A Simple Test for GARCH against a Stochastic Volatility Model*

Philip Hans Franses†
Econometric Institute
Erasmus University Rotterdam

Marco van der Leij
Tinbergen Institute
Erasmus University Rotterdam

Richard Paap
Econometric Institute
Erasmus University Rotterdam

ECONOMETRIC INSTITUTE REPORT EI 2005-41

Abstract

The GARCH model and the Stochastic Volatility [SV] model are competing but non-nested models to describe unobserved volatility in asset returns. We propose a GARCH model with an additional error term, which can capture SV model properties, and which can be used to test GARCH against SV. We discuss model representation, parameter estimation and a simple test for model selection. Furthermore, we derive the theoretical moments and the autocorrelation function of our new model. We illustrate its merits for 9 daily stock return series.

Keywords: GARCH, Stochastic volatility, Model selection

---

*Earlier versions of this paper entitled "Forecasting jumps in conditional volatility: The GARCH-IE model" have been presented at various seminars (in 2003 and 2004) in the Netherlands, Singapore and Australia. We thank Dennis Fok, Jun Wu, Asger Lunde, Peter Boswijk, Marius Ooms, Charles Bos, Michael McAleer, and Robert Engle for helpful suggestions.

†Corresponding author: Erasmus University Rotterdam, Econometric Institute, H 11-34, P.O. Box 1738, NL-3000 DR Rotterdam, The Netherlands, e-mail: franses@few.eur.nl

1 Introduction

There are two classes of models that are often used to estimate and forecast unobserved volatility in asset returns. These are (variants of) the GARCH model, see Engle (1982) and Bollerslev (1986) and (variants of) the Stochastic Volatility (SV) model, see Taylor (1986), among others. Basically, the SV model assumes two error processes, while the GARCH model allows for only a single error term. This implies that the SV model can provide a better in-sample fit, see Danielsson (1994) and Kim *et al.* (1998), and perhaps also better forecasts. On the other hand, the SV model parameters are not always easy to estimate, while GARCH parameters can easily be estimated using maximum likelihood. Hence, for practical purposes, one might want to know beforehand whether it is worth the trouble trying to estimate an SV model.

In the limits of continuous time, the GARCH and SV models bear strong similarities, see Nelson (1990) and Duan (1997), but when fitting these models to discretely-observed, say daily data, the models look rather distinct, see also Fleming & Kirby (2003). In fact, the models are non-nested, and this can complicate model comparison. In this paper, we therefore propose a simple test that can be used for selecting between GARCH and SV. The test is based on a GARCH model that is extended with an additional error term. This new model is called a stochastic GARCH model. The model is a variant of an SV type model, and it captures typical SV model properties. The test concerns only a single parameter, where under the null hypothesis the GARCH model appears. A beneficial feature is that the parameters in our stochastic GARCH model are easy to estimate, as we will demonstrate below.

The outline of our paper is as follows. In Section 2 we outline the representation of the stochastic GARCH(1,1) model. We derive the theoretical moments of this new model and consider its autocorrelation function. Furthermore, we give the details of the estimation procedure, and we discuss inference where we focus on the parameter that distinguishes standard GARCH from stochastic GARCH. In Section 3, we apply the model to nine stock markets for which we consider daily data. We see that the GARCH model gets rejected against the stochastic version for all cases, at least, based on the in-sample data. When we compare the out-of-sample fit, the regular GARCH model turns out to be a good competitor. In Section 4 we conclude and discuss potential further research areas.

2 Representation, Estimation and Inference

In this section we put forward a model for asset returns y_t that captures the features of a stochastic volatility model, and which collapses into a GARCH model when a variance parameter is equal to zero. We derive the theoretical moments and autocorrelations of the new model specification. Finally, we discuss parameter estimation and inference.

2.1 Representation

We assume that asset returns y_t for $t = 1, \ldots, T$ can be described by

$$y_t = \delta + \varepsilon_t \tag{1}$$

with $\varepsilon_t = \sqrt{h_t} z_t$ and $z_t \sim \text{NID}(0, 1)$, where

$$h_t = k_t + \eta_t \tag{2}$$

with $\eta_t \sim \text{LNID}(\mu, \sigma^2)$, where LN denotes the lognormal distribution. The lognormal distribution is used to ensure that the error contributions and hence h_t are always positive. The k_t process is defined to be a standard GARCH(1,1) specification

$$k_t = \alpha \varepsilon_{t-1}^2 + \beta k_{t-1}. \tag{3}$$

The model (1)–(3) is a variant of a stochastic volatility model, as it has two sources of uncertainty, that is, z_t for the level of the series and η_t for the conditional variance. It collapses to a standard GARCH(1,1) model when $\sigma^2 \downarrow 0$. This is easily seen as follows. Equation (2) can be rewritten as

$$k_t = h_t - \frac{\omega}{1-\beta} \, \exp(\sigma u_t), \tag{4}$$

with $u_t \sim \text{NID}(0,1)$, where $\omega \equiv (1-\beta)\exp(\mu)$. Substituting (4) into (3) and rearranging gives

$$h_t = \frac{\exp(\sigma u_t) - \beta \exp(\sigma u_{t-1})}{1-\beta}\,\omega + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1}. \tag{5}$$

As $\exp(\sigma u_t)$ converges to 1 when $\sigma^2 \downarrow 0$, it is clear that then the model converges to a standard GARCH(1,1) model where ω denotes the intercept in the volatility equation.

Our model is also related to the standard stochastic volatility model of Taylor (1986) which is defined as

$$y_t = \tilde{\delta} + \tilde{\varepsilon}_t, \tag{6}$$

with $\tilde{\varepsilon}_t = \sqrt{\tilde{h}_t} z_t$ and

$$\ln \tilde{h}_t = \gamma \ln \tilde{h}_{t-1} + \ln \tilde{\eta}_t, \tag{7}$$

where $\tilde{\eta}_t \sim \text{LNID}(\tilde{\mu}, \tilde{\sigma}^2)$.

There are two major differences between the standard SV model and our model in (1)–(3). First, our new model assumes a linear specification for h_t, while $\tilde{h}_t$ in the SV model has a loglinear specification. This does however not mean that h_t in our model can become negative. The restrictions $\alpha > 0$ and $\beta > 0$ and the fact that η_t has a lognormal distribution ensures positive values for h_t. The second difference is that in a stochastic volatility model $\ln \tilde{h}_t$ depends on all contemporaneous and lagged shocks $\tilde{\eta}_t, \tilde{\eta}_{t-1}, \tilde{\eta}_{t-2}, \ldots$, which are all unobserved. This results follows immediately if we solve (7) for $\tilde{h}_t$

$$\ln \tilde{h}_t = \gamma^t \ln \tilde{h}_0 + \sum_{k=0}^{t-1} \gamma^k \ln \tilde{\eta}_{t-k}. \tag{8}$$

On the other hand, solving for h_t in (2) gives

$$h_t = \alpha \sum_{k=1}^{t} \beta^{k-1} \varepsilon_{t-k}^2 + \beta^t k_0 + \eta_t \tag{9}$$

and hence in our new model h_t depends on past realized shocks $\varepsilon_{t-1}, \varepsilon_{t-2}, \ldots$ and only on one contemporaneous unobserved shock η_t. As we will show below, the fact that h_t in our model only depends on one contemporaneously unobserved shock facilitates parameter estimation tremendously. We like to call our model a stochastic GARCH [SGARCH] model.

2.2 Properties

The SGARCH(1,1) process in (1)–(3) contains an extra error term and hence the theoretical unconditional moments of y_t are not the same as those of a standard GARCH process. The following theorem gives expressions and existence conditions for the $2m$th unconditional moments of our SGARCH(1,1) process.

Theorem 1 *For the SGARCH(1,1) process given by (1)–(3) a necessary and sufficient condition of existence of the $2m$th centered moment is*

$$\sum_{j=0}^{m} \binom{m}{j} a_j \alpha^j \beta^{m-j} < 1, \tag{10}$$

where

$$a_k = E[z_t^{2k}] = \begin{cases} 1 & \text{for } k = 0 \\ \prod_{j=1}^{k}(2j-1) & \text{for } k = 1, 2, \ldots. \end{cases}$$

The $2m$th centered moment is given by

$$E[\varepsilon_t^{2m}] = a_m \sum_{j=0}^{m} \binom{m}{j} b_{m-j} E[k_t^j],$$

where

$$b_k = E[\eta_t^k] = \left(\frac{\omega}{1-\beta}\right)^k \exp(\sigma^2 k^2/2),$$

and the mth moment of k_t can be expressed by the recursive formula

$$E[k_t^m] = \frac{1}{1-\psi_{m,m}} \sum_{k=0}^{m-1} \binom{m}{k} b_{m-k} \psi_{m,k} E[k_t^k], \tag{11}$$

where

$$\psi_{m,k} = \sum_{j=0}^{k} \binom{k}{j} a_{m-j} \alpha^{m-j} \beta^j.$$

The proof of this theorem follows the lines of Theorem 2 of Bollerslev (1986) and is given in Appendix A. We note that the moment existence conditions are identical to that of a standard GARCH(1,1) process.

Unconditional variance and kurtosis

With Theorem 1 we derive the unconditional variance and kurtosis of ε_t (and hence y_t). The unconditional variance exists if and only if $\alpha + \beta < 1$, while the kurtosis exists if and only if $3\alpha^2 + 2\alpha\beta + \beta^2 < 1$. Under these existence conditions

$$\mathrm{E}[k_t] = \frac{\alpha\omega}{(1-\alpha-\beta)(1-\beta)} \ \exp(\sigma^2/2), \tag{12}$$

and

$$\begin{aligned}\mathrm{E}[k_t^2] &= \frac{1}{1-\psi_{1,1}} \ (b_2\psi_{2,0} + 2b_1\psi_{2,1}\mathrm{E}[k_t]) \\ &= \frac{\alpha^2\omega^2\exp(\sigma^2)}{(1-3\alpha^2-2\alpha\beta-\beta^2)(1-\beta)^2} \left(3\exp(\sigma^2) + \frac{6\alpha+2\beta}{1-\alpha-\beta}\right).\end{aligned} \tag{13}$$

Hence, the variance of ε_t is given by

$$\mathrm{E}[\varepsilon_t^2] = \frac{\omega}{1-\alpha-\beta} \ \exp(\sigma^2/2), \tag{14}$$

and the fourth moment of ε_t is given by

$$\begin{aligned}\mathrm{E}[\varepsilon_t^4] &= 3(b_2 + 2b_1\mathrm{E}[k_t] + \mathrm{E}[k_t^2]) \\ &= \frac{3\omega^2\exp(\sigma^2)}{(1-3\alpha^2-2\alpha\beta-\beta^2)(1-\beta)^2}\left((1-2\alpha\beta-\beta^2)\exp(\sigma^2) + \frac{2\alpha(1-\alpha\beta-\beta^2)}{1-\alpha-\beta}\right).\end{aligned} \tag{15}$$

Finally, the kurtosis of ε_t is given by

$$\begin{aligned}\mathrm{K}[\varepsilon_t] &= \frac{\mathrm{E}[\varepsilon_t^4]}{\mathrm{E}[\varepsilon_t^2]^2} \\ &= \frac{3(1-\alpha-\beta)\left((1-2\alpha\beta-\beta^2)(1-\alpha-\beta)\exp(\sigma^2) + 2\alpha(1-\alpha\beta-\beta^2)\right)}{(1-3\alpha^2-2\alpha\beta-\beta^2)(1-\beta)^2}.\end{aligned} \tag{16}$$

For $\sigma^2 = 0$ the kurtosis reduces to

$$\mathrm{K}[\varepsilon_t; \sigma^2 = 0] = \frac{3(1-\alpha-\beta)(1+\alpha+\beta)}{1-3\alpha^2-2\alpha\beta-\beta^2}, \tag{17}$$

which is of course equal to the kurtosis of a GARCH(1,1) process.

Autocorrelations

To derive the autocorrelation function for ε_t^2, we define

$$\nu_t = \varepsilon_t^2 - k_t - \mathrm{E}[\eta_t] = \varepsilon_t^2 - k_t - \frac{\omega}{1-\beta}\ \exp(\sigma^2/2), \tag{18}$$

which implies that $\mathrm{E}[\nu_t] = 0$. We substitute $k_t = \varepsilon_t^2 - \mathrm{E}[\eta_t] - \nu_t$ into (3) and rearrange to obtain

$$\varepsilon_t^2 = \omega \exp(\sigma^2/2) + (\alpha + \beta)\varepsilon_{t-1}^2 + \nu_t - \beta\nu_{t-1}. \tag{19}$$

Hence, the SGARCH(1,1) process can be represented as an ARMA(1,1) process for ε_t^2. In fact, exactly the same representation can be found for a standard GARCH(1,1) process, see Bollerslev (1986), although the distribution of ν_t in the SGARCH(1,1) differs for that of a GARCH(1,1) process.

The autocorrelations for ε_t^2 can be derived from (19) and these are identical to those of a GARCH(1,1) process. In particular, let ρ_k be the kth order autocorrelation of ε_t^2. Then

$$\rho_1 = \frac{\alpha(1-\alpha\beta-\beta^2)}{1-2\alpha\beta-\beta^2}, \tag{20}$$

and

$$\rho_n = (\alpha+\beta)^{n-1}\rho_1, \tag{21}$$

for $n = 2, 3, \ldots$, see Bollerslev (1986).

In sum, there are many similarities between the SGARCH and GARCH model, except for the variance and kurtosis.

2.3 Parameter Estimation

As can be seen from (9), h_t can be expressed in terms of past ε_t and the random term η_t. Hence, unlike the standard SV model, it is relatively easy to derive the likelihood function of our stochastic GARCH model.

The density function of y_t given past observations $\Omega_{t-1} \equiv \{y_{t-1}, y_{t-2}, \ldots, y_0\}$ and parameters $\theta \equiv \{\delta, \omega, \alpha, \beta, \sigma^2\}$ is given by

$$f(y_t|\Omega_{t-1};\theta) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{h_t(u_t)}}\ \phi\left(\frac{y_t-\delta}{\sqrt{h_t(u_t)}}\right) \phi(u_t)\ du_t, \tag{22}$$

for $t = 1, \ldots, T$, where $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}z^2)$. The conditional variance is defined as

$$h_t(u_t) \equiv h(u_t; \Omega_{t-1}, \theta) = k_t + \frac{\omega}{1-\beta} \exp(\sigma u_t), \tag{23}$$

with

$$k_t = k(\Omega_{t-1}, \theta) = \alpha \sum_{k=1}^{t} \beta^{k-1} (y_{t-k} - \delta)^2 + \beta^t k_0. \tag{24}$$

Similar to the GARCH model, k_t can be computed recursively using (3) where we choose k_0 as a starting value for k_t with

$$k_0 = \frac{\alpha \overline{\varepsilon_{-1}^2}}{1-\beta}, \tag{25}$$

where $\overline{\varepsilon_{-1}^2} = \frac{1}{T} \sum_{t=0}^{T-1} (y_t - \delta)^2$.

The loglikelihood function is given by

$$\ell(\theta; y) = \sum_{t=1}^{T} \ln f(y_t | \Omega_{t-1}; \theta), \tag{26}$$

where $y = (y_T, \ldots, y_0)$. The ML estimator $\hat{\theta}_{ML}$ is obtained by maximizing $\ell(\theta; y)$ with respect to θ. This can be done with standard optimization algorithms like the BFGS algorithm. Note that there is no analytic expression for the integral in (22). Hence, this integral has to be evaluated numerically for each observation t with, for example, an adaptive Simpson procedure. Since the integral is only one-dimensional, numerical integration is straightforward, and the optimization procedure typically converges without difficulties, although we note that parameter estimation is more time-consuming than for a standard GARCH model.

Asymptotic standard errors for the maximum likelihood estimator can be obtained by evaluating minus the inverse of the second-order derivative of the loglikelihood function in $\hat{\theta}_{ML}$,

$$V(\hat{\theta}_{ML}) = -\left(\frac{\partial^2 \ell(\theta; y)}{\partial \theta \partial \theta'} \right)^{-1} \Bigg|_{\theta = \hat{\theta}_{ML}}. \tag{27}$$

2.4 Inference

If we impose the restriction $\sigma^2 = 0$ in (5), our SGARCH(1,1) model simplifies to a standard GARCH(1,1) model. Hence, we can test our stochastic GARCH specification versus a standard GARCH using a likelihood ratio test. As this is a one-sided test ($\sigma^2 = 0$ versus $\sigma^2 > 0$) the likelihood ratio statistic is asymptotically $\frac{1}{2}\chi^2(0) + \frac{1}{2}\chi^2(1)$ distributed, see Wolak (1989). Hence, if one wants to test at a 5% level of significance, the critical value is the 90% percentile of the $\chi^2(1)$ distribution.

3 Illustration

We illustrate our simple test for GARCH against SV for a decade of daily data for nine stock markets. We fit GARCH(1,1) and SGARCH(1,1) models for the data for 1990-1999, while we use the year 2000 data for out-of-sample forecast evaluation. The estimation results for the Dow Jones, Nasdaq, SP500, Nikkei, FTSE appear in Table 1, while those of the Dax, Cac, AEX and the HangSeng appear in Table 2.

There are a few observations to be made from these estimation results. First, and as expected, the persistence parameter β is higher for the SGARCH model than for the GARCH model. Indeed, one may view an additional error process as a process generating additive outliers, and taking care of such outliers is known to lead to higher persistence. A second observation is that the $\hat{\sigma}^2$ parameters are always more than twice as large as the corresponding standard errors, and hence at first sight the SGARCH model seems preferable. This seems to be confirmed by the log-likelihood values in the last column of the two tables.

Table 3 substantiates the findings in Table 2 by comparing these log-likelihoods using the AIC. For all nine stock markets, the SV-like model obtains the smallest AIC value. Also, a likelihood ratio type test would indicate that the additional error process has a variance far from zero.

If an SGARCH would better describe the data, this should then also be observed from the implied properties of the data, and here notably the variance and kurtosis. From Table 4 we can observe that for some cases, the estimated SGARCH model generates empirical variance and kurtosis that are remarkably close to those of the actual data. In some other cases however, the SGARCH generates a kurtosis value which is way out of the usual range, see, for example, the kurtosis for the DAX.

Finally, when we consider the two models for forecasting the daily returns for 2000, we observe that the models perform about equally well, see Table 5. In 5 cases the SGARCH model is better, while in the other 4 the GARCH is. The largest difference between the out-of-sample log-likelihoods is found for the Nikkei, where the GARCH model is best. And, for the Nasdaq the SV-like model provides a much better out-of-sample fit.

4 Conclusion

In this paper we proposed a simple extension of a standard GARCH(1,1) model, which can capture SV-like properties of the data. The parameters in this new SGARCH model can be estimated quite easily, and we showed that its implied properties differ from those of the GARCH model in terms of variance and kurtosis. The model can be used to provide a simple and quick, though indirect, test for GARCH against SV. An illustration of the new model for nine daily returns series shows that there are gains in fit when considering an SV-like model, although it must be mentioned that the out-of-sample forecasts are not that much better.

A beneficial feature of our simple test for GARCH against SV is that it is easy to extend to the many non-linear variants of these models, see Franses & van Dijk (2000) for a survey. The main feature of our test is that it amounts to testing the constancy of the intercept term in the GARCH equation. Also, we expect that an extension to multivariate GARCH models should not be complicated either. This would be very useful, as the estimation of multivariate SV models is not easy, and one would better want to know in advance if such routes are necessary. We postpone these two extensions to our further work.

A Proof

Proof: The proof follows the lines of Theorem 2 of Bollerslev (1986). We first derive existence conditions for $\mathrm{E}[k_t^m]$. Substituting (1) and (2) into (3), we have

$$k_t = \alpha z_{t-1}^2 \eta_{t-1} + (\alpha z_{t-1}^2 + \beta) k_{t-1}. \tag{28}$$

Using the binomial theorem, k_t^m can be expressed as

$$k_t^m = \sum_{k=0}^{m} \binom{m}{k} \eta_{t-1}^{m-k} k_{t-1}^k \sum_{j=0}^{k} \binom{k}{j} \alpha^{m-j} \beta^j z_{t-1}^{2(m-j)}. \tag{29}$$

Since z_{t-1} and η_{t-1} are independent, and k_{t-1} is determined by $\Omega_{t-2} = \{y_{t-2}, y_{t-3}, \ldots\}$, the conditional expectation of k_t^m given Ω_{t-2} is

$$\mathrm{E}[k_t^m | \Omega_{t-2}] = \sum_{k=0}^{m} \binom{m}{k} b_{m-k} \psi_{m,k} k_{t-1}^k, \tag{30}$$

where

$$\psi_{m,k} = \sum_{j=0}^{k} \binom{k}{j} a_{m-j} \alpha^{m-j} \beta^j, \tag{31}$$

and

$$a_k = \mathrm{E}[z_t^{2k}] = \begin{cases} 1 & \text{for } k = 0 \\ \prod_{j=1}^{k} (2j-1) & \text{for } k = 1, 2, \ldots \end{cases} \tag{32}$$

is the kth moment of the standard normal distribution, and

$$b_k = \mathrm{E}[\eta_t^k] = \exp(\mu k + \sigma^2 k^2/2) = \left(\frac{\omega}{1-\beta} \right)^k \exp(\sigma^2 k^2 / 2) \tag{33}$$

is the kth uncentered moment of the lognormal distribution.

Let $v_t = (k_t^m, k_t^{m-1}, \ldots, k_t)'$. Then

$$\mathrm{E}[v_t | \omega_{t-2}] = d + C v_{t-1}, \tag{34}$$

where $d = (\alpha^m a_m b_m, \ldots, \alpha a_1 b_1)'$, and C is an $m \times m$ upper triangular matrix with diagonal elements

$$\mathrm{diag}(C) = (\psi_{m,m}, \psi_{m-1,m-1}, \ldots, \psi_{1,1})'. \tag{35}$$

Repeated substitution into (34) yields

$$\mathrm{E}[v_t|\Omega_{t-k-1}] = (I + C + \ldots + C^{k-1})d + C^k v_{t-k}. \tag{36}$$

This converges for $k \to \infty$ as long as all eigenvalues of C are within the unit circle, or equivalently

$$\psi_{m,m} = \sum_{j=0}^{m} \binom{m}{j} a_{m-j} \alpha^{m-j} \beta^j < 1. \tag{37}$$

This is the moment existence condition of (10). Note that $\psi_{m,m} < 1$ implies $\psi_{m-1,m-1} < 1$, see Bollerslev (1986).

If (10) holds, then $\mathrm{E}[k_t^m] = \mathrm{E}[k_{t-1}^m]$. Solving the unconditional version of (30), that is,

$$\mathrm{E}[k_t^m] = \sum_{k=0}^{m} \binom{m}{k} b_{m-k} \psi_{m,k} \mathrm{E}[k_{t-1}^k], \tag{38}$$

we obtain the recursive expression of $\mathrm{E}[k_t^m]$ in (11).

To finalize the proof we observe that the mth moment of ε_t exists if and only if mth moment of k_t exists and

$$\mathrm{E}[\varepsilon_t^{2m}] = \mathrm{E}[z_t^{2m}(k_t + \eta_t)^m] = a_m \sum_{j=0}^{m} \binom{m}{j} b_{m-j} \mathrm{E}[k_t^j]. \tag{39}$$

Q.E.D.

Table 1: ML estimates for the GARCH(1,1) and SGARCH(1,1) model for daily returns on 5 stock markets from 1/1/1990 to 12/31/1999. Estimated standard errors in parentheses.

Model	δ	ω	α	β	σ^2	log-lik.
			DOWJONES (2518 obs.)			
GARCH	0.067	0.007	0.049	0.942		-3089.579
	(0.015)	(0.003)	(0.009)	(0.011)		
SGARCH	0.074	0.002	0.034	0.952	3.443	-3027.045
	(0.014)	(0.001)	(0.006)	(0.009)	(0.918)	
			NASDAQ (2528 obs.)			
GARCH	0.097	0.035	0.118	0.855		-3534.344
	(0.018)	(0.010)	(0.021)	(0.026)		
SGARCH	0.112	0.004	0.054	0.929	3.353	-3490.344
	(0.017)	(0.002)	(0.012)	(0.015)	(0.967)	
			SP500 (2527 obs.)			
GARCH	0.062	0.006	0.053	0.941		-3032.802
	(0.014)	(0.002)	(0.009)	(0.011)		
SGARCH	0.069	0.002	0.035	0.952	3.707	-2977.213
	(0.014)	(0.001)	(0.007)	(0.009)	(0.925)	
			NIKKEI (2465 obs.)			
GARCH	0.023	0.071	0.107	0.864		-4315.032
	(0.025)	(0.016)	(0.014)	(0.018)		
SGARCH	-0.013	0.020	0.077	0.890	2.747	-4261.379
	(0.024)	(0.008)	(0.011)	(0.014)	(0.665)	
			FTSE (2524 obs.)			
GARCH	0.051	0.009	0.052	0.937		-3173.497
	(0.016)	(0.004)	(0.011)	(0.014)		
SGARCH	0.052	0.002	0.040	0.951	2.556	-3152.139
	(0.015)	(0.001)	(0.008)	(0.010)	(1.017)	

Table 2: ML estimates for the GARCH(1,1) and SGARCH(1,1) model for daily returns on 4 stock markets from 1/1/1990 to 12/31/1999. Estimated standard errors in parentheses.

Model	δ	ω	α	β	σ^2	log-lik.
			DAX (2502 obs.)			
GARCH	0.068	0.051	0.090	0.877		-3862.969
	(0.021)	(0.011)	(0.015)	(0.019)		
SGARCH	0.075	0.002	0.056	0.930	5.083	-3748.888
	(0.018)	(0.002)	(0.010)	(0.012)	(1.583)	
			CAC (2497 obs.)			
GARCH	0.054	0.072	0.079	0.872		-3919.054
	(0.022)	(0.018)	(0.014)	(0.023)		
SGARCH	0.062	0.022	0.059	0.914	1.168	-3891.759
	(0.022)	(0.011)	(0.011)	(0.018)	(0.427)	
			AEX (2524 obs.)			
GARCH	0.074	0.021	0.086	0.893		-3407.396
	(0.017)	(0.005)	(0.012)	(0.014)		
SGARCH	0.082	0.005	0.072	0.911	2.716	-3362.690
	(0.016)	(0.002)	(0.012)	(0.013)	(0.812)	
			HANGSENG (2478 obs.)			
GARCH	0.132	0.079	0.114	0.859		-4457.943
	(0.026)	(0.015)	(0.014)	(0.016)		
SGARCH	0.129	0.017	0.080	0.891	2.973	-4383.711
	(0.025)	(0.008)	(0.014)	(0.017)	(0.745)	

Table 3: Within-sample forecasting performance of the GARCH and SGARCH model for daily returns.

	log likelihood value		AIC	
Series	GARCH	SGARCH	GARCH	SGARCH
DOWJONES	-3089.579	-3027.045	2.497	2.447
NASDAQ	-3534.344	-3490.344	2.856	2.821
SP500	-3032.802	-2977.213	2.451	2.407
NIKKEI	-4315.032	-4261.379	3.486	3.443
FTSE	-3173.497	-3152.139	2.565	2.548
DAX	-3862.969	-3748.888	3.121	3.030
CAC	-3919.054	-3891.759	3.166	3.145
AEX	-3407.396	-3362.690	2.753	2.718
HANGSENG	-4457.943	-4383.711	3.601	3.542

Table 4: Sample variance and kurtosis of the nine stock returns together with the implied values from the estimated GARCH and SGARCH models

	Variance			Kurtosis		
Series	Sample	GARCH	SGARCH	Sample	GARCH	SGARCH
Dow Jones	0.797	0.851	0.803	7.825	4.154	12.593
Nasdaq	1.234	1.274	1.180	7.289	6.206	10.034
SP 500	0.790	0.900	0.777	7.916	5.510	13.941
NIKKEI	2.328	2.484	2.383	7.355	5.086	8.997
FTSE	0.834	0.831	0.813	5.157	3.986	5.243
DAX	1.575	1.559	1.796	7.937	3.998	33.286
CAC	1.492	1.461	1.443	5.435	3.452	4.264
AEX	1.172	1.056	1.033	6.889	4.797	6.925
HangSeng	3.023	2.884	2.636	14.745	5.730	9.430

Table 5: Out-of-sample forecasting performance of the GARCH and SGARCH model for daily returns in 2000.

	Log likelihood value	
series	SGARCH	GARCH
DOWJONES	-418.316	-418.676
NASDAQ	-633.571	-637.350
SP500	-438.385	-439.069
NIKKEI	-443.926	-436.570
FTSE	-397.876	-398.707
DAX	-460.682	-462.547
CAC	-459.394	-458.467
AEX	-396.470	-393.198
HANGSENG	-523.109	-518.673

References

Anderson, T., 1994, Stochastic autoregressive volatility: A framework for volatility modelling, *Mathematical Finance*, 4, 75-102.

Bollerslev, T., 1986, Generalized autoregressive conditional heteroscedasticity, *Journal of Econometrics*, 31, 307-327.

Danielsson, J., 1994, Stochastic volatility in asset prices: Estimation with simulated maximum likelihood, *Journal of Econometrics*, 64, 375-400.

Duan, J-C., 1998, Augmented GARCH(p,q) process and its diffusion limit, *Journal of Econometrics*, 79, 97-127.

Engle, R.F., 1982, Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation, *Econometrica*, 50, 987-1007.

Fleming, J. and C. Kirby, 2003, A closer look at the relation between GARCH and stochastic autoregressive volatility, *Journal of Financial Econometrics*, 1, 365-419.

Franses, P.H. and D.J.C. van Dijk, 2000, Non-linear time series models in empirical finance, Cambridge University Press, Cambridge.

Kim, S., N. Shephard, and S. Chib, 1998, Stochastic volatility: Likelihood inference and comparison with ARCH models, *Review of Economic Studies*, 65, 361-393.

Nelson, D., 1990, ARCH models as diffusion approximations, *Journal of Econometrics*, 45, 7-38.

Taylor, S.J., 1986, *Modeling Financial Time Series*, New York: Wiley.

Wolak, F.A., 1989, Local and global testing of linear and nonlinear inequality constraints in nonlinear econometric models, *Econometric Theory*, 5, 1–35.