests. Bauwens, Laurent, and Rombouts (2006) survey the latest developments in multivariate GARCH models and they also acknowledge the need for further research on multivariate diagnostic tests. There are some portmanteau statistics for neglected multivariate

conditional heteroskedasticity as in Ling and Li (1997), Tse and Tsui (1999), and Duchesne and Lalancette (2003). Some of these tests have unknown asymptotic distributions when applied to the generalized GARCH residuals. Tse (2002) proposes another type of misspecification test that is based on regressions of the standardized residuals on some explanatory variables. In that case, the usual OLS asymptotics do not apply, but it is possible to construct some statistics that are asymptotically chi-squared distributed under the null of no dynamic misspecification. None of these tests are concerned with the specification of the multivariate density. However, the knowledge of the density functional form is of paramount importance for density forecast evaluation, which is needed to assess the overall adequacy of the model. Recently, Bai and Chen (2008) adopted the empirical process based testing approach of Bai (2003), which is developed in the univariate framework, to multivariate models. They use single-indexed empirical processes to make computation feasible, but this causes loss of full consistency. Kalliovirta (2007) also takes an empirical process based approach and proposes several test statistics for checking dynamic misspecification and density functional form.

We propose a new battery of tests for dynamic specification and density functional form in multivariate time series models. We focus on the most popular models for which all the time dependence is confined to the first and second moments of the multivariate process. Multivariate dynamics in moments further than the second are difficult to find in the data and, to our knowledge, there are only a few attempts in the literature restricted to upmost bivariate systems. Our approach is not based on empirical processes, so we do not require probability integral transformations as opposed to the above mentioned

studies testing for density specification. This makes dealing with parameter uncertainty relatively less challenging on theoretical grounds. When parameter estimation is required, we will adopt a quasi-maximum likelihood procedure as opposed to strict maximum likelihood, which assumes the knowledge of the true multivariate density. If the true density were known, it would be possible to construct tests for dynamic misspecification based on the martingale difference property of the score under the null. However, if the density function is unknown, a quasi-maximum likelihood estimator is the most desirable to avoid the inconsistency of the estimator that we would have obtained under a potentially false density function. The lack of consistency may also jeopardize the asymptotic distribution of the tests. Our approach is less demanding than any score-type testing in the sense that once quasi-maximum likelihood estimates are in place, we can proceed to test different proposals on the functional form of the conditional multivariate density function.

The proposed tests are based on the concept of "autocontour" introduced in Chapter I for univariate processes. Our methodology is applicable to a wide range of models including linear and non-linear VAR specifications with multivariate GARCH disturbances. The variable of interest is the vector of generalized innovations $\varepsilon_t = (\varepsilon_{1t}, \varepsilon_{2t}, \ldots, \varepsilon_{kt})'$ in a model $y_t = \mu_t(\theta_{01}) + H_t^{1/2}(\theta_{02})\varepsilon_t$, where y_t is a $k \times 1$ vector of variables with conditional mean vector μ_t and conditional covariance matrix H_t. Under the null hypothesis of correct dynamic specification the vector ε_t must be i.i.d. with a certain parametric multivariate probability density function $f(.)$. Thus, if we

consider the joint distribution of two vectors ε_t and ε_{t-l}, then under the null we have $f(\varepsilon_t, \varepsilon_{t-l}) = f(\varepsilon_t) f(\varepsilon_{t-l})$. The basic idea of the proposed tests is to calculate the percentage of observations contained within the probability autocontour planes corresponding to the assumed multivariate density of the vector of independent innovations, i.e. $f(\varepsilon_t) f(\varepsilon_{t-l})$, and to statistically compare it to the population percentage. We develop a battery of *t*-tests based on a single autocontour and also more powerful chi-squared tests based on multiple autocontours, which have standard asymptotic distributions. Without parameter uncertainty the test statistics are all distribution free, but under parameter uncertainty there are nuisance parameters affecting the asymptotic distributions. We show that a simple bootstrap procedure overcomes this problem and yields the correct size even for moderate sample sizes. We also investigate the power properties of the test statistics in finite samples.

Since the null is a joint hypothesis, the rejection of the null begs the question on what is at fault. Thus, it is desirable to separate i.i.d-ness from density function. In the spirit of goodness-of-fit tests, we also propose an additional test that focuses on the multivariate density functional form of the vector of innovations. Following a similar approach, we construct the probability contours corresponding to the hypothesized multivariate density, $f(\varepsilon_t)$, and compare the sample percentage of observations falling within the contour to the population percentage. The goodness-of-fit tests are also constructed as *t*-statistics and chi-squared statistics with standard distributions.

The organization of the chapter is as follows. In Section 2, we describe the battery of tests and the construction of the multivariate contours and autocontours. In Section 3,

we offer some Monte Carlo simulation to assess the size and power of the tests in finite samples. In Section 4, we apply the tests to the generalized residuals of GARCH models with hypothesized multivariate Normal and multivariate Student-t innovations fitted to excess returns on five size portfolios. In Section 5, we conclude.

2 Testing Methodology

2.1 Test Statistics

Let $y_t = (y_{1t},\ldots,y_{kt})$ and suppose that y_t evolves according to the following process

$$y_t = \mu_t(\theta_{01}) + H_t^{1/2}(\theta_{02})\varepsilon_t, \quad t = 1,\ldots,T, \tag{1}$$

where $\mu_t(.)$ and $H_t^{1/2}(.)$ are both measurable with respect to time $t-1$ sigma field, $\Im_{t-1}$, $H_t(.)$ is positive definite, and $\{\varepsilon_t\}$ is an i.i.d. vector process with zero mean and identity covariance matrix. The conditional mean vector, $\mu_t(.)$, and the conditional covariance matrix, $H_t(.)$, are fully parameterized by the parameter vector $\theta_0 = (\theta_{01}', \theta_{02}')'$, which for now we assume to be known, but later on we will relax this assumption to account for parameter uncertainty.

If all the dependence is contained in the first and second conditional moments of the process y_t, then the null hypothesis of interest to test for model misspecification is

$$H_0 : \varepsilon_t \text{ is i.i.d. with density } f(.).$$

The alternative hypothesis is the negation of the null. Though we wish to capture all the dynamic dependence of y_t through the modeling of the conditional mean and conditional

covariance matrix, there may be another degree of dependence that is built in the assumed multivariate density, $f(.)$. In fact, once we move beyond the assumption of multivariate normality, for instance when we assume a multivariate Student-t distribution, the components of the vector ε_t are dependent among themselves and this information is only contained within the functional form of the density. This is why, among other reasons, it is of interest to incorporate the assumed density function in the null hypothesis.

Let us consider the joint distribution of two $k \times 1$ vectors ε_t and ε_{t-l}, $l = 1, \ldots, L < \infty$. Define a $2k \times 1$ vector $\eta_t = (\varepsilon_t', \varepsilon_{t-l}')'$ and let $\psi(.)$ denote the associated density function. Under the null hypothesis of i.i.d. and correct probability density function, we can write $\psi(\eta_t) = f(\varepsilon_t) f(\varepsilon_{t-l})$. Then, under the null, we define the α-autocontour, $C_{l,\alpha}$, as the set of vectors $(\varepsilon_t', \varepsilon_{t-l}')$ that results from slicing the multivariate density, $\psi(.)$, at a certain value to guarantee that the set contains $\alpha\%$ of observations, that is,

$$C_{l,\alpha} = \left\{ S(\eta_t) \subset \Re^{2k} \left| \int_{h_1}^{g_1} \cdots \int_{h_{2k}}^{g_{2k}} \psi(\eta_t) \mathrm{d}\eta_{1t} \ldots \mathrm{d}\eta_{2k,t} \leq \alpha \right. \right\}, \tag{2}$$

where the limits of integration are determined by the density functional form so that the shape of the probability contours is preserved under integration, e.g. when the assumed density is normal, then the autocontours are $2k$-spheres (a circle when $k = 1$). We construct an indicator process defined as

$$I_t^{l,\alpha} = \begin{cases} 1 \text{ if } \eta_t \notin C_{l,\alpha} \\ 0 \text{ otherwise} \end{cases}. \tag{3}$$

The process $\{I_t^{l,\alpha}\}$ forms the building block of the proposed test statistics. Let $p_\alpha \equiv 1-\alpha$. Since the indicator is a Bernoulli random variable, its mean and variance are given by $E[I_t^{l,\alpha}] = p_\alpha$ and $Var(I_t^{l,\alpha}) = p_\alpha(1-p_\alpha)$. Although $\{\varepsilon_t\}$ is an i.i.d. process, $\{I_t^{l,\alpha}\}$ exhibits some linear dependence because $I_t^{l,\alpha}$ and $I_{t-l}^{l,\alpha}$ share common information contained in ε_{t-l}. Hence, the autocovariance function of $\{I_t^{l,\alpha}\}$ is given by

$$\gamma_h^\alpha = \begin{cases} P(I_t^{l,\alpha} = 1,\, I_{t-h}^{l,\alpha} = 1) - p_\alpha^2 & \text{if } h = l \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

Proposition 1 Define $\hat{p}_\alpha^l = (T-l)^{-1}\sum_{t=1}^{T-l} I_t^{l,\alpha}$. Under the null hypothesis, $t_{l,\alpha} = \sqrt{T-l}(\hat{p}_\alpha^l - p_\alpha) \to_d \mathrm{N}(0, \sigma_{l,\alpha})$, where $\sigma_{l,\alpha}^2 = p_\alpha(1-p_\alpha) + 2\gamma_l^\alpha$.

Proof: See the Appendences of Chapter I for all mathematical proofs.

Now let us consider a finite number of contours, $(\alpha_1, \ldots, \alpha_n)$, jointly. Let $p_\alpha = (p_{\alpha_1}, \ldots, p_{\alpha_n})'$ where $p_{\alpha_i} = 1-\alpha_i$, and define $\hat{p}_{\alpha_i}^l = (T-l)^{-1}\sum_{t=1}^{T-l} I_t^{l,\alpha_i}$ for $i = 1, \ldots, n$. We then collect all the $\hat{p}_{\alpha_i}^l$'s in a $n \times 1$ vector, $\hat{p}_\alpha^l = (\hat{p}_1, \ldots, \hat{p}_n)'$.

Proposition 2 Under the null hypothesis, $\sqrt{T-l}(\hat{p}_\alpha^l - p_\alpha) \to_d \mathrm{N}(0,\Xi)$, where the elements of Ξ are $\xi_{ij} = \min(p_{\alpha_i}, p_{\alpha_j}) - p_{\alpha_i} p_{\alpha_j} + Cov(I_t^{l,\alpha_i}, I_{t-l}^{l,\alpha_j}) + Cov(I_t^{l,\alpha_j}, I_{t-l}^{l,\alpha_i})$. Then, it directly follows that $J_n^l = (T-l)(\hat{p}_\alpha^l - p_\alpha)'\Xi^{-1}(\hat{p}_\alpha^l - p_\alpha) \to_d \chi^2(n)$.

A complementary test to those described above can be constructed in the spirit of goodness-of-fit. Suppose that we consider only the vector ε_t and we wish to test in the direction of density functional form. We construct the probability contour sets C_α corresponding to the probability density function that is assumed under the null hypothesis. The set is given by

$$C_\alpha = \left\{ S(\varepsilon_t) \subset \Re^k \left| \int_{h_1}^{g_1} \cdots \int_{h_k}^{g_k} f(\varepsilon_t) \mathrm{d}\varepsilon_{1t} \ldots \mathrm{d}\varepsilon_{kt} \le \alpha \right. \right\}. \tag{5}$$

Then, as before, we construct an indicator process as follows

$$I_t^\alpha = \begin{cases} 1 \text{ if } \varepsilon_t \notin C_\alpha \\ 0 \text{ otherwise} \end{cases}, \tag{6}$$

for which the mean and variance are $E[I_t^\alpha] = 1-\alpha$ and $Var(I_t^\alpha) = \alpha(1-\alpha)$, respectively. The main difference between the sets $C_{l,\alpha}$ and C_α is that the latter does not explicitly consider the time-independence assumed under the null and, therefore, the following tests based on C_α will be less powerful against independence. There is also a difference in the properties of the indicator process. Now, the indicator is also an i.i.d. process, and the analogous tests to those of Propositions 1 and 2 will have a simpler asymptotic distribution.

Let $p_\alpha = 1-\alpha$ and define an estimator of p_α as $\widetilde{p}_\alpha = T^{-1}\sum_{t=1}^{T} I_t^\alpha$. Under the null hypothesis the distribution of the analogue test statistic to that of Proposition 1 is

$$t_\alpha = \frac{\sqrt{T}(\widetilde{p}_\alpha - p_\alpha)}{p_\alpha(1-p_\alpha)} \rightarrow_d \mathrm{N}(0,1). \tag{7}$$

If, as in Proposition 2, now we jointly consider a finite number of contours and define the vectors $p_\alpha = (p_{\alpha_1},\ldots,p_{\alpha_n})'$ and $\widetilde{p}_\alpha = (\widetilde{p}_{\alpha_1},\ldots,\widetilde{p}_{\alpha_n})'$, where $p_{\alpha_i} = 1-\alpha_i$ and $\widetilde{p}_{\alpha_i} = T^{-1}\sum_{t=1}^{T} I_t^{\alpha_i}$. Then $\sqrt{T}(\widetilde{p}_\alpha - p_\alpha) \rightarrow_d \mathrm{N}(0,\Xi)$ where the elements of Ξ simplify to $\xi_{ij} = \min(p_{\alpha_i}, p_{\alpha_j}) - p_{\alpha_i} p_{\alpha_j}$ and, it follows that

$$\widetilde{J}_n = T(\widetilde{p}_\alpha - p_\alpha)'\Xi^{-1}(\widetilde{p}_\alpha - p_\alpha) \rightarrow_d \chi^2(n).$$

Note that to make these tests operational we replace the covariance terms by their sample counterparts. Furthermore, the asympotic normality results established above still hold under parameter uncertainty as shown in Chapter I. However, one needs to deal with nuisance parameters in the asymptotic covariance matrices to make the statistics operational. We use a parametric bootstrap procedure in Chapter I, which imposes all restrictions of the null hypothesis to estimate asymptotic covariance matrices under parameter uncertainty. Specifically, after the model is estimated, bootstrap samples are generated by using the estimated model as the data generating process where innovation vectors are drawn from the hypothesized parametric distribution. The Monte-Carlo simulations indicate that this approach provides satisfactory results. Hence, in this chapter we take the same approach.

2.2 Multivariate Contours and Autocontours

2.2.1 Multivariate Normal Distribution

In this case the density function is $f(\varepsilon_t) = (2\pi)^{-k/2} \exp(-0.5\varepsilon_t'\varepsilon_t)$. Let $\bar{f}_\alpha$ denote the value of the density such that the corresponding probability contour contains $\alpha\%$ of the observations. Then the equation describing this contour is

$$q_\alpha = \varepsilon_t'\varepsilon_t \equiv \varepsilon_{1t}^2 + \varepsilon_{2t}^2 + \cdots + \varepsilon_{kt}^2,$$

where $q_\alpha = -2\ln(\bar{f}_\alpha \times (2\pi)^{k/2})$. Hence, the C_α contour set is defined as follows

$$C_\alpha = \left\{ S(\varepsilon_t) \subset \Re^k \left| \int_{-g_1}^{g_1} \cdots \int_{-g_k}^{g_k} (2\pi)^{-k/2} \exp(-0.5\varepsilon_t'\varepsilon_t) \mathrm{d}\varepsilon_{1t} \ldots \mathrm{d}\varepsilon_{kt} \le \alpha \right. \right\},$$

where $g_1 = \sqrt{q_\lambda}$, $g_i = \sqrt{q_\lambda - \sum_{j=1}^{i-1} \varepsilon_{jt}^2}$ for $i = 2,\ldots,k$, and $\lambda \le \alpha$. We need to determine the mapping q_α in order to construct the indicator process. Let $x_t = \varepsilon_t'\varepsilon_t$, then $x_t \sim \chi^2(k)$ and we have $q_\alpha \equiv \inf\{q : F_{x_t}(q) \ge \alpha\}$, where F_{x_t} is the cumulative distribution function of a chi-squared random variable with *k* degrees of freedom. As a result, the indicator series is obtained as follows

$$I_t^\alpha = \begin{cases} 1 \text{ if } \varepsilon_t'\varepsilon_t > q_\alpha \\ 0 \text{ otherwise} \end{cases}.$$

To construct the autocontour $C_{l,\alpha}$, we consider the joint distribution of ε_t and ε_{t-l}. Let $\eta_t = (\varepsilon_t', \varepsilon_{t-l}')'$, then the density of interest is given by $\psi(\eta_t) = (2\pi)^{-k} \exp(-0.5\eta_t'\eta_t)$. Hence, the autocontour equation is given by

$$d_\alpha = \eta_t'\eta_t \equiv \eta_{1t}^2 + \cdots + \eta_{2k,t}^2,$$

where $d_\alpha = -2\ln(\overline{\psi}_\alpha \times (2\pi)^k)$. Following the same arguments as above, the corresponding indicator process is

$$I_t^{l,\alpha} = \begin{cases} 1 \text{ if } \eta_t'\eta_t > d_\alpha \\ 0 \text{ otherwise} \end{cases},$$

where $d_\alpha \equiv \inf\{d : F_{x_t}(d) \geq \alpha\}$, $x_t = \eta_t'\eta_t$, and F_{x_t} is the cumulative distribution function of a chi-squared random variable with $2k$ degrees of freedom.

2.2.2 Multivariate Student-t Distribution

The multivariate density function is

$$f(\varepsilon_t) = G(k,v)\left[1 + \varepsilon_t'\varepsilon_t/(v-2)\right]^{-(k+v)/2},$$

where $G(k,v) = \Gamma[(v+k)/2]/\{[\pi(v-2)]^{0.5k}\Gamma(v/2)\}$. Then the equation for the α-probability contour is

$$q_\alpha = 1 + \varepsilon_t'\varepsilon_t/(v-2),$$

where $q_\alpha = [\bar{f}_\alpha / G(k,v)]^{(k+v)/2}$. As a result, the C_α contour set is defined as

$$C_\alpha = \left\{ S(\varepsilon_t) \subset \Re^k \left| \int_{-g_1}^{g_1} \cdots \int_{-g_k}^{g_k} G(k,v)(1 + \varepsilon_t'\varepsilon_t/(v-2))\mathrm{d}\varepsilon_{1t} \ldots \mathrm{d}\varepsilon_{kt} \leq \alpha \right. \right\},$$

where $g_1 = \sqrt{(q_\lambda - 1)(v-2)}$, $g_i = \sqrt{(q_\lambda - 1)(v-2) - \sum_{j=1}^{i-1}\varepsilon_{jt}^2}$ for $i = 2,\ldots,k$, and $\lambda \leq \alpha$. Now let $x_t = 1 + \varepsilon_t'\varepsilon_t/(v-2)$, then $x_t \equiv 1 + (k/v)w_t$ where w_t has an F-distribution with (k,v) degrees of freedom. Consequently, we have $q_\alpha \equiv \inf\{q : F_{w_t}[v(q-1)/k] \geq \alpha\}$. Then the indicator series is defined as

$$I_t^{\alpha} = \begin{cases} 1 \text{ if } 1+\varepsilon_t'\varepsilon_t/(v-2) > q_{\alpha} \\ 0 \text{ otherwise} \end{cases}.$$

To construct the autocontour $C_{l,\alpha}$, we consider the joint distribution of ε_t and ε_{t-l} under the null hypothesis, which is

$$\psi(\varepsilon_t, \varepsilon_{t-l}) = G(k,v)^2 \left[(1+\varepsilon_t'\varepsilon_t/(v-2))(1+\varepsilon_{t-l}'\varepsilon_{t-l}/(v-2))\right]^{-(k+v)/2}.$$

Then, the equation for the α-probability autocontour is given by

$$d_{\alpha} = 1+(\varepsilon_t'\varepsilon_t + \varepsilon_{t-l}'\varepsilon_{t-l})/(v-2) + (\varepsilon_t'\varepsilon_t)(\varepsilon_{t-l}'\varepsilon_{t-l})/(v-2)^2.$$

Let $x_t = 1+(\varepsilon_t'\varepsilon_t + \varepsilon_{t-l}'\varepsilon_{t-l})/(v-2) + (\varepsilon_t'\varepsilon_t)(\varepsilon_{t-l}'\varepsilon_{t-l})/(v-2)^2$, then we have $x_t = 1+(k/v) \times \left[(w_{1t}+w_{2t}) + (k/v)(w_{1t}w_{2t})\right]$ where w_{1t} and w_{2t} are independent random variables with an F-distribution with (k,v) degrees of freedom. Similar to the previous case, we have $d_{\alpha} \equiv \inf\{d : F_{x_t}(d) \geq \alpha\}$, but we do not have readily available results for the quantiles of x_t as before. A plausible solution is using Monte-Carlo simulation to approximate the quantiles of interest as we already know that x_t is a specific function of two independent F-distributed random variables.

As an illustration, we provide sample contour and autocontour plots under normal and Student-t (with $v=5$) distributions in Figure-2.1. Due to the graphical constraints imposed by high dimensionality, we consider $k=2$ and $k=1$ for C_{α} and $C_{l,\alpha}$ respectively. Note that while C_{α} and $C_{l,\alpha}$ are of identical shape under normality, since the product of two independent normal densities yields a bivariate normal density, this is not the case under the Student-t distribution.

3 Monte-Carlo Simulations

We investigate the size and power properties of the proposed tests in finite samples by Monte Carlo simulations for two cases: when the parameters of the model are known and when they are unknown and need to be estimated.

3.1 Size Simulations

For the size experiments we consider two alternative distributions for the innovation process: a multivariate Normal, $\varepsilon_t \sim \text{i.i.d.}\,\text{N}(0,\text{I}_k)$, and a multivariate Student-t with 5 degrees of freedom, $\varepsilon_t \sim \text{i.i.d.}\,\text{t}(0,\text{I}_k,5)$. Under parameter uncertainty, we consider a simple multivariate location-scale model: $y_t = \mu + H^{1/2}\varepsilon_t$ where we set $\mu = 0$ and $H = I_k$. We consider both distributions under parameter uncertainty and apply the tests to the estimated standardized residual vector, $\hat{\varepsilon}_t = \hat{H}^{-1/2}(y_t - \hat{\mu})$, where we obtain $H^{1/2}$ by using the Cholesky decomposition[1]. The asymptotic variance of the tests is obtained by the simple parametric bootstrap procedure outlined above (see Section 2.1). The number of Monte Carlo replications is equal to 1,000, and the number of bootstrap replications is set to 500. We consider 13 autocontours ($n = 13$) with coverage levels (%): 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, and 99, spanning the entire density function[2].

[1] Alternative decompositions can be used to calculate the square-root matrix. We conjecture that the choice of the decomposition technique is not critical for application of our tests.

[2] Our choice of the contour coverage levels is motivated by the need of covering the entire range of the density, from the tails to the very center as we do not have a theoretical result indicating the optimal choice of the number of contours to guide our practice. The flexibility of our approach permits considering different types of coverage levels depending on the purpose of application, e.g. concentrating on tails for risk models. Note also that the Monte-Carlo results presented below provide guidance as to how far one can go in the tails and the center of the denisty without losing precision in finite samples. Additional Monte-Carlo simulations, not reported here to save space, also indicate that the size and power results are robust to

We start with a sample size of 250 and consider increments of 250 up to 2,000 observations. In all experiments, the nominal size is 5%.

In Tables 2.1a and 2.1b we present the simulated size results for the J_n^l-statistics. We consider a system of 2 equations ($k = 2$) and a system of 5 equations ($k = 5$). For a small sample of 250 observations, the J_n^l-statistics are oversized for both densities and both systems. However, under parameter uncertainty, the bootstrap procedure seems to correct to some extent the oversize behavior. For samples of 1000 and more observations, the simulated size is within an acceptable range of values. There are no major differences between the results for the small versus the large systems of equations indicating that the dimensionality of the system is not an issue for the implementation of these tests.

In Tables 2.2a and 2.2b we show the simulated size for the $\widetilde{J}_n$-statistics, which should be understood primarily as goodness-of-fit tests as they do not explicitly take into account the independence of the innovations over time. The sizes reported in Table 2.1a are very good, though those in Table 2.1b tend to be slightly larger than 5% mainly for small samples. However, when we consider the tests with individual contours (see Table 2.3), the size distortion tends to disappear.

For the *t*-tests, which are based on individual contours, the simulated sizes are very good. In Table 2.3, we report these results for the case of parameter uncertainty. The major size distortions occur for small samples at the extreme contour t_{13} (99% coverage),

---

the number of contours as long as the range considered is identical, i.e. a finer grid does not change the results.

but this is not very surprising since we do not expect enough variation in the indicator series for small samples.

3.2 Power simulations

We investigate the power of the tests by generating data from a system with two equations that follows three different stochastic processes. We maintain the null hypothesis as $y_t = \mu + H^{1/2}\varepsilon_t$, where $\varepsilon_t \sim \text{i.i.d. N}(0, I_k)$, and consider the following DGP's:

DGP 1: $y_t = \mu + H^{1/2}\varepsilon_t$, where $\varepsilon_t \sim \text{i.i.d.t}(0, \text{I}_2, 5)$, $\mu = 0$, and $H = \text{I}_2$. In this case, we maintain the independence hypothesis and analyze departures from the hypothesized density function by generating i.i.d. observations from a multivariate Student-t distribution with 5 degrees of freedom.

DGP 2: $y_t = Ay_{t-1} + H^{1/2}\varepsilon_t$, where $\varepsilon_t \sim \text{i.i.d. N}(0, \text{I}_2)$, $a_{11} = 0.7$, $a_{12} = 0.1$, $a_{21} = 0.03$, $a_{22} = 0.85$, and $H = \text{I}_2$. In this case, we maintain the same density function as that of the null hypothesis and analyze departures from the independence assumption by considering a linear VAR(1).

DGP 3: $y_t = H_t^{1/2}\varepsilon_t$, $\varepsilon_t \sim \text{i.i.d. N}(0, \text{I}_2)$, with $H_t = C + A'y_{t-1}y_{t-1}'A + G'H_{t-1}G$ and parameter values $A = 0.1^{1/2} \times \text{I}_2$, $G = 0.85^{1/2} \times \text{I}_2$, and $C = V - A'VA - G'VG$ where V is the unconditional covariance matrix with $v_{11} = v_{22} = 1$ and $v_{12} = 0.5$. In this case, we analyze departures from both independence and density functional form by generating data from a system with multivariate conditional heteroscedasticity.

In Table 2.4 we report the power of the J_n^l-statistic. The test is the most powerful to detect departures from density functional form (DGP 1) as the rejection rates are almost 100% even in small samples. For departures from independence, the test has better power to detect dependence in the conditional mean (DGP 2) than in the conditional variance (DGP 3). As expected, in the case of the VAR(1) model (DGP 2), the power decreases as *l* becomes larger indicating first order linear dependence. The power is also very good (69%) for small samples of 250 observations. In the case of the GARCH model (DGP 3), the rejection rate reaches 60% for sample sizes of 500 observations and above.

As expected, in Table 2.5 we observe that the goodness-of-fit test, $\widetilde{J}_n$, has the largest power for DGP 1 and it is not very powerful for DGP 2. It has reasonable power against DGP 3 mainly for samples of 1000 observations and above.

We find a similar message in Table 2.6 when we analyze the power of the *t*-statistics. The tests are the most powerful to detect DGP 1, the least powerful to detect DGP 2, and acceptable power against DGP 3 for samples of 1000 observations and above. There is a substantial drop in power for the t_{11} test (90% contour) for the cases of DGP 1 and DGP 3. This behavior is similar to that encountered in the univariate test introduced in Chapter I. This is a result due to the specific density under the null. In the case of DGP 1, for some contour coverage levels the normal density and the Student-t are very similar. Hence it is very difficult for any test to discriminate the null from the alternative with respect to the coverage level of those contour planes. A similar argument

applies to DGP 3 as well, since the GARCH structure in the conditional covariance matrix is associated with a non-normal unconditional density.

4 Empirical Applications

In this section we apply the proposed testing methodology to the generalized residuals of multivariate GARCH models fitted to U.S. stock return data. Our data set consists of daily excess returns on five size portfolios, i.e. portfolios sorted with respect to market capitalization in an increasing order.[3] The sample period runs from January 2, 1996 to December 29, 2006, providing a total of 2770 observations. A plot of the data is provided in Figure 2.2.

Since we are working with daily data we assume a constant conditional mean vector. In terms of the multivariate GARCH specifications, we consider two popular alternatives: the BEKK model of Engle and Kroner (1995) and the DCC model of Engle (2002). Define $u_t = y_t - \mu$ where μ is the constant conditional mean vector. Then the $\text{BEKK}(1,1,K)$ specification for the conditional covariance matrix, $H_t \equiv E[u_t u_t' \mid \Im_{t-1}]$, is given by

$$H_t = C'C + \sum_{j=1}^{K} A_j' u_t u_{t-1}' A_j + \sum_{j=1}^{K} G_j' H_{t-1} G_j \ . \tag{8}$$

In our applications we set $K = 1$ and use the scalar version of the model due to parsimony considerations where $A = \alpha I_k$, $A = \beta I_k$, and α and β are scalars. We also use

[3] Data is obtained from Kenneth French's website: http://mba.tuck.dartmouth.edu/pages/faculty/ken.french. We are grateful to him for making this data publicly available.

variance targeting to facilitate estimation, i.e. we set $C'C = V - A'VA - G'VG$ where $V = E[u_t u_t']$, e.g. Ding and Engle (2001).

In the DCC specification, conditional variances and conditional correlations are modeled separately. Specifically, consider the following decomposition of the conditional covariance matrix: $H_t = D_t R_t D_t$ where $D_t = diag\{h_{11,t}^{1/2}, \ldots, h_{kk,t}^{1/2}\}$, and each element of D_t is modeled as an individual GARCH process. In our applications, we consider the standard GARCH (1,1) process:

$$h_{ii,t} = \omega_i + \alpha_i u_{i,t-1}^2 + \beta_i h_{ii,t-1}, \quad j = 1, \ldots, k .$$

Now define $z_t = D_t^{-1} u_t$, then $R_t = diag\{Q_t\}^{-1} Q_t diag\{Q_t\}^{-1}$ where

$$Q_t = (1 - \alpha - \beta)\overline{Q} + \alpha u_t u_{t-1}' + \beta Q_{t-1}, \tag{9}$$

and $\overline{Q} = E[z_t z_{t-1}']$.

Under both BEKK and DCC specifications, we consider two alternative distributional assumptions that are most commonly used in empirical applications involving multivariate GARCH models: multivariate Normal and multivariate Student-t distributions. Under multivariate normality, the sample log-likelihood function, up to a constant, is given by

$$L_T(\theta) = -\frac{1}{2}\sum_{t=1}^{T} \ln[\det(H_t)] - \frac{1}{2}\sum_{t=1}^{T} u_t' H_t u_t . \tag{10}$$

In the case of the DCC model, a two-step estimation procedure is applicable under normality as one can write the total likelihood as the sum of two parts where the former depends on the individual GARCH parameters and the latter on the correlation

parameters. Under this estimation strategy, consistency is still guaranteed to hold. For further details on two-step estimation in the DCC model, the interested reader is referred to Engle (2002), and Engle and Sheppard (2001). Under the assumption of multivariate Student-t distribution, we do not need to estimate the model with the corresponding likelihood since the estimates obtained under normality are consistent due to quasi maximum likelihood interpretation. Therefore, we obtain the standardized residual vectors under normality and then simply test the Student-t assumption on these residuals.[4] One remaining issue in the case of Student-t distribution is the choice of the degrees of freedom. We follow Pesaran and Zaffaroni (2008) and obtain estimates of the degrees of freedom parameters for all series separately and then consider an average of the individual estimates for the distributional specification in the multivariate model.

The results are summarized in Figures 2.3 through 2.6 and Table 2.7. From the figures we observe that under both GARCH specifications, the J_n^l-statistics are highly statistically significant when multivariate normality is the maintained distributional assumption. The J_n^l-Statistics of the BEKK model are larger than those obtained under the DCC specification. Furthermore, there is an obvious pattern in the behavior of the statistics as a function of the lag order, especially under the BEKK specification. This indicates that the rejection is partly due to remaining dependence in the model residuals. When we switch to the multivariate Student-t distribution with 11 degrees of freedom,[5]

[4] Note that in the specification of the multivariate Student-t distribution (see Section 2), the covariance matrix is already scaled to be an identity matrix, thus no re-scaling of residuals is necessary to implement the test, e.g. Harvey, Ruiz and Santana (1992).

[5] This value is obtained by averaging individual degrees of freedom estimates obtained from individual GARCH models under Student-t density.

the J_n^l-statistics go down substantially under both multivariate GARCH specifications. Hence, we can argue that the distributional assumption plays a greater role in the rejection of both models under normality. The J_n^l-statistics are barely significant at 5% level for only a few lag values under the DCC specification coupled with multivariate Student-t distribution. However, under the BEKK specification, J_n^l-statistics are significant at early lags, even at 1% level. Table 2.7 reports individual *t*-statistics and the $\widetilde{J}_n$-statistics. Both types of test statistics indicate that normality is very strongly rejected under both GARCH specifications. Similar to the case of J_n^l-statistics, the results dramatically change when the distributional assumption is altered to multivariate Student-t. The DCC model produces better results with respect to both types of test statistics, but especially chi-squared test strongly supports the DCC specification compared to the BEKK model. Combining the information from all test statistics we can conclude that multivariate normality is a bad assumption to make regardless of the multivariate GARCH specification. Furthermore, the DCC model with multivariate Student-t distribution does a good job in terms of capturing dependence and producing a reasonable fit with respect to density functional form.

5 Conclusion

Motivated by the relative scarcity of tests for dynamic specification and density functional form in multivariate time series models, we proposed a new battery of tests based on the concept of "autocontour" introduced Chapter I for univariate processes. We

developed *t*-tests based on a single autocontour and also more powerful chi-squared tests based on multiple autocontours, which have standard asymptotic distributions. We also developed a second type of chi-squared test statistic, which is informative as a goodness-of-fit test when combined with the first type of chi-squared test. Monte-Carlo simulations indicate that the tests have good size and power against dynamic misspecification and deviations from the hypothesized density. We applied our methodology to multivariate GARCH models and showed that the DCC specification of Engle (2002) coupled with a multivariate Student-t distribution provides a fine model for multivariate time dependence in a relative large system of stock returns.

References

Bai, J. and Z. Chen (2008), Testing Multivariate Distributions in GARCH Models, *Journal of Econometrics*, 143, 19-36.

Bai J. (2003) Testing Parametric Conditional Distributions of Dynamic Models, *Review of Economics and Statistics*, 85, 532-549.

Bauwens, L., S. Laurent, and J.V.K. Rombouts (2006), Multivariate GARCH Models: A Survey, *Journal of Applied Econometrics*, 21, 79-109.

Ding, Z and Engle, R.F. (2001), Large Scale Conditional Covariance Modeling, Estimation and Testing, *Academia Economic Papers*, 29, 157-184.

Duchesne, P. and S. Lalancette (2003), On Testing for Multivariate ARCH Effects in Vector Time Series Models, *La Revue Canadienne de Statistique*, 31, 275–292.

Engle, R.F. (2002), Dynamic Conditional Correlation-A simple Class of Multivariate GARCH Models, *Journal of Business and Economic Statistics*, 20, 339-350.

Engle, R.F. and K. Sheppard (2001), Theoretical and Empirical Properties of Dynamic Conditional Correlation Multivariate GARCH, *Mimeo*, UC San Diego.

Engle, R.F. and F.K. Kroner (1995), Multivariate Simultaneous Generalized GARCH, *Econometric Theory*, 11, 122-150.

Harvey, A., Ruiz, E., and E. Sentana (1992), Unobservable Component Time Series Models with ARCH Disturbances, *Journal of Econometrics*, 52, 129-158.

Kalliovirta, L. (2007), Quantile Residuals for Multivariate Models, *Mimeo*, University of Helsinki.

Ling, S. and W.K. Li (1997), Diagnostic Checking of Nonlinear Multivariate Time Series with Multivariate ARCH Errors, *Journal of Time Series* Analysis, 18, 447–464.

Pesaran M.H. and P. Zaffaroni (2008), Model Averaging in Risk Management with an Application to Futures Markets, *CESifo Working Paper Series No*. 1358; *IEPR Working Paper No*. 04.3.

Tse, Y.K. (2002), Residual-based Diagnostics for Conditional Heteroscedasticity Models, *Econometrics Journal* 5, 358–373.

Tse, Y.K. and A.K.C. Tsui (1999), A Note on Diagnosing Multivariate Conditional Heteroscedasticity Models, *Journal of Time Series Analysis*, 20, 679–691.

Tables and Figures

Table 2.1a: Size of the J_n^l-statistics

T	J_{13}^1	J_{13}^2	J_{13}^3	J_{13}^4	J_{13}^5	J_{13}^1	J_{13}^2	J_{13}^3	J_{13}^4	J_{13}^5
	Panel a: Normal ($k=2$)					Panel b: Student-t ($k=2$)				
250	11.3	11.3	11.6	8.8	11.8	10.5	11.0	10.5	12.3	9.4
500	6.5	6.0	5.8	5.9	8.0	7.5	5.8	5.9	7.0	6.2
1000	6.8	5.0	6.2	5.3	4.9	7.2	5.2	5.1	5.4	6.0
2000	6.4	5.1	5.7	4.1	4.8	7.2	5.8	5.5	6.4	6.4
	Panel a: Normal ($k=5$)					Panel b: Student-t ($k=5$)				
250	12.7	11.8	11.5	14.0	12.9	10.4	11.7	12.3	10.3	11.6
500	9.2	8.4	6.9	7.6	8.3	7.3	6.6	7.3	7.9	8.1
1000	6.3	7.1	5.5	6.0	6.4	5.9	4.8	6.6	5.7	7.8
2000	5.3	5.6	5.3	3.4	6.5	6.9	4.8	5.7	5.5	5.4

Notes: Simulated size (%) of the J_n^l-statistics under multivariate standard normal distribution ($\varepsilon_t \sim \text{i.i.d.} N(0, I_k)$) and multivariate student-t distribution ($\varepsilon_t \sim \text{i.i.d.t}(0, \text{I}_k, 5)$**)**. Number of MC replications is 1,000 and nominal size is 5%.

Table 2.1b: Size of the J_n^l-statistics under Parameter Uncertainty

T	J_{13}^1	J_{13}^2	J_{13}^3	J_{13}^4	J_{13}^5	J_{13}^1	J_{13}^2	J_{13}^3	J_{13}^4	J_{13}^5
	Panel a: Normal ($k=2$)					Panel b: Student-t ($k=2$)				
250	8.1	6.1	7.3	7.5	6.9	6.8	6.4	7.8	6.5	6.0
500	7.5	5.9	5.8	7.3	7.4	7.5	6.7	8.3	8.0	8.1
1000	8.1	5.8	8.0	7.3	6.6	8.5	6.9	8.8	8.3	7.6
2000	5.7	5.4	7.7	6.4	4.8	6.2	7.6	7.6	6.4	7.0
	Panel a: Normal ($k=5$)					Panel b: Student-t ($k=5$)				
250	10.5	9.3	7.7	9.2	8.1	7.1	7.3	6.3	7.2	6.3
500	7.7	6.9	6.3	6.9	7.6	6.8	5.5	6.0	6.9	6.4
1000	5.9	6.1	7.1	5.5	5.5	6.4	5.7	6.8	7.5	6.6
2000	8.0	8.0	7.4	6.8	7.1	7.0	6.5	7.3	6.3	7.9

Notes: Simulated size (%) of the J_n^l-statistics when parameter estimation is involved. The model is $y_t = \mu + H^{1/2}\varepsilon_t$ and ε_t is either multivariate standard normal ($\varepsilon_t \sim \text{i.i.d.} N(0, I_k)$) or multivariate student-t ($\varepsilon_t \sim \text{i.i.d.t}(0, \text{I}_k, 5)$) distributed. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%.

Table 2.2a: Size of the $\widetilde{J}_n$-statistics (n = 13)

	Normal		Student-t	
T	$k = 2$	$k = 5$	$k = 2$	$k = 5$
250	5.7	6.3	4.3	6.6
500	4.9	5.3	3.1	5.1
1000	5.7	5.7	5.6	5.3
2000	5.6	6.2	4.9	5.6

Notes: Simulated size (%) of the $\widetilde{J}_n$-statistics under multivariate standard normal distribution ($\varepsilon_t \sim \text{i.i.d.} N(0, I_k)$) and multivariate student-t distribution ($\varepsilon_t \sim \text{i.i.d.t}(0, \text{I}_k, 5)$). Number of MC replications is 1,000 and nominal size is 5%.

Table 2.2b: Size of the $\widetilde{J}_n$-statistics (n = 13) under Parameter Uncertainty

	Normal		Student-t	
T	$k = 2$	$k = 5$	$k = 2$	$k = 5$
250	6.9	9.1	7.3	6.8
500	7.0	6.1	6.8	6.7
1000	6.7	5.5	6.7	5.6
2000	6.4	7.4	6.8	5.7

Notes: Simulated size (%) of the $\widetilde{J}_n$-statistics when parameter estimation is involved. The model is $y_t = \mu + H^{1/2}\varepsilon_t$ and ε_t is either multivariate standard normal ($\varepsilon_t \sim \text{i.i.d.} N(0, I_k)$) or multivariate student-t ($\varepsilon_t \sim \text{i.i.d.t}(0, \text{I}_k, 5)$) distributed. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%.

Table 2.3: Size of the *t*-statistics under Parameter Uncertainty

T	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Panel a: Normal ($k = 2$)													
250	5.0	4.6	5.2	5.1	6.5	6.7	5.7	4.9	5.2	4.6	6.0	4.8	2.0
500	4.3	4.2	5.3	5.4	4.1	4.6	4.5	5.1	5.3	5.2	5.1	4.7	6.4
1000	4.7	4.2	5.2	5.8	5.4	5.5	5.2	5.7	5.7	4.6	5.9	7.6	3.7
2000	5.4	3.9	5.1	4.0	5.0	5.3	5.3	6.2	4.8	5.9	4.3	6.4	4.9
Panel b: Normal ($k = 5$)													
250	4.5	6.2	5.3	5.0	4.5	5.2	5.3	5.8	5.5	5.1	6.1	6.7	2.1
500	4.1	4.8	5.8	4.8	6.0	5.6	5.3	6.4	6.5	4.3	6.3	6.0	6.3
1000	3.8	5.3	5.7	5.3	4.9	5.2	3.8	3.3	4.6	5.3	6.0	4.7	3.9
2000	4.5	5.3	5.0	5.0	4.6	4.1	5.4	6.0	4.6	5.5	5.5	4.4	6.5
Panel c: Student-t ($k = 2$)													
250	4.5	5.1	5.3	4.9	4.9	6.0	4.8	4.6	4.5	5.4	5.7	4.3	8.7
500	4.5	6.1	5.9	4.8	4.5	4.2	4.9	5.3	4.2	5.3	6.1	5.9	4.9
1000	4.3	5.9	6.4	5.8	5.7	5.5	6.6	6.4	5.9	5.8	5.5	6.0	6.3
2000	5.7	5.0	5.2	5.4	5.5	4.7	5.4	5.9	5.5	5.0	4.9	5.2	4.8
Panel d: Student-t ($k = 5$)													
250	4.5	5.5	4.8	4.6	5.8	6.0	7.6	6.7	7.0	6.6	5.8	4.1	8.4
500	4.6	5.4	6.4	4.9	4.9	6.6	5.8	7.1	7.7	6.5	5.4	5.0	5.9
1000	3.4	4.2	4.9	5.5	4.7	6.2	5.8	5.3	5.2	6.0	5.2	4.7	3.7
2000	5.1	5.6	5.3	5.2	5.2	5.0	5.3	4.4	5.3	6.1	5.0	5.1	3.8

Notes: Simulated size (%) of the *t*-statistics when parameter estimation is involved. The model is $y_t = \mu + H^{1/2}\varepsilon_t$ and ε_t is either multivariate standard normal ($\varepsilon_t \sim$ i.i.d.$N(0, I_k)$) or multivariate student-t ($\varepsilon_t \sim$ i.i.d.$\mathrm{t}(0, \mathrm{I}_k, 5)$) distributed. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%.

Table 2.4: Power of the J_n^l-statistics under Parameter Uncertainty

T	J_{13}^1	J_{13}^2	J_{13}^3	J_{13}^4	J_{13}^5
Panel a: DGP 1					
250	98.6	98.2	98.6	97.8	98.3
500	100.0	100.0	100.0	100.0	100.0
1000	100.0	100.0	100.0	100.0	100.0
2000	100.0	100.0	100.0	100.0	100.0
Panel b: DGP 2					
250	68.9	40.2	26.6	19.3	16.5
500	93.6	60.0	38.1	27.9	20.4
1000	99.9	84.8	58.0	39.2	28.9
2000	100.0	99.4	83.7	59.8	40.6
Panel c: DGP 3					
250	35.5	36.0	32.9	31.9	31.9
500	62.8	61.6	60.5	61.4	60.3
1000	90.5	88.8	88.1	86.9	86.7
2000	99.4	99.6	99.7	98.9	99.2

Table 2.5: Power of the $\widetilde{J}_n$-statistics (n = 13) under Parameter Uncertainty

T	DGP 1	DGP 2	DGP 3
250	99.1	12.4	19.7
500	100.0	12.1	44.5
1000	100.0	12.9	70.2
2000	100.0	14.2	94.7

Notes: Simulated power (%) of the J_n^l and $\widetilde{J}_n$ statistics when parameter estimation is involved. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%. See the text for a detailed description of the alternative DGPs.

Table 2.6: Power of the t-statistics under Parameter Uncertainty

T	t_1	t_2	t_3	t_4	t_5	t_6	t_7	t_8	t_9	t_{10}	t_{11}	t_{12}	t_{13}
Panel a: DGP 1													
250	23.1	55.3	76.6	91.8	96.1	97.7	98.0	96.6	89.9	59.6	8.5	33.7	85.2
500	32.3	80.6	95.3	99.5	100.0	100.0	100.0	100.0	99.4	85.6	8.6	57.8	98.5
1000	49.7	97.4	99.9	100.0	100.0	100.0	100.0	100.0	100.0	98.9	14.0	78.7	100.0
2000	75.4	99.9	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	16.2	94.9	100.0
Panel b: DGP 2													
250	3.3	4.7	8.4	11.2	11.1	12.4	13.4	11.0	7.3	6.7	9.7	11.6	3.5
500	3.6	5.6	7.6	11.5	12.8	11.5	11.8	11.0	8.9	7.0	7.2	10.9	13.1
1000	5.1	6.4	8.4	11.2	13.5	14.0	11.7	11.9	9.6	7.1	7.9	11.9	13.2
2000	4.4	6.7	9.2	10.8	13.3	15.3	14.6	11.6	9.5	8.7	8.7	12.3	14.0
Panel c: DGP 3													
250	5.6	7.2	10.7	12.8	15.3	17.6	18.5	18.7	14.6	8.3	6.3	9.0	17.0
500	7.2	11.9	17.7	25.5	33.4	38.3	41.5	41.1	32.6	15.6	5.3	20.0	48.0
1000	8.1	20.5	31.4	46.3	58.6	64.3	68.7	67.1	59.1	32.1	8.6	34.8	70.4
2000	13.5	35.3	56.8	77.7	86.7	91.5	92.8	91.8	85.4	54.7	9.5	60.0	93.5

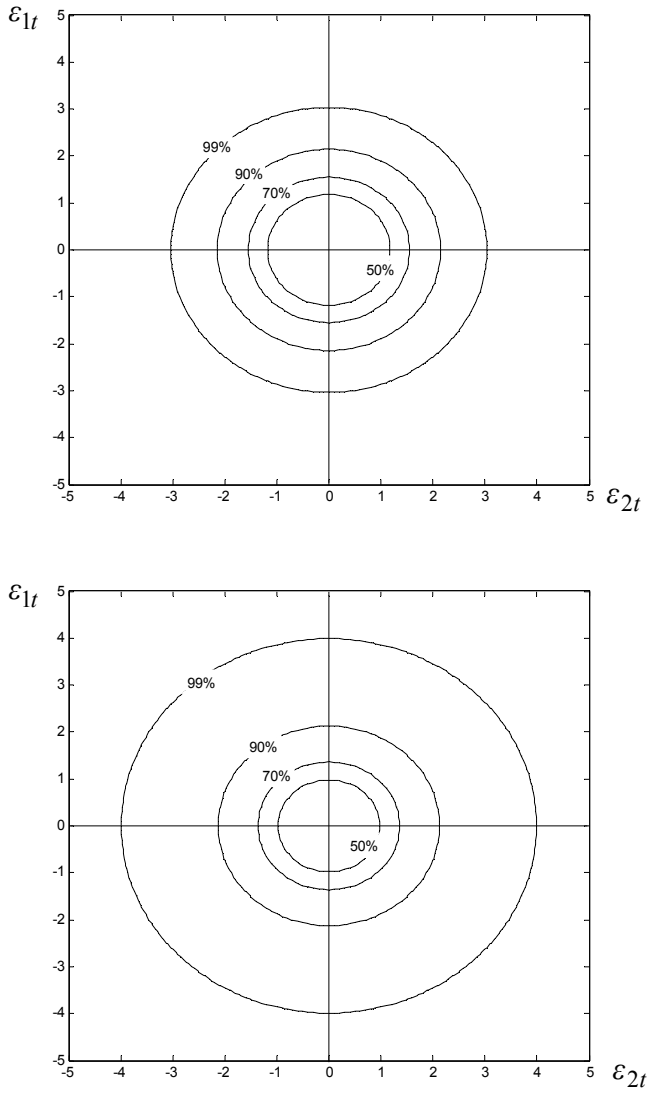
Notes: Simulated power (%) of the *t*-statistics when parameter estimation is involved. Number of MC replications is 1,000, number of bootstrap replications is 500, and nominal size is 5%. See the text for a detailed description of the alternative DGPs.

Table 2.7: Individual *t* and $\widetilde{J}_{13}$-statistics for Estimated GARCH Models

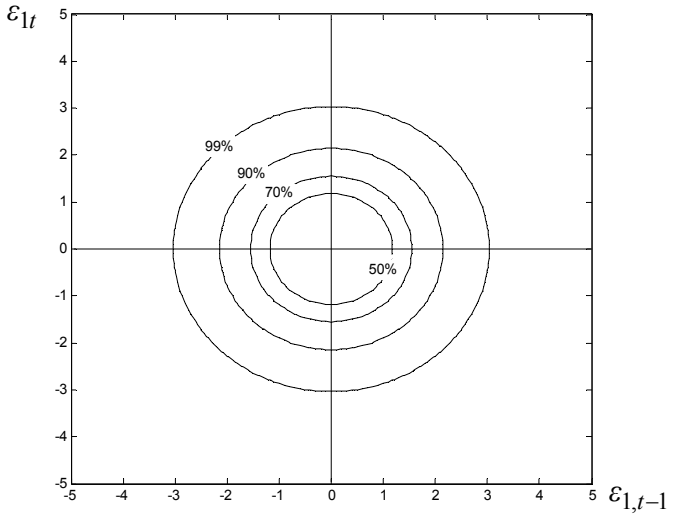
	BEKK Normal	DCC Normal	BEKK Student-t	DCC Student-t
t_1	-1.85	-2.17	2.78	2.30
t_2	-8.52	-10.18	-0.31	-0.38
t_3	-9.97	-12.26	1.00	-0.64
t_4	-9.37	-11.22	0.84	-0.10
t_5	-10.34	-11.81	2.47	0.18
t_6	-11.54	-10.95	1.13	0.95
t_7	-9.28	-10.03	0.09	0.50
t_8	-6.85	-7.19	0.25	0.59
t_9	-2.74	-5.70	0.92	-0.32
t_{10}	0.24	-1.52	0.66	-0.89
t_{11}	5.39	2.17	0.08	-3.51
t_{12}	8.23	5.58	1.00	-1.30
t_{13}	12.18	12.50	1.26	0.74
$\widetilde{J}_{13}$	351.47	388.54	30.07	24.35

Figure 2.1: Contour and Autocontour Plots under Normal and Student-t Distributions

Panel a: C_α under bivariate Normal and Student-t Distributions $\alpha \in \{0.5,0.7,0.9,0.99\}$



Panel b: $C_{l,\alpha}$ under bivariate Normal and Student-t Distributions $\alpha \in \{0.5,0.7,0.9,0.99\}$



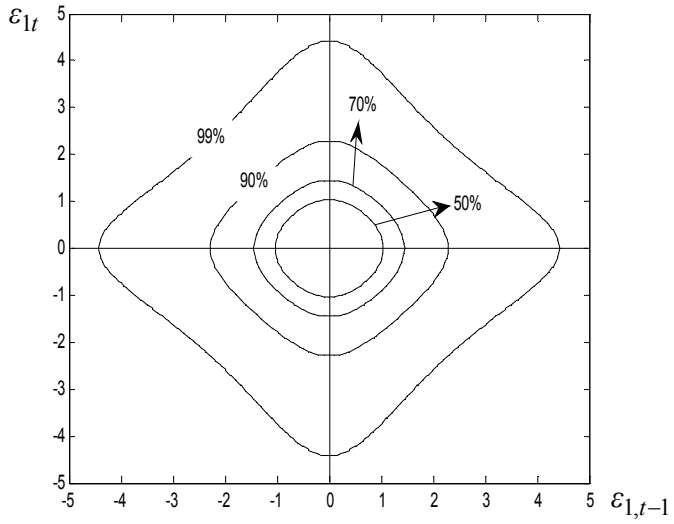


Figure 2.2: Daily Excess Returns on Five Size Portfolios (1/2/996-12/29/2006)
(From the smallest quintile portfolio to the largest quintile portfolio)

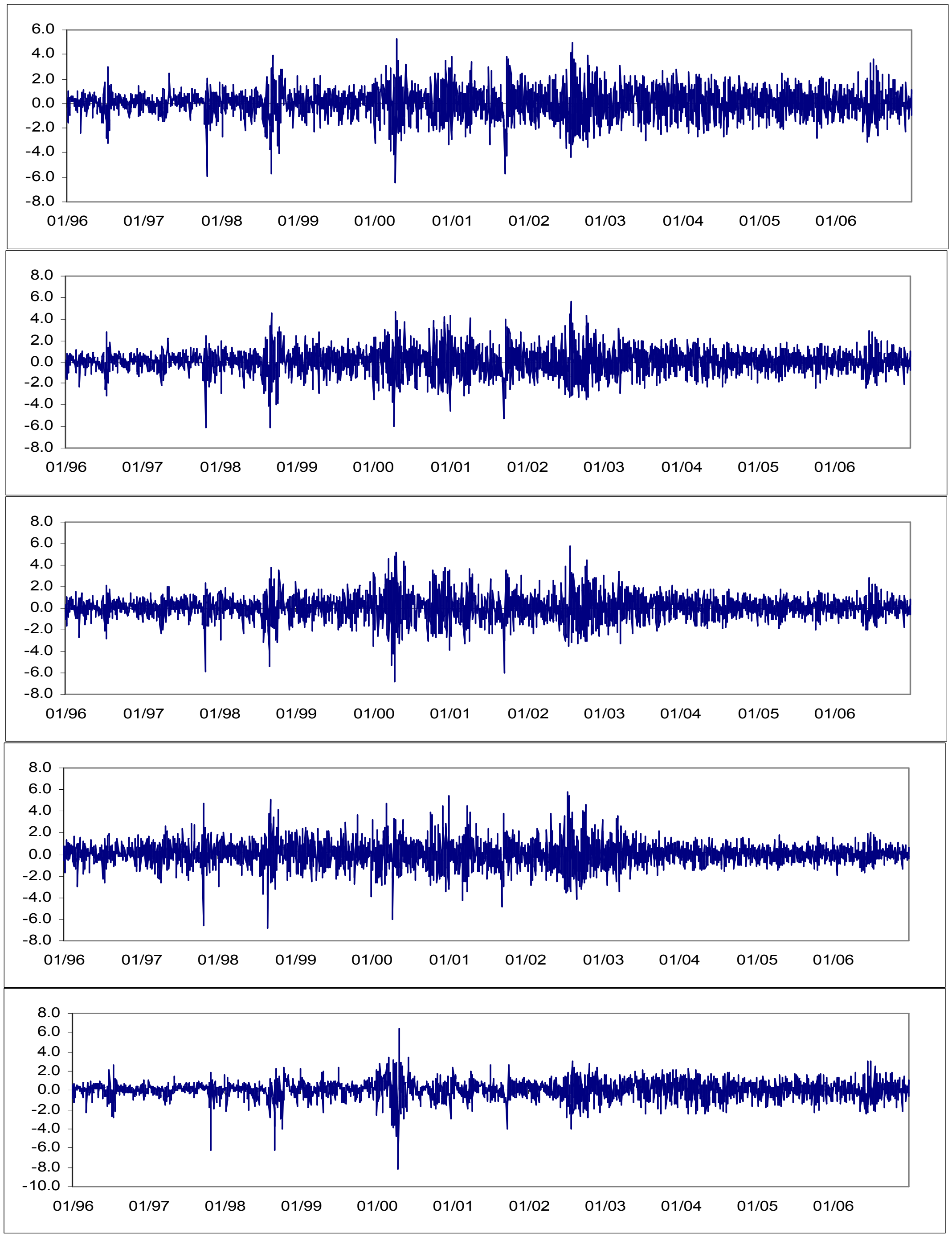


Figure 2.3: J_{13}^{l}-statistics of BEKK Model under Multivariate Normal Distribution

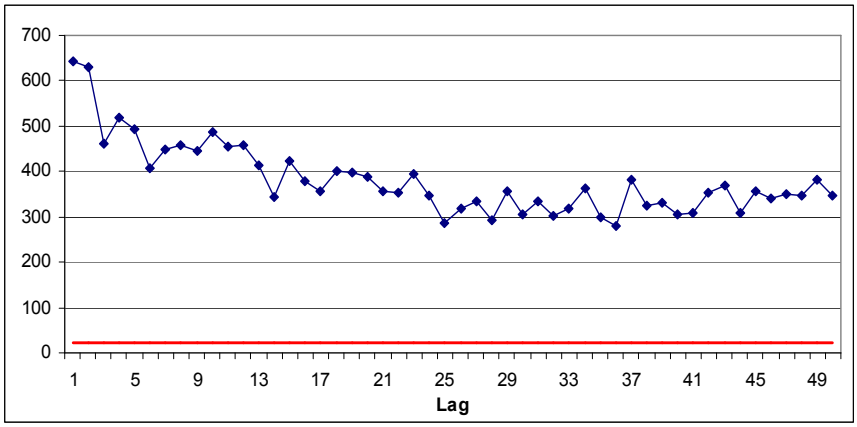


Figure 2.4: J_{13}^{l}-statistics of DCC Model under Multivariate Normal Distribution

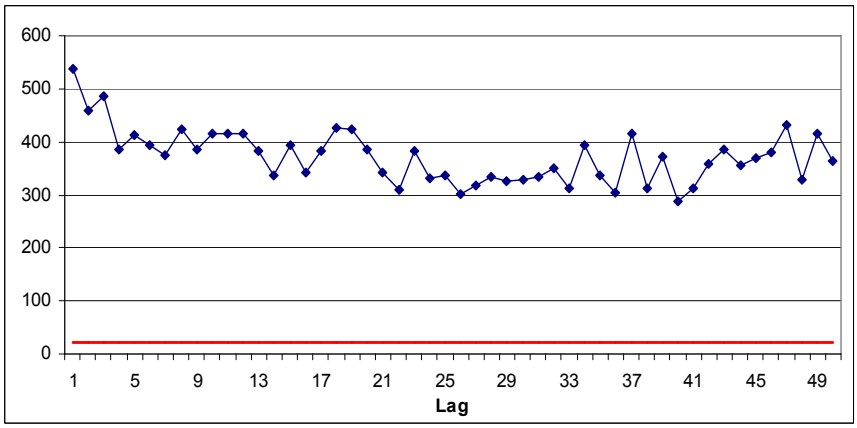


Figure 2.5: J_{13}^{l}-statistics of BEKK Model under Multivariate Student-t Distribution

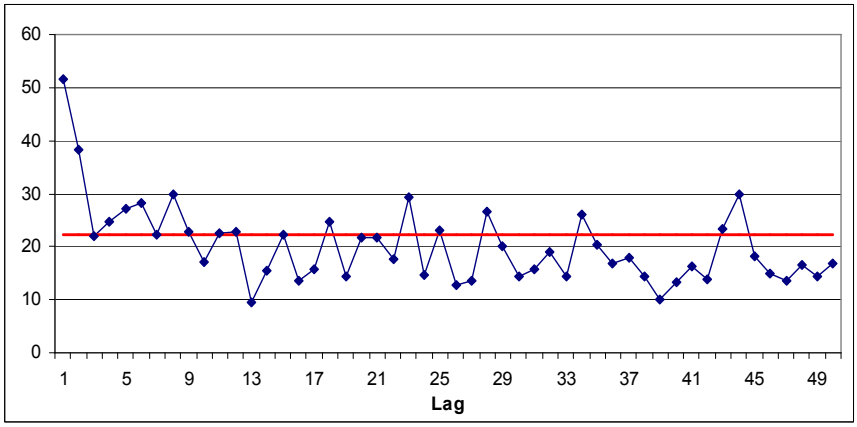
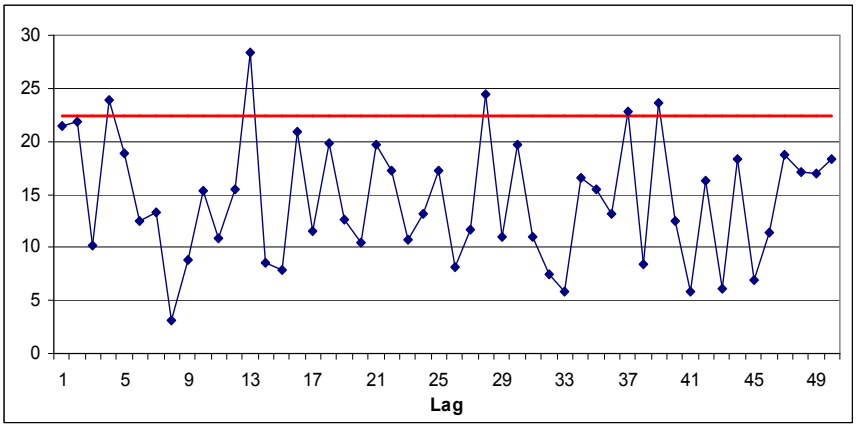


Figure 2.6: J_{13}^{l}-statistics of DCC Model under Multivariate Student-t Distribution



CHAPTER III

TESTING AND MODELING THRESHOLD ASYMMETRIES IN MULTIVARIATE DISTRIBUTIONS OF U.S. EQUITY RETURNS

1 Introduction

Modeling multivariate distributions of asset returns has crucial applications in asset pricing, portfolio allocation, and risk management. Extensive evidence in favor of various forms of asymmetries in stock return distributions has been documented in different contexts. It is well known that a broadly defined stock index has different expected return and volatility characteristics in bull versus bear markets, e.g. Maheu and McCurdy (2000). More recent empirical evidence shows that dependence between stocks also exhibit similar asymmetries. Longin and Solnik (2001) document significant asymmetric tail dependence between the U.S. stock market and international markets by using extreme value theory. Ang and Chen (2002) compare correlation asymmetry in the data with the pattern implied by a statistical model and find that U.S. equity portfolios have higher correlations with the market when returns are negative, and especially large in absolute value. Correlation asymmetries tend to be strong for small stocks, value stocks and past loser stocks. Hong et al. (2007) propose a nonparametric kernel based method to test for potential correlation and beta asymmetries in the data. According to their findings, correlations of small stocks with the market portfolio exhibit significant

asymmetries, but symmetric correlations cannot be rejected for book-to market and momentum portfolios. They also find a strong association between correlation and beta asymmetry.

The aforementioned asymmetries have important implications for financial decisions, so a realistic multivariate model of stock returns should account for them. Ang and Chen (2002) argue that a bivariate Markov-switching model performs well in terms of matching the magnitudes of asymmetric correlations and also modeling mean and volatility asymmetries.[1] On the other hand, Hong et al. (2007) use a multivariate mixture copula model to capture asymmetric dependence. Their in-sample analysis indicates that investors can obtain substantial utility gains by incorporating asymmetric dependence in their portfolio decisions. Patton (2004) adopts a copula approach to investigate the out-of-sample importance of asymmetric dependence in equity returns for portfolio decisions, using data on large and small cap stocks. He finds that modeling asymmetric dependence can be beneficial for unconstrained investors.

In this chapter we use a multivariate threshold approach to test and model asymmetries in expected returns, volatilities, correlations, and betas. The threshold approach offers a unified framework in which one can test for general regime switching dynamics with respect to observable state variables, construct a model that can incorporate the nonlinearities, and statistically assess significance of specific types of asymmetries in the data. We use monthly data on equity portfolios sorted on market capitalization, ratio of book value to market value, and industry classification.

[1] Perez-Quiros and Timmermann (2000), Tu (2007) and Guidolin and Timmermann (2008) also adopt the Markov-Switching approach to analyze various asymmetric features of equity return data.

A natural starting point is to test for the presence of threshold effects, which also guides us in terms of choosing the variable driving the regimes. We construct threshold variables based on the three stock market factors of Fama and French (1993) (market excess return, size premium, and value premium), monthly realized volatility of the market portfolio, term spread, growth in industrial production, and changes in the unemployment rate. We consider this comprehensive set of alternatives as we aim to capture regime changes with respect to systematic risk factors driving stock returns, which are also observable. Test results obtained in the arranged regression framework of Tsay (1998) indicate strong threshold effects for all three portfolio groups. The most evident rejections of linearity in favor of threshold models are obtained under the Fama-French factors, term spread, and growth in industrial production.

For model specification we refine the alternatives implied by test results with respect to Schwarz information criterion and also consider the behavior of the model sum of squares as a function of the threshold. This leads us to settle on two regime threshold models resembling the familiar bull versus bear taxonomy in the stock market. We find that market excess return is the preferred threshold variable for the size portfolios while value premium is selected for the book-to-market group. Both market excess return and term spread are suitable threshold variables for the industry portfolios. To make results comparable across the portfolio groups, we consider the model with market excess return as the threshold variable for the book-to-market group as well. Estimation results show that the model using market excess return provides a good fit for all three portfolio groups and produces results in line with previous findings obtained in different settings.

We find that small caps, value firms, and the Durables industry exhibit the strongest expected return asymmetries. Volatility asymmetry is a common feature of all stock portfolios. By focusing on the correlation with the market in each regime, we find that defensive industries, small caps, and value firms display the biggest differences across regimes. Correlations increase during downturns. Point estimates suggest relatively stronger beta asymmetry for medium sized firms, value firms, and defensive industries. In order to assess statistical significance of the above mentioned asymmetries, we use the subsampling method of Politis et al. (1999). Asymmetries in expected returns are significant for the majority of the portfolios with the exceptions of large caps and defensive industries. Volatility asymmetries are significant for all portfolio groups. In general, correlation asymmetries that are relatively big in magnitude also tend to be statistically significant. A similar observation applies to betas, but we also find that betas on average are subject to more uncertainty compared to correlations.

We evaluate out-of-sample predictive ability of the proposed threshold models relative to a linear benchmark. We consider two economic loss functions for this purpose: ex-post Sharpe ratio and realized utility of a quadratic utility investor over the forecast horizon. Our results indicate that substantial economic gains can be obtained by modeling asymmetries in the conditional distributions of returns in a threshold framework. With respect to Sharpe ratio, the threshold models are superior to the linear benchmark for all portfolio groups. These gains are also found to be statistically significant using the stationary bootstrap method. In terms of certainty equivalent utility gains, the threshold model outperforms the linear model by a substantial margin for the size portfolios, which

is also statistically significant. For book-to-market and industry groups, the threshold model performs better for highly risk averse investors but the differences are not statistically significant at conventional levels.

To our knowledge this study is the first to adopt a threshold approach to investigate various forms of asymmetries in multiple equity portfolios sorted on different characteristics. We statistically determine the drivers of regime switching behavior from a comprehensive set of variables and also provide insight on return-volatility dynamics. We go beyond the tests of general regime-switching and test significance of specific types of asymmetries across regimes, using subsampling methods. Our findings on correlation asymmetries are mostly in line with that of Ang and Chen (2002). However, we find asymmetries smaller in magnitude, which is possibly because they focus on correlations calculated conditional on contemporaneous realizations of returns and give more weight to tail behavior. In the vein of Hong et al. (2007), we also find a strong association between correlation asymmetry and beta asymmetry, but contrary to their results we do find significant asymmetries for book-to-market portfolios. Our results on portfolio allocation reinforce findings of Patton (2004) in terms of the potential benefits of modeling asymmetric dependence for unconstrained investors. However, we make use of a larger number of equity portfolios and find bigger certainty equivalent gains, which can be due to differences in model specification and choice of the loss function.

The rest of the chapter is organized as follows. Section 2 is devoted to methodological issues. A brief discussion comparing the threshold approach with the alternative modeling strategies is followed by a thorough presentation of the current

methodology. In Section 3 we provide a detailed description of the data and present the empirical results. Concluding remarks and directions for future research are provided in Section 4.

2 Methodology

2.1 Why Threshold Models?

To our knowledge, no previous study used the threshold approach to model equity returns in a multivariate framework. To motivate the threshold approach further, let us provide a brief comparison with alternative modeling strategies. Threshold models feature regime switching, which proved to be a successful modeling strategy for financial data. In this regard, the closest alternative is the Markov-switching type of models, in which regimes are driven by unobservable Markov-chains. Threshold models feature observable state variables, which allow the researcher to relate certain asymmetric features of the data with observable financial and economic factors. This makes interpretation of the results easier and also simplifies estimation and inference greatly compared to Markov-switching models.[2]

Another recently popular alternative is the Copula approach, which provides a very flexible framework for modeling asymmetric dependence. An important issue in copula based modeling is that multiple copula functions may provide similar fit to the data and

[2] Note also that it is standard practice to assume that the process driving the regimes is weakly exogenous with respect to the variables of interest in Markov-switching models. On the other hand, endogenous regime switching is straightforwardly incorporated in the threshold framework.

cannot be statistically distinguished from each other, e.g. Patton (2006). Furthermore, Copula functions are confined to the analysis of asymmetric dependence only.

High-dimensional data sets are relatively easily handled in the threshold framework while numerical optimization becomes very problematic for both Markov-switching and Copula based models as dimensionality increases. Hong et al (2007) and Tu (2007) deal with this problem by using Bayesian inference techniques, which are very computationally intensive, for copula and Markov-switching models respectively. A multivariate threshold model can be estimated in a matter of seconds without resorting to gradient-based numerical optimization or simulation methods.

2.2 The Model

The object of interest is a *k*-dimensional vector process, $y_t = (y_{1t}, \ldots, y_{kt})'$. The multivariate threshold autoregressive model is given by

$$y_t = \sum_{j=1}^{s} \Phi_j' X_t 1(\gamma_{j-1} < z_{t-d} \leq \gamma_j) + u_t \,, \tag{1}$$

where $\Phi_j = (c_j, \mathrm{A}_1^j, \ldots, \mathrm{A}_p^j)'$, $X_t = (1, y_{t-1}', \ldots, y_{t-p}')'$, z_t is the threshold variable, *d* is the delay lag, $\{\gamma_j\}_{j=0}^{s}$ are the thresholds such that $\gamma_0 = -\infty$ and $\gamma_s = \infty$, $1(.)$ is a standard indicator function, and $u_t = \sum_{j=1}^{s} \Omega_j^{1/2} \varepsilon_t 1(\gamma_{j-1} < z_{t-d} \leq \gamma_j)$ where $\{\varepsilon_t\}$ is a vector martingale difference sequence with zero mean and identity covariance matrix. Notice that there are *s* different regimes in which y_t follows a linear process, but the general dynamics of y_t over time is described by a nonlinear process.

Obviously, when $s = 1$ the threshold model boils down to a linear model, which is much simpler to deal with in terms of estimation and inference. Thus, it is of practical interest to test whether the threshold model is needed in the first place. A Wald or likelihood ratio type testing approach for the existence of threshold effects is complicated by the presence of unidentified nuisance parameters under the null. In particular, under the null hypothesis of linearity, i.e. $H_0 : \Phi_1 = \cdots = \Phi_s$, and $\Omega_1 = \cdots = \Omega_s$, the delay and the thresholds parameters are not identified. Consequently, the test statistics have nonstandard asymptotic distributions. Davies (1977) suggests using upper bounds for the critical values to deal with the problem, but Hansen (1996) argues that this is not a valid approach in the case of threshold models. He conducts Monte-Carlo simulations in the univariate threshold framework and shows that Davies' procedure behaves very conservatively. Hansen suggests using a bootstrap procedure to approximate the asymptotic distribution of the likelihood ratio test. Theoretical validity of his approach relies on local-to-null parameterization, i.e. the alternative converges to the null as sample size gets larger. Tsay (1998), on the other hand, proposes a different approach that is much less demanding in terms of computational burden and operates under standard asymptotic theory. These properties are especially important in our applications since we analyze high dimensional data sets with reasonably large number of observations. Hence, a detailed description of this approach is given below.

2.3 Threshold Nonlinearity Test based on Arranged Regression

Let us rewrite the model given in (1) under the null of linearity, i.e. when $s = 1$:

$$y_t = \Phi' X_t + u_t, \quad t = h+1, \ldots, n, \tag{2}$$

where $h = \max(p, d)$. If the null is true then the least squares estimates of the parameters of the model in (2) will be consistent under mild regularity conditions. However, under the alternative of threshold nonlinearity, least squares estimates based on (2) will be inconsistent and residuals will not be white noise. Tsay (1998) notes that if one arranges the ordering of the setup with respect to the threshold variable, the linear model is still useful under the alternative. In a given sample, the threshold variable takes values in $Z = \{z_{h+1-d}, \ldots, z_{n-d}\}$. Let $z_{(i)}$ and $t(i)$ denote the i^{th} smallest element of Z and its time index respectively. Then the arranged regression with respect to increasing order of the threshold variable is

$$y_{t(i)+d} = \Phi' X_{t(i)+d} + u_{t(i)+d}, \quad i = 1, \ldots, n-h. \tag{3}$$

Rearranging the model with respect to the threshold variable preserves the dynamics of the data while transforming the threshold problem into a change-point problem in terms of testing. Let $\hat{\Phi}_m$ denote the least squares estimate of Φ based on the first m observations and consider predictive residuals and their standardized versions based on equation (3)

$$\hat{u}_{t(m+1)+d} = y_{t(m+1)+d} - \hat{\Phi}'_m X_{t(m+1)+d},$$

$$\hat{\eta}_{t(m+1)+d} = \frac{\hat{u}_{t(m+1)+d}}{[1 + X'_{t(m+1)+d} V_m X_{t(m+1)+d}]^{1/2}},$$

where $V_m = \left[\sum_{i=1}^{m} X_{t(i)+d} X'_{t(i)+d}\right]^{-1}$. Now, consider the following regression

$$\hat{\eta}_{t(l)+d} = \Psi' X_{t(l)+d} + w_{t(l)+d}, \quad l = m_0 + 1, \ldots, n-h, \tag{4}$$

where m_0 denotes the starting point of the recursive least squares estimation.[3] Then testing for the threshold effect amounts to testing $H_0 : \Psi = 0$ against the alternative $H_1 : \Psi \neq 0$. To this end, Tsay proposes the following test statistic

$$C(d) = (n - h - m_0 - kp - 1) \times \{\ln[\det(S_0)] - \ln[\det(S_1)]\},$$

where $S_0 = \frac{1}{(n-h-m_0)} \sum_{l=m_0+1}^{n-h} \hat{\eta}_{t(l)+d} \hat{\eta}'_{t(l)+d}$, $S_1 = \frac{1}{(n-h-m_0)} \sum_{l=m_0+1}^{n-h} \hat{w}_{t(l)+d} \hat{w}'_{t(l)+d}$, and $\hat{w}_t$ is the residual of the regression given in equation (4). Under mild regularity conditions, $C(d)$ has an asymptotically chi-squared distribution with $k(kp+1)$ degrees of freedom. Conditional heteroskedasticity can be easily accommodated in this framework by modifying the way predictive residuals are standardized. Specifically,

$$\hat{\eta}_{j,t(m+1)+d} = \frac{\hat{u}_{j,t(m+1)+d}}{[\hat{\sigma}_j^2 + X'_{t(m+1)+d} V_m^* X_{t(m+1)+d}]^{1/2}},$$

where $\hat{\sigma}_j^2 = (m - kp - 1)^{-1} \sum_{i=1}^{m} \hat{u}_{j,t(i)+d}^2$, $V_m^* = V_m \left[\sum_{i=1}^{m} \hat{u}_{j,t(i)+d}^2 X_{t(i)+d} X'_{t(i)+d} \right] V_m$, and V_m is defined above. We use this version of the test statistic in our applications since time varying volatility is a renowned characteristic of stock returns, e.g. Andersen et al. (forthcoming).

[3] Choice of m_0 is critical since it determines the trade off between having good power on the one hand, and a reliable starting regression on the other. Tsay (1998) suggests $m_0 \approx 3n^{1/2}$ for stationary data and $m_0 \approx 5n^{1/2}$ for nonstationary data. We consider $m_0 = \lceil cn^{1/2} \rceil$ where $c \in \{4,5\}$ and $\lceil . \rceil$ is the ceiling function since monthly returns are subject to a high degree of uncertainty and a large number of observations are crucial for a reliable starting regression. The results are qualitatively similar for $c = 4$ and $c = 5$, so we report the results with $c = 5$ as it yields more observations for starting the regression.

2.4 Estimation

The estimation of the multivariate threshold model can be performed by conditional least squares (CLS). Given the threshold and the delay, the model reduces to *s* linear regressions for which the least squares estimation is straightforward. For ease of exposition let us focus on the case where there are only two regimes, i.e. $s = 2$. In this case the model is given by

$$y_t = \Phi_1' X_t 1(z_{t-d} \le \gamma_1) + \Phi_2' X_t 1(z_{t-d} > \gamma_1) + u_t .$$

Let $\widetilde{X}_t = (1(z_{t-d} \le \gamma_1) X_t', 1(z_{t-d} > \gamma_1) X_t')'$ and $\Theta = (\Phi_1', \Phi_2')'$, then the model can be written as $y_t = \Theta' \widetilde{X}_t + u_t$. Based on this compact form, the CLS estimates of the intercepts and the autoregressive coefficients are defined as follows

$$\hat{\Theta}(\gamma_1, d) = \left[\sum_{t=1}^{n-h} \widetilde{X}_t \widetilde{X}_t' \right]^{-1} \left[\sum_{t=1}^{n-h} \widetilde{X}_t y_t' \right]. \tag{5}$$

Let $\hat{u}_t = y_t - \hat{\Theta}'(\gamma_1, d) \widetilde{X}_t$, then the total sum of squares is given by $SSR(\gamma_1, d) = \mathrm{tr}\left(\sum_{t=1}^{n-h} \hat{u}_t \hat{u}_t' \right)$, where tr(.) denotes the trace operator. The CLS estimates of γ_1 and *d* are obtained from

$$(\hat{\gamma}_1, \hat{d}) = \arg\min_{\gamma_1, d} SSR(\gamma_1, d)$$

where $d \in [1, \overline{d}]$ and $\gamma_1 \in \Re_0$, $\Re_0 \subset \Re$, i.e. $\Re_0$ is a bounded subset of the real line. In practice, we consider $\gamma_1 \in \overline{Z}$, where $\overline{Z}$ is a trimmed version of the set *Z*, defined above. Following the usual practice in the literature we trim 15% from the top and bottom of *Z*,

e.g. Hansen (1996). The resultant least squares estimate of Θ is $\hat{\Theta}(\hat{\gamma}_1, \hat{d})$ and the covariance matrix estimates are defined analogously

$$\hat{\Omega}_j(\hat{\gamma}_1, \hat{d}) = \frac{1}{n_j - kp - 1} \sum_{t=1}^{n-h} I_{jt} \hat{u}_t \hat{u}_t', \quad j = 1, 2, \tag{6}$$

where n_j is the number of observations in regime j, $I_{1t} = 1(z_{t-d} \leq \gamma_1)$, and $I_{2t} = 1(z_{t-d} > \gamma_1)$. The following theorem establishes asymptotic properties of the CLS estimators.[4]

Tsay (1998) shows that under suitable regularity conditions the CLS estimators are consistent and the threshold estimate converges at rate *n* while the other model parameters are root-*n* consistent and asymptotically normal. This result provides a very practical first-order asymptotic approximation to the sampling distribution of the CLS estimates of the intercepts and autoregressive coefficients. One just needs to form consistent estimates of the associated asymptotic covariance matrices to make this approximation operational. However, the sampling error in the estimation of the threshold parameter is completely ignored in this setup. This poses an important problem in terms of inference since the point estimate of the threshold is likely to be different from its true value in finite samples despite consistency at rate *n*. Furthermore, Chan (1993) showed that the limiting distribution of the threshold estimate depends on several nuisance parameters, which renders inference with conventional methods unfeasible. Hansen (2000) uses local-to-null parameterization to reduce the rate of convergence, which in turn allows obtaining a nuisance parameter-free distribution for the threshold

[4] We provide a list of the regularity conditions and summarize the asymptotic results in Appendix A.

estimator. However, his assumptions are restrictive and the procedure does not readily translate into confidence intervals for other model parameters without resorting to some ad hoc rule. We deal with this problem by using the subsampling method of Politis et al. (1999) to construct asymptotically valid confidence intervals and test statistics. Hence, we account for the uncertainty in the threshold estimate when testing for the significance of the differences in other parameters across regimes. Gonzalo and Wolf (2005) take the same approach for the univariate self exciting threshold autoregressive models.[5]

3 Empirical Results

3.1 Data

We analyze a comprehensive data set consisting of equity portfolios classified with respect to different characteristics. We consider stocks sorted on market capitalization (size portfolios), ratio of book value to market value (book-to-market portfolios), and industry classification (industry portfolios). This allows us to analyze portfolio groups with different risk-return and dependence characteristics and also compare our results with those of the recent literature. Other data used in the analysis are market excess return, the size and value factors of Fama and French (1993), term spread (10-year T-Note yield minus 3-month T-Bill yield), growth rate of the industrial production index, and change in the unemployment rate.[6] All returns, except for size and value factors, are

[5] In Appendix B we describe subsampling inference for the threshold. Construction of asymptotic test statistics for other model parameters is implemented in an analogous fashion, so we present the arguments only for the threshold parameter to save space.

[6] Portfolio returns and the Fama-French factors are obtained from Kenneth French's online Data Library. All remaining data are obtained from the FRED database provided by the Federal Reserve Bank of St.

in excess of the one month T-Bill rate. We use monthly data and the sample runs from July 1963 to December 2007 providing a total of 534 observations. We hold back the last 60 observations (from January 2003 to December 2007) for out-of-sample analysis.

3.2 Alternatives for the Threshold Variable

As we are interested in systematic regime shifts in stock returns we consider functions of systematic stock market factors and variables reflecting economic fundamentals as potential threshold variables. In particular, we consider the following variables: (i) excess return on the value-weighted market portfolio (MKT); (ii) size premium from the Fama-French factor model, which is the difference between the returns on a small cap portfolio and a big cap portfolio (SMB); (iii) value premium from the Fama-French factor model, which is the difference between the returns on high versus low book-to-market portfolios (HML);[7] (iv) realized volatility of the market portfolio calculated from daily returns (MKTRV); (v) term spread implied by the Treasury yield curve, calculated as the difference between the yields on 10-year T-Notes and 3-month T-Bills (SPREAD); (vi) percentage change in the monthly industrial production index (IP); and (vii) change in the monthly unemployment rate (UR).

Now let us briefly comment on relevance of these alternatives. Using market excess return as the threshold variable can be motivated by the celebrated CAPM, which implies that the cross section of expected stock returns is determined by the return on the market

---

Louis. We are grateful to Kenneth French and the Federal Reserve Bank of St. Louis for making these data sets publicly available.

[7] Fama and French (1993) control for the ratio of book value to market value when calculating the size premium and vice versa. Hence, SMB and HML reflect the pure size and value effects respectively. See Fama and French (1993) for further details regarding the construction of these factors.

portfolio. Moreover, in the investment practice bull and bear markets are classified with respect to a broadly defined stock market index. Inclusion of size and value premiums is due to empirical success of the Fama and French (1993) factor model in explaining the cross-section of stock returns. Moreover, Fama and French (1993) argue that size and value premiums proxy for unobservable systematic risk factors reflecting economic fundamentals since small versus big firms differ in terms of access to credit and a high book-to-market ratio is associated with persistently low earnings on book equity. We consider realized volatility of the market portfolio since different phases of the stock market can potentially be identified with respect to a measure of variability that exploits information from a higher frequency.[8] We define monthly realized volatility, RV_t, as follows

$$RV_t = \left(\sum_{i=1}^{m_t} r_{i,t}^2 + 2 \frac{m_t}{m_t - 1} \sum_{i=1}^{m_t - 1} r_{i,t} r_{i+1,t} \right)^{1/2}, \quad t = 1, \ldots, n, \tag{7}$$

where m_t and $r_{i,t}$ denote the number of trading days and i^{th} daily return in month t, respectively. The second term is used to account for first order autocorrelation in the daily market return, e.g. Hansen and Lunde (2008).[9] The Treasury term spread proved to

[8] Under the assumptions that stock prices follow a special type of martingale and that there is no arbitrage, realized volatility consistently estimates the change in quadratic variation process over a certain horizon, which is the relevant measure of return variability [Andersen et al. (forthcoming)].

[9] Note that presence of long-memory in financial asset return volatility is a well established stylized fact, e.g. Andersen et al. (2003). Hence, the realized volatility measure violates the mixing assumption required for consistency of the CLS estimates in our threshold model. We deal with this problem by modeling realized volatility as a fractionally integrated process, following Andersen et al. (2001, 2003, 2006). This filters out the long memory component of the realized volatility. Specifically, we use $\xi_t = (1-L)^q \ln(RV_t)$, $q \in (0, 0.5)$ where L is the lag operator and q is the degree of fractional integration.

be successful in predicting the cycles in the aggregate economy, e.g. Estrella and Hardouvelis (1991). Finally, growth in the monthly industrial production index and change in the unemployment rate are solid indicators of the performance of the economy.

3.3 Threshold Nonlinearity Tests

Testing for the presence of threshold effects is the first step in our empirical exercise. While the test results indicate whether the threshold model is needed in the first place, they also provide an objective criterion to choose from competing definitions of the threshold variable. Intuitively, the test has the highest power when the threshold model is correctly specified. That is, a smaller p-value can be taken as an indication of better specification with respect to the threshold variable and the delay lag, e.g. Tsay (1989) and Teräsvirta (1994). Each major portfolio group contains ten portfolio sorts and the market portfolio. We include the market portfolio to explore the extent of the asymmetry in portfolio betas and correlations of the portfolios with the market. For selection of VAR order to implement the test, we consider Schwarz information criterion due to its consistency. Under the null of linearity, Schwarz criterion indicates $p = 0$ for all portfolio groups. This is not surprising given weak autocorrelations in monthly returns and difficulty of precisely estimating slope coefficients when dimensionality is high. Note also that Ang and Bekaert (2002) and Tu (2007), among others, also adopt regime-switching models with no autoregressive terms due to parsimony considerations.

Based on previous results in the literature we set $q = 0.42$. We test the null that $q = 0.42$ within the log-periodogram regression framework of Geweke and Porter-Hudak (1983) and obtain a p-value of 0.13. Moreover, the threshold non-linearity test results are not sensitive to the choice of the degree of fractional integration. We get qualitatively identical results in the empirically relevant range of [0.35, 0.45], which are available from the author upon request.

Besides the standard delay lag, z_{t-d}, we also consider first lag of average of the threshold variable over the past τ-months, i.e. $\widetilde{z}_{t-1}$, where $\widetilde{z}_t = (z_t + \cdots + z_{t-\tau+1})/\tau$.[10] Tables 3.1 and 3.2 report p-values of the threshold nonlinearity test statistics obtained from the arranged regression under the standard delay lag and the averaging scheme respectively. Table 3.1 indicates strong threshold effects shared by all portfolio groups under the standard delay lag. The null of linearity is rejected for all groups at 5% significance level under a variety of threshold variables and delay lag values. For the size portfolios the strongest rejections occur under MKT and SMB with a delay lag equal to one $(d = 1)$. HML, SPREAD and MKTRV also produce rejections, but with much larger p-values. For the book-to-market group, HML $(d = 8)$ and IP $(d = 2)$ yield the most significant rejections with p-values of similar magnitude, both smaller than 0.01. On the other hand, IP $(d = 5)$ produces the smallest p-value for the industry portfolios followed by SMB, MKT, and SPREAD. When we switch to the first lag of the τ-month average, set of variables under which the null is rejected is identical that for the size group, except for MKTRV. In addition, the null is rejected under a broad range of τ-values for MKT, SMB, and SPREAD. The results for the book-to-market group show that IP and HML still produce the smallest p-values, but MKT and SMB no longer yield rejections at 5% level. For the industry portfolios, IP produces very strong rejections for all values of τ followed by SMB $(\tau = 3)$. In general, size portfolios exhibit the strongest threshold asymmetries under both of the threshold schemes we consider.

[10] See Hansen (1997) and Tsay (1998) for similar approaches.

Interestingly, when market realized volatility is taken as the threshold variable, the null of linearity is either not rejected at all or rejections are much weaker compared to market excess return. According to our estimation results, discussed below, volatility significantly changes across regimes for most of the equity portfolios when market excess return is the variable driving regimes. This leads us to the conclusion that causality runs from returns to volatilities in the current framework. To be precise, when we consider returns and volatilities as distinct stochastic processes it is an innovation to the return process that triggers the regime shift and leads to higher/lower volatilities. In relation to the literature on the asymmetric relationship between volatility and stock returns, we can argue that a leverage type of explanation is empirically more relevant in the current context.[11]

3.4 Model Specification

The test results indicate strong threshold effects and point out to certain variables for each portfolio group. Since the results are obtained from the arranged regression, they are not informative about the number of regimes. To refine the competing alternatives further and get insight on the number of regimes, we resort to Schwarz information criterion calculated under threshold non-linearity, i.e. $SIC(p,d,s) = \sum_{j=1}^{s}[n_j \ln(\det(\hat{\Omega}_j)) + \ln(n_j)k(kp+1)]$.[12] Moreover, we consider plot of the SSR as a function of the threshold in a two-regime model as an informal criterion to determine the number of regimes, i.e.

[11] See Bollerslev et al. (2006) for a comprehensive list of references on the asymmetry between stock returns and volatility. They also provide an empirical investigation of the issue using intra-daily data and find that causality is from returns to volatility.

[12] This form of the Schwarz information criterion obtains under the assumption of mixtures of normal distributions, which proved to be very successful for financial data.

two minima distant from each other suggest a three-regime specification while a single global minimum is indicative of two-regimes. Finally, we check the sensitivity of the threshold estimate to trimming percentage by considering 10% and 7.5% besides the conventional 15% level, and abandon the specifications that are not robust.

For the size group, SMB and MKT with $d = 1$ stand out under the standard delay lag. We first consider SMB and observe that SSR as a function of the threshold suggests a three regime specification.[13] SIC goes down from 1,747.02 to 1,605.66 when we move from the two-regime specification to three regimes. However, the three-regime specification is sensitive to the level of trimming percentage. Hence, we next consider MKT with $d = 1$. The SSR plot is not very clear as to the number of regimes, but when we estimate the three-regime model we find that it is not robust to changes in the trimming percentage. While MKT (for $\tau = 2,3,4,5$) and SMB (for $\tau = 2,3$) yield strong rejections under averaging as well, the SIC values are larger than 1,711.12, which is obtained under MKT with $d = 1$ and $s = 2$. In addition SSR as a function of the threshold behaves very erratic for these alternatives. Thus, our specification for the size group is a two-regime model where first lag of MKT is the variable driving the regimes. For the book-to-market group we choose average HML with $\tau = 12$ as this yields the smallest p-value and also the smallest SIC among specifications robust to trimming percentage, with comparable p-values. Regarding the number of regimes we select $s = 2$ as SIC drops only 259.81, from 5,847.59 to 5,588.08, in the three-regime model and SSR plot also suggests $s = 2$. With regards to the industry group, IP yields the smallest p-

[13] Please note that we do not report the SSR graphs and all SIC values here to save space but they are all available from the author upon request.

value with $d = 5$ followed by $d = 1$, under the standard delay lag, but neither of these specifications are robust to the choice of the trimming level. IP produces very small p-values under averaging as well, but none of the specifications is robust to the trimming level. The next candidate is SMB with $d = 1$, but under this choice SSR behaves very erratic as a function of the threshold and does not suggest a particular value for *s*. When it comes to SMB with $\tau = 3$, SSR plot points out to a three-regime model, but the threshold estimates are sensitive to the trimming percentage. Therefore, we consider MKT with $d = 1$ and SPREAD with $d = 8$. MKT yields a smaller p-value (0.002165 versus 0.006455) but SPREAD produces a smaller SIC (10188.4 versus 10303.4), so we entertain both alternatives. In terms of the number of regimes, $s = 2$ is preferred in both cases based on the SSR plots and SIC values. As MKT with $d = 1$ turns out to be the best specification for the size group and is among the top two alternatives for the industry group, we consider this specification for the book-to-market group as well. This also facilitates comparisons across portfolio groups.

3.5 Estimation Results

Table 3.3 provides the threshold estimates and associated 95% subsampling confidence intervals for all three portfolio groups. For the size group, where the threshold variable is first lag of MKT, annualized point estimate of the threshold is -6.72%. The 95% confidence interval suggests that the threshold separating regimes can be as low as -9.94% or as high as -4.74%. Taking the uncertainty in the threshold estimate into account, approximately 37% of the observations in the estimation sample fall into the low return regime, 59% in the high return regime and 4% in the uncertain category, i.e. within

the confidence band. In this classification, sharp bear markets and short-lived downward corrections fall into the first regime while pro-longed bull markets and short and steep rallies are contained within the second one. With this point in mind we will simply refer to the first regime as the bear regime and the second one as the bull regime. When first lag of MKT is taken as the threshold variable the book-to-market and industry groups share the same point estimate for the threshold with the size group. However, the confidence band is slightly larger under the industry group, [-10.71, -3.31], and considerably larger under the book-to-market group, [-12.67 -1.76]. This is in line with the test results discussed above as the size group yielded the smallest p-values in general and under MKT with $d = 1$ in particular. On the other hand, the slight change in the confidence interval under the industry group does not give a regime classification different from the one obtained under the size portfolios while the book-to-market group implies only 2% and 1% shifts to the uncertain category from the bear and bull regimes respectively. When the twelve month average of HML is the threshold variable for the book-to-market portfolios, the annualized point estimate is 15.03% with a lower bound of 14.18% and an upper bound of 17.56%. The large positive value for the point estimate is due to the fact that the risk premium on relatively distressed firms (value firms in this case) increases substantially when times are bad.[14] Hence, low expected stock returns are generally associated with a large positive value premium. The confidence band implies that bear and bull regimes prevail 14% and 77% of the time respectively while 8% of the time we cannot classify the regime since the observations on the threshold variable fall

[14] Note that the unconditional correlation between MKT and HML is -0.42 and it is -0.51 between annual averages of them.

within the interval. For the industry group when the eight lag of SPREAD is the threshold variable, the point estimate is 0.43 (or 43 basis points) and the confidence interval is [0.27 0.5]. As the yield curve is usually inverted during economic downturns and gets flatter as the economy approaches the slowdown, plummeting of the spread signals slowdown in the economic activity and expected stock returns plunge. According to the results obtained under SPREAD, bear markets prevail 20% of the time, bull markets prevail 73% of the time, and we are uncertain about the state of the stock market 7% of the time. The commonality between the classifications with respect to the lags of SPREAD and the annual average of HML is that they imply a smaller fraction for the bear regime compared to MKT. This is probably because they are strongly linked to the underlying economic cycles, which go through downturns less frequently, but when they do this is accompanied with sharp drops in equity prices. In addition, both SPREAD and the annual average of HML are much less volatile compared to the monthly market excess return, which is also reflected by relatively tight confidence bands around the threshold estimates under these variables. This is another likely reason behind the more distinct regime classification obtained under SPREAD and HML.

Table 3.4 summarizes the estimation results for the ten size deciles and the market. The first decile exhibits the strongest expected return asymmetry. The annualized expected return difference is 41.61% and it is significant at literally any significance level. Difference in the expected returns across the regimes monotonically decreases as a function of market value and it is not statistically significant at conventional levels for the tenth decile. For the eight and ninth deciles the differences are significant only at 10%

level. Point estimate of the bull-bear spread for the market portfolio is 10.66% and it is statistically significant at 10% level with a p-value of 0.053. Firms with a large market capitalization are less affected by changing business and financial market conditions compared to the small firms, e.g. Perez-Quiros and Timmermann (2000). On the other hand, large firms have a significant impact on the market return despite their small number because the market portfolio is value weighted. These two facts explain the relatively weak expected return asymmetry in the large caps and the market portfolio compared to the small caps. For all size deciles and the market, volatility is substantially higher in the bear state implying a larger degree of uncertainty associated with considerably lower expected returns in this state. For example, the market portfolio has an annualized volatility of %19.39 in the bear regime whereas it is only %12.56 in the bull regime. The differences are highly statistically significant for all size deciles and the market. Regarding the correlations of the size deciles with the market portfolio, the point estimates indicate that they are higher in the bear regime, especially for small caps. For the first decile the difference is 0.128 while it is only 0.016 for the tenth decile. Interestingly, correlation asymmetry appears to be a monotonically decreasing function of market value, just like mean asymmetry. However, they differ in terms of statistical significance. For the first seven deciles, p-values associated with correlation asymmetry tend to be much larger compared to the ones associated with mean asymmetry, even though all correlation asymmetries are significant at 10% level. Similar to correlations, betas also tend to be higher in the bear regime, with the exception of the tenth decile. The largest difference in beta across the regimes is observed for the median sized firms (fifth

decile) with a point estimate of 0.092, which is significant at 10% level with a p-value of 0.086. Interestingly, asymmetries across regimes in the small caps' betas are of smaller magnitudes compared to their correlation counterparts and they are also insignificant. We do not observe a clear pattern for beta asymmetry as a function of market capitalization.

Tables 3.5 and 3.6 summarize the results for the book-to-market group when threshold variable is first lag of the annual average of HML and first lag of MKT respectively. When HML is the threshold, the first decile exhibits the greatest expected return asymmetry, with a highly significant annualized point estimate of 22.8%. The remaining deciles display lower levels of expected return asymmetry but there is no monotonic pattern. Differences across regimes are significant only at 10% level for deciles three to nine and for the tenth decile the point estimate of 8.52% is insignificant even at 10% level. The value firms, which are relatively distressed, carry a much greater premium relative to growth firms in the bear regime (17.80% versus 3.48%). Point estimates of volatility asymmetry are relatively small and mostly insignificant under this definition of the threshold. The largest asymmetry is found to be 3.12% for the first decile with a p-value of 0.021. Correlations tend to be higher in the bull regime but differences are usually small. The largest correlation asymmetry is 0.072 for the tenth decile, which is significant at 5% level with a p-value of 0.019. Betas are more asymmetric compared to correlations both in terms of magnitude and significance. The tenth decile's beta increases from 0.840 in the bear regime to 1.016 in the bull regime. This is in line with the fact that the premium on value firms is greater in the bear regime, i.e. their connection with the market gets stronger during upturns. When MKT is the

variable determining the regimes we get a somewhat different picture for the book-to-market group. The tenth decile has the greatest expected return asymmetry, with a highly significant point estimate of 20.40%. Interestingly, the value firms carry a bigger premium in the bull regime unlike the previous case, i.e. a higher expected return environment becomes more relevant in terms of distinguishing value versus growth firms. This is because when the value premium determines regimes, distress in the form of low earnings is reflected strongly by a large premium on the value firms. Regime dependent heteroskdeasticity is much better captured under MKT. Average of the absolute value of volatility asymmetry across all portfolios increases from 0.86% to 6.67% when we switch from HML to MKT and all volatility asymmetries become highly significant. Correlations also become more regime dependent and the largest differences are observed for value firms. For deciles eight to ten, the average correlation asymmetry is 0.09 and the differences are all significant at 1% level. Correlations become stronger during bear markets contrary to the case in which HML is the threshold variable. Finally, betas exhibit a lesser degree of asymmetry than the correlations and also their counterparts obtained under HML. In general, the model with MKT captures regime dependent heteroskedasticity much better and delivers implications on asymmetric correlations in line with previous findings in the literature.

In Tables 3.7 and 3.8 we present the results for the industry group when threshold variable is first lag of MKT and eighth lag of SPREAD respectively. When MKT drives the regimes, expected return differences across the regimes are not significant at conventional levels for NonDurables, Energy, Telecommunications, Healthcare, and

Utilities. The Utilities industry stands out as the most defensive industry for the investors with a higher expected return in the bear regime. Durables industry exhibits the strongest return asymmetry across bull and bear markets with a point estimate of 25.62%, which is highly significant. Durables industry is followed by Shops, High-Technology, Other, and Manufacturing categories in terms of the strength of expected return asymmetry. Volatility asymmetry is strong for all industry categories. The Other category experiences the biggest increase in volatility from the bull to the bear regime (14.96% to 22.38%). The strongest correlation asymmetries are associated with defensive industries. For example, point estimates suggest that the correlation of the Utilities industry with the market is 0.710 in the bear regime while it is only 0.412 in the bull state. This difference is significant at literally any significance level. Energy and Healthcare follow Utilities in terms of the strength of correlation asymmetry. The industry with the most cyclical expected returns, Durables, has a correlation asymmetry of 0.104, which is significant at 5% level. Point estimates suggest considerable Beta asymmetry but it is significant at 5% level only for Utilities and High-Technology industries. For Utilities (High-Technology) beta is 0.617 (1.386) in the bear (bull) regime while it is 0.402 (1.232) in the bull (bear) regime. When SPREAD is taken as the threshold variable, the general pattern of expected return asymmetry across the industry groups does not change, but the returns on Telecommunications and Utilities, especially the former, become much more cyclical. Energy becomes the most defensive industry with the lowest expected return asymmetry level of 6.53%, which is not significant even at 10% level. Another difference is that expected return asymmetries are on average bigger in magnitude, which is expected

because SPREAD implies less frequent but sharper low-return regimes. Similar to the case of the book-to-market group, regime dependent volatility becomes much weaker when we switch from MKT to SPREAD as the threshold variable. Volatility asymmetries tend to be insignificant for some of the industries that have significant expected return asymmetry, e.g. Durables. Correlation asymmetries continue to be relatively stronger for defensive industries with the exception of Health-care and p-values increase. Beta asymmetry tends to slightly increase for most of the industries. The model using MKT as the threshold variable performs better by capturing volatility clustering and also delivering more consistent results on correlations and expected returns.

3.6 Assessment of Predictive Ability of the Proposed Models

In this section we compare predictive ability of the proposed models with a benchmark linear model that ignores asymmetries. Evaluation is implemented in an out-of-sample setup. In particular, we use last 60 months of the sample (from January 2003 to December 2007) as the forecast evaluation period. One-step ahead predictions of the mean vectors and covariance matrices are obtained under alternative models in a recursive fashion. First forecasts are constructed by using data from July 1963 to December 2002, and then the estimation sample is increased by one observation at each step until the last forecasts for December 2007 are obtained. For comparing model performance, we consider ex-post Sharpe ratio and realized utility over the forecast horizon.

At each point in time, the investor decides how much to invest in the stocks versus the risk free asset. Minimum variance portfolio weights of stocks at time t are given by

$$\omega_t = \mu_p (\Sigma_{t+1|t}^{-1} \mu_{t+1|t}) / (\mu_{t+1|t}' \Sigma_{t+1|t}^{-1} \mu_{t+1|t}), \quad (8)$$

where μ_p is the target excess return for the portfolio, $\Sigma_{t+1|t}$ is the conditional covariance matrix, and $\mu_{t+1|t}$ is the conditional mean of the excess stock return vector at $t+1$ given information up to time *t*. Quadratic utility can be regarded as a second order approximation to the investor's true utility function, e.g. Fleming et al (2001). Then the realized utility is simply given by

$$U(W_{t+1}) = W_t R_{p,t+1} - \frac{aW_t^2}{2} R_{p,t+1}^2,$$

where W_{t+1} is the investor's wealth, $R_{p,t+1}$ is the gross portfolio return, i.e. $R_{p,t+1} \equiv (1 + r_{p,t+1})$, at $t+1$, and *a* is the coefficient of absolute risk aversion. Following Fleming et al. (2001) we consider total realized utility over the forecast horizon

$$\overline{U}(.) = W_0 \left[\sum_{t=n-T+1}^{n} \left\{ R_{p,t} - \frac{\varphi}{2(1+\varphi)} R_{p,t}^2 \right\} \right],$$

where *T* denotes the length of the forecast horizon and φ is the coefficient of relative risk aversion. The performance measure is the maximum performance fee that the investor is willing to pay to switch from the linear model to the threshold model. Let Δ denote this fee, which is defined by the following equation

$$\sum_{t=n-T+1}^{n} \left\{ (R_{p_2,t} - \Delta) - \frac{\varphi}{2(1+\varphi)} (R_{p_2,t} - \Delta)^2 \right\} = \sum_{t=n-T+1}^{n} \left\{ R_{p_1,t} - \frac{\varphi}{2(1+\varphi)} R_{p_1,t}^2 \right\}, \quad (9)$$

where $R_{p_1,t+1}$ and $R_{p_2,t+1}$ are the gross portfolio returns generated by the linear model and the threshold model respectively.

Table 3.9 summarizes the out-of-sample portfolio allocation results for annual target excess return levels of 3% and 6% and the relative risk aversion levels of 1 and 10. For the size group the threshold specification is superior to the linear benchmark with respect to both ex-post Sharpe ratio and annualized performance fees regardless of target excess return level. The difference between the Sharpe ratios is larger (1.57) for the lower target return level, but the performance fee increases with the target return. This is because certainty equivalent loss function less heavily penalizes the increased volatility resulting from a higher target return level compared to Sharpe ratio. For an investor with relative risk aversion coefficient of 10 the annualized performance fee that the investor is willing to pay is found to be 3.888% (or 388.8 basis points). On average, the linear model yields more allocation to stocks as a percentage of total wealth since it cannot capture the time varying nature of the expected returns, volatilities, and the dependence structure. This leaves the investor more vulnerable to unexpected shocks when the linear model is used for modeling the data. For book-to-market portfolios, the model using MKT as the threshold performs better than the one using HML. However, both threshold specifications beat the linear benchmark with respect to the ex-post Sharpe ratio. The model using HML is useful only to investors with high levels of risk aversion and the gains are small (18 basis points with a risk aversion level of 10). When MKT is the threshold variable the results improve substantially and performance fee can be as high as 1.548%. Similar observations apply to the industry portfolios but the model using SPREAD cannot beat the linear benchmark under any risk aversion and target return level in terms of the performance fee. The model with MKT as the threshold variable

does well for highly risk averse investors but cannot do the same at low risk aversion levels. In terms of the magnitudes of the loss differentials the size group stands out. This is expected given the stronger rejections of linearity and significance of mean and correlation asymmetries across regimes obtained with this data set. The general success of the threshold model using MKT can be attributed to the fact that it captures regime dependent heteroskedasticity and correlations better than the models using HML and SPREAD. To assess statistical significance of the loss differences we use the stationary bootstrap procedure of Politis and Romano (1994) and report the p-values for the null hypothesis that imposes equality of model performances against the alternative that the threshold model is superior. Performance fees are significantly positive under both target excess returns for the size group even at 1% level. Under the remaining two portfolio sorting schemes, the certainty equivalent gains (or losses) tend to be insignificant at conventional levels. Differences between Sharpe ratios in favor of the threshold specifications are always significant at 5% level for all three groups under both target return levels. Overall, the results of this out-of-sample comparison indicate that modeling asymmetries in a threshold framework can provide substantial economic gains that are also statistically significant.

4 Conclusion

Motivated by existing evidence on asymmetries in distributions of stock returns, we concentrate on the asymmetries in the first two moments of returns in a multivariate threshold framework. We use monthly data on size, book-to-market and industry

portfolios. Threshold nonlinearity test results and specification analysis indicate that excess return on the market portfolio, value premium, and term spread capture systematic regime shifts in the equity returns. The model using market excess return stands out as the universal threshold model providing a good fit regardless of the characteristics portfolios are sorted on. We find that small caps, value firms, and the Durables industry exhibit the strongest expected return asymmetries. Substantial asymmetry in volatility is found to be a common characteristic of all stock portfolios. Defensive industries, small firms, and value firms tend to have significantly higher correlations with the market during downturns. Correlation asymmetry usually translates into asymmetry in beta. In an out-of-sample forecasting exercise, we also find that there can be significant economic gains in incorporating asymmetries in the portfolio decisions.

A number of future research questions arise from the current study. As pointed out by Andersen et al. (forthcoming), variability in the mean component of stock returns is of a much smaller order of magnitude than the martingale component over shorter horizons. This makes modeling of covariance matrices even more critical at higher frequencies. An extension of the proposed threshold model to incorporate more sophisticated approaches to conditional second moments at higher frequencies is an interesting topic for future research. Assessing the out-of-sample economic value of modeling asymmetries under alternative preference specifications, such as loss aversion preferences of Kahneman and Tversky (1979), is another appealing direction to take since asymmetric dependence may be potentially more important for loss averse investors. Finally, investigating

performance of the proposed models with respect to multivariate density forecasts is another potential future research topic.

References

Andersen, T.G., T. Bollerslev, F.X. Diebold and H. Ebens (2001), The Distribution of Realized Stock Return Volatility, *Journal of Financial Economics*, 61, 43-76.

Andersen, T. G., T. Bollerslev, F.X. Diebold and P. Labys (2003), Modeling and Forecasting Realized Volatility, *Econometrica*, 71, 579-625.

Andersen, T. G., T. Bollerslev, F.X. Diebold and J. Wu (2006), Realized Beta: Persistence and Predictability, in T. Fomby and D. Terell (eds.), *Advances in Econometrics: Econometric Analysis of Economic and Financial Time Series in Honor of R.F. Engle and C.W.J. Granger*, Volume B, 1-40.

Andersen, T.G., T. Bollerslev and F.X. Diebold (forthcoming), Parametric and Nonparametric Volatility Measurement, in Y. Aït-Sahalia and L.P. Hansen (eds.), *Handbook of Financial Econometrics*, Amsterdam: North Holland.

Ang, A. and G. Bekaert (2002), International Asset Allocation with Regime Shifts, *Review of Financial Studies*, 1137-1187.

Ang, A. and J. Chen (2002), Asymmetric Correlations of Equity Portfolios, *Journal of Financial Economics*, 63, 443-494.

Bollerslev, T., J. Litvinova and G. Tauchen (2006), Leverage and Volatility Feedback Effects in High-Frequency Data, *Journal of Financial Econometrics*, 4, 353-384.

Chan, K. (1993) Consistency and Limiting Distribution of the Least Squares Estimator of a Threshold Autoregressive Model, *Annals of Statistics*, 21, 520-533.

Davies, R.B. (1977), Hypothesis Testing when a Nuisance Parameter is Present only under the Alternative, *Biometrika*, 64, 247-254.

Estrella, A. and G. Hardouvelis (1991), The Term Structure as a Predictor of Real Economic Activity, *Journal of Finance* 2, 555-576.

Fama, E.F. and K.R. French (1993), Common Risk Factors in the Returns on Stocks and Bonds, *Journal of Financial Economics*, 33, 3-56.

Fleming, J., C. Kirby and B. Ostdiek (2001), The Economic Value of Volatility Timing, *Journal of Finance*, 56, 329-352.

Geweke, J. and S. Porter-Hudak (1983), The Estimation and Application of Long-Memory Time Series Models, *Journal of Time Series Analysis* 4, 221-238.

Gonzalo, J. and M. Wolf (2005), Subsampling Inference in Threshold Autoregressive Models, *Journal of Econometrics*, 127, 201-224.

Guidolin, M. and A. Timmermann (2008), Size and Value Anomalies under Regime Shifts, *Journal of Financial Econometrics*, 6, 1-48.

Hansen, B.E. (1996), Inference when a Nuisance Parameter is not Identified under the Null Hypothesis, *Econometrica*, 64, 413-430.

_____ (1997), Inference in TAR Models, *Studies in Nonlinear Dynamics and Econometrics*, 2, 1-14.

_____ (2000), Sample Splitting and Threshold Estimation, *Econometrica*, 68, 575-603.

Hansen, P.R. and A. Lunde (2006), Realized Variance and Microstructure Noise, *Journal of Business and Economic Statistics*, 24, 127-161.

Hong, Y., J. Tu and G. Zhou (2007), Asymmetries in Stock Returns: Statistical Tests and Economic Evaluation, *Review of Financial Studies*, 20, 1547-1581.

Kahneman, D. and A. Tversky (1979), Prospect Theory: An Analysis of Decision under Risk, *Econometrica*, 47, 263-292.

Longin, F. and B. Solnik (2001), Extreme Correlations of International Equity Markets, *Journal of Finance*, 56, 649-676.

Maheu, J.M. and T.H. McCurdy (2000), Identifying Bull and Bear Markets in Stock Returns, *Journal of Business and Economic Statistics*, 18, 100-112.

Patton, A.J. (2004), On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation, *Journal of Financial Econometrics*, 2, 130-168.

_____ (2006), Modeling Asymmetric Exchange Rate Dependence, *International Economic Review*, 47, 527-556.

Peres-Quiros, G. and A. Timmermann, (2000), Firm Size and Cyclical Variation in Stock Returns, *Journal of Finance*, 55, 1229-1262.

Politis, D.N. and J.P. Romano (1994), The Stationary Bootstrap, *Journal of the American Statistical Association*, 89, 1303-1313.

Politis, D.N., J.P. Romano and M. Wolf (1999), *Subsampling*, New York, U.S.: Springer.

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

_____ (1998), Testing and Modeling Multivariate Threshold Models, *Journal of the American Statistical Association*, 93, 1188-1202.

Tu, J. (2007), Analyzing Regime Switching in Stock Returns: An Investment Perspective, Working Paper (available at SSRN: http://ssrn.com/abstract=954049).

Appendix A: Consistency of the CLS Estimators

First, let us define the following

$$\mathrm{D}(\gamma) = E[X_t X_t' \mid z_{t-d} = \gamma],$$

$$\mathrm{D}_2(\gamma) = E[(X_t X_t')^2 \mid z_{t-d} = \gamma],$$

$$\mathrm{V}_i(\gamma) = E[X_t X_t' \varepsilon_{it}^2 \mid z_{t-d} = \gamma],$$

$$\mathrm{V}_{2,i}(\gamma) = E[(X_t X_t')^2 \varepsilon_{it}^4 \mid z_{t-d} = \gamma],$$

for $i = 1,\ldots,k$. Now, assume that

A1. $(X_t, z_{t-d}, \varepsilon_t)$ is strictly stationary and absolutely regular (β-mixing) with mixing coefficients satisfying $\beta_j = O(j^{-\delta})$ for some $\delta > 4$.

A2. $E[\varepsilon_t \mid \Im_{t-1}] = 0$ where $\Im_{t-1}$ is the σ field generated by $(X_{t-j+1}, z_{t-d-j+1}, \varepsilon_{t-j})$ for $j \geq 1$.

A3. $E[|\, y_{it} \,|^4] < \infty$ and $E[|\, \varepsilon_{it} \,|^4] < \infty$ for $i = 1,\ldots,k$.

A4. The density function, $f(\gamma)$, of z_{t-d} is positive on a bounded subset $\Re_0 \subset \Re$, and γ_1 is an interior point of $\Re_0$.

A5. $f(\gamma)$, $\mathrm{D}(\gamma)$, $\mathrm{D}_2(\gamma)$, $\mathrm{V}_i(\gamma)$, and $\mathrm{V}_{2,i}(\gamma)$ are all continuous at $\gamma = \gamma_1$.

A6. $\Delta = \Phi_1 - \Phi_2 \neq 0$.

A7. $\Delta_j' \mathrm{D}(\gamma_1) \Delta_j > 0$, $\Delta_j' \mathrm{V}_i(\gamma_1) \Delta_j > 0$ for $j = 1,\ldots,k$, where Δ_j is the j^{th} column of Δ.

A1 rules out long-memory processes. A2 guarantees correct model specification. A4 ensures that $n_j \to \infty$ as $n \to \infty$. Note that this implies $(n_j / n) \to \lambda_j \in (0,1)$. A6 rules out local to null parameterizations and is necessary for convergence of the threshold estimate at rate *n*. A7 implies that the hyperplane of the conditional expectation, $E[y_t \mid \Im_{t-1}]$ has a discontinuity at the threshold $z_{t-d} = \gamma_1$. The following theorem provides the asymptotic results for the CLS estimators under these assumptions.

Theorem A.1 [Tsay (1998)] *Under A1-A7* $\hat{d} \to d$, $\hat{\gamma} \to \gamma_1$, $\hat{\Theta} \to \Theta$, *and* $\hat{\Omega}_j \to \Omega_j$ *almost surely as* $n \to \infty$. *Furthermore,* $n(\hat{\gamma}_1 - \gamma_1) = O_p(1)$, $\sqrt{n}\,\mathrm{vec}(\hat{\Theta} - \Theta) \to_d N(0, \Gamma)$.

This theorem is a slightly different version of Theorem 3 of Tsay (1998). It is essentially a generalization of the results given in Chan (1993) and Hansen (2000) in the univariate framework.

Appendix B: Subsampling Inference for the Threshold

The basic idea of subsampling is to reconstruct the statistic of interest on smaller blocks (subsamples) of the observed sample $\{y_1, \ldots, y_n\}$ and use the resulting values to approximate the associated sampling distribution. Let *b* denote the block size, such that $1 < b < n$, and $\hat{\gamma}_j^{b,t}$ the threshold estimate on the block $\{y_t, \ldots, y_{t+b-1}\}$ for $t = 1, \ldots, n - b + 1$. Thus, with reference to our previous notation, we have $\hat{\gamma}_j^{n,1} = \hat{\gamma}_j$.

Define $J_n(x,P) = \text{Prob}_P(n\,|\,\hat{\gamma}_j - \gamma_j\,| \le x)$ where P is the probability law governing $\{y_t\}$.[15] The subsampling approximation to $J_n(x,P)$ is defined as follows:

$$L_{n,b}(x,P) = \frac{1}{n-b+1}\sum_{t=1}^{n-b+1} 1(n\,|\,\hat{\gamma}_j^{b,t} - \hat{\gamma}_j\,| \le x).$$

Let $c_{n,b}(1-\alpha)$ be the α quantile of $L_{n,b}(x,P)$, then the corresponding symmetric subsampling confidence interval is

$$I_{b,\alpha} = [\hat{\gamma}_j \pm n^{-1}c_{n,b}(1-\alpha)].$$

The idea behind subsampling is actually quite intuitive. Each block is a valid sample, so the exact distribution of $b\,|\,\hat{\gamma}_j^{b,t} - \gamma_j\,|$ is $J_b(P)$. If b increases at a suitable rate as n gets large then the empirical distribution of the values $b\,|\,\hat{\gamma}_j^{b,t} - \gamma_j\,|$ for $t = 1,\ldots,n-b+1$ provides a good approximation to $J_n(P)$. Finally, replacing γ_j by $\hat{\gamma}_j$ has an asymptotically negligible effect assuming that $b/n \to 0$. The following proposition formalizes these ideas in the present context.

Proposition B.1 *Assume that* $b \to \infty$ *and* $b/n \to 0$ *as* $n \to \infty$. *Then the confidence interval given above has asymptotic coverage probability of* $(1-\alpha)$.

Proof. From A1 given in Appendix A, $\{y_t\}$ is β-mixing, which implies strong mixing. Theorem 1 guarantees that $J_n(x,P)$ converges weakly to a continuous limiting distribution. Then the result follows from Corollary 3.2.1 of Politis et al. (1999). ■

[15] Here we consider symmetric intervals only. Equal tailed two-sided intervals and one-sided intervals can are treated along the same lines.

The only issue that remains is the choice of the block size. This problem closely resembles the issue of bandwidth selection in nonparametric analysis. The requirements for the block size given in Proposition 1 are satisfied by a wide range of alternatives, so we need to use some specific criterion to determine the block size in practice. In our applications we use the algorithm proposed by Politis et al. (1999), which minimizes confidence interval volatility as a function of the block size. The steps of the algorithm can be summarized as follows:

1. For $b \in \{\underline{b},\ldots,\overline{b}\}$ form subsampling confidence intervals for which endpoints are denoted by $I_{b,\alpha}^{low}$ and $I_{b,\alpha}^{up}$.

2. For $b \in \{\underline{b}+l,\ldots,\overline{b}-l\}$, compute a volatility index, VI_b, which is the sum of the standard deviations of $\{I_{b-l,\alpha}^{low},\ldots,I_{b+l,\alpha}^{low}\}$ and $\{I_{b-l,\alpha}^{up},\ldots,I_{b+l,\alpha}^{up}\}$.

3. Choose *b* associated with the smallest value of the volatility index.

Based on the simulation results, Politis et al. (1999) argue that the choice of *l* is not critical and suggest using $l \in \{2,3\}$. To satisfy the requirements on block size, Romano and Wolf (2001) set $\underline{b} = c_1 n^{\kappa}$ and $\overline{b} = c_2 n^{\kappa}$ where $0 < c_1 < c_2$, $0 < \kappa < 1$ and recommend using $c_1 \in [0.5,1]$, $c_2 \in [2,3]$ and $\kappa = 0.5$. In our applications we set $c_1 = 0.5$ and $c_2 = 3$, to consider the broadest possible range. Even though this algorithm does not have a theoretical optimality property its performance is found to be satisfactory in finite samples according to the simulation results of Politis et al. (1999) and Romano and Wolf (2001).

Table 3.1: Threshold Non-linearity Test Results under the Standard Delay Lag

$d =$	1	2	3	4	5	6	7	8	9	10	11	12
Panel a: Size Portfolios and the Market												
MKT	**2.E-07**	**0.0050**	0.2582	0.8766	0.0732	0.6408	0.1667	**0.0393**	0.4820	0.8917	0.6386	0.9344
SMB	**5.E-11**	**0.0069**	0.1112	0.1030	**0.0406**	0.8735	0.1506	0.6328	0.3803	0.7468	0.5310	**0.0006**
HML	0.0697	0.4570	0.1206	0.2874	0.2267	**0.0157**	0.7927	0.8939	0.5187	0.5715	0.6124	0.1293
MKTRV	0.3129	0.4968	0.5473	0.5956	0.2844	**0.0427**	0.2812	0.6874	0.6848	0.4192	0.3182	0.8139
SPREAD	0.0943	**0.0333**	**0.0219**	0.0547	**0.0440**	0.1264	0.0757	0.1674	0.1971	0.1595	0.3014	0.6478
IP	0.3140	0.7050	0.6257	0.3208	0.0989	0.7310	0.8955	0.7998	0.1425	0.0676	0.5127	0.1990
UR	0.2095	0.4386	0.2622	0.7769	0.5092	0.3280	0.3307	0.4514	0.5175	0.2925	0.4051	0.2768
Panel b: Book-to-Market Portfolios and the Market												
MKT	0.2979	0.7203	0.4076	0.5055	0.2098	0.8752	0.3740	**0.0254**	0.2875	**0.0339**	0.1485	0.9254
SMB	0.2493	0.6207	0.2282	0.1484	0.1548	**0.0141**	**0.0246**	0.1741	0.1752	0.9311	0.1041	0.1032
HML	0.2223	0.5286	0.6864	**0.0308**	**0.0256**	0.1951	0.2910	**0.0061**	**0.0186**	**0.0437**	0.0755	0.5389
MKTRV	0.6712	0.9216	0.9603	0.4971	0.6083	0.1422	0.4145	0.8031	0.4703	0.5625	0.2187	**0.0402**
SPREAD	0.3162	0.4172	0.3779	0.5296	0.6680	0.8136	0.2419	0.3268	0.1704	0.1184	0.1972	0.3082
IP	0.1108	**0.0090**	0.5097	**0.0333**	0.2191	0.1384	0.1884	0.0592	0.5754	0.0657	0.9750	0.9941
UR	0.9657	0.0619	0.4434	**0.0298**	0.9361	0.3188	0.5339	0.3874	0.3669	0.2580	0.7254	0.9497
Panel c: Industry Portfolios and the Market												
MKT	**0.0022**	0.9352	0.5684	0.0732	0.2843	0.9167	0.6059	**0.0270**	0.6153	**0.0483**	0.8766	0.6455
SMB	**0.0009**	0.1131	0.1263	0.8672	0.2230	**0.0322**	**0.0496**	**0.0286**	0.2420	0.0958	0.1250	0.0916
HML	0.7292	0.1845	0.4692	0.2126	0.8545	0.1961	0.1578	0.2313	0.7198	0.9792	0.1362	0.7086
MKTRV	0.2036	0.1412	0.1926	0.9173	**0.0411**	0.0754	0.3703	0.6304	0.4044	0.2657	0.4479	0.0773
SPREAD	**0.0260**	**0.0458**	0.1731	0.4721	0.1788	0.3567	**0.0302**	**0.0065**	**0.0405**	0.0554	**0.0142**	**0.0422**
IP	**0.0007**	**0.0091**	0.3271	**0.0139**	**2.E-06**	0.0541	0.1589	0.0627	0.3046	**0.0316**	**0.0034**	0.4480
UR	0.7481	0.2248	0.8834	0.0800	0.0789	0.1143	0.9835	0.5127	0.9326	0.5195	0.4050	0.7007

Notes: This table reports p-values of the threshold nonlinearity test based on the arranged regression of Tsay (1998). The threshold variable is subject to a standard delay lag, i.e. z_{t-d} determines the regimes, where d is the delay. Test statistics are asymptotically chi-squared distributed with 11 degrees of freedom. Threshold variables are given in the first column. MKT is the excess return on the value weighted market portfolio, SMB and HML are size and value factors of Fama and French (1993), MKTRV is the realized volatility of the market portfolio, SPREAD is the yield on the 10-year T-Bond minus the 3-month T-Bill rate, IP is the growth rate of monthly industrial production index, and UR is the change in monthly unemployment rate.

Table 3.2: Threshold Non-linearity Test Results under the Averaging Scheme

$\tau =$	2	3	4	5	6	7	8	9	10	11	12
Panel a: Size Portfolios and the Market											
MKT	**2.E-06**	**1.E-07**	**2.E-06**	**4.E-06**	**0.0006**	**0.0002**	**0.0021**	**0.0005**	**0.0034**	**0.0124**	**0.0046**
SMB	**6.E-09**	**2.E-06**	**2.E-05**	**0.0007**	**0.0218**	**0.0060**	**0.0020**	**0.0380**	**0.0254**	0.1020	**0.0177**
HML	0.5091	**0.0219**	**0.0158**	0.0780	0.0509	0.1649	0.1814	0.2022	0.4070	0.7739	0.3864
MKTRV	0.0839	0.9547	0.8812	0.6929	0.9236	0.8098	0.5735	0.8624	0.5818	0.8280	0.7166
SPREAD	**0.0176**	**0.0100**	**0.0160**	**0.0232**	**0.0435**	**0.0420**	**0.0169**	**0.0246**	**0.0245**	0.0725	0.0660
IP	0.1814	0.2023	0.1349	0.1877	0.5640	0.4945	0.6855	0.7060	0.5351	0.5671	0.5616
UR	0.7249	0.6710	0.9631	0.9589	0.8917	0.9633	0.9519	0.9750	0.8908	0.7336	0.7105
Panel b: Book-to-Market Portfolios and the Market											
MKT	0.6225	0.2074	0.2138	0.3848	0.4785	0.3522	0.6369	0.3299	0.2521	0.3663	0.2019
SMB	0.6880	0.3232	0.6979	0.2421	0.5017	0.5652	0.5021	0.4457	0.5498	0.4874	0.1583
HML	0.2588	0.4100	0.8132	0.2648	0.0739	**0.0297**	**0.0142**	**0.0254**	**0.0128**	**0.0013**	**0.0002**
MKTRV	0.8790	0.9103	0.7987	0.5380	0.7818	0.4374	0.5838	0.6459	0.5923	0.7731	0.7914
SPREAD	0.3705	0.2960	0.3880	0.3505	0.3038	0.4843	0.4734	0.4897	0.4338	0.5604	0.5163
IP	**0.0208**	**0.0106**	**0.0137**	**0.0115**	**0.0048**	**0.0077**	0.1184	0.2015	0.0860	0.1690	0.2129
UR	0.3860	0.2098	0.0732	**0.0154**	0.0714	0.1558	0.4217	0.4476	0.4741	0.5751	0.5506
Panel c: Industry Portfolios and the Market											
MKT	0.4223	0.7762	0.8632	0.7668	0.5239	0.6464	0.9154	0.8514	0.2898	0.7301	0.4673
SMB	**0.0361**	**0.0031**	0.2501	**0.0100**	0.0804	0.2141	0.1233	0.2775	0.1338	0.0849	**0.0180**
HML	0.2116	0.4994	0.5865	0.9431	0.7664	0.2515	0.0665	0.1591	0.0572	0.0585	0.0727
MKTRV	0.2000	0.2591	0.3904	0.2084	0.2874	0.3480	0.4874	0.4854	0.7291	0.5853	0.5770
SPREAD	0.0618	0.0562	0.1309	0.1241	0.1007	0.1086	0.0927	0.0900	0.1085	0.0846	0.0573
IP	**0.0016**	**0.0085**	**0.0015**	**0.0012**	**0.0022**	**0.0009**	**0.0008**	**0.0015**	**0.0003**	**0.0001**	**0.0005**
UR	0.6806	0.3989	0.0791	**0.0242**	**0.0119**	0.1070	0.0930	0.2665	0.1437	0.0873	**0.0333**

Notes: This table reports p-values of the threshold nonlinearity test based on the arranged regression of Tsay (1998). The threshold variable is the τ-month average and the delay is set to be one, i.e. $\tilde{z}_{t-1}$ derives the regimes, where $\tilde{z}_t = (z_t + \cdots + z_{t-\tau+1})/\tau$. Test statistics are asymptotically chi-squared distributed with 11 degrees of freedom. Threshold variables are given in the first column. MKT is the excess return on the value weighted market portfolio, SMB and HML are size and value factors of Fama and French (1993), MKTRV is the realized volatility of the market portfolio, SPREAD is the yield on the 10-year T-Bond minus the 3-month T-Bill rate, IP is the growth rate of monthly industrial production index, and UR is the change in monthly unemployment rate.

Table 3.3: Threshold Estimates and Subsampling Confidence Intervals

	CI_L	$\hat{\gamma}_1$	CI_U
Size Group (MKT, $d = 1$)	-9.94	-6.72	-4.24
Book-to-Market Group (HML, $\tau = 12$)	14.18	15.03	17.56
Book-to-Market Group (MKT, $d = 1$)	-12.67	-6.72	-1.76
Industry Group (MKT, $d = 1$)	-10.71	-6.72	-3.31
Industry (SPREAD, $d = 8$)	0.27	0.43	0.50

Notes: This table reports the threshold estimates and 95% subsampling confidence intervals for all portfolio groups. The results under MKT and HML are annualized by multiplying the original results by 12 (% per annum). CI_L and CI_U denote the lower and upper bounds of the confidence interval respectively.

Table 3.4: Estimation Results for the Size Group (MKT, $d = 1$)

	Expected Returns			Volatilities			Correlations with the Market			Betas		
	Bear	Bull	Bull-Bear	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull
S1	-16.94	24.67	41.61 (0.000)	25.59	19.28	6.31 (0.013)	0.815	0.687	0.128 (0.051)	1.076	1.054	0.022 (0.430)
S2	-10.88	20.12	31.00 (0.000)	26.73	18.03	8.70 (0.000)	0.869	0.792	0.077 (0.095)	1.198	1.136	0.062 (0.295)
S3	-9.04	19.30	28.34 (0.000)	25.56	17.14	8.42 (0.000)	0.906	0.830	0.076 (0.038)	1.194	1.132	0.062 (0.275)
S4	-7.56	17.59	25.15 (0.000)	24.76	16.63	8.13 (0.000)	0.916	0.847	0.069 (0.038)	1.170	1.122	0.048 (0.255)
S5	-4.29	16.24	20.53 (0.000)	24.22	15.47	8.75 (0.000)	0.937	0.875	0.062 (0.051)	1.170	1.078	0.092 (0.086)
S6	-4.72	14.09	18.81 (0.000)	22.97	14.68	8.29 (0.000)	0.948	0.908	0.040 (0.064)	1.124	1.061	0.063 (0.151)
S7	-3.22	13.98	17.20 (0.000)	22.61	14.28	8.33 (0.000)	0.964	0.935	0.029 (0.044)	1.124	1.063	0.061 (0.089)
S8	0.25	10.90	10.65 (0.071)	22.23	13.89	8.34 (0.000)	0.970	0.948	0.022 (0.058)	1.112	1.048	0.064 (0.051)
S9	-0.09	9.45	9.54 (0.073)	20.28	12.69	7.59 (0.000)	0.981	0.961	0.020 (0.020)	1.027	0.971	0.056 (0.053)
S10	0.83	6.90	6.07 (0.142)	18.33	12.60	5.73 (0.002)	0.978	0.962	0.016 (0.064)	0.925	0.965	-0.040 (0.095)
MKT	-1.54	9.12	10.66 (0.053)	19.39	12.56	6.83 (0.000)						

Notes: This table reports estimation results for the two-regime model of 10 size portfolios (S1,...,S10) and the market portfolio (MKT), where the threshold variable is the first lag of MKT. Expected returns and volatilities are annualized (% per annum). The numbers in parentheses are subsampling p-values for testing the null of equality of parameters across regimes against one-sided alternatives.

Table 3.5: Estimation Results for the Book-to-Market Group (HML, $\tau = 12$)

	Expected Returns			Volatilities			Correlations with the Market			Betas		
	Bear	Bull	Bull-Bear	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull
BTM1	-14.63	8.21	22.84 (0.000)	21.11	17.99	3.12 (0.021)	0.937	0.925	0.012 (0.089)	1.173	1.078	0.095 (0.028)
BTM2	-7.90	7.90	15.80 (0.002)	17.85	16.69	1.16 (0.094)	0.954	0.957	-0.003 (0.418)	1.010	1.035	-0.025 (0.221)
BTM3	-3.78	7.50	11.28 (0.059)	18.01	16.63	1.38 (0.056)	0.944	0.953	-0.009 (0.230)	1.008	1.027	-0.019 (0.251)
BTM4	-3.33	7.69	11.02 (0.077)	16.23	16.78	-0.55 (0.479)	0.906	0.930	-0.024 (0.087)	0.872	1.012	-0.140 (0.000)
BTM5	-1.96	7.10	9.06 (0.099)	15.51	15.57	-0.06 (0.622)	0.881	0.912	-0.031 (0.066)	0.810	0.920	-0.110 (0.005)
BTM6	-1.77	9.01	10.78 (0.085)	15.43	15.61	-0.18 (0.552)	0.886	0.919	-0.033 (0.049)	0.811	0.929	-0.118 (0.002)
BTM7	-2.41	10.21	12.62 (0.075)	14.93	15.41	-0.48 (0.484)	0.850	0.871	-0.021 (0.157)	0.752	0.870	-0.118 (0.080)
BTM8	-2.19	10.31	12.50 (0.070)	15.59	15.13	0.46 (0.216)	0.849	0.868	-0.019 (0.120)	0.785	0.851	-0.066 (0.134)
BTM9	-1.07	11.03	12.10 (0.087)	16.80	16.33	0.47 (0.308)	0.817	0.867	-0.050 (0.012)	0.814	0.917	-0.103 (0.094)
BTM10	3.17	11.69	8.52 (0.275)	18.88	19.07	-0.19 (0.606)	0.750	0.822	-0.072 (0.019)	0.840	1.016	-0.176 (0.077)
MKT	-7.93	7.74	15.67 (0.005)	16.86	15.44	1.42 (0.047)						

Notes: This table reports estimation results for the model of 10 book-to-market portfolios (BTM1,...,BTM10) and the market portfolio (MKT), where the threshold variable is the first lag of the 12 month average of HML. Expected returns and volatilities are annualized (% per annum). The numbers in parentheses are subsampling p-values for testing the null of equality of parameters across regimes against one-sided alternatives.

Table 3.6: Estimation Results for the Book-to-Market Group (MKT, $d = 1$)

	Expected Returns			Volatilities			Correlations with the Market			Betas		
	Bear	Bull	Bull-Bear	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull
BTM1	-2.17	8.04	10.21 (0.087)	22.26	15.77	6.49 (0.000)	0.945	0.906	0.039 (0.066)	1.085	1.137	-0.052 (0.133)
BTM2	-1.47	9.31	10.78 (0.016)	20.33	14.09	6.24 (0.000)	0.965	0.943	0.022 (0.021)	1.012	1.058	-0.046 (0.141)
BTM3	-2.58	10.78	13.36 (0.002)	20.51	13.63	6.88 (0.000)	0.955	0.942	0.013 (0.016)	1.010	1.022	-0.012 (0.536)
BTM4	-0.44	9.71	10.15 (0.105)	20.50	13.28	7.22 (0.000)	0.936	0.903	0.033 (0.016)	0.990	0.955	0.035 (0.098)
BTM5	-0.93	9.70	10.63 (0.063)	19.14	12.35	6.79 (0.000)	0.926	0.866	0.060 (0.014)	0.915	0.852	0.063 (0.103)
BTM6	2.17	10.24	8.07 (0.073)	19.18	12.41	6.77 (0.000)	0.935	0.874	0.061 (0.019)	0.925	0.864	0.061 (0.080)
BTM7	2.28	11.69	9.41 (0.007)	19.03	12.09	6.94 (0.000)	0.896	0.820	0.076 (0.021)	0.879	0.789	0.090 (0.035)
BTM8	3.24	11.68	8.44 (0.103)	18.35	12.74	5.61 (0.000)	0.906	0.800	0.106 (0.000)	0.857	0.812	0.045 (0.152)
BTM9	1.65	13.55	11.90 (0.073)	20.06	13.19	6.87 (0.000)	0.892	0.799	0.093 (0.009)	0.923	0.839	0.084 (0.068)
BTM10	-1.96	18.44	20.40 (0.000)	22.48	15.76	6.72 (0.002)	0.835	0.763	0.072 (0.028)	0.968	0.957	0.011 (0.349)
MKT	-1.54	9.12	10.66 (0.019)	19.39	12.56	6.83 (0.000)						

Notes: This table reports estimation results for the model of 10 book-to-market portfolios (BTM1,...,BTM10) and the market portfolio (MKT), where the threshold variable is the first lag of MKT. Expected returns and volatilities are annualized (% per annum). The numbers in parentheses are subsampling p-values for testing the null of equality of parameters across regimes against one-sided alternatives.

Table 3.7: Estimation Results for the Industry Group (MKT, $d = 1$)

	Expected Returns			Volatilities			Correlations with the Market			Betas		
	Bear	Bull	Bull-Bear	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull
NoDur	3.39	10.32	6.93 (0.170)	19.25	12.96	6.29 (0.000)	0.870	0.776	0.094 (0.018)	0.864	0.800	0.064 (0.072)
Durbl	-11.28	14.34	25.62 (0.000)	22.65	17.23	5.42 (0.020)	0.853	0.749	0.104 (0.032)	0.997	1.027	-0.030 (0.538)
Mnf	-1.85	10.25	12.10 (0.018)	20.59	14.17	6.42 (0.014)	0.945	0.918	0.027 (0.045)	1.004	1.036	-0.032 (0.276)
Enrgy	5.17	7.15	1.98 (0.326)	21.72	15.34	6.38 (0.000)	0.752	0.561	0.191 (0.002)	0.843	0.686	0.157 (0.156)
HiTec	-4.07	12.25	16.32 (0.002)	27.28	21.15	6.13 (0.007)	0.876	0.823	0.053 (0.109)	1.232	1.386	-0.154 (0.014)
Telcm	2.82	5.44	2.62 (0.251)	18.64	14.57	4.07 (0.000)	0.749	0.659	0.090 (0.100)	0.720	0.765	-0.045 (0.337)
Shops	-3.32	13.15	16.47 (0.000)	22.44	16.19	6.25 (0.014)	0.877	0.838	0.039 (0.158)	1.015	1.080	-0.065 (0.310)
Hlth	6.39	8.69	2.30 (0.407)	21.30	15.50	5.80 (0.005)	0.832	0.683	0.149 (0.007)	0.914	0.843	0.071 (0.265)
Util	4.69	2.60	-2.09 (0.378)	16.85	12.27	4.58 (0.000)	0.710	0.412	0.298 (0.000)	0.617	0.402	0.215 (0.014)
Other	-3.54	12.70	16.24 (0.000)	22.38	14.96	7.42 (0.000)	0.945	0.891	0.054 (0.009)	1.091	1.061	0.030 (0.235)
MKT	-1.54	9.12	10.66 (0.032)	19.39	12.56	6.83 (0.000)						

Notes: This table reports estimation results for the model of 10 industry portfolios and the market portfolio (MKT), where the threshold variable is the first lag of MKT. Expected returns and volatilities are annualized (% per annum). The numbers in parentheses are subsampling p-values for testing the null of equality of parameters across regimes against one-sided alternatives.

Table 3.8: Estimation Results for the Industry Group (SPREAD, $d = 8$)

	Expected Returns			Volatilities			Correlations with the Market			Betas		
	Bear	Bull	Bull-Bear	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull	Bear	Bull	Bear-Bull
NoDur	-4.84	11.78	16.62 (0.049)	17.25	15.14	2.11 (0.072)	0.874	0.808	0.066 (0.063)	0.841	0.832	0.009 (0.510)
Durbl	-13.49	10.05	23.54 (0.009)	20.36	19.45	0.91 (0.253)	0.807	0.804	0.003 (0.548)	0.916	1.063	-0.147 (0.038)
Mnf	-7.82	9.75	17.57 (0.034)	18.99	16.29	2.70 (0.038)	0.956	0.922	0.034 (0.069)	1.012	1.022	-0.010 (0.450)
Enrgy	1.42	7.95	6.53 (0.186)	22.83	16.29	6.54 (0.000)	0.783	0.608	0.175 (0.020)	0.997	0.674	0.323 (0.020)
HiTec	-8.37	10.20	18.57 (0.058)	25.63	23.24	2.39 (0.089)	0.842	0.854	-0.012 (0.430)	1.204	1.351	-0.147 (0.076)
Telcm	-16.20	11.12	27.32 (0.000)	15.54	16.18	-0.64 (0.248)	0.665	0.718	-0.053 (0.157)	0.576	0.790	-0.214 (0.018)
Shops	-8.10	11.61	19.71 (0.049)	20.34	18.50	1.84 (0.132)	0.862	0.858	0.004 (0.423)	0.977	1.080	-0.103 (0.087)
Hlth	-1.21	10.64	11.85 (0.094)	20.33	17.12	3.21 (0.029)	0.769	0.761	0.008 (0.425)	0.872	0.887	-0.015 (0.416)
Util	-6.60	6.87	13.47 (0.034)	16.45	13.37	3.08 (0.002)	0.721	0.493	0.228 (0.002)	0.661	0.448	0.213 (0.027)
Other	-10.49	11.89	22.38 (0.009)	21.46	17.02	4.44 (0.016)	0.937	0.915	0.022 (0.161)	1.121	1.060	0.061 (0.130)
MKT	-8.79	9.41	18.20 (0.031)	17.93	14.70	3.23 (0.029)						

Notes: This table reports estimation results for the model of 10 industry portfolios and the market portfolio (MKT), where the threshold variable is the 8th lag of SPREAD. Expected returns and volatilities are annualized (% per annum). The numbers in parentheses are subsampling p-values for testing the null of equality of parameters across regimes against one-sided alternatives.

Table 3.9: Out of Sample Portfolio Allocation Results

μ_p	$\bar{r}_{p_1}$	$\bar{r}_{p_2}$	$\sigma_{r_{p_1}}$	$\sigma_{r_{p_2}}$	SR_1	SR_2	$SR_2 - SR_1$	Δ_1	Δ_{10}
Panel a: Size Portfolios and the Market (MKT, $d = 1$)									
3.00	2.94	4.02	3.35	1.64	0.88	2.44	1.57 (0.000)	1.116 (0.044)	1.512 (0.008)
6.00	2.99	5.15	6.72	3.25	0.45	1.59	1.14 (0.000)	2.328 (0.036)	3.888 (0.000)
Panel c: Book-to-Market Portfolios and the Market (HML, $\tau = 12$)									
3.00	4.52	4.50	2.09	1.85	2.16	2.44	0.28 (0.034)	-0.013 (0.540)	0.036 (0.422)
6.00	6.16	6.12	4.17	3.63	1.48	1.69	0.21 (0.072)	-0.014 (0.534)	0.180 (0.388)
Panel d: Book-to-Market Portfolios and the Market (MKT, $d = 1$)									
3.00	4.52	5.16	2.09	1.72	2.16	3.00	0.84 (0.048)	0.648 (0.180)	0.708 (0.172)
6.00	6.16	7.43	4.17	3.47	1.48	2.14	0.67 (0.058)	1.296 (0.184)	1.548 (0.168)
Panel f: Industry Portfolios and the Market (MKT, $d = 1$)									
3.00	5.91	5.30	3.75	2.05	1.58	2.59	1.01 (0.006)	-0.564 (0.206)	-0.096 (0.424)
6.00	8.93	7.72	7.41	4.03	1.21	1.91	0.71 (0.026)	-1.020 (0.226)	0.816 (0.298)
Panel f: Industry Portfolios and the Market (SPREAD, $d = 8$)									
3.00	5.91	4.99	3.75	2.33	1.58	2.14	0.56 (0.004)	-0.876 (0.072)	-0.468 (0.248)
6.00	8.93	7.10	7.41	4.59	1.21	1.55	0.34 (0.014)	-1.668 (0.088)	-0.048 (0.520)

Notes: This table summarizes the performance of the threshold models relative to a linear benchmark with respect to portfolio allocation. μ_p is the target excess return on the portfolio (% per annum), $\bar{r}_{p_1}$ and $\bar{r}_{p_2}$ denote the average realized returns (% per annum) on the portfolios formed from linear and threshold models respectively, $\sigma_{r_{p_1}}$ and $\sigma_{r_{p_2}}$ are the standard deviations of the portfolio returns, SR_1 and SR_2 are the ex-post Sharpe ratios, and Δ_1 and Δ_{10} denote the annualized performance fees (% per annum) the investor is willing to pay to switch from the linear model to the threshold model under constant relative risk aversion levels of 1 and 10 respectively. The numbers in parentheses are stationary bootstrap p-values for testing the null of equality of model performances across models against one-sided alternatives.