

EKONOMI OCH SAMHÄLLE

Skrifter utgivna vid Svenska handelshögskolan
Publications of the Swedish School of Economics
and Business Administration

Nr 162

ANDERS WILHELMSSON

ESSAYS ON THE MODELING AND PREDICTION
OF VOLATILITY AND HIGHER MOMENTS OF
STOCK RETURNS

Helsingfors 2006

Essays on the Modeling and Prediction of Volatility and Higher Moments of Stock Returns

Key words: GARCH models, Volatility forecasts, Density forecasts, Forecast Evaluation, intra-day data, realized volatility, Time varying kurtosis, Time varying skewness, Value at Risk, Normal Inverse Gaussian distribution.

© Swedish School of Economics and Business Administration & Anders Wilhelmsson

Anders Wilhelmsson
Swedish School of Economics and Business Administration
Department of Finance and Statistics

Distributor:

Library
Swedish School of Economics and Business Administration
P.O.Box 479
00101 Helsinki, Finland

Telephone: +358-9-431 33 376, +358-9-431 33 265
Fax: +358-9-431 33 425
E-mail: publ@hanken.fi
<http://www.hanken.fi>

ISBN 951-555-928-6 (printed)
ISBN 951-555-929-4 (PDF)
ISSN 0424-7256

Edita Prima Ltd, Helsingfors 2006

ACKNOWLEDGEMENTS

I would first like to thank my thesis supervisors Professor Johan Knif and Anders Ekholm. I would also like to thank Professor Eva Liljebloom for accepting me to pursue my doctoral studies at Hanken and for keeping the department running.

As a complement to the courses at Hanken I also studied mathematics at the Helsinki University where I received help, guidance and encouragement far beyond what can be expected from Doctor Leif Andersson.

I spent the last year of my doctoral studies as a visiting scholar at the Aarhus School of Business where I received supervision from Professor Asger Lunde that I believe was very beneficial for my thesis. The thesis has also improved greatly from the comments of my pre-examiners Professor Markku Lanne and Hossein Asgharian.

For making the processes of writing the thesis not only interesting but also fun I would like to thank all the colleagues at Hanken and Aarhus School of Business. Especially Peter Nyberg – whom I could not get rid of even in Denmark -, Valeri Voev, Bernard Ben Sita, Arto Thurlin, Niklas Ahlgren, Jan Antell, Alexander von Nandelstadh, Mats Rosenberg, Annika Sandström, Olga Karakozova and Sofie Kulp.

Anette requires a special thank you, not only for having to live with another researcher but also for having to read, comment, correct and discuss several versions of this dissertation.

For financial support I'm indebted to Hanken's fund for young doctoral students, Stiftelsen för främjande av värdepappers marknaden i Finland and to the Microstructure of Financial Markets in Europe network.

Lastly I wish to thank my family for conversations that do not revolve around the relative efficiency of asymptotically equivalent estimators.

Table of Contents

PART 1 – RESEARCH FOCUS AND CENTRAL FINDINGS	1
1 Introduction	3
2 Modeling and prediction of financial risk	4
3 Summary of the essays	11
4 Joint results and contributions of the essays.....	14
REFERENCES	16
PART 2 – THE ESSAYS	23
GARCH FORECAST PERFORMANCE UNDER DIFFERENT DISTRIBUTION ASSUMPTIONS	25
ABSTRACT	25
1 Introduction	26
2 Related articles	27
3 Presentation of Models	28
4 Data and Estimation	32
5 Discussion on forecast evaluation	36
6 Results	39
7 Conclusions	49
REFERENCES	51
DOES ALLOWING FOR CONDITIONAL SKEWNESS AND KURTOSIS IMPROVE DENSITY FORECASTS?	55
ABSTRACT	55
1 Introduction	56
2 Evaluating density forecasts	57
3 Models	60
4 Data and estimation	65
5 Results	67
6 Conclusions	78
REFERENCES	79
AN EXTENSION OF THE NIG-S&ARCH MODEL WITH AN APPLICATION TO VALUE AT RISK	83
ABSTRACT	83
1 Introduction	84
2 A theoretical motivation for the NIG-distribution.....	86
3 Presentation of models	87
4 Value at risk.....	92
5 Data and estimation	95
6 Value at risk results	106
7 Conclusions	110
REFERENCES	112

LIST OF FIGURES

Figure 1 - Density plot.....	30
Figure 2 - Variance signature plot.....	35
Figure 3 - Conditional skewness over time.....	42
Figure 4 - Conditional kurtosis over time.....	43
Figure 1 - Skewness-Kurtosis bounds.....	64
Figure 2 - Q tests in-sample 1-6.....	69
Figure 3 - Q tests out-of-sample 1-7.....	71
Figure 4 - Histograms of the probability integral transforms.....	73
Figure 1 - Skewness-Kurtosis bounds.....	92
Figure 2 - The Normal Inverse Gaussian shape triangle.....	97
Figure 3 - Conditional kurtosis from the NIG-S&ARCH-tv model.....	101
Figure 4 - Conditional skewness from the NIG-S&ARCH-tv model.....	102
Figure 5 - Q-Tests insample.....	105
Figure 6 - Q-Tests out of sample.....	106
Figure 7 - 1% Value at Risk on a 100\$ investment on the SP500 index.....	107

LIST OF TABLES

Table 1 - Error Distributions.....	32
Table 2 - Descriptive statistics of the Standard and Poor's 500 Future Index daily returns.....	33
Table 3 - Results from in-sample estimation.....	40
Table 4 - Model diagnostics.....	45
Table 5 - Forecast evaluation - 1, 5 and 20 day forecast horizons.....	47
Table 1 - Presentation of the models used in the density forecast evaluation.....	63
Table 2 - Descriptive statistics of the daily Standard and Poor 500 Index returns.....	66
Table 3 - Parameter estimates on the Standard and Poor's 500 index 1980-1989.....	68
Table 4 - Model diagnostics.....	70
Table 5 - Density forecast results from the M(m,l) test for independence.....	77
Table 1 - Descriptive statistics of the Standard and Poor's 500 Index.....	96
Table 2 - Estimation results for daily percentage Standard and Poor's 500 returns.....	99
Table 3 - 1 day-ahead Value at Risk predictions at the 1% level.....	108
Table 4 - 1 day-ahead Value at Risk predictions at the 5% level.....	109

PART 1 – RESEARCH FOCUS AND CENTRAL FINDINGS

1 Introduction

One of the most fundamental and widely accepted ideas in finance is that investors are compensated through higher returns for taking on non-diversifiable risk. Hence the quantification, modeling and prediction of risk have been, and still are one of the most prolific research areas in financial economics. For example, construction of optimal portfolios in the sense of Markowitz (1952) requires forecasts for the expected returns and covariances between all assets under consideration, and pricing of options according to Black and Scholes (1973) requires an estimate of the variance of the underlying asset.

It was recognized early on by Mandelbrot (1963) that there are predictable patterns in the variance of speculative prices¹. A model that accounts for these patterns which has turned out to be tremendously successful is the autoregressive conditional heteroscedasticity (ARCH) model of Engle (1982) later generalized by Bollerslev (1986).

Mandelbrot (1963) and Fama (1965) noted that the normal distribution is not sufficiently flexible to give an accurate description of the distribution of speculative returns. Consequently if returns are drawn from a distribution, which is not completely described by its first two moments an interest in higher moments naturally arises. Early studies documenting skewness in equity returns include Beedles (1979) and Schwert (1990). Studies documenting predictable variation in the skewness includes Badrinath and Chatterjee (1991), Aggarwal and Aggarwal (1993), Alles and Kling (1994) as well as Jondeau and Rockinger (2003). Extended asset pricing models that include skewness and/or kurtosis are suggested by e.g. Rubinstein (1973), Kraus and Litzenberger (1976), Harvey and Siddique (2000) as well as Dittmar (2002). An option pricing formula that accounts for conditional skewness has recently been derived by Christoffersen et al. (2006).

To be able to correctly measure risk and return, to facilitate improvements in asset pricing models and for efficient risk sharing in the economy, statistical models that accurately describe asset returns are needed. If not only variance but also skewness and kurtosis vary in a predictable fashion this needs to be accounted for by the models. Such models have been proposed by e.g. Hansen (1994), Harvey and Siddique (1999), Premaratne and Bera (2001), Mittnik and Paolella (2003), Brännäs and Nordman (2003a,b) as well as Lanne and Saikkonen (2005).

Lacking in the literature so far, is an out-of-sample forecast evaluation of the potential benefits of these new more complicated models with time-varying higher moments. Recent advances in the testing methodology of density forecasts (Diebold et al. (1998), Hong and Li (2005)), on Value-at-Risk forecasts (Christoffersen (1998), Christoffersen and Pelletier (2004)) as well as in the construction of less noisy proxies for *ex post* variance (e.g. Andersen et al. (2003, 2005)) have made the time ripe for such an evaluation. This thesis therefore evaluates the volatility and density forecast

¹ Even though Mandelbrot appears to be the first to put the observed clustering of variance in print, it was the Harvard professor Hendrik S. Houthakker who made Mandelbrot aware of the fact (Mandelbrot 1963).

performance of several of these new models in essays 1 and 2 respectively. In essay 3 the ability of the models to correctly forecast the leftmost tail of the return distribution is investigated using a Value-at-Risk framework. Also, Jensen and Lunde's (2001) NIG-S&ARCH model is extended to allow for conditional modeling of higher moments.

The results of the thesis show that kurtosis appears not to exhibit predictable time variation, whereas there is found some predictability in the skewness. However, the dynamic properties of the skewness are not completely captured by any of the models. The findings are consistent in that none of the variance or density forecasts are improved by allowing for time variation in higher moments. This result is also consistent for the different models evaluated. These findings have important implications for future improvements in the modeling of conditional moments. A further implication is that asset pricing and risk management models that incorporate higher return moments should treat the variance and possibly the skewness as time-varying but, when variance is time varying, kurtosis needs not to be modeled as time varying separately.

The structure of the dissertation is as follows: the relevant literature and methodology used in the essays will be presented. Thereafter, the main findings and contributions of the individual essays are summarized and also the results and contributions of the dissertation taken as a whole will be discussed. The last part of the dissertation presents the three individual essays.

2 Modeling and prediction of financial risk

Since the literature on financial risk is vast, the focus here will necessarily be narrow. The strand of literature most closely connected to the essays of the dissertation is that on univariate parametric models. In particular, much of the focus will be on the autoregressive conditional heteroscedasticity (ARCH) model developed by Engle (1982) and on the numerous extensions of this model. Also the more recent non-parametric approach often called realized volatility will be discussed².

2.1 Stylized facts of equity returns

When modeling equity market volatility there are certain stylized facts that have to be accounted for. Large price changes are often followed by large price changes and small price changes are often followed by small price changes as noted by Mandelbrot (1963). This is often referred to as volatility clustering. Another finding, first observed by Black (1976), is that price changes due to bad news tend to be larger in magnitude than changes due to good news. A third finding is that the return distribution is not only skewed but also shows excess kurtosis (Mandelbrot (1963), Fama (1965)), commonly referred to as fat tails³. More recently, in inter alia Harvey and Siddique (1999, 2000), there is emerging evidence of predictability in the skewness of stock returns. The

² The terms realized volatility and realized variance are used interchangeably in the literature.

³ This is, of course, because a distribution with a kurtosis higher than 3 has more probability mass in the tails (and around the peak, but less on the flanks) than the normal distribution. Such a distribution is also called leptokurtic.

development in modeling stock returns can to a great extent be seen as being driven by the inability of previous models to capture all of the salient features of the data described above.

2.2 Popular model specifications

Some of the most commonly used ARCH type specifications will be presented and discussed below⁴. To fix notation, let P_t with $t \in \{T, T-1, T-2, \dots, 0\}$ denote the time t price of a financial asset. The continuously computed one period return will be given by $r_t = \log(P_t / P_{t-1})$ assuming dividends being added to the price.

All the models are to be estimated jointly with a specification for the mean equation given by $r_t = \mu + \varepsilon_t$. Here $\varepsilon_t = \sqrt{h_t} z_t$, $h_t | \Omega_{t-1}$ is the variance conditional on the information set Ω_{t-1} and z is a random variable assumed to follow some distribution, often the normal. Empirically, an autoregressive or moving average term is often included in the mean equation to account for possible serial correlation in returns.

The ARCH(q) model of Engle (1982) with a conditional variance given by

$$h_t = c + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 \quad (1)$$

captures the persistence in volatility by making the variance conditional on lagged squared innovations from the mean equation. A sufficient but not necessary condition for h_t to be positive is $c > 0$ and $\alpha_1, \alpha_2, \dots, \alpha_q > 0$. Some of the excess kurtosis in returns is also captured in this models since the time variation in the variance will produce an unconditional fourth moment that is higher than the fourth moment of the error distribution⁵. This specification is thus able to capture the tendency of volatility to cluster over time but requires a rather large value of q to do so. In response to this Bollerslev (1986) suggests a more parsimonious way of modeling the dependence by adding lags of the conditional variance giving the GARCH (q,p) model

$$h_t = c + \sum_{i=1}^q \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^p \beta_j h_{t-j}, \quad (2)$$

with $c > 0$ and $\alpha_1, \alpha_2, \dots, \alpha_q, \beta_1, \beta_2, \dots, \beta_p > 0$ sufficient for h_t to be positive. For empirical studies Bollerslev, Chou and Kroner (1992) among others have shown that the GARCH (1,1) model is usually sufficient to capture the persistence effect. This model has been very successful and is still after almost 20 years frequently used in both academia and industry.

⁴ For a broader introduction to the subject see for example Bollerslev et al. (1994) and the collection of articles in Engle (1995).

⁵ The unconditional fourth moment for the (G)ARCH(1,1) model when the error distribution is normal is derived in Bollerslev (1986) and for more general class of distributions in He and Teräsvirta (1999).

Neither of the above models however, responds asymmetrically to positive and negative return innovations and hence they ignore the so called leverage, or volatility feedback effect⁶. Alternative specifications were therefore proposed by Nelson (1991) as well as Glosten et al (1993). The model of Glosten et al. with $q=p=1$ is given below in (3).

$$h_t = c + \alpha \varepsilon_{t-1}^2 + \delta I_{t-1} \varepsilon_{t-1}^2 + \beta h_{t-1} \text{ with } I_t = \begin{cases} 1 & \text{for } \varepsilon_t < 0 \\ 0 & \text{Otherwise} \end{cases} \quad (3)$$

As can be seen in this model, the effect of positive innovations is given by α and the effect of negative returns is given by $\alpha + \delta$. The regular GARCH(1,1) model is obtained as a special case by setting $\delta = 0$.

All of the above specifications parameterize the conditional variance of the model. However, there is no clear reason why, for example, the conditional standard deviation should not be modeled instead. Ding et al. (1993) suggest to let the data decide on what power of the conditional standard deviation ($h_t^{1/2}$) to model. This results in the asymmetric power ARCH given below

$$\left(h_t^{1/2}\right)^\nu = c + b\left(h_{t-1}^{1/2}\right)^\nu + a\left[|\varepsilon_{t-1}| - \tau\varepsilon_{t-1}\right]^\nu \quad (4)$$

The model also allows for an asymmetric response by the τ parameter, and it nests the GARCH (1,1) model as can be seen by setting $\tau = 0$ and $\nu = 2$.

To be able to capture all the excess kurtosis in, especially daily, equity returns different error distributions have been proposed. Probably the most frequently used, besides the normal distribution, are the t-distribution (Bollerslev (1987)) and the Generalized error distribution (Nelson (1991)). A less used distribution, introduced to finance by Eberlein and Keller (1995) and Barndorff-Nielsen (1995), is the normal inverse Gaussian distribution (NIG). This distribution, in addition to being flexible enough to provide a good fit to the data, can be theoretically motivated from Clark's (1973) mixture of distribution hypothesis⁷. Barndorff-Nielsen (1997) suggested a model based on the NIG distribution to model financial volatility. Extensions of Barndorff-Nielsen's model have been made by Forsberg (2002) as well as Forsberg and Bollerslev (2002) who allow for a richer dependence structure in the variance. The model has also been extended to allow for skewness by Jensen and Lunde (2001).

⁶ Different economic explanations have been proposed for why volatility responds differently to positive and negative innovations of the same magnitude. Black's (1976) leverage effect states that a negative return decreases the value of the firm's assets and consequently increases the risk by increased leverage. The volatility feedback effect discussed by for example Campbell and Hentschel (1992) states that since both good and bad news will imply a higher volatility in the future, the required return increases and consequently prices fall. This effect dampens the impact of positive news and increases the impact of negative news, creating asymmetry in returns.

⁷ The NIG distribution may be represented as a normal-Inverse Gaussian mixture distribution. This means that if Clark's (1973) assumption of a log normally distributed mixing variable is replaced by the empirically very close inverse Gaussian distribution, the unconditional return distribution will be NIG. Forsberg (2002) shows the inverse Gaussianity assumption to be well founded.

All of the models above assume that (conditional) moments higher than the second are constant. Since this assumption may not be supported by the data, Hansen (1994) proposed the autoregressive conditional density model. This model allows for not only a non-central and leptokurtic error distribution but also for time variation in the skewness and kurtosis. The time dependence in skewness and kurtosis is modeled in a GARCH like fashion by making the higher moments conditional on the residuals from the mean equation. Other models with time variation in skewness and/or kurtosis include for example Harvey and Siddique (1999), Mittnik and Paolella (2003) as well as Brännäs and Nordman (2003a,b).

2.3 Realized variance and volatility forecast evaluation

The concept of realized variance (RV) is a relatively new but rapidly growing research area. Here the focus will be on the role of realized variance for forecast evaluation. For a formal treatment of realized variance see Andersen et al. (2003, 2005). One of the major complications of volatility forecast evaluation is that the forecasted variable is non observable. To see this, consider a (zero mean)⁸ return given by $r_t = \sigma_t z_t$ with σ_t being the standard deviation and z being an independent random variable with mean zero and unit variance. *Ex post* the return r_t is of course observable but not the variance since some kind of assumption (model) is needed to divide the return into a standard deviation and a white noise part. For example, the GARCH models in the previous section make the conditional variance h_t observable with respect to the information set (Ω_{t-1}) at time $t-1$. However, without a model assumption the variance can not be observed and a proxy is needed.

The unobservable variance can be replaced by the squared return which is an unbiased estimate since $E_{t-1}[r_t^2] = E_{t-1}[\sigma_t^2 z_t^2] = E_{t-1}[\sigma_t^2] E_{t-1}[z_t^2] = E_{t-1}[\sigma_t^2] \cdot 1 = \sigma_t^2$. However, the day to day variability in z_t^2 is several orders of magnitude larger than the variability in σ_t^2 making the squared return a very noisy proxy for the variance. The effects of replacing the true latent variance by a proxy was largely ignored in the variance forecast literature until Andersen and Bollerslev (1998) showed the detrimental impact this will have on the perceived forecast ability.

To get a less noisy proxy of the *ex post* variance it is possible to use the RV which is the idea of using the sum of squared returns of a higher frequency than the desired frequency of the variance. For example to use intra daily returns to construct daily variances or daily returns to construct monthly variances⁹. Formally the theory of quadratic variation states that for serially uncorrelated returns of frequency m $r_{(m),t} = \log(P_t / P_{t-1/m})$ we have:

⁸ The zero mean assumption is not necessary but notationally more convenient.

⁹ This idea was used before 1998 by for example Poterba and Summers (1986), French et al. (1987), Schwert (1990) and Hsieh (1991) but without any theoretical foundation.

$$p \lim_{m \rightarrow \infty} \left(\int_0^1 \sigma_{t+\tau}^2 d\tau - \sum_{j=1, \dots, m} r_{(m)t+j/m}^2 \right) = 0 \quad (5)$$

The quantity $\int_0^1 \sigma_{t+\tau}^2$ is the integrated variance, which of course is unobservable in practice. This result shows that as the sample frequency increases an increasingly less noisy proxy of the integrated variance is obtained.¹⁰ However, this theoretical results is sensitive to autocorrelation in the returns produced by market microstructure effects as analyzed in for example Aït-Sahalia *et al.* (2005) and Oomen (2005). Reducing the noise in the proxy is essential for gauging the true degree of predictability of a volatility model but also for examining the relative performance of competing models as shown in Hansen and Lunde (2006) as well as in Patton (2006).

2.4 Density forecast evaluation

Methods for evaluating density forecasts have received considerable attention in the last few years following the seminal paper by Diebold *et al.* (1998).¹¹ Akin to variance, the true return distribution is unobservable and can hence not be directly compared to the forecasted density. Pointing to a further complication, Diebold *et al.* (1998) show that it is impossible to find a ranking of density forecasts that is independent of individual loss functions, the only exception being that the true model is preferred by everyone.

Nevertheless, the result that the probability integral transform (PIT) of the predicted density will form a series that is iid uniform on the unit interval if the predicted density equals the true density has proved very useful. Using this result as a tool for model assessment was concurrently proposed by Diebold *et al.* (1998) and Kim *et al.* (1998).

The probability integral transform of a return, r_t , is given by $\hat{z}_t = \int_{-\infty}^{r_t} \hat{p}_t(u) du$, where \hat{p}_t is the predicted density. Diebold *et al.* (1998) suggest drawing histograms of the PIT series and constructing confidence bands based on the binomial distribution¹². This however, is only a test for half of the hypothesis. In order to investigate the dependence it is suggested to plot the autocorrelations of $\left\{ \left(\hat{z}_t - E[z_t] \right)^p \right\}_{t=0}^T$ for $p=1,2,3,4$.

A second major approach in the literature is to compare the predicted density with a kernel based estimate of the true density as proposed by e.g. Sarno and Valente (2004). However, this approach is problematic when the true density has time-varying moments since the kernel estimator of the density will not be consistent. As a solution, Hong and

¹⁰ Results on how the difference between the integrated and realized variance depends on the sample frequency are available in Andersen *et al.* (2004) and an asymptotic distribution theory for the difference is derived in Barndorff-Nielsen and Shephard (2002).

¹¹ For a comprehensive review of the density forecasting literature see Corradi and Swanson (2006).

¹² Under the null, the probability for an observation being in a certain bin is simply $1/\#\text{bins}$.

Li (2005) combine the two approaches and compare a kernel estimate of the PIT-series and lagged PIT-series from the proposed models with a bivariate uniform distribution¹³.

2.5 Value at Risk and “Backtesting”

Value at Risk (VaR) is the maximum dollar loss that will occur for a given probability α over a certain time horizon h . A VaR *model* is a theoretical construct that is used to compute the VaR.

The practical relevance of “backtesting” VaR models is highlighted in The Basel committee on banking supervision who state in their 2004 “International convergence of capital measurement and capital standards” (page 39) that

“Internal models will only be accepted when a bank can prove the quality of its model to the supervisor through the backtesting of its output using one year of historical data.”

The Basel committee is only interested in the so-called unconditional coverage of the VaR model. The unconditional coverage is the percentage of exceptions, i.e. when a loss is larger than predicted by the model, which should on average be equal to the VaR level (α) of the model. Christoffersen (1998) argues that it is also of importance that the exceptions are independently distributed over time, since a clustering of exceptions could very well lead to bankruptcy. Christoffersen suggests a test for correct conditional coverage which can be seen as the sum of two individual tests, one for a correct unconditional coverage and one test for the independence of the exceptions.

Define the indicator variable I_t with t being a time subscript according to

$$I_t = \begin{cases} 1, & \text{if } r_t > F_t^{-1}(\alpha) | \Omega_{t-1} \\ 0 & \text{Otherwise} \end{cases}, \quad (6)$$

where $F_t^{-1}(\alpha) | \Omega_{t-1}$ is the conditional VaR forecast (the inverse of the cumulative distribution function evaluated at α) from the particular model being evaluated. This means that a series that constitutes of zeros whenever there is an exception, i.e. when a return is below the value at risk given by the model, and ones otherwise is constructed. To test if the number of exceptions is correct is called to test for correct unconditional coverage. If we have correct unconditional coverage alpha percent of the returns will be lower than the VaR prediction, $F_t^{-1}(\alpha) | \Omega_{t-1}$, so under the null we will have

$$E \left[\frac{1}{T} \sum_{t=1}^T I_t \right] = 1 - \alpha. \quad (7)$$

The test for correct conditional coverage can be divided into two separate parts; one part tests for correct unconditional coverage and one part tests for independence in the

¹³ If the proposed model is correct the PIT-series will be iid and uniform, the purpose of using the bivariate distribution of the PIT- and lagged PIT-series is to jointly test for both independence and uniformity.

sequence of exceptions. This is very useful since it can then be investigated if a model rejection is due to unconditional coverage failure, clustering of the exceptions, or both.

The null hypothesis for correct unconditional coverage gives that $I_t \sim \text{Bern}(1-\alpha)$ which can be tested by a likelihood ratio test of the form

$$LR_{UC} = 2 \left(\log \left(\hat{\pi}_1^{T_1} (1-\hat{\pi}_1)^{T-T_1} \right) - \log \left((1-\alpha)^{T_1} \alpha^{T-T_1} \right) \right). \quad (8)$$

The number of observations is given by T , the number of ones is given by T_1 and $\hat{\pi}_1 = T_1 / T$.

To see if the exceptions tend to cluster together over time Christoffersen (1998) suggests testing for independence with first order Markov dependence used as an alternative. The test statistic is given by

$$LR_{IND} = 2 \left(\log \left((1-\hat{\pi}_{01})^{T_0-T_{01}} \hat{\pi}_{01}^{T_{01}} (1-\hat{\pi}_{11})^{T_1-T_{11}} \hat{\pi}_{11}^{T_{11}} \right) - \log \left(\hat{\pi}_1^{T_1} (1-\hat{\pi}_1)^{T-T_1} \right) \right), \quad (9)$$

T_{ij} is the number of observations valued i followed by observations valued j . The maximum likelihood estimates of $\hat{\pi}_{ij}$ are simply $\hat{\pi}_{01} = T_{01} / T_0$ and $\hat{\pi}_{11} = T_{11} / T_1$.

The joint test of correct conditional coverage means that $I_t \sim iid \text{Bern}(1-\alpha) \forall t$. The test statistic is simply given as the sum of the two individual tests in (8) and (9).

$$LR_{CC} = LR_{UC} + LR_{IND} \quad (10)$$

However, Christoffersen and Pelletier (2004) argue that this alternative may be too restrictive and suggest duration based tests for independence as an alternative. Duration is in this context defined as the time in days between two exceptions. For a correctly specified model, with no dependence between exceptions, this should always be $1/\alpha$. Furthermore, Christoffersen and Pelletier point out that the effective sample size is rather small in most VaR applications. They therefore suggest simulating the distribution of the test statistics under the null instead of using the result that asymptotically the test statistics will be χ^2 distributed.

2.6 Summary of research focus

Lacking in the literature so far is an out-of-sample forecast evaluation of the potential benefits of models with time-varying higher moments. Such an evaluation of the volatility and density forecast performance is the subject of essays 1 and 2, respectively. In essay 3 the ability of the models to correctly forecast the leftmost tail of the return distribution is investigated using a Value-at-Risk framework. Also an extension of Jensen and Lunde's (2001) NIG-S&ARCH model is proposed.

3 Summary of the essays

In this section each essay will be summarized. The focus will be on the methods, results and contributions of the essays. Lastly, the joint contribution of the essays will be discussed.

3.1 Essay 1

In this essay named “*GARCH forecast performance under different distribution assumptions*” the out-of-sample forecast performance of the GARCH (1,1) model estimated with 9 different error distributions is investigated. The different distributions can all be seen as special cases of Hansen’s (1994) autoregressive conditional density model. By restricting the model, the effects of allowing for skewness, leptokurtosis and time variation in the higher moments are isolated.

The models are estimated on daily Standard and Poor 500 index Future returns, and intra-daily 1-minute returns are used to construct the *ex post* measure of variance against which the forecasts are judged. The time period investigated is January 2, 1996 to December 30, 2002. The sample is divided into a five year estimation sample and a two year forecast sample. Forecasts are computed for 1, 5 and 20 day horizons, and the predictive ability of the different models is compared with the Diebold and Mariano (1995) test and with the model confidence set (MCS) test of Hansen et al. (2003) to investigate if the differences are statistically significant.

The results show the GARCH model with Student t distributed errors to be the best performing model. It produces superior forecasts on 5 out of 6 comparisons (2 loss functions times 3 forecast horizons). The GARCH model with Student t distributed errors is also the only model that is always in the set of superior models given by the Hansen et al. (2003) MCS test. These results are in partial agreement with Hamilton and Susmel (1994), who found the GARCH model with a t-distribution to perform the best when the logarithmic loss criterion was used. In contradiction to the results of this essay, Franses and Ghijssels (1999) found the out-of-sample performance for the GARCH model with a t-distribution to be worse than for a GARCH model estimated with the normal distribution. This contradictory results can probably be explained by the noisy *ex post* variance proxy (squared weekly returns) used by Franses and Ghijssels.

It is further found that incorporating constant or time-varying skewness or time-varying kurtosis does not improve the forecasts further. In contrast to this, the in-sample estimation reported significant improvements when skewness was allowed in the error distribution.

The contribution of this essay over the previous volatility forecast literature is firstly to allow for a very flexible error distribution that nests several well-known distributions and allows for higher order dependence in the data. Secondly, a much less noisy proxy than commonly used for the *ex post* variance is constructed to facilitate meaningful

comparison of the forecast results. In addition, the forecast horizon is extended¹⁴ to include one, five and twenty day forecasts and statistical methods are used to gauge if the differences in forecast accuracy are significant.

3.2 *Essay 2*

The aim of this essay called “*Does allowing for conditional skewness and kurtosis improve density forecasts*” is to ascertain the potential benefits of some recently developed models with dependence structures in skewness and kurtosis. Related work on the predictive ability of various density forecast models have been done by, for example, Níguez and Perote (2004) as well as Sarno and Valente (2004) on exchange rates, and by Clements and Smith (2000) on US output growth and changes in unemployment rates. Evaluations on stock return data have been done by de Raaij and Raunig (2002) as well as Bao et al. (2006). However, none of the existing work has examined the out-of-sample predictive ability of models with time-varying skewness and kurtosis on equity returns.

In this essay, Hansen’s (1994) autoregressive conditional density model as well as Brännäs and Nordman’s (2003) GARCH model with generalized log gamma errors and restrictions of these models are estimated and their forecast ability is evaluated. The data consists of 20 years of daily Standard and Poor 500 index returns from 1980-1999. The sample is divided into a ten year estimation sample and a ten year prediction sample. One step density forecasts are computed recursively so that the estimation sample is extended by one observation and the parameters re-estimated after each forecast is made. The forecasts are evaluated by applying Hong and Li’s (2005) non-parametric specification test for continuous time models to the density forecast field.

The results show that the t-distribution without higher moment dependence seems sufficiently flexible to produce decent density forecasts. Allowing for non-zero skewness does not improve forecasts further. Neither does allowing the skewness and kurtosis to be time-varying improve forecasts. This may, however, be due to poorly specified dynamics for the skewness since the forecast results show skewness predictability in the data. Furthermore, all the GARCH models explain the dependencies in kurtosis rather well, which questions the need for explicitly modeling the kurtosis as time-varying.

Of course, only a small subset of the proposed models with time-varying skewness and kurtosis have been investigated in this essay but the tentative evidence is that at least kurtosis needs not be modeled as time dependent. Further, the specification of the conditional skewness used in conjunction with the generalized t distribution is found unsatisfactory.

The contribution of this essay is to broaden the currently rather meager literature on out-of-sample density forecast performance by being the first study to consider the effect of allowing for conditional skewness and kurtosis. Additionally, density forecasts are

¹⁴ The forecast horizon is extended compared to the previous studies that specifically tested the effects of different error distributions (Hamilton and Susmel (1994), Lopez (2001) as well as Franses and Ghijssels (1999)).

evaluated by using a recent test originally proposed for continuous time models developed by Hong and Li (2005). This test is especially well suited to deal with possible time dependence in the higher moments of the return distribution.

3.3 *Essay 3*

In this essay, “*An extension of the NIG-S&ARCH model with an application to Value at Risk*”, a new model that allows for time variation in skewness and possibly kurtosis is proposed. The model builds on the NIG distribution that has been used to model financial assets by inter alia Barndorff-Nielsen (1997), Jensen and Lunde (2001), Forsberg (2002) as well as Forsberg and Bollerslev (2002).

The new model as well as the NIG-S&ARCH model of Jensen and Lunde (2001) and the GARCH-NIG model of Forsberg and Bollerslev (2002) are estimated and then used to construct Value at Risk predictions on the SP500 index during the period 3rd of July 1962 to 20th of September 2005¹⁵. The VaR forecasts are evaluated with consideration given to both the models’ abilities to produce the right number of exceptions and to the independence of the exceptions. This is done by using the test for correct conditional coverage of Christoffersen (1998) as well as the duration based test for independence suggested in Christoffersen and Pelletier (2004).

The results show that the NIG based models perform very well compared with extant models. Moreover, the results show that only the new model produces satisfactory VaR forecasts for both 1% and 5% VaR. The models also perform better than the mixture GARCH models with stable Paretian error distributions proposed and evaluated by Haas et al. (2005) on a German Dax-30 sample of comparable size. Furthermore, in terms of correct conditional coverage for VaR at the 1% level, the GARCH-NIG and NIG-S&ARCH models in this essay beat all 13 models evaluated in Kuester et al. (2006). The models in Kuester et al. include GARCH models with different error distributions and the CAViaR model of Engle and Manganelli (2004). Further, it is shown that allowing for constant or conditional skewness does improve the in-sample fit but it does not improve the VaR forecast. This is in line with previous variance forecast results reviewed in Poon and Granger (2003).

The contribution of this essay is to extend the NIG-S&ARCH model to allow for conditional skewness and kurtosis. This is motivated by the recent interest in conditional higher moments (see e.g. Harvey and Siddique (1999, 2000), Dittmar (2002), Guermat and Harris (2002) as well as Christoffersen et al. (2006)). In addition, the essay contributes by examining both the conditional and unconditional VaR estimates, using the testing methodology of Christoffersen (1998) and the recent advances by Christoffersen and Pelletier (2004), on a long time series of Standard and Poor 500 returns. The models considered have not previously been used for VaR calculations on stock return data.

¹⁵ The 3rd of July 1962 is the start date for the daily Center for research in security prices (CRSP) data. However, in February 2006 the database will be extended to include daily data from as early as the 31st of December 1925.

4 Joint results and contributions of the essays

Taken together this dissertation examines the variance, density as well as the VaR forecasts of a number of extant models, as well as of one model proposed in this dissertation. The models all allow for time variation in higher moments. Modeling the excess kurtosis is found essential for computing correct VaR estimates. It is also found that allowing for constant excess kurtosis is critical for improving both variance and density forecasts. However, inclusion of non-zero skewness, constant or conditional, does not improve variance or density forecasts. Neither does incorporating time variation in the kurtosis.

These findings have important implications for future improvements of conditional density models. A further implication is that asset pricing and risk management models that incorporate higher return moments should treat the variance and possibly the skewness as time-varying but the kurtosis (when variance is time-varying) need not be modeled as time-varying. Also, a very important implication is that GARCH type models are greatly improved by using leptokurtic error distributions, failing to do so will severely underestimate the true risk of an asset. It is not an acceptable solution to rely on the normal distribution, with the justification of the quasi maximum likelihood estimator being consistent, since the interval and density forecasts produced by the model will depend on the distribution assumption also asymptotically.

The previous results on variance forecasting are rather mixed with some studies (e.g. Franses and Ghijssels (1999)) finding that forecasts are not improved by using leptokurtic error distributions while some (e.g. Hamilton and Susmel (1994)) find the opposite result. These mixed results are largely due to the noisy proxy commonly used for *ex post* variance, worsened by the use of inconsistent loss functions. Both of these issues are addressed in essay 1 improving the reliability of the results.

The density forecasts are evaluated by the recent test procedure developed by Hong and Li (2005). Using this test is an important enhancement over previous studies, that rely solely on Diebold et al. (1998) or Berkowitz (2001), in that it allows for formal testing of the *joint* hypothesis of the PIT being \sim iid $U(0,1)$. To the best of the author's knowledge no previous research has been conducted on the density forecast ability of models with time-varying higher moments on equity returns. Hence the finding that time variation in higher moments does not improve density forecasts is novel and cannot yet be compared to results in other studies. For the VaR evaluation in essay 3 the results can be compared to Guermat and Harris (2002) who do find that time variation in kurtosis improves the VaR estimates. However, the improvement they find over a GARCH-t model is very slight and not consistent for all markets and VaR levels investigated.

One possible limitation of the generalization of the results in this dissertation, that show time-varying higher moments not to improve variance, density or VaR forecasts, is that all the empirical analysis has been conducted only on Standard and Poor 500 data. It would be of interest to see if the results can be replicated on other asset classes and for other markets. However, the Standard and Poor's 500 index is probably the most interesting single choice to investigate, representing a market value of over 11 000

billion dollars as of 30th of September 2005¹⁶. A further limitation is that it cannot generally be said that time variation in skewness and kurtosis does not improve forecasts; this conclusion can only be drawn for the particular model specifications used in this dissertation. To alleviate this limitation, a number of different models have been investigated in the dissertation. Furthermore, the specification tests conducted in essay 2 show that there is no apparent structure in the fourth power of the residuals for the models with constant kurtosis. However, there appears to be some predictability in the skewness not captured by the investigated models. Developing new models that allow for both conditional variance and skewness could hence be a fruitful area for further research and an attempt of this is made in essay 3.

In further research, perhaps more general results could be established by investigating higher moments without being restricted to a particular model. One possible avenue to achieve this is to build on the results from realized variance to compute other realized objects, such as realized skewness and possible realized kurtosis.

¹⁶ Source: Standard an Poor 500 fact sheet available at <http://www2.standardandpoors.com/spf/pdf/index/500factsheet.pdf>

REFERENCES

- Ait-Sahalia, Y., Mykland, P., Zhang, L. 2005. How often to sample a continuous-time process in the presence of market microstructure noise. *Review of Financial Studies*, **18**: 351-416.
- Aggarwal, R., Aggarwal, R. 1993. Security Return Distributions and Market Structure: Evidence from the NYSE/AMEX and the NASDAQ Markets *Journal of Financial Research*, **Fall**: 209-220.
- Alles, L., Kling, J. 1994. Regularities in the Variation of Skewness in Stock Index Returns. *Journal of Financial Research*, **Fall**: 427-438.
- Andersen, T.G., Bollerslev, T. 1998. Answering the skeptics: yes standard variance models do provide accurate forecasts. *International Economic Review*, **39**: 885-905.
- Andersen, T.G., Bollerslev, T., Diebold F.X 2005. Parametric and Nonparametric Volatility Measurement. In: *Handbook of Financial Econometrics*, Hansen LP, Ait-Sahalia Y (eds.); North-Holland: Amsterdam. Forthcoming
- Andersen, T.G., Bollerslev, T., Diebold F.X., Labys P. 2003. Modeling and Forecasting Realized Volatility. *Econometrica*, **71**: 529-626.
- Andersen, T.G., Bollerslev, T., Meddahi N. 2004. Analytic Evaluation of Variance Forecasts. *International Economic Review*, **45:4** 1079-1110.
- Badrinath, S.G., Chatterjee S. 1991. A Data-Analytic Look at Skewness and Elongation in Common Stock Return Distributions. *Journal of Business and Economic Statistics*, **April**: 223-233.
- Bao Y., Lee T.H., Saltoglu B. 2006. Comparing Density Forecasts. *Journal of Forecasting*, forthcoming.
- Barndorff-Nielsen, O.E. 1995. Normal inverse Gaussian process and the modelling of stock returns. Research Report 300. Department of Theoretical Statistics, University of Aarhus.
- Barndorff-Nielsen, O.E. 1997. Normal inverse Gaussian distributions and stochastic volatility modeling. *Scandinavian Journal of Statistics*, **24**: 1-13.

- Barndorff-Nielsen, O.E., Shephard N. 2002. Econometric analysis of realized volatility and its use in estimating stochastic volatility models. *Journal of the Royal Statistical Society, Series B* **64**: 253-280.
- Beedles, W.L. 1979. Return, Dispersion and Skewness: Synthesis and Investment Strategy. *Journal of Financial Research*, **Spring**: 71-80.
- Berkowitz, J. 2001. Testing Density Forecasts, With Applications to Risk Management. *Journal of Business & Economic Statistics*, **19**: 465-474.
- BIS 2004. International Convergence of Capital Measurement and Capital Standards. Basel Committee on Banking Supervision
- Black, F. 1976. Studies of Stock Market volatility changes. *Proceedings of the American Statistical Association, Business and Economic Statistics Section*, 177-181.
- Black, F., Scholes M. 1973. The Pricing of Options and Corporate Liabilities. *Journal of Political Economy*, **81**: 637-654.
- Bollerslev, T. 1986. Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics*, **31**: 307-327.
- Bollerslev, T. 1987. A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics*, **69**: 542-547.
- Bollerslev, T., Chou R., Kroner K. 1992. ARCH modeling in finance. *Journal of Econometrics*, **52**: 5-59.
- Bollerslev, T., Engle, R., Nelson D. 1994. ARCH Models. In: Engle, R., McFadden, D. (Eds.) *Handbook of Econometrics*. Vol. 4. North Holland, Amsterdam, 2961-3038.
- Brännäs, K., Nordman N. 2003a. An alternative conditional asymmetry specification for stock returns. *Applied Financial Economics*, **13**: 537-541.
- Brännäs, K., Nordman N. 2003b. Conditional skewness modeling for stock returns. *Applied Economics Letters*, **10**: 725-728.
- Campbell, J., Hentschel L. 1992. No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns. *Journal of Financial Economics*, **31**: 281-331.

- Christoffersen, P. 1998. Evaluating Interval Forecasts. *International Economic Review*, **39**: 841-862.
- Christoffersen, P., Pelletier, D. 2004. Backtesting Value-at-Risk: A Duration-Based Approach. *Journal of Financial Econometrics*, **2**: 84-108.
- Christoffersen, P., Heston S., Jacobs K. 2006. Option valuation with conditional skewness. *Journal of Econometrics*, **131**: 253-284.
- Clark, P.K. 1973. A subordinated stochastic process model with finite variance for speculative prices. *Econometrica*, **41**: 135–155.
- Clements, M., Smith, J. 2000. Evaluating the forecast densities of linear and non-linear models: applications to output growth and unemployment. *Journal of Forecasting*, **19**: 255-276.
- Corradi, V., Swanson N. 2006. Predictive Density Evaluation. In: Handbook of Economic Forecasting, Granger, C., Elliot G., Timmerman A. (Eds.), Elsevier, Amsterdam.
- de Raaij, G., Raunig B. 2002. Evaluating Density Forecasts with an Application to Stock Market Returns. *Discussion paper 08/02 Economic Research Centre of the Deutsche Bundesbank*
- Diebold, F.X., Mariano, R. 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics*, **13**: 253-263.
- Diebold, F.X., Gunther T., Tay S. 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review*, **39**, 863-883.
- Ding, Z., Granger, C., Engle R. 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, **1**: 83-106.
- Dittmar, R. 2002. Nonlinear Pricing Kernels, Kurtosis Preference, and the Cross-Section of Equity Returns. *Journal of Finance*, **57**: 369-403.
- Eberlein, E., Keller U. 1995. Hyperbolic distributions in finance. *Bernoulli*, **1**: 281-299.
- Engle, R. 1982. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica*, **50**: 987-1008.
- Engle, R. (ed.) 1995. *ARCH. Selected Readings*. Oxford: Oxford University Press.

- Engle, R., Manganelli S. 2004. CAViaR: Conditional Autoregressive Value at Risk by Regression Quantiles. *Journal of Business and Economic Statistics*, **22**: 367–381.
- Fama, E. 1965. The Behaviour of Stock Market Prices. *Journal of Business*, **38**: 34–105.
- Forsberg, L. 2002. On the Normal inverse Gaussian Distribution in Modeling Volatility in the Financial Markets. Doctoral thesis, *Studia Statistica Upsaliensia* **5**, Christofferson, A., Jöreskog, K.G. (eds). Uppsala University Library.
- Forsberg, L., Bollerslev T. 2002. Bridging the gap between the distribution of realized (ECU) volatility and ARCH modelling (of the Euro): The GARCH-NIG model. *Journal of Applied Econometrics*, **17**: 535-548.
- Franses, P., Ghijssels, H. 1999. Additive outliers, GARCH and forecasting variance. *International Journal of Forecasting*, **15**: 1-9.
- French, K., Schwert, G.W., Stambaugh, R.F. 1987. Expected Stock Returns and Volatility. *Journal of Financial Economics*, **19**: 3-29.
- Glosten, L., Jagannathan, R., Runkle, D. 1993. On the Relation Between Expected Value and the Volatility of the Nominal Excess Return on Stocks. *Journal of Finance*, **48**: 1779-1801
- Guermat, C., Harris, R. 2002. Forecasting Value-at-Risk: Allowing for Time Variation in the Variance and Kurtosis of Portfolio Returns. *International Journal of Forecasting*, **18:3** 409-419
- Haas, M., Mittnik, S., Paoletta, M., Steude, S. 2005. Stable Mixture GARCH Models. NCCR FINRISK Working paper series no 257.
- Hamilton, J., Susmel, R. 1994. Autoregressive conditional heteroscedasticity and changes in regime. *Journal of Econometrics*, **64**: 307-333.
- Hansen, B. 1994. Autoregressive conditional density estimation. *International Economic Review*, **35**: 705-730.
- Hansen, P., Lunde A. 2006. Consistent Ranking of Variance Models. *Journal of Econometrics*, **131**: 97-121.
- Hansen, P., Lunde A., Nason J. 2003. Choosing the Best Volatility Models: The Model Confidence Set Approach. *Oxford Bulletin Of Economics and Statistics*, **65**: 839-861.

- Harvey, C. R., Siddique, A. 1999. Autoregressive Conditional Skewness. *Journal of Financial and Quantitative Analysis*, **34**: 465-488.
- Harvey, C. R., Siddique, A. 2000. Conditional skewness in asset pricing tests. *Journal of Finance*, **55**: 1263-1295.
- He, C., Teräsvirta T. 1999. Forth moment structure of the GARCH(p,q) process. *Econometric Theory*, **15**: 824-846.
- Hong, Y., Li, H. 2005. Nonparametric Specification Testing for Continuous-Time Models with Applications to Term Structure of Interest Rates. *The review of Financial Studies*, **18**: 37-84.
- Hsieh, D.A. 1991. Chaos and nonlinear dynamics: Applications to financial markets. *Journal of Finance*, **46**: 1839-1877.
- Jensen, M.B., Lunde A. 2001. The NIG-S&ARCH model: a fat-tailed, stochastic, and autoregressive conditional heteroscedastic volatility model. *Econometrics Journal*, **4**: 319-342.
- Jondeau, E., Rockinger, M. 2003. Conditional volatility, skewness, and kurtosis: existence, persistence, and comovements. *Journal of Economic Dynamics & Control*, **27**: 1699-1737.
- Kim, S., Shephard N., Chib S. 1998. Stochastic volatility: Likelihood Inference and Comparison with ARCH Models. *Review of Economics Studies*, **65**: 361-393.
- Kraus, A., Litzenberger, R.H. 1976. Skewness preference and the valuation of risk assets. *Journal of Finance*, **31**:4 1085-1100.
- Kuester, K., Mittnik S., Paolella M. 2006. Value-at-Risk Prediction: A Comparison of Alternative Strategies. *Journal of Financial Econometrics*, **4**:1 53-89.
- Lanne, M., Saikkonen, P. 2005. Modeling conditional skewness in stock returns. Unpublished Working paper, European University Institute Working Paper ECO No. 2005/14.
- Lopez, J. 2001. Evaluating the Predictive Accuracy of Variance Models. *Journal of Forecasting*, **20**: 87-109.
- Markowitz, H. 1952. Portfolio Selection. *Journal of Finance*, **7**:1 77-91.
- Mandelbrot, B. 1963. The variation of certain speculative prices. *Journal of Business*, **36**: 394-417.

- Mittnik, S., Paolella, M. 2003. Prediction of financial downside risk with heavy tailed conditional distributions. CFS Working paper, 2003/04
- Nelson, D. 1991. Conditional heteroskedasticity in asset returns: A new approach. *Econometrica*, **69**: 347-370.
- Ñíguez, T.M., Perote, J. 2004. Forecasting the Density of Asset Returns. Discussion paper EM/04/479, London School of Economics and Political Science
- Oomen, R. 2005. Properties of bias-corrected realized variance under alternative sampling schemes. *Journal of Financial Econometrics*, **3**: 555-577.
- Patton, A. 2006. Volatility Forecast Valuation and Comparison Using Imperfect Volatility Proxies. Quantitative Finance Research Papers 175, University of Technology, Sidney.
- Poon, S. Granger, C. 2003. Forecasting Variance in financial markets: A review. *Journal of Economic Literature*, **41**: 478-540.
- Premaratne, G., Bera, A.K. 2001. Modeling asymmetry and excess kurtosis in stock return data. Working paper, Department of Economics, University of Illinois.
- Poterba, J. Summers, L. 1986. The persistence of volatility and stock market fluctuations. *American Economic Review*, **76**: 1124-1141.
- Rubinstein, M.E. 1973. The fundamental theorem of parameter-preference security valuation. *Journal of Financial and Quantitative Analysis*, **8**:1 61-69.
- Sarno, L. Valente G. 2004. Comparing the Accuracy of Density Forecasts from Competing Models. *Journal of Forecasting*, **23**: 541-557.
- Schwert, G.W. 1990. Stock volatility and the crash of '87. *Review of Financial Studies*, **3**: 77-102.