Empirical Economics (1994) 19:397–418

German Stock Market Dynamics

Rob M. M. J. Bauer

RWTH Aachen, Lehrstuhl für Betriebliche Finanzwirtschaft, Ahornstraße 55, 52056 Aachen, Germany

Frederick G. M. C. Nieuwland

Limburg Institute of Financial Economics, University of Limburg, P.O. Box 616, 6200 MD Maastricht, The Netherlands

Willem F. C. Verschoor[1]

Kempen & Co. N.V., Amsterdam, The Netherlands

Abstract: This article estimates generalized ARCH (GARCH) models for German stock market indices returns, using weekly and monthly data, various GARCH specifications and (non)normal error densities, and a variety of diagnostic checks. German stock return series exhibit significant levels of second-order dependence. Our results clearly demonstrate that for both weekly as well as monthly return series the Student-*t* distribution is superior to the standard normal distribution. In particular, the estimated GARCH-*t* models appear to be reasonably successful in accounting for both observed leptokurtosis and conditional heteroskedasticity from German stock return movements.

JEL Classification System-Numbers: C32, G15

Introduction

The debate regarding the validity of empirical asset return distributions continues to be an issue of central concern in the financial economics literature – see Fama (1991) and Bollerslev, Chou, and Kroner (1990), for instance. These empirical distributional properties have been tested for stock indices returns by Baillie and DeGennaro (1990), French, Schwert, and Stambaugh (1987), Chou (1988), Akgiray (1989), Jorion (1988), Nelson (1989, 1991), Nieuwland (1992), and Wiggins (1989) among others. Probably the most important factor that has

[1] Both Nieuwland and Verschoor gratefully acknowledge research support from the Foundation for the Promotion of Research in Economic Sciences (Ecozoek), the Hague, The Netherlands. The authors are grateful to two anonymous referees and the Editor of this Journal for helpful comments and suggestions.
Bauer gratefully acknowledges research support from Inquire Europe.

generated this considerable interest, is the fact that these distributional properties have a direct impact on the validity of theoretical models in financial economics. Many standard asset pricing models imply the martingale difference model in which price changes are uncorrelated and hence unpredictable in the mean. Furthermore, the Black and Scholes (1973) option pricing model assume a continous time stochastic process for the representation of the price behavior in the form of a geometric Brownian motion where the log of the price relatives are assumed to be independent, and identically normally distributed. Also, the assumptions on the distributional properties of the price process are critical in many tests of the efficient market hypothesis.

As is well documented, empirical distributions of asset returns exhibit leptokurtic behavior and clusters of high and low volatility. The Autoregressive Conditional Heteroskedasticity (ARCH) class of models, introduced by Engle (1982) and generalized (GARCH) by Bollerslev (1986), have been shown to provide a good fit for many financial return time series.[2] GARCH imposes an autoregressive structure on conditional variance, parameterized as a linear function of past squared innovations and lagged conditional variances, allowing volatility shocks to persist over time. This persistence captures the propensity of large absolute returns to cluster in time and can explain the well documented nonnormality and nonstability of empirical asset return distributions – see especially the pioneering works of Mandelbrot (1963) and Fama (1965). Studies by Baillie and Bollerslev (1989), Boothe and Glassman (1987), Hsieh (1989), Jorion (1988), Meese and Rogoff (1983) and Wolff (1987) provide extensive statistical evidence on US Dollar exchange rates and US stock market returns. Overall, the findings overwhelmingly favor the conclusion that the assumption of conditional normality does not capture all the excess kurtosis observed in both high frequency stock and exchange rate returns. Several alternative conditional distributions have consequently been employed in the literature, for instance the Student-*t*, normal-Poisson, generalized error, and normal-lognormal distributions [see, e.g., Akgiray and Booth (1988), Baillie and Bollerslev (1989), Jorion (1988), Hsieh (1989), Nelson (1991), and Nieuwland (1992)].

Interestingly, this branch of literature pays little attention to observed statistical distributions of European asset returns. In the current article we aim to provide extensive statistical evidence for German stock market indices. The observed leptokurtosis may be explained by several classes of models. Therefore, alternative time series processes and distributional specifications characterizing the German stock market are considered. In this article we estimate GARCH models and GARCH-in-mean models for three German stock market indices, using weekly and monthly data over the 1973–1992 period. The (G)ARCH-M model developed by Engle, Lilien and Robins (1987) which we propose can be used in addressing questions regarding the risk-return tradeoff

---

[2] See, e.g., Bollerslev (1987), Lamoureux and Lastrapes (1990), Baillie and Bollerslev (1989), and Lastrapes (1989).

in a time series context where the conditional variance may be time-varying. In addition to exploring normal conditional densities, the Student-*t* distribution is employed. The research will form a body of evidence which can serve as a frame of reference for further research.

The article is organized as follows. In Section 1 we describe our dataset and summary statistics are provided. The methodology and models employed to describe the patterns followed by German stock market indices returns are explained in Section 2. Section 3 presents the main empirical results of the article and Section 4 contains our concluding comments.

I Data and Summary Statistics

Our database contains three weekly and monthly German stock market indices. One of the indices used is the Datastream Total Market Index, which contains all stocks quoted on the Frankfurt stock exchange. In addition, two alternative German stock market indices were included, the Frankfurter Allgemeine Zeitung (FAZ) index and the Commerzbank (COMM) index. The indices are value weighted and adjusted for stock dividends, capital modifications and unusually large dividends payments. The data were obtained from Datastream, a U.K. incorporated data service company. Even though daily stock indices returns are available, we choose to employ weekly and monthly data in order to avoid issues surrounding the day-of-the-week effect with regard to stock return volatility (on which, see Hsieh, 1988, for example). Our sample includes 1003 weekly and 231 monthly observations, ranging from 12 January 1973 through 25 March 1992.[3]

In Tables 1a and 1b we present summary statistics for respectively weekly and monthly stock indices returns. Continuously compounded returns are used, defined as the difference in logarithmic value of the two consecutive observations:

$$R_t = \ln(P_t) - \ln(P_{t-1}) \quad (1)$$

For the period analyzed (January 12th, 1973 through March 25th, 1992) all stock returns series are subject to substantial skewness and kurtosis. Under the normality hypothesis the corresponding measures would have asymptotic distributions of $N(0, 6/T)$ and $N(3, 24/T)$, respectively.[4] Furthermore, to assess the

[3] The weekly quotes used are wednesday closing prices. Wednesdays were chosen because very few holidays occur on that day, and there is no problem of "weekend effects". When the quotes fall on a holiday or weekend (monthly quotes), the next business day is chosen.

[4] T is equal to the number of observations.

Table 1a. Summary statistics of weekly returns (1/12/73–3/25/92)

	Comm-index	FAZ-index	DSindex-DL
mean	0.0010	0.0011	0.0010
st. dev.	0.0222	0.0215	0.0203
t-test	1.3722	1.6104	1.6102
skewness	−0.8395**	−0.9398**	−1.0490**
kurtosis	8.6112**	8.6631**	9.3436**
BJ-test	1430.80**	1485.00**	1862.00**
KS-1	117.57**	147.37**	183.57**
KS-2	1313.20**	1337.61**	1678.41**
LB(50) R	72.22*	74.64*	88.22**
D(50) R	44.94	44.70	47.53
LB(50) $\|R\|$	648.03**	604.55**	621.79**
LB(50) R^2	285.67**	310.97**	332.57**
LB(25) R	28.47	47.11**	60.40**
D(25) R	11.03	24.47	27.03
LB(25) $\|R\|$	213.09**	445.30**	460.02**
LB(25) R^2	176.44**	241.87**	267.12**
Autocor(1)	0.0979	0.1151	0.1457
	(0.0535)	(0.0545)	(0.0571)
Autocor(2)	0.1113	0.1089	0.1307
	(0.0565)	(0.0584)	(0.0622)
Autocor(3)	0.0435	0.0480	0.0448
	(0.0454)	(0.0454)	(0.0484)
Autocor(4)	−0.0231	−0.0125	−0.0051
	(0.0386)	(0.0386)	(0.0378)
Autocor(5)	−0.0260	−0.0354	−0.0342
	(0.0398)	(0.0411)	(0.0404)

* Denotes statistical significance at the 5 % level, ** denotes statistical significance at the 1% level, whereas BJ-test gives the Bera-Jarque test for normality and KS-1 and KS-2, respectively, give the Kiefer-Salmon Normality test for skewness and kurtosis. LB(p) denotes the Ljung-Box test for serial correlation using p lags. D(p) gives the Diebold test for serial correlation using p lags. Autocor(p) denotes the p-th order autocorrelation; robust standard errors are given in parentheses.

Comm-index = Commerzbank-index
FAZ-index = Frankfurter Allgemeine Zeitung-index
DSindex-D1 = Datastream-index Germany

distributional properties of stock indices returns, the Bera Jarque (1982) Normality test and the Kiefer Salmon (1983) Lagrange multiplier normality test are reported in the Tables, where the former represents a joint test using both skewness and kurtosis and the latter being a LM test for normal skewness (KS-1) and normal kurtosis (KS-2), respectively.[5] Overall, the evidence presented

[5] The Bera Jarque test is asymptotically Chi-square(2) distributed, and the Kiefer Salmon normality tests are asymptotically Chi-square(1) distributed.

Table 1b. Summary statistics of monthly returns (1/12/73–3/25/92)

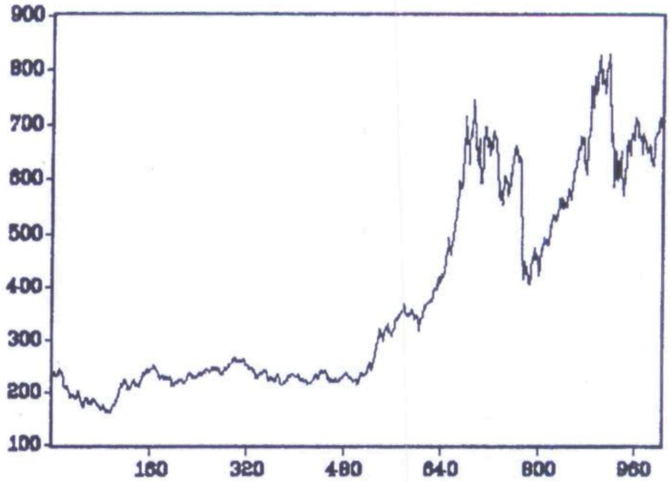
	Comm-index	FAZ-index	DSindex-DL
mean	0.0043	0.0049	0.0047
dev.	0.0495	0.0481	0.0471
t-test	1.3216	1.5359	1.4959
skewness	−1.4692**	−1.6464**	−1.6644**
kurtosis	12.3786**	13.3160**	14.0230**
BJ-test	921.65**	1118.90**	1265.11**
KS-1	82.38**	103.46**	105.73**
KS-2	839.27**	1015.40**	1159.38**
LB(50) R	54.17	56.78	56.38
D(50) R	48.21	51.04	52.72
LB(50) $\|R\|$	77.09**	77.82**	83.50**
LB(50) R^2	24.83	20.25	19.04
LB(25) R	27.22	30.02	29.42
D(25) R	25.66	28.37	28.11
LB(25) $\|R\|$	43.19*	40.33*	45.27**
LB(25)R^2	14.99	10.01	10.80
Autocor(1)	0.0840	0.0969	0.1092
	(0.0681)	(0.0693)	(0.0675)
Autocor(2)	0.0549	0.0550	0.0515
	(0.0646)	(0.0627)	(0.0635)
Autocor(3)	0.0428	0.0572	0.0464
	(0.0731)	(0.0705)	(0.0774)
Autocor(4)	−0.0412	−0.0254	−0.0254
	(0.0768)	(0.0741)	(0.0772)
Autocor(5)	−0.0219	−0.0397	−0.0358
	(0.0576)	(0.0570)	(0.0592)

* Denotes statistical significance at the 5 % level, ** denotes statistical significance at the 1% level, whereas BJ-test gives the Bera-Jarque test for normality and KS-1 and KS-2, respectively, give the Kiefer-Salmon Normality test for skewness and kurtosis. LB(p) denotes the Ljung-Box test for serial correlation using p lags. D(p) gives the Diebold test for serial correlation using p lags. Autocor(p) denotes the p-th order autocorrelation; robust standard errors are given in parentheses.

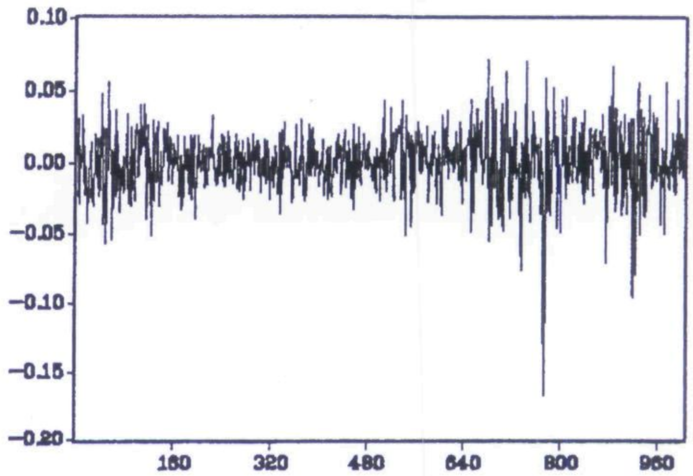
suggests a consistent rejection of the normality hypothesis. Thus, in spite of the notion that leptokurtic unconditional densities of ARCH processes approach normality by temporal aggregation – see Diebold (1988) and Baillie and Bollerslev (1989) – it appears that the monthly series used here may be characterized as highly leptokurtic. In addition, Tables 1a and 1b report the Ljung-Box (1978) and Diebold (1988) test statistics for kth-order serial correlation in stock returns, R_t, absolute returns, $|R_t|$, and squared returns, R_t^2, respectively. The Ljung-Box (LB) statistic is given by:

$$LB(p) = T(T+2) \sum_{\tau=1}^{p} (T-\tau)^{-1} \hat{\rho}^2(\tau) \overset{a}{\sim} \chi^2(p) \ , \tag{2}$$

where $\hat{\rho}(\tau)$ is the τ-th order autocorrelation coefficient. As the choice of the appropriate laglength is somewhat arbitrary, both LB(50) and LB(25) were calculated. Diebold (1988) proposes an adjusted Ljung-Box test statistic to allow for heteroskedasticity. Diebold (1988) showed that in the presence of ARCH effects, the Ljung-Box test has a larger empirical size than a nominal test size of 5%, because the asymptotic variance of the autocorrelations under ARCH is larger than under the null of Gaussian white noise. Diebold (1988) proposed a for conditional heteroskedasticity adjusted Ljung-Box statistic which preserves the proper size:



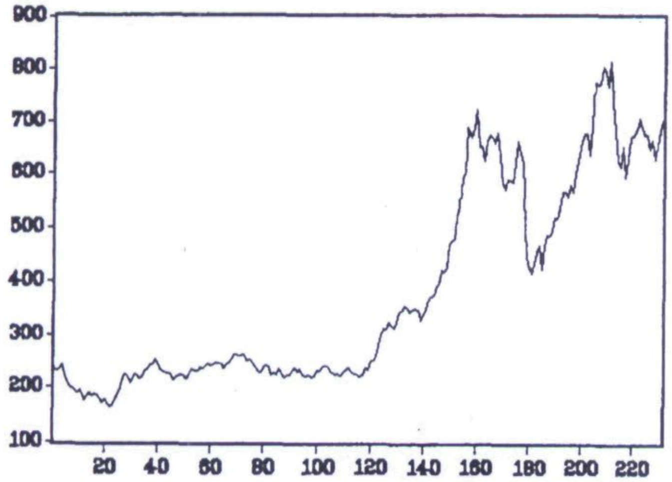
Graph 1. Weekly prices FAZ index (1/12/1973–3/25/1992)



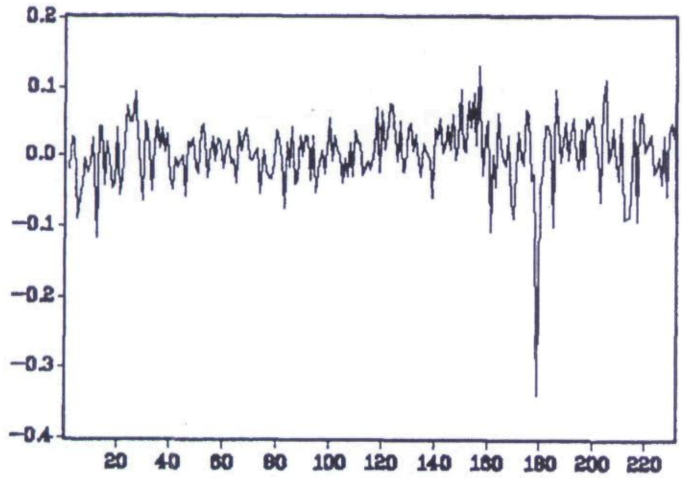
Graph 2. Weekly returns FAZ index (1/12/1973–3/25/1992)

$$D(p) = T(T+2) \sum_{\tau=1}^{p} (1 + (\gamma^2(\tau)/\sigma^4))^{-1} (T-\tau)^{-1} \hat{\rho}^2(\tau) \overset{a}{\sim} \chi^2(p) \ , \tag{3}$$

where $\gamma^2(\tau)$ is the τ-th order sample autocovariance of the squared process, and σ is the sample standard deviation of the data. Any evidence on serial correlation in German stock returns using the standard Ljung-Box test vanishes, when allowing for conditional heteroskedasticity. In contrast to the monthly stock returns, weekly stock indices returns exhibit substantial first (and second) order autocorrelation. The absolute and squared weekly stock returns exhibit sub-



Graph 3. Monthly prices FAZ index (1/1973–3/1992)



Graph 4. Monthly returns FAZ index (1/1973–3/1992)

stantially more autocorrelation than the raw data, R_t, which is indicative of strong non-linear dependence, conditional heteroskedasticity, and the clustering phenomenon. One notes that for monthly stock indices returns, only significant serial autocorrelations are presented in the absolute return series, which could be indicative of no conditional heteroskedasticity.

In Graphs 1 and 3 examples of the weekly and monthly FAZ index are presented. Graphs 2 and 4 display corresponding figures of log differences in these index levels, corresponding (approximately) to percentage changes in these levels. The figures indicate substantial variation in the FAZ index series. Note that the variation of the indices (return) series is significantly smaller for the period 1973–1982 than for the period 1982–1992. The shifts of the FAZ index due to both stock market crashes in the late eighties can be clearly detected in these figures.[6] Graphs 2 and 4 capture another feature of stock market volatility: volatility comes in waves. That is, large (small) changes in share prices tend to be followed by large (small) changes and this phenomenon is more marked for higher frequency series.

In order to remove possible first order autocorrelation and to differentiate between correlation effects and heteroskedasticity effects, the following OLS regression was fitted:

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t \tag{4}$$

In order to test for the presence of heteroskedasticity in German stock returns, two different approaches are employed. First the Lagrange multiplier tests for autoregressive conditional heteroskedasticity – see Breusch and Pagan (1979) – are performed, and secondly a non-parametric test based on finite-state homogeneous Markov chains – see Gregory (1989) – is applied.[7] Using Monte Carlo analysis Gregory (1989) concludes that under other distributions than the Normal the LM test is biased towards the null hypothesis of no ARCH, and that the Markov Chain test is superior to the LM test in terms of better finite sample properties. Both tests only require estimation under the null hypothesis of no heteroskedasticity and are appropriate under all distributional assumptions.[8] The results of the LM and Markov chain tests for the presence of heteroskedasticity are given in Tables 2a and 2b.

Overall, the evidence presented suggests a fairly consistent rejection of the hypothesis of no heteroskedasticity for weekly stock indices return series. In con-

---

[6] As a suggestion for future research, it might be interesting to replicate all model estimations up to the time the first big crash occured and compare these results with our initial ones.

[7] For a more detailed description, see Appendix A, Nieuwland (1992), and Nieuwland and Verschoor (1992), for instance.

[8] Weiss (1986) has shown that the proposed LM-test is appropriate for non-normal distributions, provided some moment conditions are satisfied. The Markov chain test is completely distribution free.

Table 2a. ARCH tests for weekly returns

	Comm-index	FAZ-index	DSindex-DL
LM(1)	53.82**	59.20**	68.52**
LM(2)	106.88**	123.44**	141.98**
LM(5)	113.05**	129.69**	154.62**
LRIM1	24.25**	22.93**	18.40**
LRIM2	36.97**	27.31**	28.91**
LRM1M2	12.72**	4.38	10.51**

Table 2b. ARCH tests for monthly returns

	Comm-index	FAZ-index	DSindex-DL
LM(1)	0.02	0.02	0.05
LM(2)	0.02	0.04	0.85
LM(5)	0.43	0.29	1.03
LRIM1	3.48	0.60	0.91
LRIM2	3.53	1.02	1.68
LRM1M2	0.05	0.42	0.77

*, ** Denote significance at respectively the 5% and the 1% level. The LM(p) tests are computed as TR^2 from a regression of squared residuals on a constant and p lags, and are asymptotically $\chi^2(p)$-distributed. LRIM 1 is a Likelihood Ratio test of independence against a first order Markov Chain, and is distributed as $\chi^2(1)$. LRIM2 is a Likelihood Ratio test of independence against a second order Markov Chain, and is distributed as $\chi^2(3)$. LRM1M2 is a Likelihood Ratio test of a first order against a second order Markov Chain, and is distributed as $\chi^2(2)$.

trast, for monthly stock returns we cannot reject the null hypothesis of no heteroskedasticity in all cases.

II Modeling German Stock Market Dynamics: Methodology

In the academic literature it is agreed upon that empirical distributions of stock market returns exhibit fatter tails than one expects from a normal distribution. See, for example, Akgiray (1989) and Bollerslev et al. (1990). In Tables 1a and 1b we have provided evidence confirming this fact. All series, on a weekly and monthly basis, suffer from substantial leptokurtosis. These fat tails, or observed leptokurtosis, may be explained by several classes of models. In this article we will concentrate on two possible explanations. A first explanation suggests that German stock market indices returns can be described by a normal distribution with time-varying parameters (see Hsieh, 1989). Second, we consider the possibility that German stock return series are generated by a conditionally leptokurtic distribution. These considerations lead us to the maximum-likelihood estimation of the following stochastic processes:

$$R_t = \phi_0 + \phi_1 R_{t-1} + \varepsilon_t , \qquad \varepsilon_t = \eta_t h_t^{1/2} , \qquad \eta_t \sim D(0, 1) \tag{5}$$

$$h_t = \alpha_0 + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j h_{t-j} \qquad \alpha_0 > 0 , \qquad \alpha_i, \beta_j \geq 0 \tag{6}$$

The empirical distribution of variables generated by these processes are fat-tailed, compared to the normal distribution. Here $D(0, 1)$ can be any symmetrical distribution with zero mean and unit variance. The conditional variance, h_t, is a linear function of squared lagged residuals and lagged conditional variances. The fact that conditional variances are allowed to depend on past realized variances is particularly consistent with the actual volatility pattern of the stock market where there are both stable and unstable periods. In the remainder of this article, we restrict our attention to a GARCH (1, 1) specification since it has been shown to be a parsimonious representation of conditional variance that adequately fits many economic time series – see Bollerslev (1986), Chou (1988), Akgiray (1989), Baillie and DeGennaro (1990), Nieuwland (1992), and Poon and Taylor (1992). Equation (6) now reduces to:

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{7}$$

The model, given by equations (5) and (7), will be estimated under two different assumptions for the conditional error distribution. Most commonly used in GARCH applications is the standard normal distribution. Under conditional normality, the conditional standard deviation can be seen as the stochastic volatility of the process, see Taylor (1990). This has the unfavorable implication that the volatility during a particular time period is known at the end of the previous time period, thereby reducing the impact on volatility, of news during the period, to zero. An empirical shortcoming of the assumption of conditional normality pertains to the frequent leptokurtosis of standardized GARCH-normal residuals. If conditional normality is appropriate this phenomenon should not occur. The conditional density function for ε_t under normality reads:

$$f_n(\varepsilon_t|\psi_{t-1}) = (\sqrt{2\pi h_t})^{-1} \exp(-\varepsilon_t^2/2h_t) \tag{8}$$

$$\psi_{t-1} = \{\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots\} \tag{9}$$

The associated log-likelihood function becomes:

$$LL_n(\theta) = \frac{1}{2}\left(-T \ln 2\pi - \sum_{t=1}^{T} \ln h_t - \sum_{t=1}^{T} \varepsilon_t^2/h_t \right) \tag{10}$$

$$\theta = (\phi_0, \phi_1, \alpha_0, \alpha_1, \beta_1) \tag{11}$$

As an alternative to conditional normality we propose a scaled Student t-distribution. Theoretically and, as will become clear, also empirically this distribution should be preferred to a conditional normal distribution. The Student t-distribution is conditionally fat-tailed, a feature which can explain – next to conditional heteroskedasticity – unconditional fat tails. Furthermore it can be seen as a continuous variance mixture of normals, where the mixing variable followes an inverted Gamma-1 distribution. Such a mixture distribution is desirable when information arrives randomly during a time period, see Tauchen and Pitts (1983), because the random character of information is conveyed by the unobservable mixing variable. In this case the stochastic volatility for period t is dependent on unexpected news, and is in general not equal to the conditional standard deviation for period t, which is known at the end of period $t-1$. Finally, the observed leptokurtosis of GARCH residuals does not pose a problem under a conditional t-distribution. The conditional density function for ε_t depends on a degrees-of-freedom parameter ν and is given by:

$$f_\nu(\varepsilon_t|\psi_{t-1}) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\sqrt{\pi}\,\Gamma\left(\frac{\nu}{2}\right)}((\nu-2)h_t)^{-1/2}\left(1+\frac{\varepsilon_t^2}{h_t(\nu-2)}\right)^{-(\nu+1)/2}, \qquad \nu > 2 \quad (12)$$

The associated log-likelihood function reads:

$$\begin{aligned} LL_\nu(\theta) = {} & T\left(\ln\Gamma\left(\frac{\nu+1}{2}\right) - \ln\Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2}\ln(\pi(\nu-2))\right) \\ & + \sum_{t=1}^{T}\left(-\frac{1}{2}\ln h_t - \left(\frac{\nu+1}{2}\right)\ln\left(1+\frac{\varepsilon_t^2}{(\nu-2)h_t}\right)\right) \end{aligned} \quad (13)$$

$$\theta = (\phi_0, \phi_1, \alpha_0, \alpha_1, \beta_1, \nu) \quad (14)$$

The scaled t-distribution approaches normality when $\nu \to \infty$. In addition, we estimate a GARCH(1, 1)-in-mean model, initially developed by Engle, Lilien and Robins (1987), which can be used in adressing questions regarding the risk-return tradeoff in a time series context where the conditional variance may be time-varying. In order to provide empirical evidence for a relationship between stock index returns and conditional variances of the underlying process, the following model was fitted:

$$R_t = \phi_0 + \phi_1 R_{t-1} + \delta h_t + \varepsilon_t \quad (15)$$

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \quad (16)$$

Again, estimation will be performed under the normal and the Student t-distribution. Prominent examples of applications of this model to stock index returns are French, Schwert and Stambaugh (1987), Chou (1988), and Baillie and DeGennaro (1990). Although the model has been used extensively, Backus, Gregory and Zin (1989) challenge the usefulness of the GARCH-in-mean models by stating that there is no explicit theoretical relationship between the risk premium and the conditional variance. Furthermore Pagan and Ullah

(1988) pointed out that the estimates for the parameters in the conditional mean equation are not asymptotically independent from the estimates of the parameters in the conditional variance equation. Misspecification of the GARCH part of the model, therefore, leads to biases and inconsistencies in the mean part of the model. Consistent with the first criticism, Baillie and DeGennaro (1990) obtained insignificant estimates for δ, in their analysis for daily and monthly returns.

In the next section we present and discuss our estimation results for the GARCH and the GARCH-in-mean models, respectively. Likelihood Ratio Tests and diagnostic checks will also be included.

III Estimation Procedures and Empirical Results

Maximum likelihood estimates of the parameters and their heteroskedasticity consistent asymptotic standard errors were obtained by numerical methods using the Berndt, Hall, Hall and Hausman (1974) (BHHH) algorithm.[9] In Tables 3a through 4b, the estimation results are reported for the stochastic processes described in the previous section.

First, we present the results for weekly stock indices return series. Inspection of the Tables reveals several notable facts. GARCH parameters and autocorrelation coefficients are highly significant for all weekly stock indices and model specifications. For the conditional t-distribution, the degrees of freedom parameter, v, is significant in all cases. Note that $\alpha_1 + \beta_1$ is near unity for all indices, indicating infinite persistence in the volatility shocks, or IGARCH behavior (see Engle and Bollerslev (1986)]. As conjectured by Lamoureux and Lastrapes (1990), this may be the result of not accounting for discrete shifts in regimes which affect the level of the unconditional variances, and, therefore can lead to misspecification of the GARCH model.[10] Comparing the results of the standard normal GARCH model to the results of the GARCH-t model, we notice that the volatility persistence has been reduced. For the GARCH-in-mean model, all the estimated models result in statistically insignificant δ coefficients.[11] Tables 4a and 4b present the estimation results for monthly stock index return series. It is interesting to note that the monthly estimates of the GARCH models demonstrate less significant parameter coefficients. This is consistent with the results of

[9] All calculations were performed with the software package GAUSS.

[10] Lastrapes (1989) finds that persistence of exchange rate volatility decreases when regime shifts are accounted for, diminishing the likelihood of integrated-in-variance processes.

[11] The Tables are omitted from the text. For the interested readers, however, the estimation results for the GARCH-in-mean models are available upon request.

Table 3a. GARCH(1, 1)-normal estimates of weekly returns

	Comm-index	FAZ-index	DSindex-D1
φ_0	0.000871	0.000982	0.000786
	(0.000599)	(0.000585)	(0.000555)
φ_1	0.094064*	0.111538**	0.136151**
	(0.039800)	(0.038225)	(0.037266)
α_0	0.000008	0.000008	0.000006
	(0.000006)	(0.000007)	(0.000005)
α_1	0.107042*	0.101380*	0.088571*
	(0.044798)	(0.047520)	(0.038680)
β_1	0.880662**	0.884134**	0.899572**
	(0.051068)	(0.058402)	(0.047429)
$\alpha_1 + \beta_1$	0.987704	0.985514	0.988143
LL	2505.96	2535.94	2599.62

Table 3b. GARCH(1, 1) *t*-distribution estimates of weekly returns

	Comm-index	FAZ-index	DSindex-D1
φ_0	0.001170*	0.001393**	0.001186**
	(0.000557)	(0.000547)	(0.000507)
φ_1	0.074317*	0.087549**	0.120357**
	(0.033799)	(0.033343)	(0.033371)
α_0	0.000011*	0.000011*	0.000009*
	(0.000005)	(0.000005)	(0.000004)
α_1	0.100409**	0.098237**	0.089402**
	(0.030853)	(0.030876)	(0.032872)
β_1	0.877022**	0.876119**	0.886785**
	(0.036790)	(0.037891)	(0.032872)
$\alpha_1 + \beta_1$	0.977431	0.974356	0.976187
V	8.734596**	9.085585**	8.280524**
LL	2521.34	2550.64	2618.05

* Denotes statistical significance at the 5% level, ** denotes statistical significance at the 1% level. Heteroskedasiticity consistent standard errors are given in parentheses. *LL* denotes the log-Likelihood Values.

Table 2b, which indicates insignificant ARCH effects in monthly stock index return series. However, the β_1 estimates are statistically significant for all stock indices. The degree of freedom parameter, *v*, is significant in all cases and, contrary to what one might expect, does not increase when moving from a weekly to a monthly time interval. This is surprising as the general notion in the finan-

Table 4a. GARCH(1, 1)-normal estimates of monthly returns

	Comm-index	FAZ-index	DSindex-D1
φ_0	0.002231 (0.002990)	0.003178 (0.002987)	0.002977 (0.002745)
φ_1	0.062990 (0.062275)	0.077143 (0.062665)	0.087397 (0.061478)
α_0	0.000025 (0.000042)	0.000029 (0.000045)	0.000027 (0.000045)
α_1	0.067639 (0.037761)	0.064116 (0.040911)	0.078193 (0.047609)
β_1	0.929123** (0.016713)	0.929959** (0.015613)	0.918033** (0.021475)
$\alpha_1 + \beta_1$	0.996762	0.994075	0.996226
LL	373.05	377.83	384.87

Table 4b. GARCH(1, 1) *t*-distribution estimates of monthly returns

	Comm-index	FAZ-index	DSindex-D1
φ_0	0.004833 (0.002954)	0.005698 (0.002866)	0.005154 (0.002758)
φ_1	0.043347 (0.062208)	0.062848 (0.063220)	0.070670 (0.058937)
α_0	0.000050 (0.000027)	0.000052 (0.000028)	0.000047 (0.000026)
α_1	0.033715 (0.019878)	0.031446 (0.021668)	0.033186 (0.021848)
β_1	0.942998** (0.018894)	0.940728** (0.023211)	0.940262** (0.022837)
$\alpha_1 + \beta_1$	0.975723	0.972174	0.973448
V	6.326147*	6.445831*	6.391206*
LL	386.13	393.19	399.20

* Denotes statistical significance at the 5% level, ** denotes statistical significance at the 1% level. Heteroskedasiticity consistent standard errors are given in parentheses. *LL* denotes the log-Likelihood Value.

cial economics literature suggests that conditional distributions of asset returns, when aggregated, approach normality – see Drost and Nijman (1991), and Nieuwland (1992), for instance. If this were indeed true we would expect to see a significant increase in the value of v. This is however not the case. For the GARCH-in-mean model, the estimated δ coefficients are greater than those for

Table 5a. Generalized likelihood ratio tests, weekly data

	[a] $\chi^2_{(1)}$	[b] $\chi^2_{(1)}$	[c] $\chi^2_{(1)}$	[d] $\chi^2_{(1)}$
Commerz-index	30.76	9.86	0.00	20.74
	(0.000)	(0.002)	(1.000)	(0.000)
FAZ-index	29.40	10.06	0.04	19.38
	(0.000)	(0.002)	(0.840)	(0.000)
DS-index	36.86	10.36	0.06	26.56
	(0.000)	(0.001)	(0.810)	(0.000)

[a] GARCH Normal Model against GARCH *t*-Distribution Model; [b] GARCH Normal Model against GARCH in Mean Normal Model; [c] GARCH *t*-Distribution Model against GARCH in Mean *t*-Distribution Model; [d] GARCH in Mean Normal Model against GARCH in Mean *t*-Distribution Model; *P*-values are given in parentheses.

weekly indices returns, however, as before, the GARCH-*M* parameter remains insignificantly different from zero.[12]

Given the above results, it is interesting to compare the relative fit of the alternative models. We employ generalized likelihood ratio tests to compare nested models. Such nested models can be tested using the generalized likelihood ratio:

$$\Lambda = \frac{\sup_{\phi \in \Phi 0} L(\phi; x)}{\sup_{\phi \in \Phi} L(\phi; x)} \tag{17}$$

of the maximized likelihood values under the null and under the encompassing parameter space, Φ, which also includes the alternative hypothesis. Here, $L(.;.)$ is the likelihood function, ϕ is the parameter vector and x is the relevant set of observations. Under the null Φ_0, the statistic $-2 \ln \Lambda$ has a Chi-square distribution with degrees of freedom equal to the difference in the number of parameters between the two models. Thus, the improvement in the maximized likelihood indicates to what extent an enlarged specification helps in fitting the data.

Tables 5a and 5b present the generalized likelihood ratio tests to compare the relative fit of the models employed. All but nine of the *p*-values associated with the chi-square statistics are close to zero. Thus, the generalized likelihood ratio tests reject the simpler model in favor of the more complicated model in most of the cases. It is clear that for both weekly and monthly return series the Student *t*-distribution is superior to the standard normal distribution. For the stock indices, the generalized likelihood ratio tests cannot reject the simpler (GARCH)

[12] In addition to German stock market indices, we estimated GARCH models for individual stock returns. The individual asset dataset comprises weekly and monthly returns for five actively traded stocks quoted on the Frankfurt stock exchange. The results are available upon request.

Table 5b. Generalized likelihood ratio tests, monthly data

	[e] $\chi^2_{(1)}$	[f] $\chi^2_{(1)}$	[g] $\chi^2_{(1)}$	[h] $\chi^2_{(1)}$
Commerz-index	26.16	0.30	0.52	26.38
	(0.000)	(0.580)	(0.470)	(0.000)
FAZ-index	30.72	0.36	0.58	30.94
	(0.000)	(0.550)	(0.450)	(0.000)
DS-index	28.56	0.34	0.44	28.76
	(0.000)	(0.560)	(0.510)	(0.000)

[e] GARCH Normal Model against GARCH *t*-Distribution Model; [f] GARCH Normal Model against GARCH in Mean Normal Model; [g] GARCH *t*-Distribution Model against GARCH in Mean *t*-Distribution Model; [h] GARCH in Mean Normal Model against GARCH in Mean *t*-Distribution Model; *P*-values are given in parentheses.

Table 6a. Diagnostics for GARCH *t*-distribution models, weekly data

	Commerz-index	FAZ-index	DS-index
Skewness	−0.38**	−0.47**	−0.49**
Kurtosis	4.60**	4.64**	4.94**
BJ-Test	133.01**	148.54**	196.38**
KS-1	25.34**	36.65**	40.11**
KS-2	107.67**	111.89**	156.27**
F(1)	8.68**	9.26**	5.82**
F(2)	5.91**	6.41**	4.52**
F(5)	3.13**	3.53**	2.61**

The BJ-test denotes the Bera Jarque test for normality; KS-1 and KS-2 pertain to the Kiefer Salmon Normality test for respectively skewness and kurtosis.

Table 6b. Diagnostics for GARCH *t*-distribution models, monthly data

	Commerz-index	FAZ-index	DS-index
Skewness	−1.20**	−1.51**	−1.39**
Kurtosis	10.75**	12.63**	12.27**
BJ-Test	628.64**	971.73**	893.44**
KS-1	55.48**	86.47**	73.73**
KS-2	573.16**	885.26**	819.71**
F(1)	0.059	0.021	0.027
F(2)	0.050	0.027	0.025
F(5)	0.062	0.046	0.045

The BJ-test denotes the Bera Jarque test for normality; KS-1 and KS-2 pertain to the Kiefer Salmon Normality test for respectively skewness and kurtosis.

model in favor of the more complicated (GARCH-in-mean) model in most of the cases.

In order to determine the adequacy of the statistical specification, the models are subjected to diagnostic checks on the standardized residuals:

$$z_t = \frac{\hat{\varepsilon}_t}{\sqrt{\hat{h}_t}} , \tag{18}$$

where $\hat{\varepsilon}_t$ is the residual from equation (5) and $\sqrt{\hat{h}}$ is the estimated conditional variance from equations (7). From Jensen's inequality it follows that the standardized residuals, z_t, should demonstrate less absolute skewness and should be thinner tailed than their unconditional raw data counterparts. Any strong violation of this rule should be regarded as evidence of model misspecification – see Hsieh (1989).

The diagnostics for the estimated GARCH-*t* models are presented in Tables 6a and 6b. Overall, the evidence presented suggests a less consistent rejection of the normality hypotheses as compared with the results of Tables 1a and 1b. In addition, we find that in most cases the estimated statistics – the BJ-test, KS-1, and KS-2 – are smaller than those reported by Tables 1a and 1b thus supporting our model specifications. In particular, for weekly stock indices the GARCH-*t* model is reasonably successful at removing excess kurtosis and skewness in all cases. However, for the monthly indices skewness and kurtosis of the standardized residuals are only marginally smaller than their raw data counterparts. In order to test for remaining heteroskedasticity, a residual-based test of the models may be carried out by regressing $(\hat{\varepsilon}_t^2 - \hat{h}_t^2)/\hat{h}_t^2$ on $1/\hat{h}_t^2$ and on one to two and one to five lags, respectively, of the dependent variable and testing whether the estimated coefficients are significantly different from zero by a conventional F-test.[13] The results are reported under F(1), F(2) and F(5). These statistics follow F(2,998), F(3,996) and F(6,990) distributions for weekly return series and F(2,226), F(3,224) and F(6,218) for monthly return series, respectively. For the GARCH-*t* models, rejection of the null hypothesis of no heteroskedasticity occurs for all of the weekly stock indices return series, whereas all of the monthly stock indices return series result in statistically insignificant test statistics at the 1% significance level.

IV Conclusions

In this paper we have extensively studied the statistical properties of German stock market returns. These properties are important for the pricing of stock

[13] See Domowitz and Hakkio (1985) and Pagan and Hall (1983) for a discussion of the general principles leading to such a test statistic.

options and for international asset pricing models. GARCH models may be used to further understand the relationship between volatility and expected returns.

Our results demonstrate clearly that for both weekly as well as monthly return series the Student *t*-distribution is superior to the standard normal distribution. GARCH-*t* models fit the data significantly better than the standard normal GARCH using a variety of Goodness-of-fit diagnostics. In particular, the estimated GARCH-*t* models appear to be reasonably successful in accounting for both observed leptokurtosis and conditional heteroskedasticity from German stock return movements. This does not preclude the possibility that the GARCH-*t* model for German stock indices is misspecified, but it is the best alternative considered. Furthermore, the estimated GARCH-in-mean models do not support a statistically significant relationship between a stock portfolio's return and its own volatility. Moreover, the empirical finding of near integrated GARCH behavior suggests that the conditional variances have very long memories and are highly sensitive to initial conditions. Exponential GARCH models may perform better in this respect. In the Student's Autoregressive model with dynamic heteroskedasticity (STAR) the memory of the conditional variance is as long as the sample while the question of unit roots does not arise.[14] The STAR approach seems to be a promising alternative specification to the GARCH and may be a fruitful route for further research into the dynamics of German stock market returns.

Many issues in asset pricing and portfolio allocation decisions can only be meaningfully analyzed in a multivariate context. The evidence presented suggests that for multivariate analysis, the assumption of multivariate conditional normality fails to be a reasonable empirical working hypothesis. This issue requires further investigation in future research.

Appendix A

This appendix outlines the construction of the ARCH test based on finite-state Markov Chains. A detailed description of the construction of this test can be found in Gregory (1989) and Nieuwland (1992). Basically this test is equivalent to a test for independence in a two way contingency table. The first step in the construction of the test is to obtain the squared residuals from the estimation of equation (4) the main text. The next step is to apply a discretization rule, by which the squared residuals are divided into different categories or states. The

[14] Recently, the STAR model has been developed by Spanos (1991, 1992) and successfully applied to exchange rate data by McGuirk, Robertson and Spanos (1993).

rule applied here is to mark the residuals as being high or low, with the sample median as the boundary. This results in a two-state definition: the squared residual are either low (state 1) or high (state 2). Next it is assumed that that $\{\varepsilon_t^2\}$, which values are now either 1 or 2, possesses the Markov property that says that the probability distribution of ε_t^2 conditional on its entire past equals the probability distribution conditional on its first previous value only:

$$P(e_t^2 = j | e_{t-1}^2 = i, e_{t-2}^2, e_{t-3}^2, \ldots) = P(e_t^2 = j | e_{t-1}^2 = i) = \lambda_{ijt} \qquad \text{A.1}$$

Where $i, j = 1, 2$; and $t = 2, \ldots, T$. λ_{ijt} is the probability of being in state j at time t, coming from state i at time $t-1$. The transition probabilities are assumed to be time independent (the subscript t can now be deleted), allowing us to define $\{\varepsilon_t^2\}$ as a homogeneous first order Markov chain. A second order homogeneous Markov chain will also be used to test for second order ARCH effects. In this case the transition probabilities are defined as:

$$\lambda_{ijk} = P(e_t^2 = k | e_{t-1}^2 = j, e_{t-2}^2 = i) \qquad i, j = 1, 2 \qquad t = 2, \ldots, T \qquad \text{A.2}$$

λ_{ijk} is the probability of being in state k at time t, coming from state j at time $t-1$ and from state i at time $t-2$. The ML estimates of the transition probabilities for the first order Markov chain are determined as $\lambda_{ij} = n_{ij}/(n_{i1} + n_{i2})$, where n_{ij} is the number of times that a transition from i to j is observed. For the second order Markov chain they are determined as $\lambda_{ijk} = n_{ijk}/(n_{ij1} + n_{ij2})$, where n_{ijk} is the number of transitions from i to j to k. For the first order Markov chain the log-likelihood value at the ML estimates becomes:

$$\overline{LLM1} = \sum_{i=1}^{2} \sum_{j=1}^{2} n_{ij} \ln(n_{ij}/N_i) \qquad \text{A.3}$$

where N_i is the number of times that state i is observed. For the second order Markov chain the log-likelihood value at the ML estimates becomes:

$$\overline{LLM2} = \sum_{i=1}^{2} \sum_{j=1}^{2} \sum_{k=1}^{2} n_{ijk} \ln(n_{ijk}/N_{ij}) \qquad \text{A.4}$$

where N_{ij} is the number of times that a transition from i to j occurs. These likelihoods are necessary to develop tests for the null hypothesis of homoskedasticity. If this hypothesis were true then there would be no serial dependence in the squared residuals and the transition probabilities would be without any structure. This means that the probability of observing a current state k, is independent of previous realizations (i and j). This translates into:

$$H_0\colon \lambda_{ij} = \lambda_j \qquad \text{A.5}$$

when testing independence against a first order Markov chain, and into:

$$H_0\colon \lambda_{ijk} = \lambda_k \qquad \text{A.6}$$

when testing independence against a second order Markov chain. The ML estimates under both null hypotheses are: $\lambda_j = N_j/N$; N is the total number of observations used in the estimation procedure. The log-likelihood value under the null at the ML estimates is:

$$\overline{LLI} = \sum_{i=1}^{2} N_i \ln(N_i/N) \tag{A.7}$$

Two likelihood ratio tests can now be constructed: The first one is a test of independence against a first order homogeneous Markov chain:

$$LR(I/M1) = -2(\overline{LLI} - \overline{LLM1}) \overset{a}{\sim} \chi^2(1) \tag{A.8}$$

The second one is a test of independence against a second order homogeneous Markov chain:

$$LR(I/M2) = -2(\overline{LLI} - \overline{LLM2}) \overset{a}{\sim} \chi^2(3) \tag{A.9}$$

These are the tests that are used in the main text.

References

Akgiray V (1989) Conditional heteroskedasticity in time series of stock returns: Evidence and forecasts. Journal of Business 1:55–80

Akgiray V, Booth GG (1988) Mixed diffusion-jump process modelling of exchange rate movements. Review of Economics and Statistics 70:631–37

Backus DK, Gregory AW, Zin SE (1989) Risk premiums in the term structure: Evidence from artificial economies. Journal of Monetary Economics 24:371–99

Baillie RT (1987) Inference in dynamic models containing "surprise" variables. Journal of Econometrics 35:101–17

Baillie RT (1989) Econometric tests of rationality and market efficiency. Econometric Review 8: 151–86

Baillie R, Bollerslev T (1989) The message in daily exchange rates: A conditional variance tale. Journal of Business and Economic Statistics 7:297–305

Baillie R, DeGennaro RP (1990) Stock return and volatility. Journal of Financial and Quantitative Analysis 2:203–14

Bera AK, Jarque CM (1982) Model specification tests: A simultaneous approach. Journal of Econometrics 20:59–82

Berndt EK, Hall B, Hall R, Hausman J (1974) Estimation and inference in nonlinear structural models. Annals of Economic and Social Measurement 4:653–65

Black F, Scholes M (1973) The pricing of options and corporate liabilities. Journal of Political Economy 81:637–59

Bollerslev T (1986) Generalized autoregressive conditional heteroskedasticity. Journal Econometrics 31:307–27

Bollerslev T, Wooldridge J (1989) Quasi maximum likelihood estimation of dynamic models with time varying covariances. Unpublished manuscript, Department of Economics, MIT

Bollerslev T, Chou R, Jayaraman N, Kroner KF (1992) ARCH modeling in finance: A selective review of the theory and empirical evidence, with suggestions for future research. Journal of Econometrics 52:5–59

Boothe P, Glassman D (1987) The statistical distributions of exchange rates: Empirical evidence and economic implications. Journal of International Economics 22:153–67

Breusch TS, Pagan AR (1979) A simple test for heteroskedasticity and random coefficient variation. Econometrica 47:1287–94

Chou RY (1988) Volatility persistence and stock valuations: Some empirical evidence using GARCH. Journal of Applied Econometrics 3:279–294

Diebold FX (1988) Empirical modeling of exchange rate dynamics. Springer Verlag New York

Domowitz I, Hakkio CS (1985) Conditional variance and the risk premium in the foreign exchange market. Journal of International Economics 19:47–66

Drost FC, Nijman TE (1993) Temporal aggregation of GARCH processes. Econometrica 61(4): 909–927

Engle RF (1982) Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. Econometrica 50:987–1008

Engle RF, Bollerslev T (1986) Modeling the persistence of conditional variances. Econometric Reviews 5:1–50

Engle RF, Lilien DM, Robbins RP (1987) Estimating time-varying risk premia in the term structure: The ARCH-M model. Econometrica 55:391–407

Fama EF (1965) The behavior of stock market prices. Journal of Business 38:34–105

Fama EF (1991) Efficient capital markets: II. Journal of Finance 5:1575–1617

French KR, Schwert GW, Stambaugh RF (1987) Expected stock returns and volatility. Journal of Financial Economics 19:3–29

Gregory A (1989) A non-parametric test for autoregressive conditional heteroskedasticity: A markov-chain approach. Journal of Business and Economic Statistics 7:307–17

Hsieh DA (1989) Modeling heteroskedasticity in daily foreign-exchange rates. Journal of Business and Economic Statistics 7:307–17

Jorion P (1988) On jump processes in the foreign exchange and stock markets. Review of Financial Studies 1:427–45

Kiefer N, Salmon M (1983) Testing normality in econometric models. Economics Letters 11:123–27

Lamoureux C, Lastrapes W (1990) Persistence in variance, structural change and the GARCH model. Journal of Business and Economic Statistics 225–34

Lastrapes WD (1989) Exchange rate volatility and U.S. monetary policy: An ARCH application. Journal of Money, Credit, and Banking 21:66–77

Ljung G, Box G (1978) On a measure of lack of fit in time series models. Biometrika 65:297–303

McGuirk A, Robertson J, Spanos A (1993) Modeling exchange rate dynamics: Nonlinear dependence and thick tails. Econometric Review 12:33–63

Mandelbrot B (1963) The variation of certain speculative prices. Journal of Business 36:394–419

Meese RA, Rogoff K (1983) Empirical exchange rate models of the seventies: Do they fit out of sample? Journal of International Economics 14:3–24

Nelson DB (1989) Modeling stock market volatility changes. Proceedings from the American Statistical Association

Nelson DB (1991) Conditional heteroskedasticity in asset returns: A new approach. Econometrica 2:347–370

Nieuwland FGMC (1992) Conditional and unconditional distributions for stock returns. Working Paper, University of Limburg Maastricht

Nieuwland FGMC, Verschoor WFC (1992) EMS exchange rate expectations and time-varying risk premia. Working Paper, University of Limburg

Pagan AR, Hall AD (1983) Diagnostic tests as residual analysis. Australian National University

Pagan AR, Ullah A (1988) The econometric analysis of models with risk terms. Journal of Applied Econometrics 3:87–105

Poon SH, Taylor SJ (1992) Stock returns and volatility: An empirical study of the UK stock market. Journal of Banking and Finance 16:37–59

Spanos A (1991) A parametric approach to heteroskedasticity: The student's *t* and elliptic linear regression models. Unpublished manuscript, Virginia Tech

Spanos A (1992) The student's *t* autoregressive model with dynamic heteroskedasticity. Unpublished manuscript, Virginia Tech

Tauchen GE, Pitts M (1983) The price variability-volume relationship on speculative markets. Econometrica 2:485–505

Taylor SJ (1990) Modeling stochastic volatility. Unpublished Manuscript, University of Lancester

Weiss AA (1986) Asymptotic theory for ARCH models: Estimation and testing. Econometric Theory 2:107–131

Wiggins JB (1989) Risk and excess return on the dow jones index: 1897–1925. Journal of Financial and Quantitative Analysis

Wolff CCP (1987) Forward foreign exchange rates, expected spots rates and premia: A signal – extraction approach. Journal of Finance 42:395–406

First version received: February 1993
Final version received: August 1993