Working Paper 01-08
Statistics and Econometrics Series 05
March 2001

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (34) 91 624-98-49

IS STOCHASTIC VOLATILITY MORE FLEXIBLE THAN GARCH?

M. Angeles Carnero, Daniel Peña and Esther Ruiz*

Abstract

This paper compares the ability of GARCH and ARSV models to represent adequately the main empirical properties usually observed in high frequency financial time series: high kurtosis, small first order autocorrelation of squared observations and slow decay towards zero of the autocorrelation coefficients of squared observations. We show that the ARSV(1) model is more flexible than the GARCH(1,1) model in the sense that it is able to generate series with higher kurtosis and smaller first order autocorrelation of squares for a wider variety of parameter specifications. Our results may help to clarify some puzzles raised in the empirical analysis of real financial time series.

Keywords: ARSV; GARCH; high persistence; excess kurtosis.

*Department of Statistics and Econometrics; Universidad Carlos III de Madrid, Getafe (Madrid), Carnero: e-mail: acarnero@est-econ.uc3m.es; Peña: e-mail: dpena@est-econ.uc3m.es; Ruiz: e-mail: ortega@est-econ.uc3m.es

Is Stochastic Volatility more flexible than GARCH?

M. Angeles Carnero, Daniel Peña and Esther Ruiz*
Departamento de Estadistica y Econometria
Universidad Carlos III de Madrid
C/ Madrid 126, 28903 Getafe (Spain)

Abstract

This paper compares the ability of GARCH and ARSV models to represent adequately the main empirical properties usually observed in high frequency financial time series: high kurtosis, small first order autocorrelation of squared observations and slow decay towards zero of the autocorrelation coefficients of squared observations. We show that the ARSV(1) model is more flexible than the GARCH(1,1) model in the sense that it is able to generate series with higher kurtosis and smaller first order autocorrelation of squares for a wider variety of parameter specifications. Our results may help to clarify some puzzles raised in the empirical analysis of real financial time series.

Keywords: ARSV, GARCH, high persistence, excess kurtosis.

*Corresponding author. Tel: 34 91 6249851, Fax: 34 91 6249849. e-mail: ortega@est-econ.uc3m.es. Financial support from projects BEC2000-0167 and PB98-0026 by the Spanish Government is acknowledge. We are grateful to Ana Pérez and Andrés Alonso for helpful comments and suggestions. Any remaining errors are our own.

1 Introduction

The main objective of this paper is to compare the ability of Generalized Autoregressive Conditional Heteroscedastic (GARCH) and Autoregressive Stochastic Volatility (ARSV) models to represent adequately the observed properties of real time series with conditional heteroscedasticity.

Financial series of returns are mainly characterized by having: (1) high kurtosis, (2) small first order autocorrelation of squared observations and (3) high persistence in the autocorrelation of squared observations. These stylized characteristics have been documented by a large number of authors; see, for example, Liesenfeld and Jung (2000) and Loudon et al. (2000) as two recent references.

The simplest model able to generate these effects is given by

$$y_t = \varepsilon_t \sigma_t \tag{1}$$

where ε_t is a serially independent and identically distributed (i.i.d.) process with zero mean, unit variance and finite fourth order moment, that is assumed to be independent of the process σ_t that is known as volatility in the financial literature. Under these conditions and if the conditional expectation of σ_t is finite, Ghysels et al. (1996) show that the process y_t, defined in (1), is a martingale difference and it can explain volatility clustering via autoregressive dynamics in the conditional expected value of σ_t^2. Finally, excess kurtosis can be obtained in either of the following ways:

(i) Heavy tails in the marginal distribution of the white noise ε_t.

(ii) Conditional heteroscedasticity. The conditional variance of y_t is given by $E(y_t^2 \mid Y_{t-1}) = E(\sigma_t^2 \mid Y_{t-1})$ where $Y_{t-1} = (y_1, ..., y_{t-1})$, and the excess kurtosis of y_t depends on the dynamic evolution of its conditional variance.

(iii) An unexpected component in the volatility which does not depend on the past.

A wide spectrum of models have been proposed in the literature to represent the dynamic evolution of σ_t. However, in this paper, we concentrate our attention on the two basic parametric models widely used in the empirical analysis of high frequency financial time series, the GARCH(1,1) model originally proposed by Bollerslev (1986) and Taylor (1986) and the ARSV(1) model proposed by Taylor (1986). Although, both models are able to explain volatility clustering and excess kurtosis, they are different in the measurability properties of the volatility process with respect to certain benchmark information sets; see Andersen (1992). As Taylor (1994) points out, the fundamental difference between both types of models is that the volatility of ARSV(1) models is a latent variable with an unexpected noise while in GARCH(1,1) models, σ_t is observable given Y_{t-1}. Consequently, if ε_t is Gaussian, the excess kurtosis of y_t in GARCH(1,1) models can only be explained by the evolution of its conditional variance while in ARSV(1) models, the excess kurtosis may depend on both the conditional variance and on the unexpected component of volatility. Ghysels et al. (1996) suggest that this additional source of kurtosis could make the ARSV(1) model more flexible to represent the observed properties of real time series. In this paper, we show that this is effectively the case.

The ability of each model to represent the empirical properties observed in real time series has only been analyzed separately. Teräsvirta (1996) shows that the basic GARCH(1,1) model is not able to account for the simultaneous presence of excess kurtosis and low first order autocorrelation of squared observations, even if ε_t has a Student-t distribution. With respect to ARSV(1) models, Liesenfeld and Jung (2000) conclude that the ARSV(1) model with ε_t being Gaussian does not adequately account for the simultaneous presence of leptokurtic returns and the low autocorrelations of squared observations.

In this paper, we compare the capacity of GARCH(1,1) and ARSV(1) models to generate series not only with excess kurtosis and autocorrelations of squares characterized by a small order one autocorrelation, but also by a slow decay. We show that the ARSV(1) model is more flexible than the GARCH(1,1) model in the sense that it is able to generate series with larger excess kurtosis and smaller order one autocorrelation of squares for a wider variety of parameter specifications implying different degrees of persistence. GARCH(1,1) models can only generate series with high kurtosis and low first order autocorrelation of squares if the persistence is high. Consequently, we show that ARSV(1) models are in closer conformance with the properties usually observed in real data than GARCH(1,1) models. Our results may also help to clarify some puzzles raised in empirical studies comparing both models. For example, we explain why in the empirical analysis of financial series, it has often been found that when the GARCH(1,1) specification is chosen, ε_t requires a distribution with heavy tails, while when σ_t is modelled with an ARSV(1) model, the assumption of Gaussianity seems to be adequate; see, for example, Ghysels et al. (1996), Shephard (1996) and Kim et al. (1998). We can also explain why in empirical applications, the persistence of volatility implied by GARCH(1,1) models is usually higher than that implied by ARSV(1) models; see Taylor (1994), Shephard (1996) and Kim et al. (1998). Finally, the results obtained in this paper may explain why the choice of the conditional distribution has systematic effects on the parameter estimates of the volatility process; see, for example, Mahieu and Schotman (1998) and Liesenfeld and Jung (2000).

The paper is organized as follows. In section 2, we describe the main properties of the GARCH(1,1) model. Section 3 describes the basic ARSV(1) model and shows that it is more flexible than the GARCH(1,1) model to represent simultaneously high kurtosis, small order one autocorrelation and high persistence of autocorrelations of squared observations. In section 4, we analyze twelve daily financial time series to illustrate the performance of both models. Finally, section 5 concludes the paper and gives some suggestions for future research.

2 The GARCH(1,1) model

The AutoRegressive Conditional Heteroscedasticity (ARCH) model was introduced by Engle (1982) to model the conditional variance of UK inflation. ARCH type models have been very popular in the financial econometrics literature generating a huge number of papers. Some useful reviews on these models are Bollerslev et al. (1992), Bollerslev et al. (1994), Bera and Higgins (1995), Diebold and López (1995) and Palm (1997). Engle (1995) is a survey of some of the main papers related with ARCH models. If y_t follows a GARCH(1,1) model, then

$$\begin{aligned} y_t &= \varepsilon_t \sigma_t \\ \sigma_t^2 &= \omega + \alpha y_{t-1}^2 + \beta \sigma_{t-1}^2 \end{aligned} \tag{2}$$

where $\varepsilon_t \sim NID(0,1)$, and ω, α and β are parameters such that $\omega > 0$ and $\alpha, \beta \geq 0$. The positivity conditions are needed to guaranty the positivity of the conditional variance[1] and ω has to be strictly positive for the process y_t not to degenerate. Finally, if $\alpha + \beta < 1$, the marginal variance of y_t is finite and the process is covariance stationary. Nelson (1990) shows that y_t is strictly stationary if $E[\log(\beta + \alpha\varepsilon_t^2)] < 1$[2]. Notice that if ε_t is Gaussian, this condition is satisfied even if $\alpha + \beta = 1$. Therefore, when $\alpha + \beta = 1$, the GARCH(1,1) process is strictly stationary although the marginal variance is not finite.

Notice that, once Y_{t-1} is observed, σ_t^2 is known. Consequently, σ_t^2 is the conditional variance of y_t and

$$y_t \mid Y_{t-1} \sim N(0, \sigma_t^2) \tag{3}$$

All GARCH(1,1) processes are martingale differences. If $\alpha + \beta < 1$, the marginal variance of y_t is given by

[1]The positivity conditions for the general GARCH(p,q) model have been given by Nelson and Cao (1992).

[2]The conditions for strict stationarity of general GARCH(p,q) processes have been derived by Bougerol and Picard (1992).

$$\sigma_y^2 = \frac{\omega}{1-\alpha-\beta} \tag{4}$$

The condition for the existence of the fourth order moment is $3\alpha^2+2\alpha\beta+\beta^2<1$; see Bollerslev (1986). If this condition is satisfied, then the kurtosis is given by

$$\kappa_y = \frac{E(y_t^4)}{[E(y_t^2)]^2} = 3 + \frac{6\alpha^2}{1-3\alpha^2-2\alpha\beta-\beta^2} \tag{5}$$

which is greater than 3. Therefore, the marginal distribution of returns has fat tails. All odd moments of y_t can be seen to be zero.

The dynamics of the GARCH(1,1) process appear in the acf of the squared observations. Bollerslev (1988) shows that the autocorrelations of y_t^2 are given by

$$\begin{aligned} \rho_2(1) &= \frac{\alpha(1-\alpha\beta-\beta^2)}{1-2\alpha\beta-\beta^2} \\ \rho_2(\tau) &= (\alpha+\beta)^{\tau-1}\rho_2(1), \tau>1 \end{aligned} \tag{6}$$

The acf of squared observations has the same pattern as an ARMA(1,1) process with autoregressive parameter $\alpha+\beta$. Therefore, the GARCH(1,1) model is able to generate volatility clustering. Notice that the persistence of the volatility process depends on the value of $\alpha+\beta$. In table 1, we report the acf of squared observations of GARCH(1,1) process for different values of the kurtosis coefficient. Notice that, for a given kurtosis coefficient, the order one autocorrelation $\rho_2(1)$ is smaller the larger the persistence. These results are illustrated in Figure 1, which plots the acf of squared observations corresponding to four selected cases. Teräsvirta (1996) observed that, even if the persistence of the volatility is high, high kurtosis is associate with values of $\rho_2(1)$ bigger than the values often observed in practice. It is possible to obtain the relationship between κ_y, $\rho_2(1)$ and persistence ($p = \alpha + \beta$), given by

$$\rho_2(1) = \frac{\sqrt{\frac{(\kappa_y - \kappa_\varepsilon)(1-p^2)}{(\kappa_\varepsilon - 1)\kappa_y}}\left[1 - p^2 + p\sqrt{\frac{(\kappa_y - \kappa_\varepsilon)(1-p^2)}{(\kappa_\varepsilon - 1)\kappa_y}}\right]}{1 - p^2 + \frac{(\kappa_y - \kappa_\varepsilon)(1-p^2)}{(\kappa_\varepsilon - 1)\kappa_y}} \qquad (7)$$

where, if ε_t is Gaussian, $\kappa_\varepsilon = 3$. Figure 2 represents this relationship, that introduces the persistence of shocks to volatility, measured by $\alpha + \beta$, a component not considered by Teräsvirta (1996). This Figure shows that low values of $\rho_2(1)$ and high kurtosis cannot be simultaneously generated by GARCH(1,1) models with normal errors. It is also possible to observe why $\alpha + \beta$ could be estimated close to one, even if there is no high persistence. The GARCH(1,1) model is only able to increase the value of κ_y with low values of $\rho_2(1)$ by forcing $\alpha + \beta$ to be very close to 1. Since in many empirical studies, the estimates of α and β are such that $\widehat{\alpha} + \widehat{\beta} \simeq 1$, Engle and Bollerslev (1986) proposed the Integrated GARCH (IGARCH) process given by model (2) with $\alpha + \beta = 1$. However, Teräsvirta (1996) shows that even IGARCH models are unlikely to provide an adequate characterization of high kurtosis and low first order autocorrelation of squared observations and suggests that substituting the normal distribution of ε_t by a heavy-tailed distribution like, for example, the Student-t distribution, may improve the adequacy of the GARCH(1,1) model to characterize the stylized facts observed in practice.

The Gaussian assumption on ε_t has been relaxed by several authors. For example, a Student-t distribution was suggested by Bollerslev (1987), the normal-poisson mixture distribution is used by Jorion (1988), the power exponential distribution in Baillie and Bollerslev (1989), the normal-lognormal mixture distribution in Hsieh (1989) and the Generalized error distribution (GED) in Nelson (1991). Bollerslev et al. (1994) used the Generalised-t distribution which includes both the Student-t and the GED distributions as particular cases. Finally, Granger and Ding (1995) and González-Rivera (1997) also consider the use of the Laplace distribution in conjunction with GARCH models. In this paper, we will focus on the properties of the GARCH(1,1) model with ε_t having a Student-t distribution with ν degrees of freedom because it is the most popular one in empirical applications. Given that ε_t is standardized to have variance one, the marginal variance of y_t is given by (4), like in the Gaussian case. The condition for the existence of the fourth order moment is $\kappa_\varepsilon \alpha^2 + 2\alpha\beta + \beta^2 < 1$; see He and Teräsvirta (1999). If this condition is satisfied, then the kurtosis is given by

$$\kappa_y = \kappa_\varepsilon \frac{1 - \alpha^2 - \beta^2 - 2\alpha\beta}{1 - \kappa_\varepsilon \alpha^2 - 2\alpha\beta - \beta^2} \tag{8}$$

where κ_ε is the kurtosis of ε_t, which if $\nu > 4$ is given by $\kappa_\varepsilon = 3(\nu - 2)/(\nu - 4)$.

He and Teräsvirta (1999) show that the autocorrelation function of y_t^2 is the same as in the Gaussian case. The relationship between κ_y, $\rho_2(1)$ and the persistence of the volatility, is also given by (7) and it has been represented in Figure 2 for GARCH(1,1) models with ε_t having Student-t distributions with 5 and 7 degrees of freedom. We can see that, as pointed out by Teräsvirta (1996), GARCH(1,1) models with a conditional Student-t distribution, seem to be better than the Gaussian ones at explaining simultaneously the three stylized facts observed in real time series. However, notice that the kurtosis of y_t is heavily linked to the kurtosis of ε_t. Finally, looking at Figure 2, it could be expected that the estimated parameters $\widehat{\alpha}$ and $\widehat{\beta}$ depend on the assumed distribution for ε_t.

3 The ARSV(1) model

ARSV models assume that σ_t is a latent variable that usually follows an autoregressive process after being transformed into logarithms. Surveys on the properties of ARSV models are given by Taylor (1994), Ghysels et al. (1996) and Shephard (1996).

The simplest case is the ARSV(1) [3] given by:

$$\begin{aligned} y_t &= \sigma_\star \varepsilon_t \sigma_t \qquad (9) \\ \log \sigma_t^2 &= \phi \log \sigma_{t-1}^2 + \eta_t \end{aligned}$$

where ε_t and η_t are assumed to be white noise processes mutually independent and normally distributed with zero mean and variances one and σ_η^2 respectively. The parameter $\sigma_\star$ is a scale factor that removes the necessity of including a constant term in the equation of $\log \sigma_t^2$ and the restriction $|\phi| < 1$ guarantees the stationarity of y_t. Although the assumption of Gaussianity of η_t can seem *ad hoc* at first sight, Andersen et al. (1999) show that the daily log-volatility distribution of real financial series may be well approximated by a normal distribution. Notice that σ_η^2 is the variance of the volatility disturbance. When σ_η^2 is zero, the model in (9) is no longer identified. So et al. (1999) present an interesting interpretation of the ARSV(1) model in

[3] The specification of the log-volatility process as an ARMA(p,q) process has been considered, for example, by Hwang and Satchell (2000)

(9) by decomposing the overall volatility of y_t into a baseline volatility, which represents the volatility in a typical day, and the volatility due to fluctuating information arrivals to the market.

In general, the distribution of y_t conditional on past observations has an unknown form. With respect to the marginal distribution, using the properties of the log-normal distribution, it can be seen that the variance of y_t is given by

$$\sigma_y^2 = \sigma_\star^2 \exp\left(0.5\sigma_h^2\right) \tag{10}$$

where $\sigma_h^2 = \sigma_\eta^2/(1-\phi^2)$. The kurtosis of y_t is given by

$$\kappa_y = 3\exp\left(\sigma_h^2\right) \tag{11}$$

which is bigger than 3. Therefore, the ARSV(1) model is also able to generate series with heavy tails. Notice that, while in the ARSV(1) model, the condition for the existence of the fourth order moment is the stationarity condition, i.e. $|\phi| < 1$, a GARCH(1,1) model can be stationary without having a finite fourth order moment.

All the odd moments of y_t can be easily seen to be zero. Finally, notice that y_t is an uncorrelated process although it is not independent. The autocorrelations of squared observations have been derived by Taylor (1986) and are given by

$$\rho_2(\tau) = \frac{\exp(\sigma_h^2\phi^\tau) - 1}{3\exp(\sigma_h^2) - 1} \tag{12}$$

Taylor (1986) shows that when σ_h^2 is small and/or the autocorrelations of $\log \sigma_t^2$ are close to one, the shape of the acf of squared observations is approximately the same as the shape of the acf of $\log \sigma_t^2$ multiplied by a factor of proportionality, i.e.

$$\rho_2(\tau) \simeq \frac{\exp(\sigma_h^2) - 1}{3\exp(\sigma_h^2) - 1}\phi^\tau \tag{13}$$

Consequently, the acf of squared observations is similar to that of an ARMA(1,1) process characterized by an exponentially decaying rate determined by the parameter ϕ. Therefore, the persistence of shocks to volatility depends on ϕ. Notice that, the acf of squared observations generated by the ARSV(1) model behaves apparently similarly to the acf of squares of the GARCH(1,1) model in (6). It is also important to notice that, in the ARSV(1) model, the parameter σ_η^2 governs the degree of kurtosis independently of the persistence of volatility measured by ϕ. There could be excess kurtosis even when $\phi = 0$. As we will see later, this property makes ARSV(1) models more flexible than GARCH(1,1) models to represent the stylized facts observed in real data. Remember that in a GARCH(1,1) process, kurtosis and persistence are heavily tied up.

Given that ϕ plays a role similar to that of $\alpha + \beta$, in table 1, we report the acf of squared observations of ARSV(1) models for several values of the kurtosis coefficient and for $\phi = \alpha + \beta$. Comparing the acf of the squared observations implied by the GARCH(1,1) and the ARSV(1) model both with ε_t having a Gaussian distribution, we can see that, when $\phi = \alpha + \beta$ and for the same implied kurtosis, the autocorrelations of squared observations implied by the ARSV(1) process are smaller than the autocorrelations implied by the GARCH(1,1) process, except when the volatility approaches the non stationary region, where both correlations are the same; see also Figure 1.

The relationship between kurtosis, persistence and $\rho_2(1)$ does not depend on the parameter σ_η^2 and is given by

$$\rho_2(1) = \frac{\left(\frac{\kappa_y}{3}\right)^\phi - 1}{\kappa_y - 1} \tag{14}$$

This relationship has been plotted in figure 3 for the ARSV(1) model together with the corresponding relationship for the GARCH(1,1) model both with normal errors. In this figure it is possible to observe that the ARSV(1) model is able to generate series with higher kurtosis and lower $\rho_2(1)$ than the GARCH(1,1) model. Introducing the noise η_t makes the ARSV(1) model more flexible in the sense that is able to generate higher kurtosis than the GARCH(1,1) model without increasing $\rho_2(1)$ and without forcing the persistence of volatility to be close to the non stationarity region. This fact could explain why in practice when the volatility is represented by a GARCH(1,1) model, usually it is necessary to specify a fat-tailed distribution of ε_t while when an ARSV(1) model is used, the assumption of Gaussianity of ε_t may be adequate; see Ghysels et al. (1996), Shephard (1996) and Kim et al. (1998). Figure 3 can also explain why the persistence estimated in ARSV(1) models is usually lower than in GARCH(1,1) models; see, Taylor (1994), Shephard (1996) and Kim et al. (1998). Notice that, contrary to what happens in GARCH(1,1) models, in an ARSV(1) model, given κ_y, higher persistence implies higher order one autocorrelation of squared observations; see also table 1. Therefore, it is possible to have ARSV(1) models with high kurtosis, low $\rho_2(1)$ and persistence far from the non stationary region. However, in a GARCH(1,1) model, the persistence should be high because it is the only way to have both, high kurtosis and low $\rho_2(1)$. Notice also that, as we have seen in table 1, the properties of both models are similar when the volatility is close to the non stationarity region.

In figure 4, we plot the relationship between κ_y, $\rho_2(1)$ and persistence for ARSV(1) models with normal errors and GARCH(1,1) models with Student-t with 7 degrees of freedom errors. Notice that the ARSV(1) model is still able to generate series with higher kurtosis and smaller $\rho_2(1)$ for most of the parameter values.

As in the case of GARCH(1,1) models, ARSV(1) models may capture higher kurtosis by allowing ε_t to have a leptokurtic distribution; see Gallant et al. (1994), Harvey et al. (1994), Ruiz (1994), Sandmann and Koopman (1998) and Chib et al. (1998). If a Student-t distribution is assumed for ε_t in (9), the variance of y_t is the same as in the Gaussian case and it is given by (10). However the kurtosis of y_t is now given by

$$\kappa_y = \kappa_\varepsilon \exp\left(\sigma_h^2\right) \tag{15}$$

where κ_ε is the kurtosis of ε_t. Notice that, the condition needed for the existence of the kurtosis is that κ_ε is finite and $|\phi| < 1$. The parameters that govern the dynamic evolution of the volatility are not restricted as far as the model is stationary. This makes the ARSV(1) model even more flexible than the GARCH(1,1) model, since no adicional conditions need to be satisfied.

The acf of squared observations is now given, approximately, by

$$\rho_2(\tau) \simeq \frac{\exp(\sigma_h^2) - 1}{\kappa_\varepsilon \exp(\sigma_h^2) - 1}\phi^\tau \tag{16}$$

where κ_ε is the kurtosis of ε_t. The acf in (16) is equal to the acf of the $\log \sigma_t^2$ process multiplied by a factor of proportionality that depends on the distribution of ε_t. The smaller the kurtosis of ε_t, the bigger the factor of proportionality. Therefore, considering distributions of ε_t with heavy tails, the acf of the squared observations is higher the bigger is the number of degrees of freedom and, consequently, the acf of squares is a maximum for a normal distribution; see Ghysels et al. (1996). Finally, the relationship between kurtosis and $\rho_2(1)$ is given by

$$\rho_2(1) = \frac{\left(\frac{\kappa_y}{\kappa_\varepsilon}\right)^{\phi} - 1}{\kappa_y - 1} \tag{17}$$

Figure 5 plots the relationship between κ_y, $\rho_2(1)$ and persistence for ARSV(1) models with normal and Student-t errors. Observe that, when the Student-t distribution is assumed for the error, the ARSV(1) model can imply negative first order autocorrelations of squared observations if the ϕ parameter is small. This could explain why estimates of ϕ under the ARSV(1)-t specification are, usually, greater than those under ARSV(1)-normal; see Mahieu and Schotman (1998) and Liesenfeld and Jung (2000). It is rather clear that substituting a Gaussian noise by a Student-t noise, allows to have higher kurtosis without increasing the order one autocorrelation of squared observations, introducing even more flexibility in the model. Figure 5 also illustrates the result of Ghysels et al. (1996) previously mentioned that the correlogram of y_t^2 is at a maximum under Normality.

4 Empirical application

In order to illustrate the main empirical properties often observed in high frequency financial time series, table 2 contains descriptive statistics of twelve series observed daily. If we denote by p_t the observed price at time t, we are considering as the series of interest, the returns defined as $r_t = 100(\log(p_t) - \log(p_{t-1}))$. The series considered are returns of the US Dollar against the Canadian Dollar, the Spanish Peseta, the German Mark, the Japanese Yen, the Swiss Franc, the Swedish Krona and the British Pound exchange rates observed from January 1993 to October 2000 and returns of five international stock market indexes, the Amsterdam E.O.E. index and the Bombay stock market index observed from October 1995 to October 2000, the Dow Jones from January 1990 to October 2000, the IBEX 35 of the Madrid Stock Exchange observed from January 1992 to December 1999 and, finally, the S&P 500 index observed from November 1987 to December 1998. All the series have been filtered when necessary to get rid of a small first order autocorrelation in the levels and the presence of outliers [4]. In this table, it is possible to observe that all the series have zero mean and excess kurtosis. It is also important to note that, although the series are not autocorrelated, the squared observations are correlated. Therefore, the variables are not serially independent. Finally, note that the autocorrelations of squared observations start at low levels.

Figures 6 and 7 contain plots of daily returns of all the series. It is possible to observe volatility clustering with days of large movements in prices followed by days with large returns in absolute value. These Figures also give kernel estimates of the marginal density of returns together with the corresponding normal density. The density plots confirm the results reported in table 2 about the returns being heavy-tailed. Finally, correlograms of

[4] See Carnero et al. (2001) for a review on the effects of the simultaneous presence of outliers and conditional heteroscedasticity on the diagnostic and estimation of GARCH models.

the series y_t^2 are also plotted. The volatility clustering is reflected in the significant correlations of squared returns. In particular, in the correlogram of y_t^2 the autocorrelations start at low values but are significant even for very large lags.

Table 3 reports the ML estimates of the parameters of the Normal GARCH(1,1) model for all the series considered [5]. In this table it is possible to observe that all the series considered have significant ARCH effects and high persistence measured by $\widehat{\alpha}+\widehat{\beta}$. Model diagnostics are based on the standardized observations defined as $\widehat{\varepsilon}_t = y_t/\widehat{\sigma}_t$, where $\widehat{\sigma}_t$ is obtained substituting the estimated parameters in the corresponding expression of the conditional variance. In table 3, we also report several sample moments of $\widehat{\varepsilon}_t$. Note that the standardized observations have still heavy tails. However, the autocorrelations of squares are not any longer significant. Therefore, the GARCH(1,1) model is able to represent adequately the dynamics of squares of the financial series considered although it is not able to explain the excess kurtosis present in the data. The last row of table 3 reports the number of standardized observations bigger than 3.5 standard deviations. These observations, that could be considered as "conditional" outliers, may explain why the standardized observations have excess kurtosis.

Given that the standardized observations by the Gaussian GARCH(1,1) model are leptokurtic, in table 4 we report the estimation results of GARCH(1,1)-t models fitted to the same time series. There are not big differences between the parameter estimates of GARCH(1,1)-normal models in table 3. However, notice that, except for Bombay, the estimated persistence is even greater than

[5] The estimation has been carried out with EViews, version 3.1.

in table 3. Also, it is important to point out that the condition for the existence of the four order moment is violated by the estimated parameters of US-SW, US-UK, AMST, DWJ and S&P 500. Finally, the standardized observations have, as expected, excess kurtosis, since we are assuming a Student-t distribution for the conditional distribution.

Table 5 reports the estimates of the parameters of the ARSV(1) model. The estimates have been obtained using the QML method proposed independently by Harvey et al. (1994) and Nelson (1988)[6]. The asymptotic standard errors of the QML estimators of the parameters ϕ, σ_η^2 and σ_ξ^2 have been computed using the results in Ruiz (1994). The scale parameter $\sigma_\star^2$ can be estimated using the sample variance of the heteroscedasticity corrected observations. Under normality of ε_t and for large sample size, T, the variance of this estimator is $4.93\sigma_\star^4/T$; see Harvey and Shephard (1993). Given that under normality, $\sigma_\xi^2 = \pi^2/2$, a natural test for normality in this framework is to test the null hypothesis $H_0 : \sigma_\xi^2 = \pi^2/2$ using a Wald test. The test only rejects the null for four of the twelve time series: US-CA, US-JA, US-UK and BOMBAY. However, as in table 3, the standardized observations still have excess kurtosis, although smaller than when the GARCH(1,1) model is fitted. This fact can also be observed in the number of standardized observations greater than 3.5 standard deviations, which, for most of the series, is smaller than in table 3.

Figure 8 plots density estimates of standardized observations with GARCH(1,1)

[6]The estimation of the parameters is based on obtaining the prediction error decomposition of the Gaussian likelihood of the log-squared observations given by $\log y_t^2 = \mu + \log\sigma_t^2 + \xi_t$ where $\mu = \log\sigma_\star^2 + E(\log\varepsilon_t)$ and $\xi_t = \log\varepsilon_t^2 - E(\log\varepsilon_t)$. Notice that when ε_t is Normal, ξ_t has a $\log(\chi^2_{(1)})$ distribution with variance $\pi^2/2$.

and ARSV(1) models together with the Normal density for the US Dollar/British Pound exchange rate and the Amsterdan E.O.E. index. As we can see, the Gaussian ARSV(1) specification seems to be more adequate than the GARCH(1,1). For the rest of the series both specifications seem to be very similar in terms of estimated densities for the standardized observations.

Notice that for the series US-CA, US-GE, US-UK and BOMBAY the persistence estimated with the ARSV(1) model is smaller than the one estimated with the GARCH(1,1). This could be due to the fact that the GARCH(1,1) model need to have a persistence very close to one to explain high kurtosis and low $\rho_2(1)$ and therefore, the high persistence found in these series with the GARCH(1,1) model could be spurious. For the rest of the series the persistence is estimated very close to one in both models, indicating that, in fact, there is persistence in variance.

Figure 9 plots the sample kurtosis, first order autocorrelation and $\widehat{\alpha} + \widehat{\beta}$ estimated together with the moments implied by the Gaussian ARSV(1) and GARCH(1,1)-t_{10} models. Observe that the ARSV(1) model is closer to most of the empirical points than the GARCH(1,1) model. Therefore, it seems that ARSV(1) models are in closer conformance with real data than GARCH(1,1) models. To illustrate more clearly this point, table 6 reports the sample moments implied by the Gaussian GARCH(1,1), GARCH(1,1)-t and Gaussian ARSV(1) models estimated for the US Dollar/Spanish Peseta exchange rate and the Dow Jones index. Notice that the kurtosis coefficients implied by the GARCH(1,1)-t models are either not defined or too big compared with the sample kurtosis in table 2. The same has been observed for all other returns considered. On the other hand, comparing the moments implied by the

Gaussian GARCH(1,1) and Gaussian ARSV(1) models, the latter is usually closer to the sample moments in table 2. For some of the other series, the evidence is mixed.

Finally, since in finance there is also a big interest for the estimates of the volatility itself we have compared the GARCH(1,1) and ARSV(1) models with respect to the estimated volatilities. Notice that the ARSV(1) specification of the volatility allows to obtain smoothed estimates of σ_t using the whole sample, $(y_1, \ldots, y_T)$; see, for example, Harvey et al. (1994). However, in GARCH(1,1) models, since σ_t is observable at time $t-1$, the observations at time t and later do not modify the estimate of σ_t. Figure 10 plots GARCH(1,1) and ARSV(1) estimates of the volatility for the Dow Jones and the Ibex 35 indexes. It can be seen that the ARSV(1) specification produces smoother volatility estimates than the GARCH(1,1). ARSV(1) estimates are less sensitive to large movements in prices than GARCH(1,1) estimates. The dynamic shape of the two series of estimates is very similar. The only point worth to notice is that, as can be seen in the plot od the Dow Jones estimates, for most of the returns considered, the volatilities estimated by the ARSV(1) model are, usually, over the GARCH(1,1) ones.

5 Conclusions

In this paper we have shown that ARSV(1) models are more flexible than GARCH(1,1) models to explain the excess kurtosis, low first order autocorrelation and high persistence of volatility often observed in real high frequency financial time series. The properties of both models are similar when volatility is close to the non stationarity region. Our results can explain why in the empirical analysis of real financial series, it has often been observed that when the GARCH(1,1) specification is chosen for the volatility, the conditional distribution of returns needs to have heavy tails while when the volatility is modelled with an ARSV(1) model, the assumption of Gaussianity seems to be adequate. We also show why the persistence of volatility estimated in GARCH(1,1) models is usually higher than that estimated in ARSV(1) models and why the conditional distribution assumed has systematic effects on the parameter estimates of the volatility process.

Most of the papers that compare empirically ARSV(1) and GARCH(1,1) models suggest that ARSV(1) models are more adequate than GARCH(1,1) models to explain the stylized facts observed in real time series and the empirical analysis presented in this paper confirms these results. However, the estimates of the volatility obtained with each of the models considered are very similar. Therefore, it seems that the definite proof for these models is to analyze their performance in predicting future returns and volatilities. So et al. (1999) compare the predictive performance of both models analyzing returns of five exchange rates. They conclude that both models have similar performance in terms of the Mean Square Prediction Error and Mean Absolute Prediction Error. Only in two of the five series considered, the ARSV predictions of volatility outperform the GARCH predictions. However, it could be interesting to compare confidence intervals for the predicted volatility generated by each of the models. In this sense, the procedure proposed by Pascual et al. (2000) seems to be very promising.

References

Andersen, T.G. (1992). Volatility. Discussion paper, Northwestern University.

Andersen, T.G., T. Bollerslev, F.X. Diebold, and P. Labys (1999). The Distribution of Exchange Rate Volatility. W.P. 99-08 and 6961, Wharton Financial Institutions Center and NBER.

Baillie, R.T. and T. Bollerslev (1989). The Message in Daily Exchange Rates: A Conditional Tale. *Journal of Business and Economic Statistics*, **7**, 297–305.

Bera, A.K. and M.L. Higgins (1995). On ARCH Models: Properties, Estimation and Testing. In: *Surveys in Econometrics,* L. Oxley, D.A.R. George, C.J. Roberts and S. Sayer (eds). Oxford:Blackwell.

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, **31**, 307–327.

Bollerslev, T. (1987). A Conditionally Heteroskedasticity Time Series Model for Speculative Prices and Rates of Return. *The Review of Economics and Statistics*, **69**, 542–547.

Bollerslev, T. (1988). On the Correlation Structure for the Generalized Autoregressive Conditional Heteroskedastic Process. *Journal of Time Series Analysis*, **9**, 121–131.

Bollerslev, T., R.Y. Chou, and K.F. Kroner (1992). ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence. *Journal of Econometrics*, **52**, 498–505.

Bollerslev, T., R. F. Engle, and D. B. Nelson (1994). ARCH models. *The Handbook of Econometrics*, **4**, 2959–3038.

Bougerol, P. and N. Picard (1992). Stationarity of GARCH Processes and Some Nonnegative Time Series. *Journal of Econometrics*, **52**, 115–128.

Carnero, M.A., D. Peña, and E. Ruiz (2001). Outliers and Conditional Autoregressive Heteroscedasticity in Time Series. W.P. 01-07, Universidad Carlos III de Madrid.

Chib, S., F. Nardari, and N. Shephard (1998). Markov Chain Monte Carlo Methods for Generalized Stochastic Volatility Models. W.p., John M. Olin School of Business, Washington University.

Danielsson, J. (1994). A Comment on Bayesian Analysis of Stochastic Volatility Models. *Journal of Business and Economic Statistics*, **12**, 393–395.

Danielsson, J. (1998). Multivariate Stochastic Volatility Models: Estimation and a Comparison with VGARCH Models. *Journal of Empirical Finance*, **5**, 155–173.

Diebold, F.X. and J.A. López (1995). ARCH Models. In: *Macroeconomics: Developments, Tensions and Prospects* , K. Hoover (eds). Oxford:Blackwell.

Engle, R.F. (1995). *ARCH*. Oxford Universtity Press.

Engle, R. F. (1982). Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation. *Econometrica*, **50**, 987–1007.

Engle, R. F. and T. Bollerslev (1986). Modelling the Persistence of Conditional Variances. *Econometric Reviews*, **5**, 1–50.

Gallant, A.R., D. Hsieh, and G. Tauchen (1994). Estimation of Stochastic Volatility Models with Diagnostics. W.P. 95-36, Duke Economics Working Paper.

Ghysels, E., A. Harvey, and E. Renault (1996). Stochastic Volatility. *Handbook of Statistics*, **14**.

González-Rivera, G. (1997). A Note on Adaptation in GARCH Models. *Econometric Reviews*, **16**, 55–68.

Granger, C.W.J. and Z. Ding (1995). Some Properties of Absolute Return, an Alternative Measure of Risk. *Annales d'Economie et de Statistique*, **40**, 67–90.

Harvey, A., E. Ruiz, and N. Shephard (1994). Multivariate Stochastic Variance Models. *Review of Economic Studies*, **61**, 247–264.

Harvey, A. and N. Shephard (1993). The Econometrics of Stochastic Volatility. W.P. 166, London School of Economics.

He, C. and T. Teräsvirta (1999). Properties of Moments of a Family of GARCH Processes. *Journal of Econometrics*, **92**, 173–192.

Hsieh, D.A. (1989). Modelling Heteroskedasticity in Daily Foreign Exchange Rates. *Journal of Business and Economic Statistics*, **7**, 307–317.

Hwang, S. and S.E. Satchell (2000). Market Risk and the Concept of Fundamental Volatility: Measuring Volatility across Asset and Derivative Markets and Testing for the Impact of Derivatives Markets on Financial Markets. *Journal of Banking and Finance*, **24**, 759–785.

Jacquier, E., N.G. Polson, and P.E. Rossi (1994). Bayesian Analysis of Stochastic Volatility Models (with discussion). *Journal of Business and Economic Statistics*, **12**, 371–417.

Jorion, P. (1988). On Jump Processes in the Foreign Exchange and Stock Markets. *Review of Financial Studies*, **1**, 427–445.

Kim, S., N. Shephard, and S. Chib (1998). Stochastic Volatility: Likelihood Inference and Comparison with ARCH Model. *Review of Economic Studies*, **65**, 361–393.

Liesenfeld, R. and R.C. Jung (2000). Stochastic Volatility Models: Conditional Normality versus Heavy-Tailed Distributions. *Journal of Applied Econometrics*, **15**, 137–160.

Loudon, G.F., W.H. Watt, and P.K. Yadav (2000). An Empirical Analysis of Alternative Parametric ARCH Models. *Journal of Applied Econometrics*, **15**, 117–136.

Mahieu, R.J. and P.C. Schotman (1998). An Empirical Application of Stochastic Volatility Models. *Journal of Applied Econometrics*, **13**, 333–360.

Nelson, D.B. (1988). *Time Series Behavior of Stock Market Volatility and Returns*. Ph. D. thesis, M.I.T.

Nelson, D. B. (1990). Stationary and Persistence in the GARCH(1,1) Model. *Econometric Theory*, **6**, 318–344.

Nelson, D. B. (1991). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Econometrica*, **59**, 347–370.

Nelson, D. B. and C.Q. Cao (1992). Inequality Constraints in the Univariate GARCH Model. *Journal of Business and Economic Statistics*, **10**, 229–235.

Palm, F.C. (1997). GARCH Models of Volatility. *Handbook of Statistics*, **14**.

Pascual, L., J. Romo, and E. Ruiz (2000). Forecasting Returns and Volatilities in GARCH Processes using the Bootstrap. W.P. 00-68, Universidad Carlos III de Madrid.

Ruiz, E. (1994). Quasi-Maximum Likelihood Estimation of Stochastic Volatility Models. *Journal of Econometrics*, **63**, 289–306.

Sandmann, G. and S.J. Koopman (1998). Estimation of Stochastic Volatility Models via Monte Carlo Maximum Likelihood. *Journal of Econometrics*, **87**, 271–301.

Shephard, N. (1996). Statistical Aspects of ARCH and Stochastic Volatility. In: *Time Series Models in Econometrics, Finance and Other Fields,* D.R. Cox, D.V. Hinkley and O.E. Bardorff-Nielsen (eds). Chapman & Hall, London.

So, Mike K.P., K. Lam, and W.K. Li (1999). Forecasting Exchange Rate Volatility using Autoregressive Random Variance Model. *Applied Financial Economics*, **9**, 583–591.

Taylor, S.J. (1986). *Modelling Financial Time Series*. John Wiley.

Taylor, S.J. (1994). Modelling Stochastic Volatility: A Review and Comparative Study. *Mathematical Finance*, **4**, 183–204.

Teräsvirta, T. (1996). Two Stylized Facts and the GARCH(1,1) Model. W.P. Series in Finance and Economics 96, Stockholm School of Economics.

Figure 1: Autocorrelation function of squared observations of GARCH(1,1) and ARSV(1) processes with the same kurtosis and approximately the same persistence

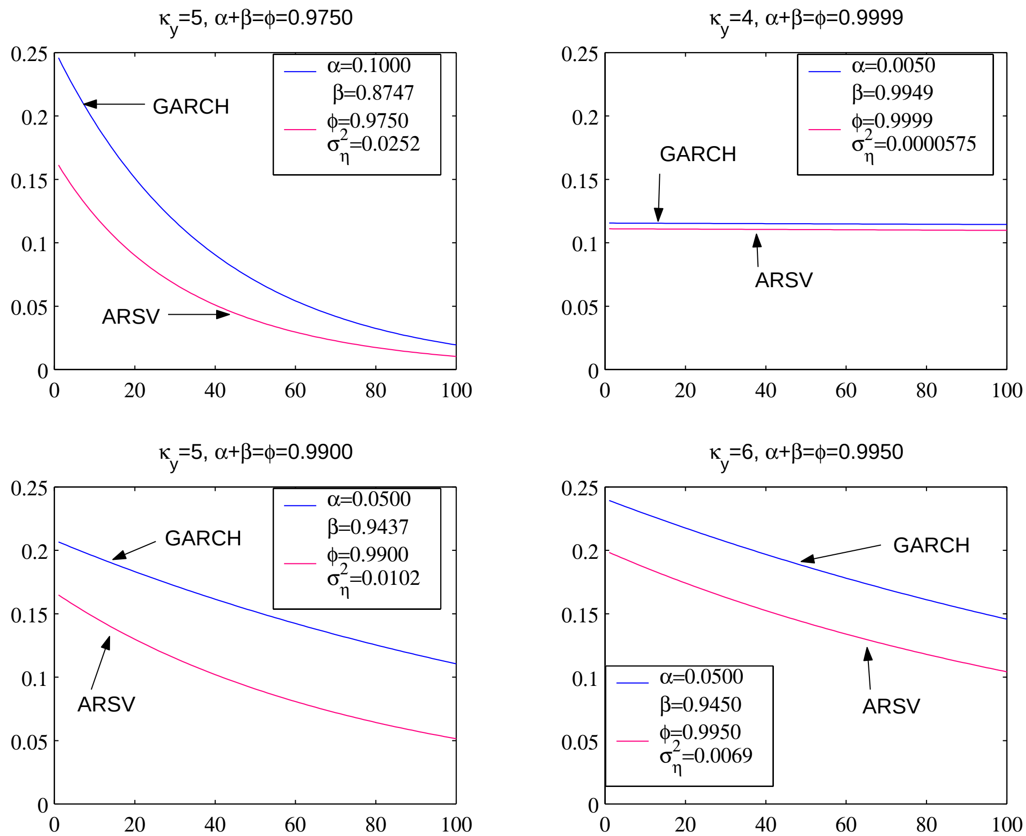

Table 1: Autocorrelation function of squared observations of GARCH(1,1) and ARSV(1) processes

	$\kappa_y = 4$		$\kappa_y = 5$		$\kappa_y = 6$		$\kappa_y = 7$	
	GARCH	ARSV	GARCH	ARSV	GARCH	ARSV	GARCH	ARSV
	$\alpha = 0.2000$ $\beta = 0.6246$	$\sigma_\eta^2 = 0.0919$ $\phi = 0.8250$	$\alpha = 0.2000$ $\beta = 0.6944$	$\sigma_\eta^2 = 0.1016$ $\phi = 0.8950$	$\alpha = 0.2000$ $\beta = 0.7165$	$\sigma_\eta^2 = 0.1109$ $\phi = 0.9165$	$\alpha = 0.2000$ $\beta = 0.7411$	$\sigma_\eta^2 = 0.1401$ $\phi = 0.9400$
$\rho_2(1)$	0.2694	0.0893	0.3157	0.1449	0.3433	0.1775	0.3921	0.2334
$\rho_2(2)$	0.2222	0.0721	0.2824	0.1264	0.3146	0.1580	0.3691	0.2108
$\rho_2(3)$	0.1832	0.0584	0.2526	0.1106	0.2884	0.1410	0.3473	0.1909
$\rho_2(4)$	0.1511	0.0475	0.2259	0.0970	0.2643	0.1262	0.3269	0.1733
$\rho_2(5)$	0.1246	0.0387	0.2021	0.0852	0.2422	0.1131	0.3076	0.1577
	$\alpha = 0.1000$ $\beta = 0.8592$	$\sigma_\eta^2 = 0.0231$ $\phi = 0.9590$	$\alpha = 0.1000$ $\beta = 0.8747$	$\sigma_\eta^2 = 0.0252$ $\phi = 0.9750$	$\alpha = 0.1000$ $\beta = 0.8798$	$\sigma_\eta^2 = 0.0277$ $\phi = 0.9798$	$\alpha = 0.1000$ $\beta = 0.8856$	$\sigma_\eta^2 = 0.0344$ $\phi = 0.9856$
$\rho_2(1)$	0.1955	0.1059	0.2458	0.1614	0.2760	0.1944	0.3296	0.2529
$\rho_2(2)$	0.1875	0.1010	0.2396	0.1563	0.2704	0.1891	0.3249	0.2467
$\rho_2(3)$	0.1798	0.0963	0.2335	0.1514	0.2649	0.1839	0.3202	0.2408
$\rho_2(4)$	0.1725	0.0918	0.2276	0.1467	0.2596	0.1788	0.3156	0.2350
$\rho_2(5)$	0.1654	0.0876	0.2218	0.1421	0.2543	0.1740	0.3110	0.2293
	$\alpha = 0.0500$ $\beta = 0.9399$	$\sigma_\eta^2 = 0.0057$ $\phi = 0.9900$	$\alpha = 0.0500$ $\beta = 0.9437$	$\sigma_\eta^2 = 0.0102$ $\phi = 0.9900$	$\alpha = 0.0500$ $\beta = 0.9450$	$\sigma_\eta^2 = 0.0069$ $\phi = 0.9950$	$\alpha = 0.0500$ $\beta = 0.9464$	$\sigma_\eta^2 = 0.0086$ $\phi = 0.9964$
$\rho_2(1)$	0.1544	0.1098	0.2073	0.1645	0.2390	0.1986	0.2954	0.2577
$\rho_2(2)$	0.1529	0.1086	0.2060	0.1624	0.2378	0.1972	0.2943	0.2561
$\rho_2(3)$	0.1514	0.1073	0.2047	0.1604	0.2366	0.1959	0.2933	0.2545
$\rho_2(4)$	0.1498	0.1061	0.2034	0.1584	0.2354	0.1945	0.2922	0.2529
$\rho_2(5)$	0.1483	0.1049	0.2021	0.1564	0.2342	0.1932	0.2912	0.2514
	$\alpha = 0.0050$ $\beta = 0.9949$	$\sigma_\eta^2 = 5.75 \times 10^{-5}$ $\phi = 0.9999$	$\alpha = 0.0050$ $\beta = 0.9949$	$\sigma_\eta^2 = 10^{-4}$ $\phi = 0.9999$	$\alpha = 0.0050$ $\beta = 0.9949$	$\sigma_\eta^2 = 1.38 \times 10^{-4}$ $\phi = 0.9999$	$\alpha = 0.0050$ $\beta = 0.9949$	$\sigma_\eta^2 = 2.4 \times 10^{-4}$ $\phi = 0.9999$
$\rho_2(1)$	0.1155	0.1111	0.1708	0.1666	0.2040	0.2000	0.2630	0.2592
$\rho_2(2)$	0.1155	0.1111	0.1708	0.1666	0.2040	0.1999	0.2629	0.2592
$\rho_2(3)$	0.1155	0.1111	0.1708	0.1666	0.2040	0.1999	0.2629	0.2591
$\rho_2(4)$	0.1151	0.1111	0.1708	0.1666	0.2040	0.1999	0.2629	0.2591
$\rho_2(5)$	0.1150	0.1110	0.1708	0.1666	0.2039	0.1999	0.2629	0.2590

Table 2: Descriptive statistics of daily returns

	US CA	US ES	US GE	US JA	US SF	US SW	US UK	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
T	1963	1963	1963	1963	1963	1963	1963	1261	1262	2728	1982	2888
Mean	0.0078	0.0224	0.0137	-0.0004	0.0097	0.0149	0.0009	0.0818	0.0047	-0.0001	-0.0001	-0.0001
S.D.	0.3042	0.6167	0.6089	0.7568	0.6808	0.6267	0.5013	1.3468	1.7867	0.8905	1.2031	0.8582
Skew.	-0.1462*	-0.2858*	-0.3201*	-0.2605*	-0.2865*	-0.1535*	-0.0109	-0.3101*	-0.1602*	-0.1678*	-0.0871*	-0.1826*
kurt.	5.0285*	4.4661*	4.5399*	5.6028*	4.7137*	4.4174*	4.8952*	5.7362*	5.9110*	4.9289*	5.4521*	5.2678*
r(1)	0.0000	-0.0300	0.0200	-0.0200	0.0100	0.0050	0.0100	0.0300	0.0000	0.0000	0.0000	0.0000
$Q(20)$	25.9	22.4	21.0	20.4	17.1	20.07	28.3	39.7*	31.3	30.5	23.4	33.3*
Autocorrelations of y_t^2												
$r_2(1)$	0.1000*	0.0900*	0.1100*	0.1300*	0.1300*	0.0600*	0.1200*	0.1800*	0.2050*	0.0800*	0.1900*	0.0800*
$r_2(2)$	0.0900*	0.0600*	0.0300	0.0800*	0.0100	0.0800*	0.1000*	0.2700*	0.1900*	0.1000*	0.2000*	0.1000*
$r_2(5)$	0.0800*	0.0300	0.0500*	0.0900*	0.0700*	0.0300	0.0900*	0.1300*	0.1230*	0.1000*	0.1300*	0.1300*
$r_2(10)$	0.1100*	0.0600*	0.0500*	0.0400*	0.0600*	0.0400*	0.0900*	0.1600*	0.0550*	0.1200*	0.1200*	0.1500*
$Q_2(20)$	314*	111*	129*	235*	157*	141*	273*	749*	223*	503*	980*	854*

T: Sample size.
* Significant at the 5% level.
r(k): Autocorrelation of order k of the original observations y_t.
$r_2(k)$: Autocorrelation of order k of the squared observations y_t^2.
$Q(20)$ and $Q_2(20)$: Box-Ljung statistic for y_t and y_t^2 respectively (31.4 is the 5% critical value).

Table 3: Estimated Gaussian GARCH(1,1) models and diagnostics for the series

	US CA	US ES	US GE	US JA	US SF	US SW	US UK	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
ω	0.0009 (0.0003)	0.0049 (0.0016)	0.0047 (0.0013)	0.0113 (0.0017)	0.0099 (0.0031)	0.0054 (0.0019)	0.0047 (0.0012)	0.0162 (0.0051)	0.1874 (0.0367)	0.0049 (0.0013)	0.0409 (0.0085)	0.0033 (0.0007)
α	0.0394 (0.0049)	0.0342 (0.0060)	0.0312 (0.0053)	0.0477 (0.0060)	0.0274 (0.0057)	0.0320 (0.0064)	0.0453 (0.0062)	0.0646 (0.0092)	0.0864 (0.0126)	0.0379 (0.0042)	0.0867 (0.0130)	0.0302 (0.0035)
β	0.9514 (0.0058)	0.9537 (0.0092)	0.9565 (0.0079)	0.9329 (0.0073)	0.9507 (0.0116)	0.9537 (0.0099)	0.9359 (0.0095)	0.9276 (0.0091)	0.8558 (0.0214)	0.9562 (0.0049)	0.8844 (0.0166)	0.9646 (0.0037)
$\alpha+\beta$	0.9908	0.9879	0.9862	0.9806	0.9781	0.9857	0.9812	0.9922	0.9422	0.9941	0.9711	0.9948
log L	-364.27	-1791.12	-1765.32	-2138.15	-1992.17	-1814.00	-1348.02	-2012.55	-2453.65	-3376.43	-3013.28	-3421.48
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$												
Mean	0.0332	0.0360	0.0220	-0.0040	0.0135	0.0250	-0.0028	0.0737	0.0040	-0.0039	-0.0009	-0.0082
S.D.	0.9965	0.9972	1.0004	1.0002	1.0013	1.0015	1.0026	0.9980	1.0001	1.0001	0.9988	1.0004
Skew	-0.0779	-0.2561*	-0.3574*	-0.4587 *	-0.3588 *	-0.1099*	-0.0350	-0.4011*	-0.1004*	-0.3290*	-0.1040*	-0.2930*
Kurtosis	4.3787*	4.5055*	4.4384*	5.6339*	4.4000*	4.1535*	4.4192*	4.4960*	6.2941*	4.7227*	4.7616*	4.5502*
r(1)	0.0130	-0.0300	0.0300	0.0060	0.0200	0.0100	0.0300	0.0100	0.0300	0.0100	0.0300	0.0100
$Q(20)$	22.23	18.32	21.03	20.22	16.74	16.95	21.85	21.73	37.25	16.92	18.98	25.31
$r_2(1)$	0.0250	0.0200	0.0300	0.0200	0.0200	0.0200	0.0300	-0.0200	0.0200	-0.0100	-0.0200	-0.0200
$r_2(2)$	-0.0100	-0.0100	-0.0200	-0.0100	-0.0200	-0.0100	-0.0100	0.0200	-0.0100	0.0100	0.0000	0.0100
$r_2(5)$	0.0100	-0.0100	-0.0100	-0.0100	0.0100	-0.0200	0.0200	-0.0100	0.0000	-0.0100	0.0000	0.0100
$r_2(10)$	-0.0100	0.0200	0.0100	-0.0300	0.0100	0.0100	0.0300	-0.0200	0.0000	0.0200	-0.0100	0.0100
$Q_2(20)$	10.05	9.41	14.61	13.02	10.80	18.23	18.40	11.98	13.71	6.02	10.21	14.49
Obs> 3.5	7(0.35%)	11 (0.56%)	8 (0.41%)	18 (0.92 %)	7 (0.35%)	7 (0.35%)	6 (0.31%)	3 (0.24%)	8 (0.63%)	15 (0.55%)	6 (0.30%)	14 (0.48%

* Significant at the 5% level.
r(k): Autocorrelation of order k of the original observations y_t.
$r_2(k)$: Autocorrelation of order k of the squared observations y_t^2.
$Q(20)$ and $Q_2(20)$: Box-Ljung statistic for y_t and y_t^2 respectively (31.4 is the 5% critical value).

Table 4: Estimated GARCH(1,1)-t_ν models and diagnostics for the series

	US CA	US ES	US GE	US JA	US SF	US SW	US UK	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
ω	0.0008 (0.0014)	0.0016 (0.0010)	0.0022 (0.0015)	0.0041 (0.0022)	0.0074 (0.0039)	0.0029 (0.0043)	0.0007 (0.0028)	0.0162 (0.0023)	0.2033 (0.0373)	0.0050 (0.0006)	0.0301 (0.0104)	0.0034 (0.0010)
α	0.0467 (0.0283)	0.0324 (0.0065)	0.0346 (0.0072)	0.0371 (0.0072)	0.0289 (0.0075)	0.0570 (0.0409)	0.0420 (0.0011)	0.0658 (0.0015)	0.1239 (0.0162)	0.0370 (0.0048)	0.0897 (0.0172)	0.0300 (0.0043)
β	0.9457 (0.0355)	0.9646 (0.0070)	0.9607 (0.0086)	0.9535 (0.0100)	0.9554 (0.0129)	0.9426 (0.0409)	0.9579 (0.0024)	0.9340 (0.0011)	0.8059 (0.0039)	0.9623 (0.0044)	0.8905 (0.0206)	0.9689 (0.0040)
$\alpha+\beta$	0.9924	0.9970	0.9953	0.9906	0.9843	0.9996	0.9999	0.9998	0.9298	0.9993	0.9802	0.9989
ν	7	7	7	5	7	5	7	9	5	9	9	7
log L	-324.1491	-1748.4	-1725.0	-2056.5	-1953.7	-1799.6	-1314.0	-2066.8	-2353.5	-3354.6	-2988.7	-3368.9
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$												
Mean	0.0338	0.0362	0.0218	-0.0048	0.0133	0.0238	-0.0062	0.0641	0.0048	-0.0024	-0.0012	-0.0076
S.D.	1.0012	0.9987	1.0000	1.0352	0.9988	0.9534	0.9744	0.9198	1.0379	0.9259	1.0040	0.9393
Skew	-0.0770	-0.2501*	-0.3843*	-0.4989*	-0.3714*	-0.0891	-0.0580	-0.4258*	-0.0930	-0.3252*	-0.1041*	-0.2934*
Kurtosis	4.4376	4.6745	4.6200	5.8181	4.4507	4.2814	4.7005	4.6373	6.6781	4.7414	4.9352	4.5242
r(1)	0.0100	-0.0300	0.0400	0.0100	0.0250	0.0100	0.0300	0.0050	0.0500	0.0100	0.0300	0.0100
$Q(20)$	23.56	17.37	20.63	21.15	16.93	17.36	25.75	23.37	35.18*	19.58	18.92	25.33
$r_2(1)$	0.0200	0.0100	0.0200	0.0300	0.0200	0.0200	0.0400	0.0000	0.0030	0.0050	-0.0200	-0.0100
$r_2(2)$	-0.0100	-0.0200	-0.0200	-0.0100	-0.0300	-0.0200	-0.0100	0.0400	-0.01300	0.0200	-0.0030	0.0080
$r_2(5)$	0.0100	-0.0200	-0.0200	0.0050	0.0100	-0.0300	0.0100	0.0200	-0.0060	0.0000	-0.0080	0.0080
$r_2(10)$	-0.0100	0.0100	0.0100	-0.0300	0.0030	0.0020	0.0200	-0.0100	-0.0100	0.0200	-0.0080	0.0200
$Q_2(20)$	9.83	9.46	12.77	12.29	10.04	16.84	18.08	19.53	17.58	8.95	10.43	15.43
Obs$>q$	0(0.00%)	0 (0.00%)	1 (0.05 %)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	0 (0.00%)	1 (0.04%)	2 (0.10%)	0(0.00 %)

q: Quantile equivalent to 3.5 in N(0,1), which is 7 for $\nu = 5$, 5.5 for $\nu = 7$ and 5 for $\nu = 9$.

Table 5: Estimated ARSV(1) models and diagnostics for the series

	US CA	US ES	US GE	US JA	US SF	US SW	US UK	AMST. E.O.E	BOMBAY S.M.I.	DOW JONES	IBEX 35	S&P 500
ϕ	0.9868 (0.0067)	0.9920 (0.0050)	0.9218 (0.0347)	0.9876 (0.0063)	0.9805 (0.0123)	0.9875 (0.0072)	0.9252 (0.0281)	0.9911 (0.0053)	0.8844 (0.0387)	0.9973 (0.0019)	0.9913 (0.0049)	0.9974 (0.0017)
σ^2_η	0.0103 (0.0054)	0.0041 (0.0026)	0.0548 (0.0336)	0.0102 (0.0053)	0.0066 (0.0052)	0.0063 (0.0040)	0.0811 (0.0403)	0.0142 (0.0069)	0.2773 (0.1212)	0.0020 (0.0010)	0.0063 (0.0034)	0.0032 (0.0013)
σ^2_ξ	5.6432 (0.2913)	5.3298 (0.2824)	4.8691 (0.2854)	5.6722 (0.2923)	4.9429 (0.2762)	5.3184 (0.2823)	5.7088 (0.3046)	5.4201 (0.3566)	6.5674 (0.4347)	5.2652 (0.2801)	5.2602 (0.2801)	5.3211 (0.2324)
$\sigma^2_\star$	0.0809 (0.0041)	0.3514 (0.0176)	0.3173 (0.0159)	0.4847 (0.0243)	0.4297 (0.0215)	0.3563 (0.0179)	0.1999 (0.0100)	1.3148 (0.0822)	2.2931 (0.1433)	0.6795 (0.0289)	1.1773 (0.0587)	0.6060 (0.0250)
log L	2710.4	2643.4	2590.7	2715.7	2570.9	2645.3	2757.2	1723.3	1899.3	3654.9	2663.1	3890.9
$\hat{\varepsilon}_t = \frac{y_t}{\hat{\sigma}_t}$												
Mean	0.0039	-0.0017	0.0051	0.0098	0.0022	-0.0031	-0.0076	0.0224	0.0078	0.0029	0.0059	-0.0032
S.D.	1.0003	1.0003	1.0003	1.0003	1.0003	1.0003	1.0003	1.0004	1.0004	1.0002	1.0003	1.0002
Skew	-0.0558	-0.2052*	-0.2904*	-0.3104*	-0.2865*	-0.0997*	-0.0311	-0.1406*	-0.095*3	-0.2086*	-0.1034*	-0.1699*
Kurtosis	4.3793*	4.5661*	4.1130*	5.1451*	4.2344*	3.7591*	3.7847*	3.9182*	5.3939*	4.3336*	4.7265*	4.2201*
r(1)	0.0040	-0.0400	0.0200	-0.0100	0.0200	0.0000	0.0200	-0.0100	0.0250	-0.0040	0.0100	-0.0050
$Q(20)$	29.28	21.85	18.59	18.72	18.39	17.13	23.76	25.73	25.46	15.37	14.01	25.79
$r_2(1)$	0.0400	0.0500	-0.0030	0.0100	0.0250	0.0300	0.0020	0.0000	-0.0380	0.0200	0.0300	-0.0200
$r_2(2)$	-0.0200	-0.0100	-0.0500	-0.0100	-0.0400	-0.0200	-0.0500	0.0200	-0.0180	0.0200	0.0600	0.0100
$r_2(5)$	0.0300	-0.0100	-0.0200	-0.0100	0.0040	-0.0300	-0.0040	-0.0200	-0.0270	0.0100	0.0400	0.0200
$r_2(10)$	0.0100	-0.0010	0.0040	-0.0300	0.0040	-0.0100	0.0100	-0.0200	0.0050	0.0200	0.0060	-0.0010
$Q_2(20)$	26.22	13.31	20.11	16.77	14.91	22.61	33.35	14.64	21.22	17.24	27.62	20.87
Obs$>$ 3.5	12 (0.61%)	7 (0.35%)	7 (0.35%)	10 (0.51%)	5 (0.25%)	6 (0.31%)	2 (0.10%)	7 (0.55%)	6 (0.47%)	13 (0.48%)	4 (0.20%)	10 (0.35%)

Table 6: Implied moments by the estimated models

	Implied by GARCH(1,1)	Implied by GARCH(1,1)-t	Implied by ARSV(1)
US-ES			
Mean	0.0000	0.0000	0.0000
Variance	0.4050	0.5333	0.3996
Kurtosis	3.3232	16.7163	3.8802
$r_2(1)$	0.0784	0.1762	0.1009
$r_2(2)$	0.0775	0.1757	0.10000
$r_2(5)$	0.0747	0.1741	0.0973
$r_2(10)$	0.0703	0.1715	0.0930
DOW JONES			
Mean	0.0000	0.0000	0.0000
Variance	0.8305	7.1429	0.8179
Kurtosis	3.9692	$\nexists$	4.3470
$r_2(1)$	0.1419	0.5128	0.1337
$r_2(2)$	0.1411	0.5125	0.1333
$r_2(5)$	0.1386	0.5114	0.1320
$r_2(10)$	0.1346	0.5096	0.1299

Figure 2: Relationship between kurtosis, first order autocorrelation of squared observations and persistence for GARCH(1,1) models

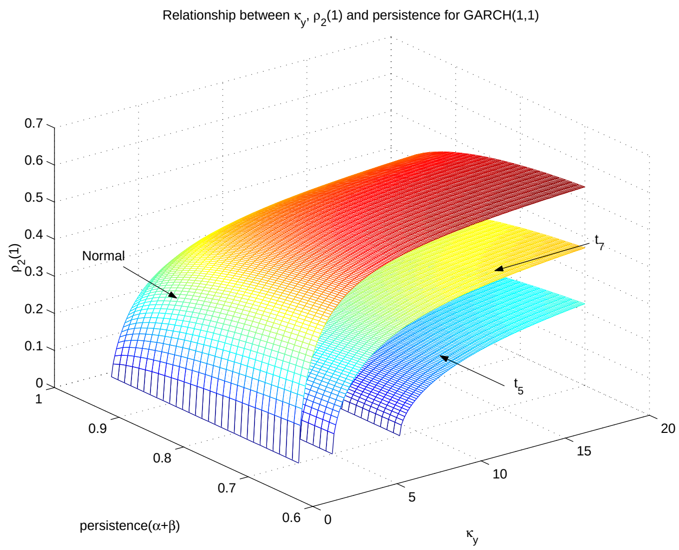

Figure 3: Relationship between kurtosis, first order autocorrelation of squared observations and persistence for GARCH(1,1) and ARSV(1) models with Gaussian errors

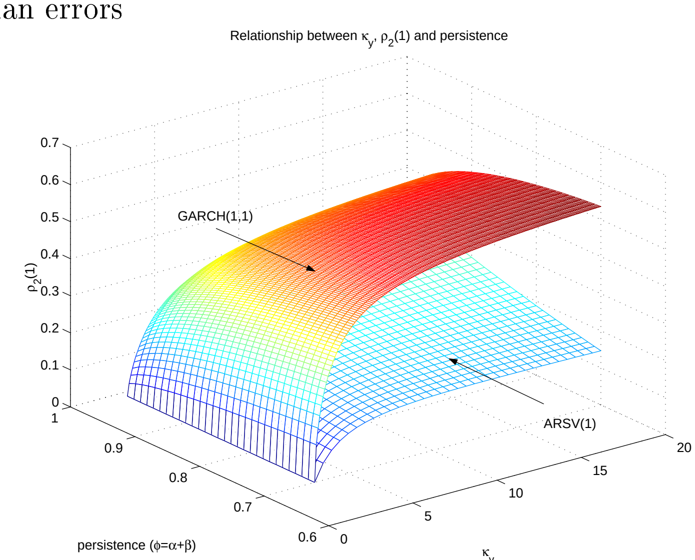

Figure 4: Relationship between kurtosis, first order autocorrelation of squared observations and persistence for GARCH(1,1)-t and Gaussian ARSV(1) models

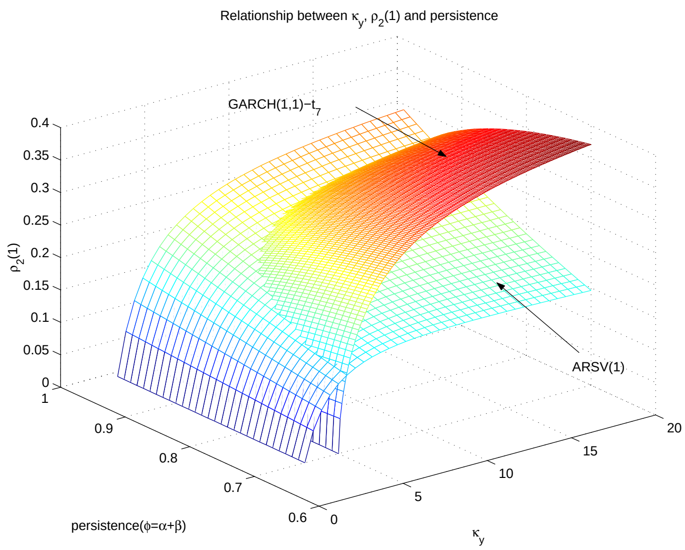

Figure 5: Relationship between kurtosis, first order autocorrelation of squared observations and persistence for ARSV(1)-t models

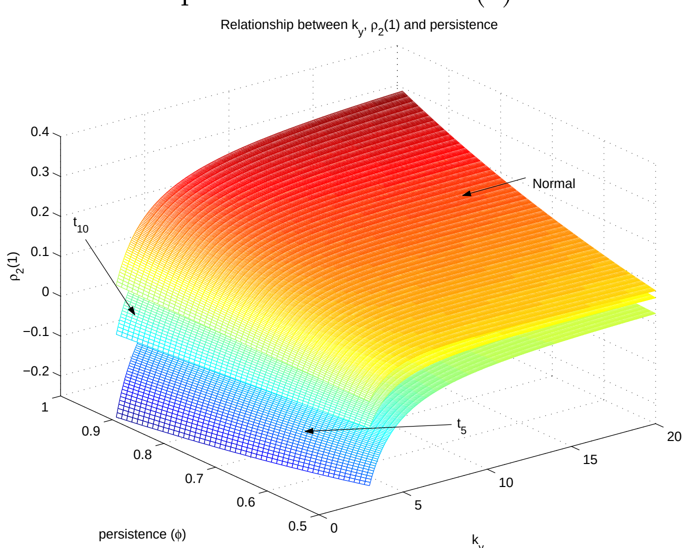

Figure 6: Daily exchange rates

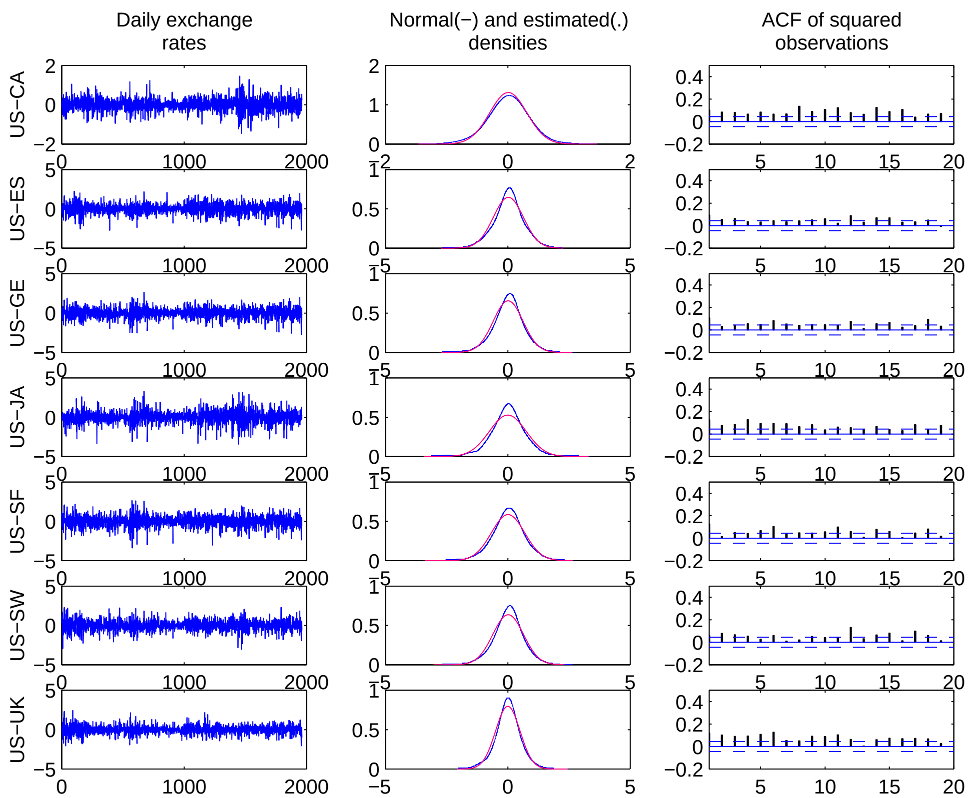

Figure 7: Daily financial indexes

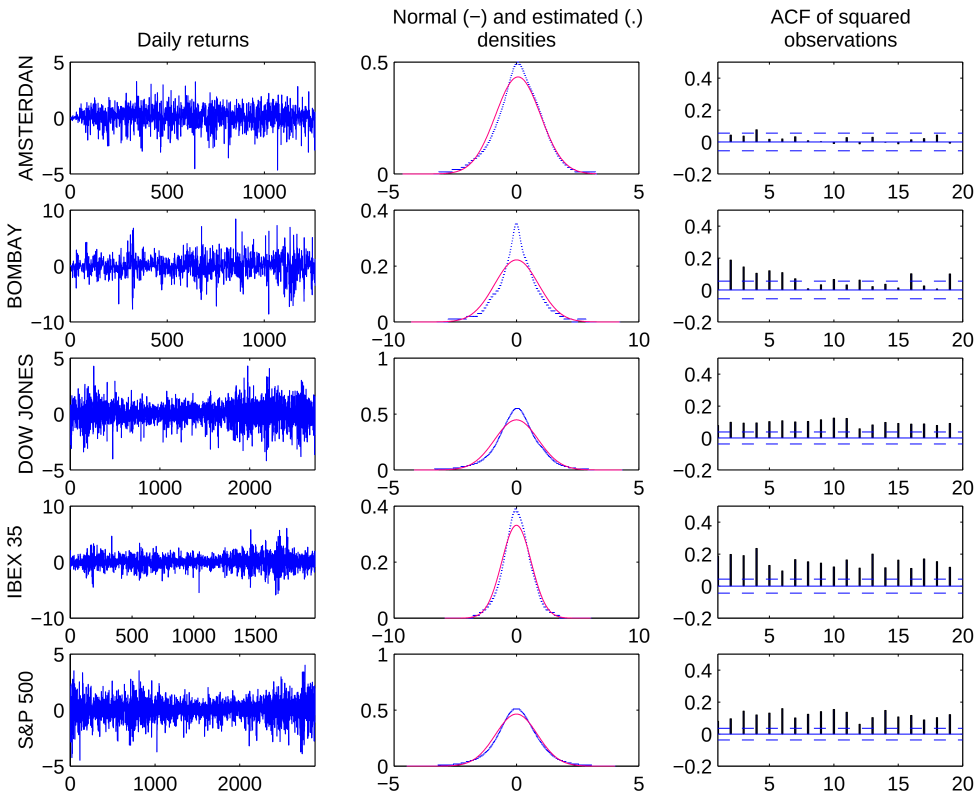

Figure 8: Densities of standardized observations for US Dollar/British Pound exchange rate and AMST. E.O.E.

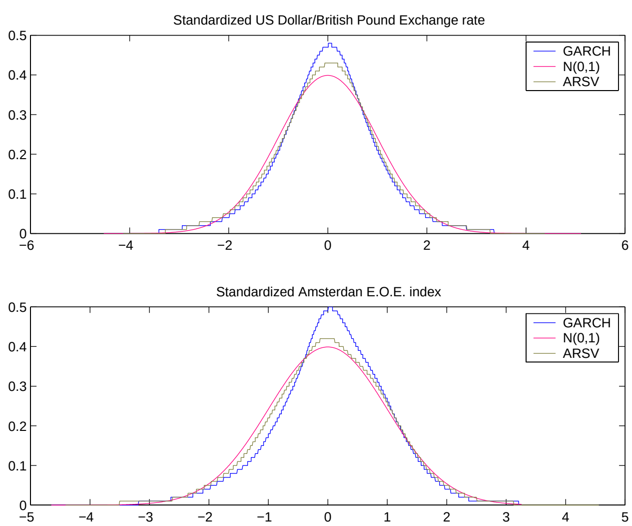

Figure 9: Relationship between κ_y, $\rho_2(1)$ and persistence for ARSV(1) and GARCH(1,1)-t models together with the sample values

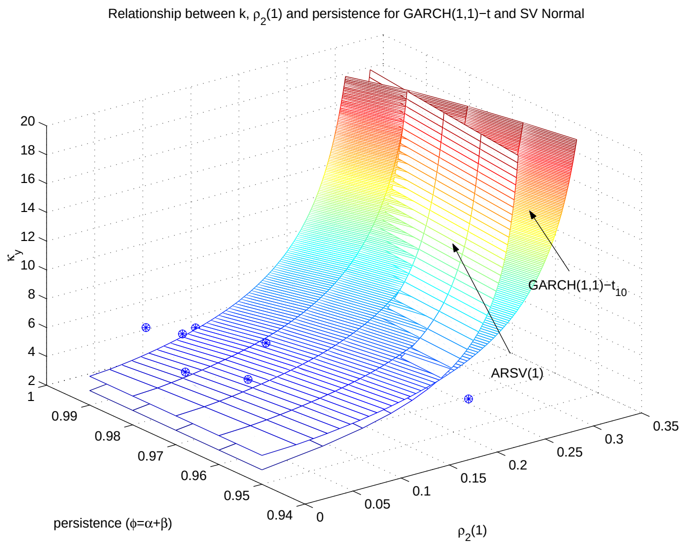

Figure 10: Estimated volatilities with GARCH(1,1) and ARSV(1) models for DOW JONES and IBEX 35 indexes