Working Paper 93-44
Statistics and Econometrics Series 26
December, 1993

Departamento de Estadística y Econometría
Universidad Carlos III de Madrid
Calle Madrid, 126
28903 Getafe (Spain)
Fax (341) 624-9849

STOCHASTIC VOLATILITY VERSUS AUTOREGRESSIVE CONDITIONAL HETEROSCEDASTICITY

Esther Ruiz*

Abstract

Several alternative models have been proposed in the literature to model time varying volatilities. There are two main classes of parametric models: ARCH (autoregressive conditional heteroscedasticity) and stochastic volatility (SV) models. In this paper, we fit three models, GARCH(1,1), EGARCH(1,0) and ARV(1), to daily exchange rates with the aim of investigating the different implications each model might have for the predictability of volatility. We will show how the SV within-sample estimates of volatility can be improved by using subsequent observations and therefore have better fits. When forecasting out-sample volatilities, the ARCH based volatilities can have severe biases which the SV volatilities do not have.

Key Words
EGARCH, exchange rates, GARCH, stochastic volatility.

*Departamento de Estadística y Econometría, Universidad Carlos III de Madrid. A preliminar version of this paper was presented at the Econometric Society Meeting in Uppsala, at the Financial Markets Dynamics and Forecasting Conference in Paris and at the ASSET Meeting in Barcelona. The author is grateful to all the participants and in particular to R. Engle and C. Granger. Support from the Applied Econometrics Program of the Argentaria Chair at the University Carlos III de Madrid and from the DGICYT, grant no. PB92-0244, is gratefully acknowledged.

1. Introduction

In the last decade there has been an increasing interest in modelling time varying volatilities. These models are particulary useful when dealing with high frequency financial time series. In the simplest set up, the series of interest is a white noise multiplied by σ_t, the volatility, that is

$$y_t = \sigma_t \varepsilon_t, \qquad \varepsilon_t \sim \text{IID}(0,1). \tag{1}$$

In this paper, we consider two alternative ways of modeling σ_t: models based on autoregressive conditional heteroscedasticity (ARCH) processes proposed by Engle (1982) and stochastic volatility (SV) processes proposed by Taylor (1986).

At present there is no way to identify the best type of discrete-time model for volatility. SV and ARCH models both can explain the following "stylized facts" often found in high frequency financial time series: (i) excess kurtosis, (ii) small autocorrelations and (iii) significant autocorrelations of the squared series. The objective of this paper is to compare empirically both alternatives. As the statistical properties of y_t implied by both kinds of models may be different, we will analyse which model is in closer conformance with the observed sample moments of the data. Also, we will investigate the different implications each model might have for the properties and predictibility of volatility both within-sample and out-sample.

The rest of the paper is organised as follows. In section 2 we analyse the statistical properties of two models based on the ARCH methodology: GARCH(1,1) and EGARCH(1,0). Section 3 deals with SV models and their properties, with special attention to the stationary AR(1) case. In section 4 we fit each model to the same data set, daily exchange rates of four international currencies against the dollar. Comparisons are made between the corresponding univariate models. Finally, section 5 presents the conclusions.

2. ARCH Based Models

The most popular processes for modelling σ_t are based on the ARCH models introduced by Engle (1982); see Bollerslev et al. (1992) and Bera and Higgins (1993) for detailed surveys on ARCH models. These models share the property of specifying σ^2_t as the conditional variance of y_t given $Y_{t-1}=\{y_1,\ldots,y_{t-1}\}$. As a result, if ε_t is assumed to be normally distributed, the ARCH based processes are conditionally Gaussian and, therefore, maximum likelihood (ML) estimation of the parameters is straightforward.

2.1 GARCH(1,1) Processes

The GARCH(1,1) model, proposed independently by Bollerslev (1986) and Taylor (1986), is given by

$$y_t = \varepsilon_t \, \sigma_t \qquad (2.a)$$

$$\sigma^2_t = \omega + \alpha \, y^2_{t-1} + \beta \, \sigma^2_{t-1} \qquad (2.b)$$

where the restrictions $\omega > 0$, $\alpha \geq 0$ and $\beta \geq 0$ are imposed to ensure the positivity of σ^2_t. All GARCH models are martingale differences, and if $\alpha+\beta<1$, they have constant finite variance and so they are white noise. If $\alpha+\beta<1$, the unconditional variance of y_t is given by

$$\sigma^2_y = \omega/(1-\alpha-\beta). \qquad (3)$$

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

$$\kappa_y = 3 + 6 \, \alpha^2 \, (1 - \beta^2 - 2 \, \alpha \, \beta - 3 \, \alpha^2)^{-1}. \qquad (4)$$

The dynamics of a GARCH model show up in the autocorrelation function (acf) of the squared observations. In the GARCH(1,1) case, the acf of the squares is like that of an ARMA(1,1) process. Bollerslev (1988) shows that the autocorrelations of y^2_t are given by

$$\rho_1 = \alpha\ (1-\alpha\beta-\beta^2)/(1-2\alpha\beta-\beta^2) \qquad (5.a)$$

$$\rho_h = (\alpha + \beta)^{h-1}\ \rho_1, \qquad h>1. \qquad (5.b)$$

The partial autocorrelation function will in general be infinite, but dominated by a damped exponential. Bollerslev also shows that the first two autocorrelations must lie in the region defined by

$$\rho_2 \geq \rho^2_1 \qquad (6.a)$$

$$\rho_2 < \rho_1, \qquad 0 \leq \rho_1 \leq 1/3 \qquad (6.b)$$

$$\rho_2 < (2/3 + 2^{3/4}\ \rho_1\ (1/3-\rho^2_1)^{1/2} - \rho^2_1)^{1/2}\ \rho_1 \qquad 1/3 \leq \rho_1 \leq (1/3)^{1/2}. \qquad (6.c)$$

2.2 EGARCH(1,0) Processes

Nelson (1991) points out some important limitations of the GARCH processes: i) the nonnegativity constrains on the parameters, which some times are not satisfyed in empirical analysis, ii) GARCH processes are not able to model the assymetry of volatility movements often observed in real data ("leverage" effect) and iii) the interpretation of persistence in the volatility of GARCH processes is not clear. Consequently, Nelson proposes the exponential GARCH process (EGARCH). Assuming normality of ε_t an EGARCH(1,0) process is given by

$$y_t = \varepsilon_t\ \sigma_t \qquad (7.a)$$

$$\log \sigma^2_t = \omega + \gamma\ \varepsilon_{t-1} + \alpha\ [|\varepsilon_{t-1}| - (2/\pi)^{1/2}] + \beta \log \sigma^2_{t-1} \qquad (7.b)$$

The moments of y_t have been derived by Nelson (1991), and have quite complicated expressions. The moments of $\log(y^2_t)$ have simpler expressions given by

$$E[\log(y^2_t)] = -1.27 + \omega/(1-\beta) \qquad (8.a)$$

$$Var[\log(y^2_t)] = \pi^2/2 + (\gamma^2+\alpha^2(1-2/\pi))/(1-\beta^2) \qquad (8.b)$$

$$\rho_h=\frac{\dfrac{\gamma^2+\alpha^2(1-\frac{2}{\pi})}{(1-\beta^2)}+1.1058\dfrac{\alpha}{\beta}}{\dfrac{\pi^2}{2}+\dfrac{\gamma^2+\alpha^2(1-\frac{2}{\pi})}{(1-\beta^2)}}\beta^h, \quad h\geq 1 \qquad (8.c)$$

see appendix 1 for the derivation of these expressions.

2.3 Estimation of GARCH(1,1) and EGARCH(1,0) processes

As we mentioned before, if ε_t is assumed to be normally distributed, y_t is conditionally Gaussian and, therefore, ML estimation of the parameters of any model based on ARCH is straightforward; see Bollerslev *et al.* (1993) for a extensive review of estimation of ARCH models.

It is also possible to assume other distributions for ε_t with fatter tails than the normal. For example, Bollerslev (1987) assumes a t-Student distribution in the case of a GARCH(1,1) process. Nelson (1991) considers ε_t having a standard general distribution (GED), so that ε_t may have a kurtosis which could be smaller or bigger than the normal kurtosis. The reason why the GED family of distributions is atractive is because it includes the normal as an special case when the parameter which regulates the thickness of the tails of the distribution, c, is equal to 2. In these cases is still possible to obtain analitical expressions for the likelihood. Recently, Bollerslev *et al.* (1993) have used the Generalized-t distribution, which includes both the Student-t and the GED distributions as particular cases. The Generalized-t distribution has two parameters to control the shape of the density.

3. Stochastic Volatility Models

In a SV process the volatility, σ_t, is modelled as an unobserved variable, the logarithm of which follows a linear stationary process, usually an autoregression. These models are denoted by Taylor (1993) as autoregressive random variance (ARV) models.

3.1 Properties of ARV(1) Processes

A simple stationary AR model is given by

$$y_t = \varepsilon_t \, \sigma_t, \qquad \varepsilon_t \sim \text{IID}(0,1) \tag{9.a}$$

$$\log \sigma^2_t = \gamma + \phi \log \sigma^2_{t-1} + \eta_t, \qquad \eta_t \sim \text{NID}(0,\sigma^2_\eta) \tag{9.b}$$

with η_t generated independently of ε_t. Model (9) will be called a ARV(1) model. In (9), we may observe that the mean of y_t and the volatility have separate noises. The restrictions needed to ensure stationarity of y_t are just the standard restrictions needed to ensure stationarity of $\log(\sigma^2_t)$, i.e. $|\phi| < 1$. The fact that y_t is white noise follows almost inmediately when ε_t and η_t are mutually independent. Even when ε_t and η_t are not mutually independent, y_t is a white noise; see Taylor (1993). Under normality of ε_t the odd moments of y_t are all zero. The variance and kurtosis of y_t can easily shown to be :

$$\text{Var}(y_t) = \exp\{\gamma_h + 0.5\, \sigma^2_h\} \tag{10.a}$$

$$\kappa_y = 3 \exp\{\sigma^2_h\} \tag{10.b}$$

where $\gamma_h = \gamma/(1-\phi)$ and $\sigma^2_h = \sigma^2_\eta/(1-\phi^2)$. Notice that the condition for the fourth order moment to exist is the same as the stationarity condition, i.e. $|\phi| < 1$. Furthermore, the σ^2_η parameter governs the degree of kurtosis independently of the degree of persistence in the volatility equation.

The dynamic properties of the model appear in the logarithm of the squared observations. The acf of $\log(y^2_t)$ is equivalent to that of an ARMA(1,1); see Harvey *et al.* (1994). The moments of $\log(y^2_t)$ can be easily seen to be

$$E(\log y^2_t) = -1.27 + \gamma_h$$

$$\mathrm{Var}(\log y^2_t) = \pi^2/2 + \sigma^2_h$$

$$\rho_k = \sigma^2_h \, \phi^k \, / \mathrm{Var}(\log y^2_t).$$

Finally, Jacquier *et al.* (1994) show that the acf of y^2_t is given by

$$\rho_i = \frac{\exp[\sigma^2_h \phi^i] - 1}{3 \exp[\sigma^2_h] - 1}. \tag{11}$$

3.2 Estimation of SV Models

Even assuming normality of ε_t, model (9) is not conditionally Gaussian and, therefore, its estimation may present some difficulties. There are three main estimation methods proposed in the literature. The methods used traditionally to estimate model (9) were based on the Method of Moments principle; see, for example, Melino and Turnbull (1990). These methods have the disadvantage of having their efficiency depending on the moments used for estimation. Also, numerical problems can occur when (9) is close to the nonstationarity, i.e. when ϕ is close to one, which is often the case in empirical applications. The second class of estimation methods use the Maximum Likelihood (ML) principle; see Danielsson (1992), Jacquier *et al.* (1994) and Shephard (1992). There are clear advantages to applying ML procedures, but these methods are, in general, very time consuming and difficult to apply. Finally, a Quasi-Maximum Likelihood (QML) estimator of model (9) has been proposed independently by Nelson (1988) and Harvey *et al.* (1994). This latter approach for estimation of SV models is the one adopted in this paper.

The QML approach is based on transforming y_t by taking logarithms of the squares and obtaining the following linear state space

$$\log y^2_t = h_t + \xi_t \tag{12.a}$$

$$h_t = \gamma^* + \phi\, h_{t-1} + \eta_t \qquad (12.b)$$

where $h_t=\log(\sigma^2_t)+E(\log(\varepsilon^2_t))$, $\xi_t = \log(\varepsilon^2_t) - E[\log(\varepsilon^2_t)]$ and $\gamma^*=\gamma+(1-\phi)E(\log(\varepsilon^2_t))$.

If ε_t is, for example, a GED variate then

$$E(\log \varepsilon^2_t) = (2/c)[\psi(1/c)+\log(2)] \qquad (13.a)$$

$$E(\xi^2_t) = (2/c)^2\, \psi'(1/c) \qquad (13.b)$$

where $\psi(\cdot)$ and $\psi'(\cdot)$ are the digamma and trigamma functions respectively.

Even under normality of ε_t, we can estimate more efficiently model (12) by leaving c as an unknown parameter, estimating σ^2_ξ and computing the value of c implied by (13.b); see Ruiz (1994). Then, using (13.a) we can compute $E(\log(\varepsilon^2_t))$ and obtain the estimate of γ by

$$\hat{\gamma} = \hat{\gamma}^* - (1 - \hat{\phi})\, E(\log(\varepsilon^2_t)).$$

The asymptotic distribution of the QML estimator can be easily obtained using the results in Ruiz (1994) by adequately substituting the corresponding moments of $\log(\varepsilon^2_t)$.

Given that under normality, $\sigma^2_\xi=\pi^2/2$, a natural test for normality in this framework is to test the null hypothesis, H_0: $\sigma^2_\xi=\pi^2/2$, using a Wald test.

4. Empirical Comparison

The ARCH and SV approaches to model volatility are different in the measurability properties of the volatility processes with respect to certain benchmark information sets; see Andersen (1992). As Taylor (1993) points out, the fundamental difference between the two types of models is that volatility news exclusively explain changes in volatility for the SV models whilst past prices are the mayor determinant of volatility changes for ARCH models. However, both EGARCH(1,0) and ARV(1) models are discrete approximations to the same diffusion process of interest in the continuos time asset pricing literature. In this sense Dassios (1992) shows that, when ε_t is assumed to be normal, the ARV(1) model converges at a more rapid rate than the EGARCH(1,0) model. On the other hand, in practice it is much more time consuming to estimate EGARCH models than to estimate SV models. Also, the convergence of the optimization algorithm for the maximization of the likelihood function for EGARCH models, is not always easy, beign highly dependent on the starting values for such algorithm.

At the moment, there are no studies comparing the performance of both approaches to model volatility. Only Danielsson (1992) presents goodness of fit evidence that SV models may perform better than some variants of ARCH models including the EGARCH model. In this section we compare both approaches empirically.

First, the fitted GARCH, EGARCH and ARV models may imply different moments of y_t. In particular, the implied acf's of y^2_t and $\log(y^2_t)$ could be quite different, being worth to investigate which one is in closer conformance with the data.

Also, researches are often interested in the conditional variance sequence or in the distribution of future values of y_t implied by the model. This is the reason why we will analyse the properties of the with-in sample one-step-ahead estimates of volatility obtained with each model and the different out-sample multi-step-ahead volatility estimates.

We estimate both types of models using daily data on four exchange rates: Pound/Dollar, Deustschemark/Dollar, Yen/Dollar and Swiss-Franc/Dollar. The data has been previously used in Harvey *et al.* (1994), and consist of daily observations of weekdays close exchange rates from 1/10/81 to 28/6/85. The sample size is 946.

For each exchange rate, the analysed series is the first differences of the logarithms of the spot price, p_t, i.e. the rates of return. For convenience, the rates of return have been centered about the sample mean prior to analysis. In consequence, the analysed series is given by

$$y_t = (\Delta \log(p_t) - \frac{\sum_2^T \Delta \log(p_t)}{(T-1)}) \times 100.$$

Table 1 shows several descriptive statistics of the series y_t, t=1,845. Because the Box-Ljung statistic for 10 lags is significant at the 5% level for the Swiss-Franc, we fit a MA(2) model and work with the residuals. The Box-Ljung statistic is not significant for any of the other exchange rates. However, when we look at the Box-Ljung statistics of y^2_t, we may observe that they are highly significant for all series. We should note that the conditions for the first and second autocorrelations of the squares of a GARCH(1,1) process given in (6) are only satisfyed by the Yen and Swiss-franc series. The sample moments of $\log(y^2_t)$ are also reported to make comparisons with the moments implied by each of the models analysed. Finally, table 1 shows the sample statistics $Q^3(-1)$ and $Q^3(1)$, where

$$Q^3(\tau) = \frac{\sum (y_t - \bar{y})(y^2_{t-\tau} - \bar{y}_2)}{\sqrt{\sum (y_t - \bar{y})^2 \sum (y^2_t - \bar{y}_2)^2}}$$

where $\bar{y}_2$ is the sample mean of y^2_t. For $\tau > 0$, $Q^3(\tau)$ measures volatility effects in the mean and, therefore the expected sign of the statistic is positive. On the other hand, for $\tau < 0$, $Q^3(\tau)$ measures the "leverage" effect and the expected sign is negative. In table 1, we observe significative volatility effects in the mean for the Yen and Swiss-Franc but with the wrong sign. The "leverage" effect is significative for the Deustschemark and the Swiss-Franc, but in the latter case has the wrong sign. As we will see later, there could be structural breaks in volatility in the Yen and Swiss-Franc exchange rates.

Figure 1 represents the recursive estimates of the standard deviation of each series given by

$$\hat{\sigma}_t=\sqrt{\sum_1^t \frac{y^2_i}{t}} \qquad (14)$$

It is quite clear that the standard deviation is not constant over time for any of the series. Also, it seems that there is a change in the level of the Yen volatility. This could be also the case for the Swiss-Franc.

4.1 Model Estimation

To estimate each of the three models considered in this paper we are using the first 846 observations of each series leaving the last 100 observations to make comparisons out sample.

Table 2 shows the estimation results of the GARCH(1,1) model and some implied moments of y_t. First of all, we should note that the GARCH estimates imply high persistence in volatility.

In table 3, we report the estimates of the EGARCH(1,0) model. As expected dealing with exchange rates, the estimates of the assymetry parameter, γ, are very close to zero. The order of magnitud of β, the parameter which measures the volatility persistence, is similar to the order implied by the GARCH estimates.

Finally, in table 4 we present the estimates of the ARV(1) model. Once more, we can observe the high persistence of volatility implied by the estimates. It is also possible to observe that the Wald test of normality (t-ratio) do not reject the normality hypothesis for any of the series considered.

The estimates of the parameters in all three models are close to the values which define non-stationary models. In the Yen case, the estimated unit root could be due to the existence of a change in the volatility level and consequently, we will focus on the other three exchange rates.

Another shared characteristic of the models is that all of them can explain both the high kurtosis of daily returns and the statistically significant positive autocorrelations of squared daily returns. However, comparing the acf of y^2_t implied by the three estimated models with the observed acf, we may observe that, for the Pound, the acf implied by the ARV model is in closer conformance with the data, while for the Deustschemark and the Swiss-Franc, none of the implied acf's seem to be in close conformance with the observed acf. With respect to the kurtosis, the kurtosis implied by the GARCH models is closer to the sample kurtosis for all exchange rates. In the case of the Swiss-Franc, all three models imply kurtosis much smaller than the sample kurtosis. With respect to the variances, all three models imply variances which are quite close to the observed ones, although the EGARCH variances are slightly closer.

Finally, comparing the moments of $\log(y^2_t)$ implied by the EGARCH and ARV estimates both are quite similar, with the ARV estimates being slightly better for the Deustschemark and the Swiss-Franc and the EGARCH estimates being better for the Pound.

Summarizing, it seems that for the data sets analyzed, the moments implied by all three models are rather similar and none of the models seem to clearly overperform the others in this sense.

4.2 Estimates of volatility with-in sample

Figure 2 represents the estimates of the with-in sample one-step ahead estimates of volatility, $\hat{s}_{t/t-1}$, for the GARCH(1,1), EGARCH(1,0) and ARV(1) models. In the case of the ARV(1) model, the volatility is an unobserved component, and a better estimator can be obtained by making use of subsequent observations; see, for example, Harvey (1989). Therefore, in figure 2, we also represent smoothed estimates of volatility, $\hat{s}_{t/T}$, for the ARV(1) model. We may observe that all four series of volatility estimates behave very similarly.

To analyse the properties of the one-step-ahead estimates of volatility, the moments of the standardized observations, $e_t = y_t/\hat{s}_{t/t-1}$, have been computed for each exchange rate and each model considered. For the ARV(1) model, we also report the moments of $y_t/\hat{s}_{t/T}$. The moments of e_t appear in tables 2 to 4 for the GARCH, EGARCH and ARV models respectively. First of all, from table 4, we should note that the moments of the observations standarized using the smoothed estimates of volatility are much satisfactory than using the one-step-ahead estimates, with the kurtosis and the $Q^2(10)$ statistic being smaller for all exchange rates. Therefore, the smoothed estimates should be used as estimates of volatility. In the Swiss-Franc case, it seems that there are important problems with the ARV estimates of volatility with $Q^2(10)$ being highly significant. For the Pound and the Deustschemark, the results are very similar for the three models, with the kurtosis of the observations standarized using the ARV(1) volatility estimates being smaller than for the other models.

Figure 3 represents kernel estimates of the densities of the standarized observations[1]. The estimated densities are quite similar for all three models. In all cases, there are more negative observations close to zero than expected under normality.

Finally, following Pagan and Schwert (1990), we will test if the one-step-ahead estimates of volatility are unbiased, i.e. if $E_{t-1}(y^2_t) = s^2_{t/t-1}$ by running the regressions $y^2_t = \alpha + \beta\, \hat{s}^2_{t/t-1} + \nu_t$ and testing $\alpha=0$ and $\beta=1$. The results appear in table 5. Standard errors using Newey and West (1987) autocorrelation-heteroscedasticity correction are in

[1]I am grateful to M. Delgado and C. del Río for providing me with the subroutines needed to compute the density.

parentheses under the coefficient estimates. R^2 is the coefficient of determination. We may observe that in the Swiss-Franc case, the residuals, $\hat{\nu}_t$, are highly correlated for all three models. The serial correlation in the residuals show that there are additional dynamics in volatility that are not capture by these models. As shown in figure 1, it is possible that there could be a structural change in this case. Consequently, we will compare the out-sample performance of the GARCH, EGARCH and ARV models only for the Pound and Deustschemark exchange rates. With respect to the Pound and Deustschemark series, the biggest R^2's values are obtained for the ARV(1) model. Except the EGARCH model of the Deustschemark exchange rate, all estimates of volatility have slight biases.

4.3 Out-sample predictions of volatility

Engle and Bollerslev (1986) show that in the GARCH(1,1) model, the prediction of σ^2_t τ-steps ahead is given by

$$E_t(\sigma^2_{t+\tau}) = \omega/(1-\alpha-\beta) + (\alpha+\beta)^{\tau-1}(\sigma^2_{t+1} - \omega/(1-\alpha-\beta)), \tau\geq 2. \qquad (15)$$

If y_t is an EGARCH(1,0) process as (7), then Nelson (1991) shows that

$$E_t(\sigma^2_{t+\tau})=\exp\{\omega(1-\beta^{\tau-1})/(1-\beta)+\beta^{\tau-1}\log\sigma^2_{t+1}\} \prod_{1}^{s-1} E\{\exp[\beta^{i-1}(\gamma\varepsilon_{t-i}+\alpha(|\varepsilon_{t-i}|-(2/\pi)^{½}))]\},\tau\geq 2 \qquad (16)$$

where denoting by z the expression inside the square brackets, its expected value is given by

$$E(z)=[\Phi(\beta^{i-1}(\alpha+\gamma))\exp\{0.5\beta^{2(i-1)}(\gamma+\alpha)^2\}+\Phi(\beta^{i-1}(\alpha-\gamma))\exp\{0.5\beta^{2(i-1)}(\alpha-\gamma)^2\}]\exp\{-\beta^{i-1}\alpha(2/\pi)^{½}\}$$

where Φ is the distribution function of a standard normal variable.

Finally, for the ARV(1) process in (9)

$$E_t(\sigma^2_{t+\tau}) = \exp\{\gamma(1-\phi^{\tau-1})/(1-\phi) + \phi^{\tau-1}\log \hat{s}^2_{t+1/t}\}, \qquad s\geq 2. \qquad (17)$$

In all three cases we estimate the τ-steps ahead predictions of σ^2_t by substituting the parameters in (15) to (17) by their estimates.

Tables 6 to 8 report the moments of the standardized observations $y_t/\hat{s}_{t/t-\tau}$ for $t=845+\tau,\ldots,945$ and $\tau=1,5$ and 20, where $\hat{s}_{t+\tau/t}$ are the volatilities estimated by the recursive standard deviations, the GARCH, EGARCH and ARV models. The results of the regressions of y^2_t on $\hat{s}^2_{t+\tau/t}$ for each procedure are also reported in tables 6 to 8. Figure 4 shows the volatilities estimated by each procedure and forecasting horizon. Looking at the results for the Pound, we may observe that the volatilities estimated by all procedures present very severe biases. This could be due to the big changes in the absolute Pound returns which occur at the end of the prediction period; see figure 1. The results for the Deustschemark are more encouraging. All the models for volatility sucess in reducing the variance of the standardized observations using the recursive estimates $\hat{\sigma}_t$ in (14). Looking at the results of the regressions, we may observe that the recursive, GARCH and EGARCH estimates of volatility present important biases, but that the ARV estimates only have slight biases.

5. Conclusions

After analysing empirically four exchange rates using three simple models for volatility, GARCH(1,1), EGARCH(1,0) and ARV(1), we can conclude that all three models can imply similar properties of the time series analysed, being hard to decide in this sense which model is in better conformance with the data.

On the other hand, looking at the within-sample estimates of volatility, we may observe that when one of the models has problems in fitting the dynamics of y^2_t, the other models also may have problems (cases of Yen and Swiss-Franc). However, when the models are able to capture the dynamics of y^2_t, cases of the Pound and Deustschemark, the best within-sample fit is provided by the ARV model due to its capability of using future observations to estimate the volatility.

Finally, in relation with the out-sample estimates of volatility, all the models can overperform the recursive estimates. As we could expect, when the dynamics of the squares are different in the post-sample period (case of the Pound), all the models fail to properly forecast future movements in volatility. However, even when the dynamics are the same (Deustschemark), GARCH and EGARCH models may estimate the volatility with severe biases, while the ARV models only have slight biases.

Of course all models analysed can be extended in a great number of ways but I think that the general conclusions will still apply. In any case, it is interesting to analyse what happens in more complicated cases as, for example, when there are "leverage" effects in volatility, as it is the case of some stock returns series, or when considering models with ε_t having heavily tailed distributions as the GED or the Generalized-t distributions.

Summarizing, SV models have the atractive of having two separate disturbances for the mean and the volatility. Allowing the volatility to have its own noise, it is possible to use subsequent observations to improve the volatility estimates. Also, when forecasting future volatilities, ARCH based models seem to put too much weight on the latest observed data, and this may imply severe biases in the forecasted volatilities, which the volatilities forecasted using SV models may not have.

Table 1.- Sample Moments of Centered First Differences of Logged Daily Exchange Rates

	Pound	DM	Yen	Swiss-Franc*
y_t				
Variance	0.3610	0.4128	0.3649	0.5794
Skewness	-0.1350	-0.1099	-0.5199	-0.1748
Kurtosis	3.5310	3.7778	4.9849	8.0567
Q(10)	7.22	8.79	16.64	6.27
y^2_t				
ρ_1	0.0804	0.1537	0.0974	0.5268
ρ_2	0.0473	0.0470	0.0216	0.2800
ρ_3	0.0969	0.0541	0.1802	0.1494
ρ_4	0.0409	0.0926	0.0938	0.0307
ρ_5	0.0792	0.0835	0.0830	0.0463
$Q^2(10)$	40.57	49.53	97.93	324.84
$\log(y^2_t)$				
Mean	-2.3961	-2.2646	-2.6139	-2.0895
Variance	4.9700	5.1552	6.1020	5.3227
ρ_1	0.0244	0.0144	0.0725	0.0493
ρ_2	0.0496	0.0495	0.0723	0.0214
ρ_3	0.0553	0.0542	0.0745	0.1484
ρ_4	0.0515	0.0891	0.1067	0.1142
ρ_5	0.0459	0.0782	0.1060	0.0547
$Q^3(\tau)$				
-1	0.0370	-0.0757	-0.0098	0.1008
1	0.0691	-0.0285	-0.0867	-0.1964

* The Swiss-Franc estimates have being computed using residuals from the model

$$y_t = a_t - 0.0748\, a_{t-1} + 0.1326\, a_{t-2}$$

where a_t is a white noise.

Table 2. GARCH(1,1) Models

a) Estimates and implied moments

	Pound	DM	Yen	Swiss Franc
ω	0.0147	0.0234	0.0032	0.0926
α	0.0828	0.0854	0.0273	0.1194
β	0.8811	0.8588	0.9637	0.7058
y_t				
σ^2	0.4072	0.4194	0.3556	0.5297
Kurtosis	3.7193	3.4660	3.2722	3.2944
y^2_t				
ρ_1	0.1605	0.1395	0.0658	0.1496
ρ_2	0.1547	0.1317	0.0652	0.1234
ρ_3	0.1491	0.1244	0.0646	0.1019
ρ_4	0.1437	0.1174	0.0640	0.0841
ρ_5	0.1385	0.1109	0.0634	0.0693

b) Standardized residuals: $e_t = y_t / \hat{s}_{t/t-1}$

Variance	0.9811	1.0029	0.9764	1.0097
Skewness	-0.2307	-0.0308	-0.5789	0.3872
Kurtosis	3.5458	3.8144	5.4365	6.5286
$Q^2(10)$	3.90	12.35	6.59	6.05

Table 3. EGARCH(1,0) Models

a) Estimates and implied moments

	Pound	DM	Yen	Swiss-Franc
ω	-0.0495	-0.0795	-0.0100	-0.1184
γ	-0.0090	-0.0131	-0.0044	-0.0259
α	0.1568	0.1865	0.0500	0.2953
β	0.9530	0.9155	0.9905	0.8113
y_t				
Variance	0.3668	0.4065	0.3577	0.5611
Kurtosis	3.3320	3.2690	3.1542	3.3452
y^2_t				
ρ_1	0.1088	0.1133	0.0435	0.1650
ρ_2	0.1030	0.1025	0.0431	0.1288
ρ_3	0.0976	0.0928	0.0427	0.1014
ρ_4	0.0924	0.0842	0.0423	0.0802
ρ_5	0.0876	0.0764	0.0418	0.0639
$\log(y^2_t)$				
Mean	-2.3232	-2.2108	-2.3226	-1.8975
Variance	5.0330	5.0139	4.9839	5.0295
ρ_1	0.0530	0.0556	0.0208	0.0802
ρ_2	0.0506	0.0509	0.0206	0.0651
ρ_3	0.0482	0.0466	0.0205	0.0528
ρ_4	0.0459	0.0426	0.0203	0.0428
ρ_5	0.0438	0.0390	0.0201	0.0347

b) Standardized residuals: $e_t = y_t / \hat{s}_{t/t-1}$

Variance	1.0027	1.0074	0.9895	1.0071
Skewness	-0.2428	-0.0058	-0.6274	0.4318
Kurtosis	3.4680	3.7880	5.7020	6.8269
$Q^2(10)$	3.51	13.89	12.68	8.23

Table 4. Stochastic Volatility Models

a) Estimates and implied moments

	Pound	DM	Yen	Swiss Franc
γ	-0.0442	-0.0344	-0.0057	-0.0321
ϕ	0.9518 (0.0362)	0.9619 (0.0260)	0.9971 (0.0033)	0.9562 (0.0253)
σ^2_η	0.0254 (0.0254)	0.0249 (0.0214)	0.0032 (0.0025)	0.0411 (0.0289)
σ^2_ξ	4.6972 (0.4186)	4.8316 (0.4217)	5.5951 (0.4404)	4.8491 (0.4246)
c	2.408	2.155	1.404	2.127
y_t				
σ^2	0.4575	0.4789	0.1847	0.6108
Kurtosis	3.9299	4.1859	5.2129	4.8467
y^2_t				
ρ_1	0.1000	0.1186	0.1744	0.1513
ρ_2	0.0946	0.1133	0.1738	0.1431
ρ_3	0.0895	0.1083	0.1731	0.1354
ρ_4	0.0847	0.1036	0.1725	0.1282
ρ_5	0.0801	0.0991	0.1718	0.1215
$\log(y^2_t)$				
Mean	-2.1870	-2.1729	-3.2355	-2.0029
Variance	5.2048	5.2679	5.4873	5.4145
ρ_1	0.0494	0.0608	0.1004	0.0847
ρ_2	0.0470	0.0585	0.1001	0.0810
ρ_3	0.0448	0.0562	0.0998	0.0775
ρ_4	0.0426	0.0541	0.0995	0.0741
ρ_5	0.0405	0.0520	0.0992	0.0708

b) Standardized residuals: $e_t = y_t / \hat{s}_{t/t-1}$

Variance	0.8749	0.9876	2.1985	1.1622
Skewness	-0.2107	-0.0873	-0.8543	0.2938
Kurtosis	3.4290	4.2689	8.0241	9.4166
$Q^2(10)$	7.84	46.44	26.93	257.69

c) Standardized residuals: $e_t = y_t / \hat{s}_{t/T}$

Variance	0.8121	0.9052	2.1449	0.9814
Skewness	-0.2415	-0.0085	-0.6148	0.0639
Kurtosis	3.4114	3.5533	5.7716	4.6391
$Q^2(10)$	5.97	16.86	18.92	128.79

Table 5. Regression of y^2_t on $\hat{s}^2_{t/t-1}$

	Pound			Deustschemark			Yen			Swiss Franc		
	G	E	S	G	E	S	G	E	S	G	E	S
α	0.104 (0.02)	0.058 (0.02)	-0.17 (0.02)	0.089 (0.03)	-0.02 (0.04)	-0.13 (0.03)	-0.04 (0.04)	0.007 (0.03)	0.023 (0.02)	-0.30 (0.67)	-0.59 (0.85)	-0.39 (0.68)
β	0.682 (0.04)	0.833 (0.06)	1.285 (0.05)	0.776 (0.09)	1.062 (0.11)	1.249 (0.08)	1.099 (0.11)	0.971 (0.10)	1.885 (0.13)	1.602 (1.33)	2.147 (1.68)	1.806 (1.36)
R^2	0.026	0.035	0.103	0.029	0.048	0.097	0.063	0.051	0.078	0.158	0.164	0.105
Q_{10}	5.77	5.97	6.89	14.40	14.41	14.01	24.41	30.45	26.90	151.6	100.8	251.92

G: GARCH(1,1)
E: EGARCH(1,1)
S: ARV(1). In this case the regression is of y^2_t on $\hat{s}^2_{t/T}$.

Table 6. Moments of standardized observations one-step ahead.

a) Pound

	$\hat{\sigma}_t$	GARCH	EGARCH	ARV
Variance	3.9741	1.6453	2.0752	2.5195
Skewness	0.8272	0.3469	0.5192	0.6971
Kurtosis	5.1123	4.4519	4.3513	4.7691
Regressions				
α	4.7887	1.4948	1.8666	2.6337
β	-6.9175	0.1521	-0.2098	-1.3527
R^2	0.0100	0.0009	0.0003	0.0028

b) Deustschemark

	$\hat{\sigma}_t$	GARCH	EGARCH	ARV
Variance	2.3333	1.4879	1.6236	1.5469
Skewness	-0.4990	-0.3881	-0.4134	-0.5400
Kurtosis	3.9302	3.7837	3.6825	4.1865
Regressions				
α	4.5478	0.9989	1.1707	0.5757
β	-7.7870	0.0388	-0.2112	0.6523
R^2	0.0113	0.0001	0.0005	0.0066

Table 7. Moments of Standardized Observations 5 steps ahead

a) Pound

	$\hat{\sigma}_t$	GARCH	EGARCH	ARV
Variance	4.2012	2.0632	2.5594	2.9117
Skewness	0.8183	0.8977	0.7634	0.7104
Kurtosis	4.9710	6.3580	5.1923	4.7498
Regressions				
α	6.4075	2.1448	3.3427	2.8871
β	-10.4699	-0.3493	-2.0974	-1.8765
R^2	0.0219	0.0034	0.0131	0.0030

b) Deustschemark

	$\hat{\sigma}_t$	GARCH	EGARCH	ARV
Variance	2.4297	1.7808	1.9228	1.8241
Skewness	-0.4513	-0.7505	-0.5095	-0.5920
Kurtosis	3.8033	6.0782	4.4221	4.8819
Regressions				
α	5.9229	1.1999	1.4850	0.5761
β	-10.7766	-0.1913	-0.7175	0.7379
R^2	0.0207	0.0007	0.0019	0.0053

Table 8. Moments of Standardized Observations 20 steps ahead

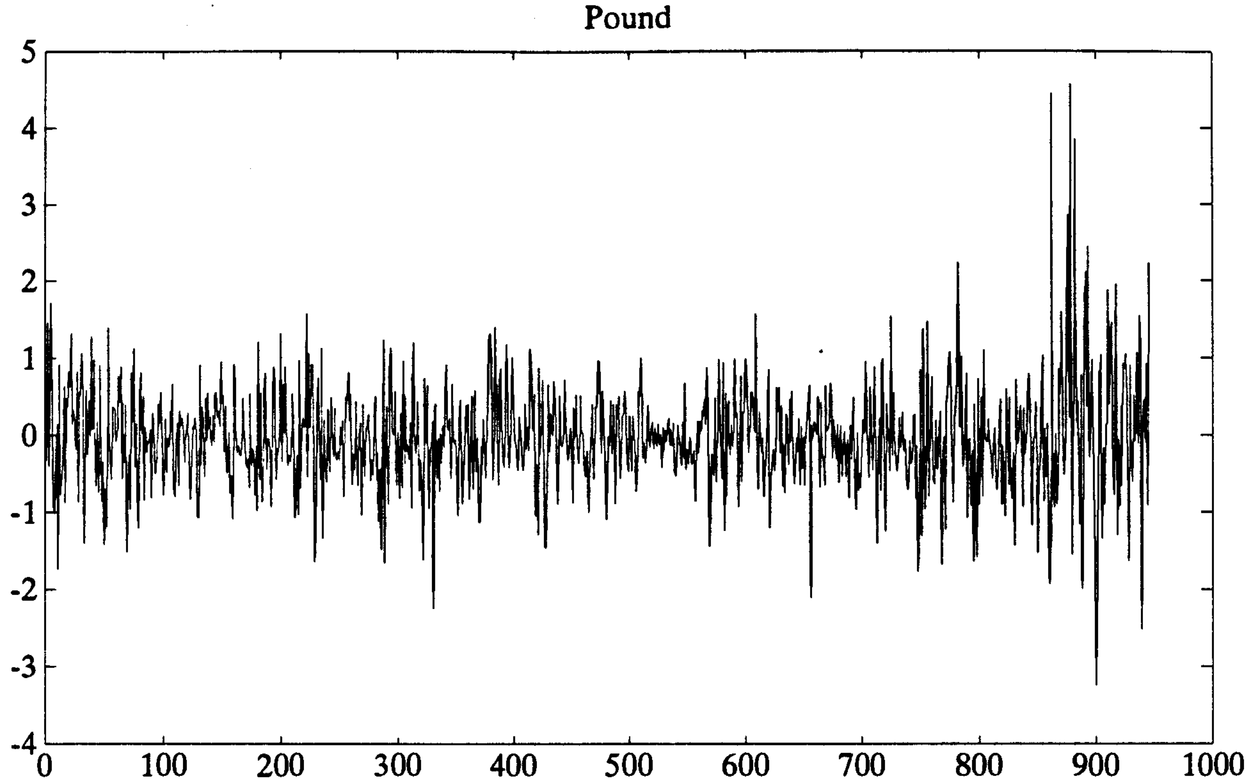
a) Pound

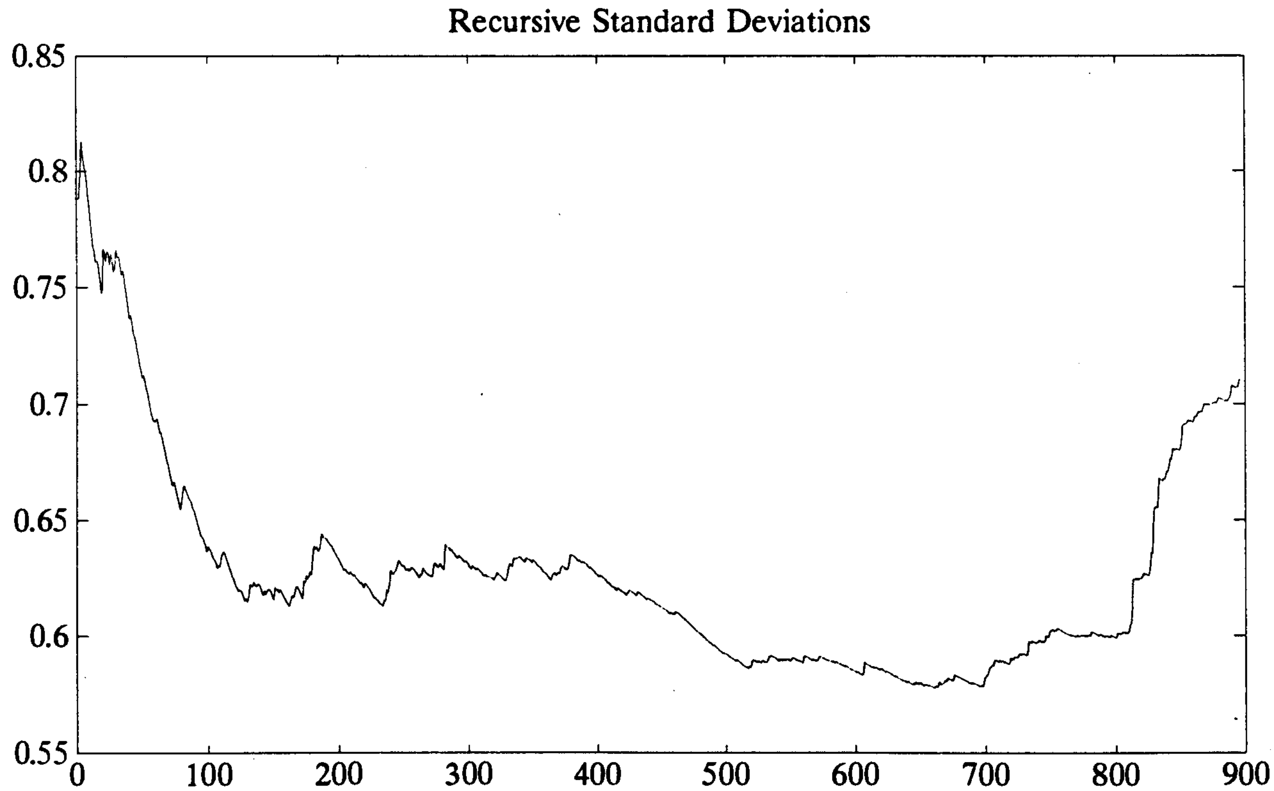
	$\hat{\sigma}_t$	GARCH	EGARCH	ARV
Variance	4.0493	2.2905	3.1938	3.4142
Skewness	0.5344	1.0063	0.5360	0.4968
Kurtosis	4.6971	6.2869	4.6292	4.5900
Regressions				
α	6.9744	1.5444	2.7235	3.2056
β	-12.1842	0.1177	-2.0174	-2.8416
R^2	0.0333	0.0002	0.0014	0.0010

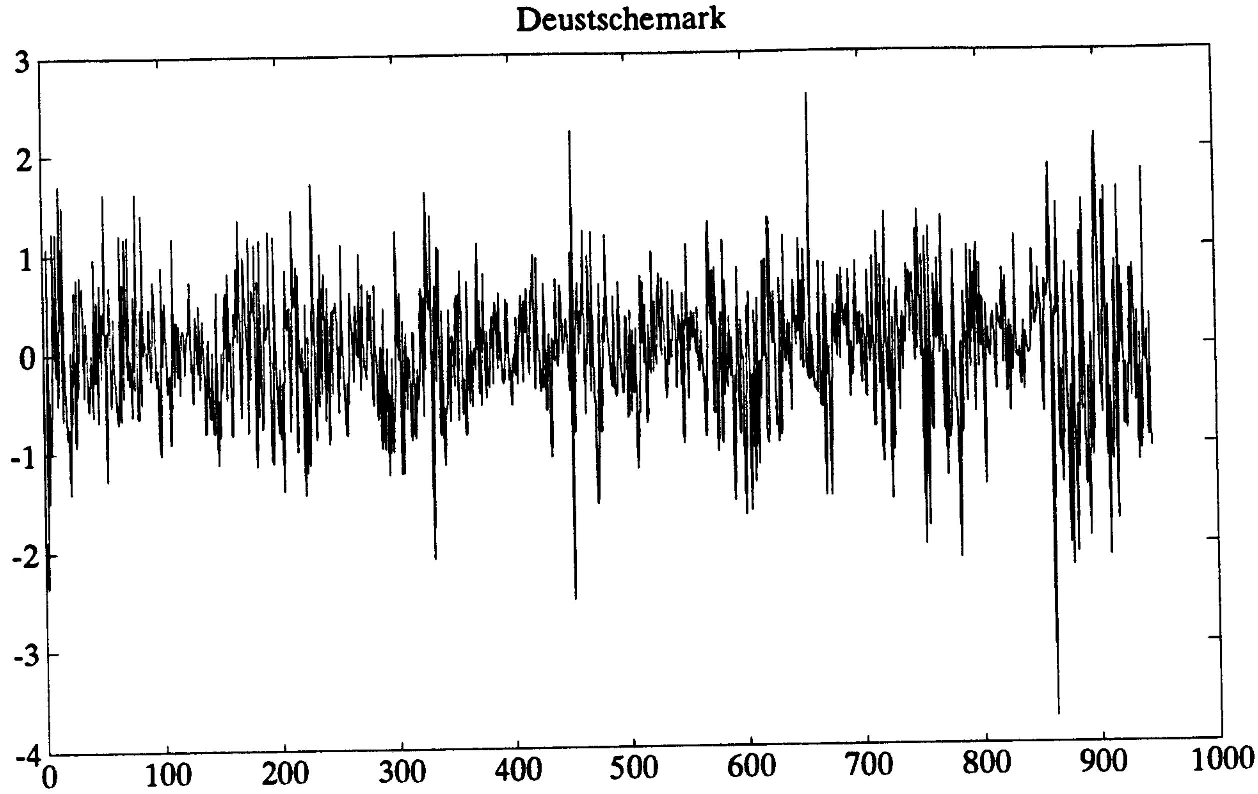
b) Deustschemark

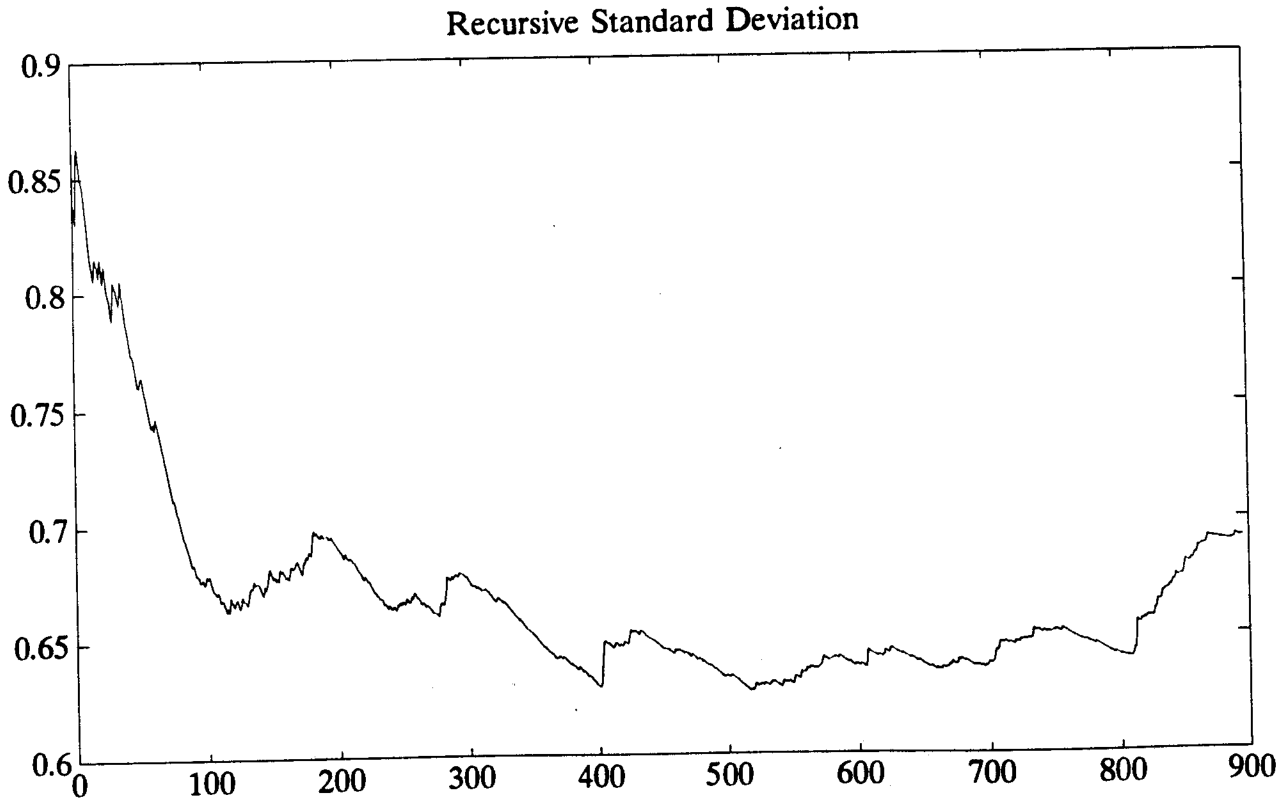
	$\hat{\sigma}_t$	GARCH	EGARCH	ARV
Variance	2.2552	1.8301	2.1789	1.9507
Skewness	0.0735	0.0134	0.0814	0.0293
Kurtosis	2.5629	2.6350	2.5406	2.5687
Regressions				
α	5.3008	0.6166	-1.0009	0.4127
β	-9.6602	0.6571	4.3858	1.3817
R^2	0.0311	0.0029	0.0044	0.0044

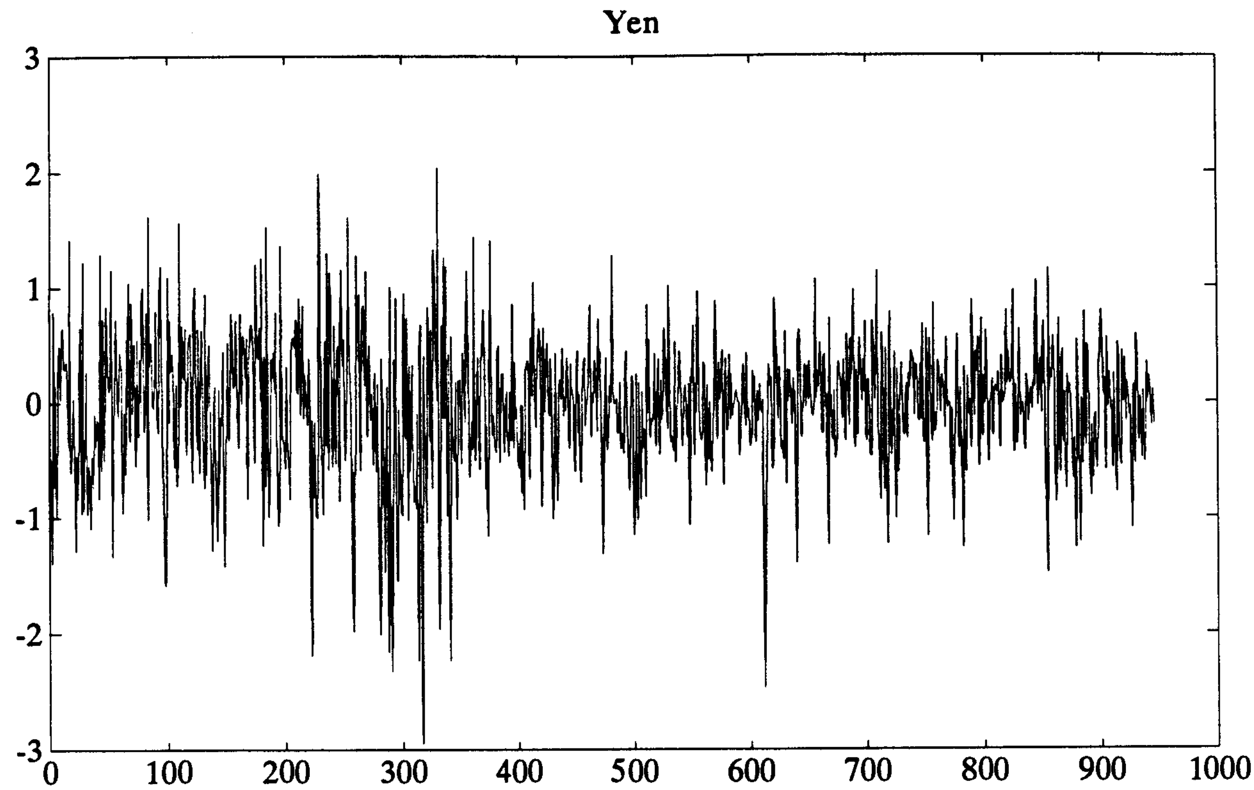
FIGURE 1. FIRST DIFFERENCES OF LOGGED EXCHANGE RATES AND RECURSIVE STANDARD DEVIATIONS

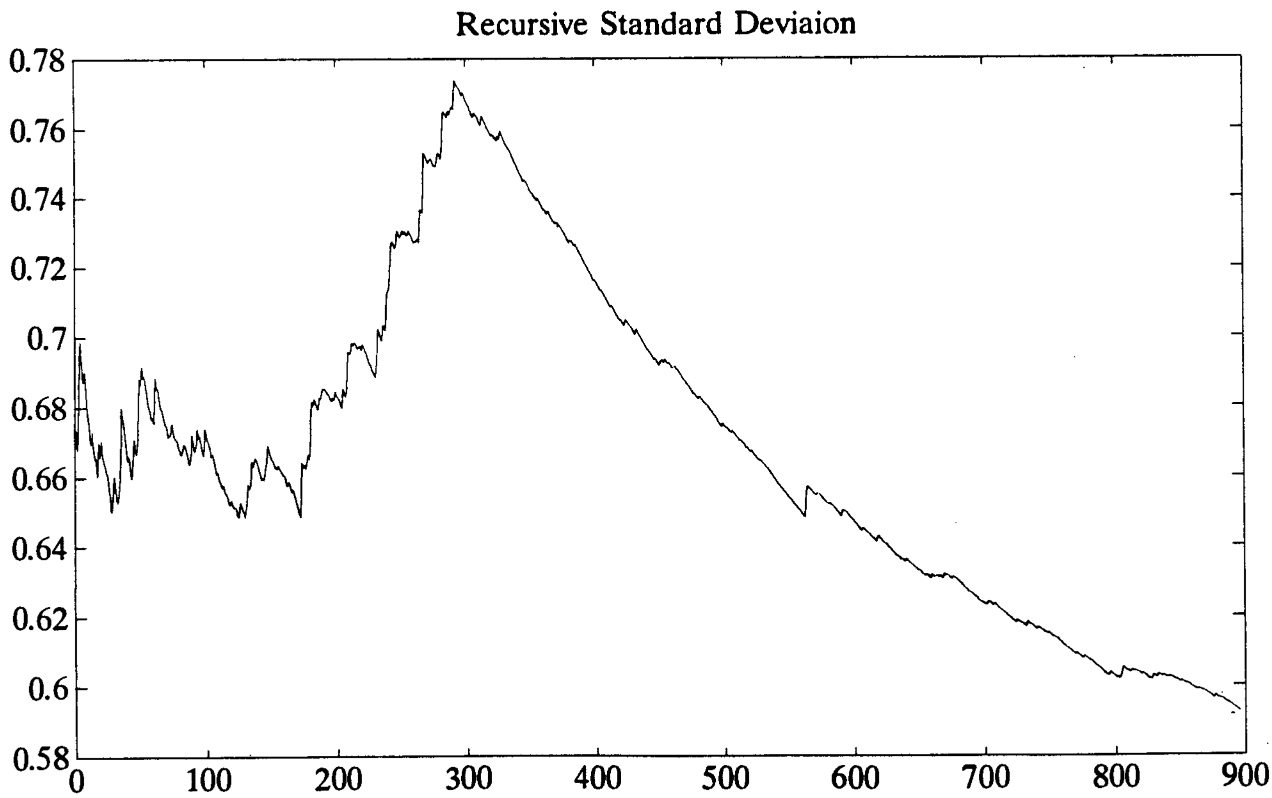


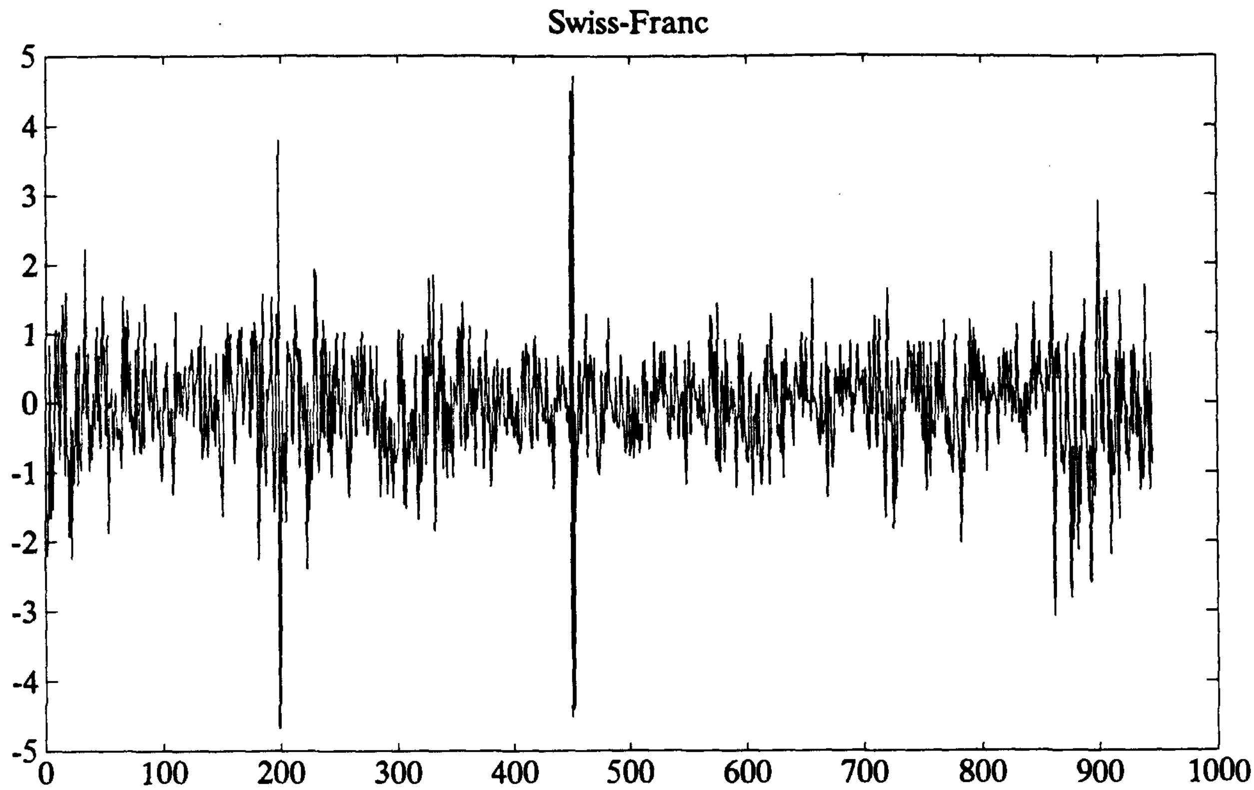



Deustschemark
3
2
1
0
-1
-2
-3
-4
0
100
200
300
400
500
600
700
800
900
1000


Recursive Standard Deviation
0.9
0.85
0.8
0.75
0.7
0.65
0.6
0
100
200
300
400
500
600
700
800
900


Yen
3
2
1
0
-1
-2
-3
0
100
200
300
400
500
600
700
800
900
1000


Recursive Standard Deviaion
0.78
0.76
0.74
0.72
0.7
0.68
0.66
0.64
0.62
0.6
0.58
0
100
200
300
400
500
600
700
800
900


Swiss-Franc
5
4
3
2
1
0
-1
-2
-3
-4
-5
0
100
200
300
400
500
600
700
800
900
1000

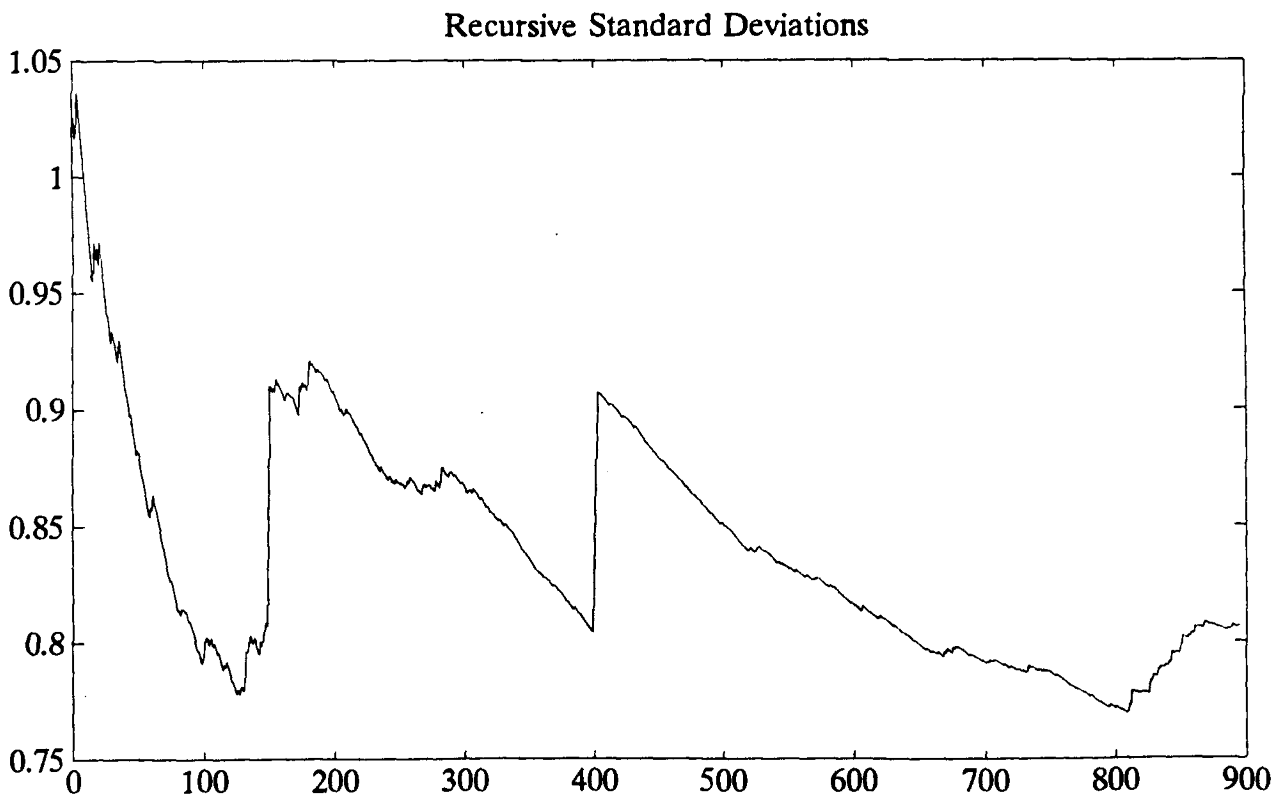
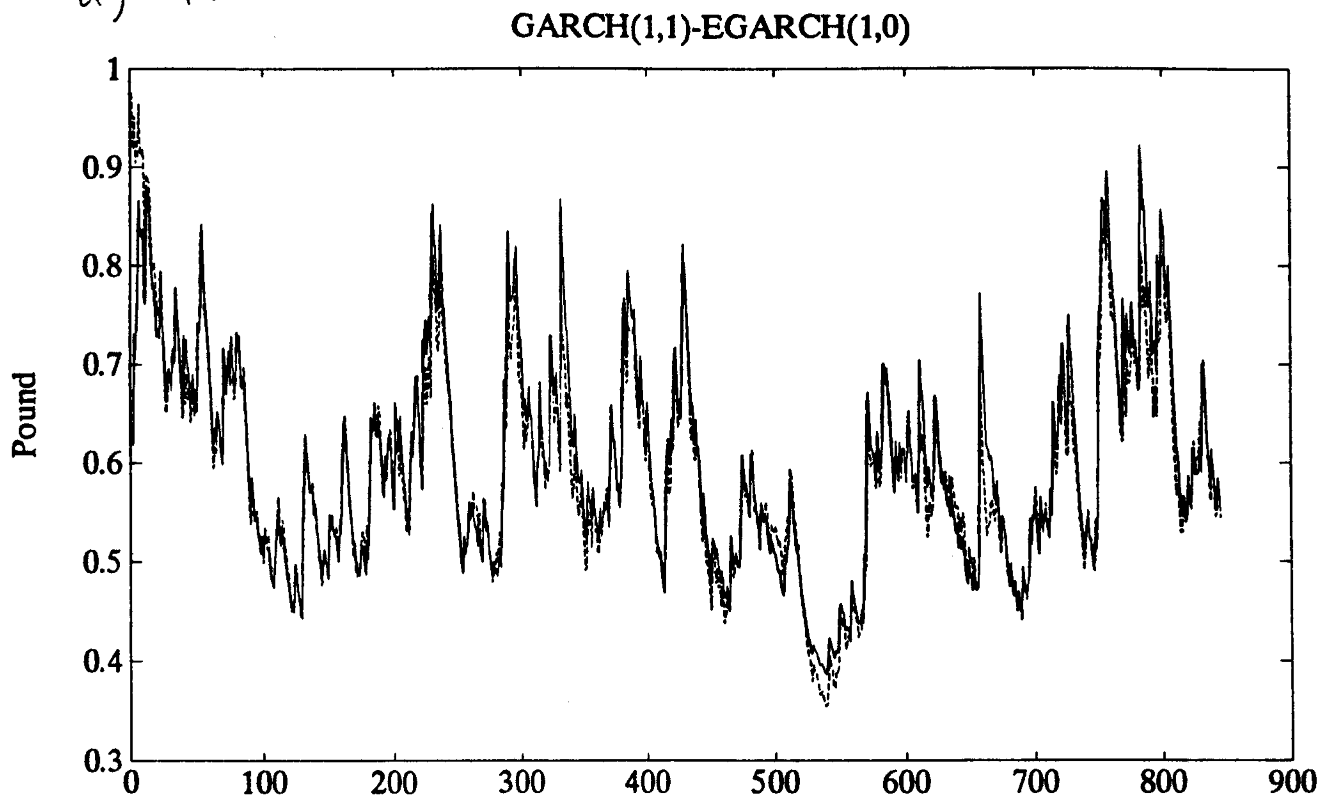
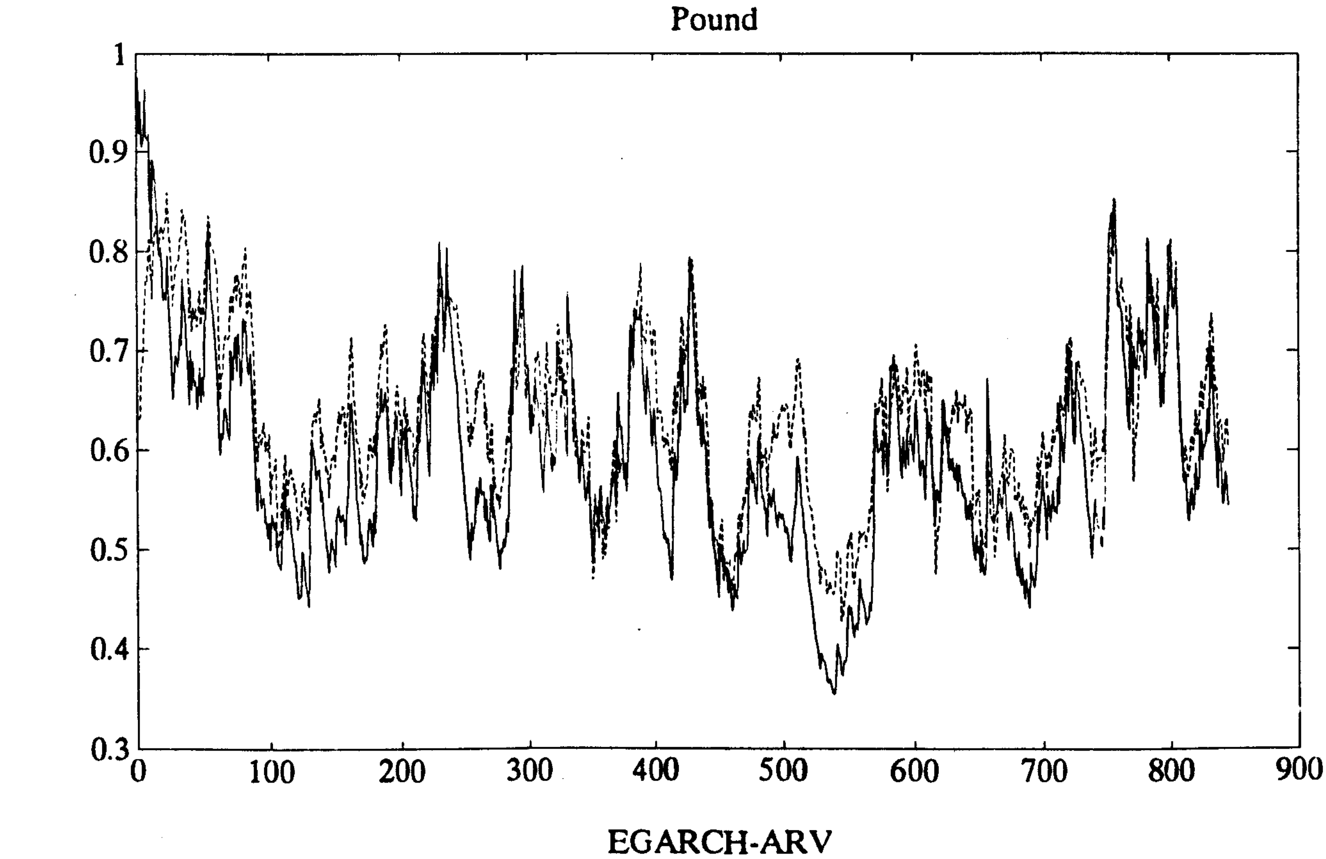

Recursive Standard Deviations
1.05
1
0.95
0.9
0.85
0.8
0.75
0
100
200
300
400
500
600
700
800
900

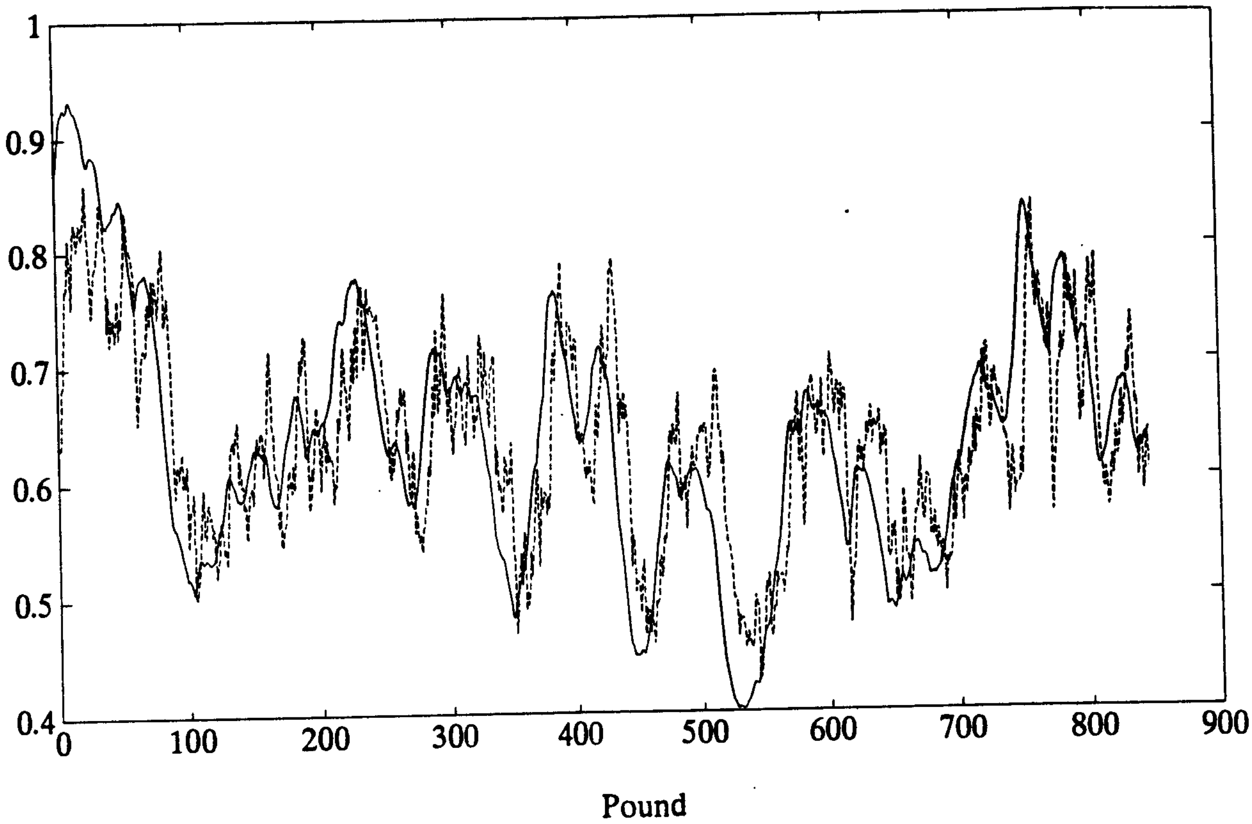
FIGURE 2. ESTIMATES OF VOLATILITY

a) POUND



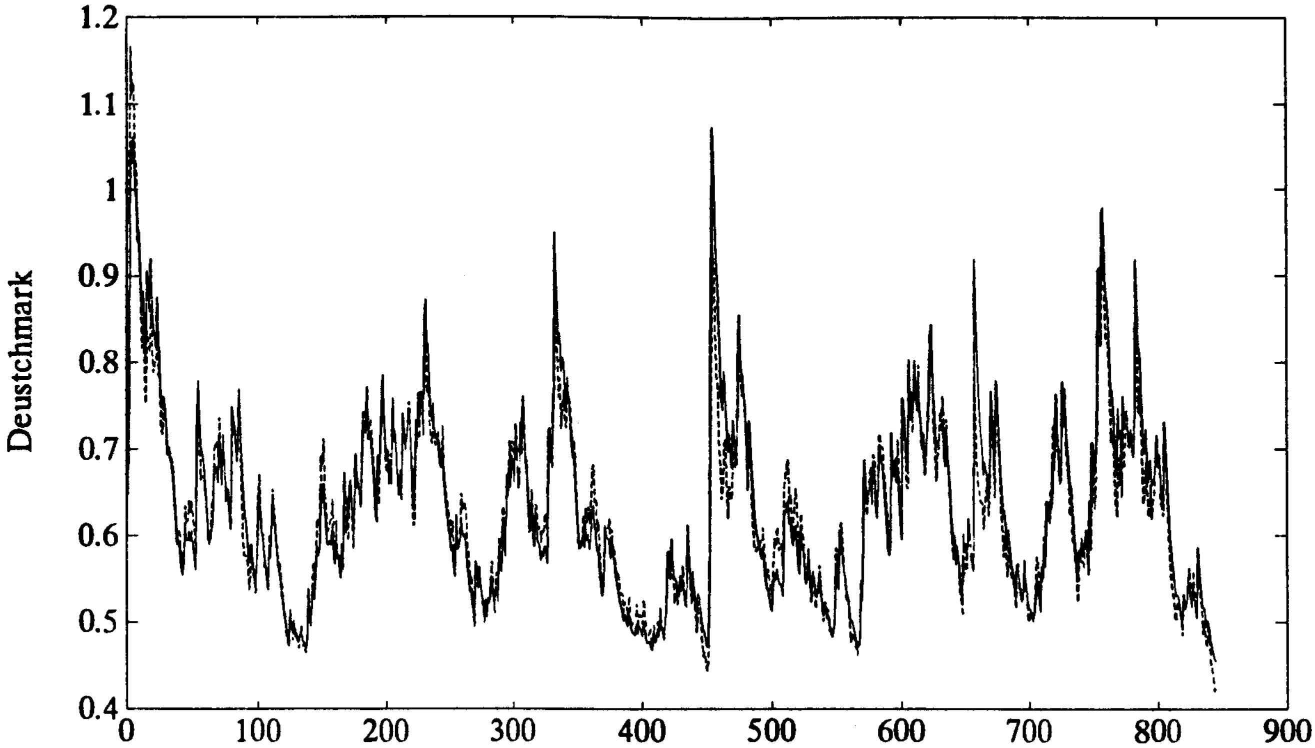


ARV One-step-ahead and Smoothed Volatility Estimates

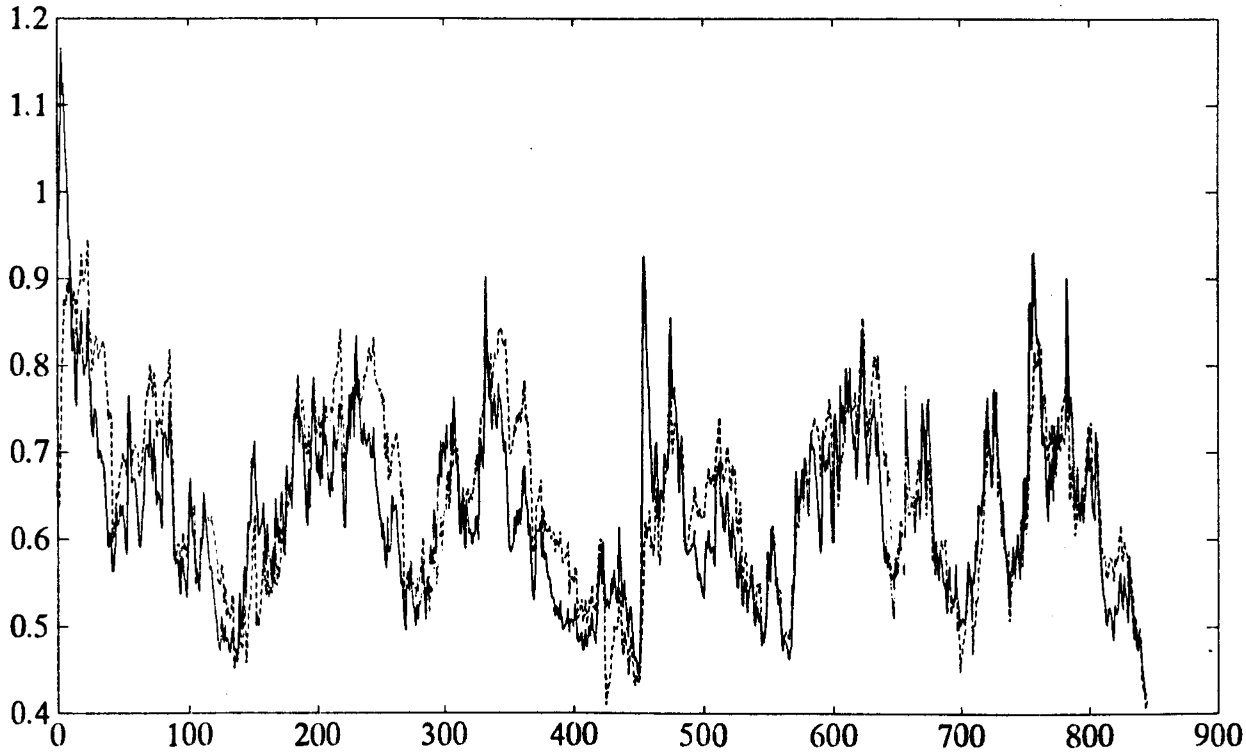


b) DEUSTSCHEMARK

GARCH(1,1)-EGARCH(1,0)

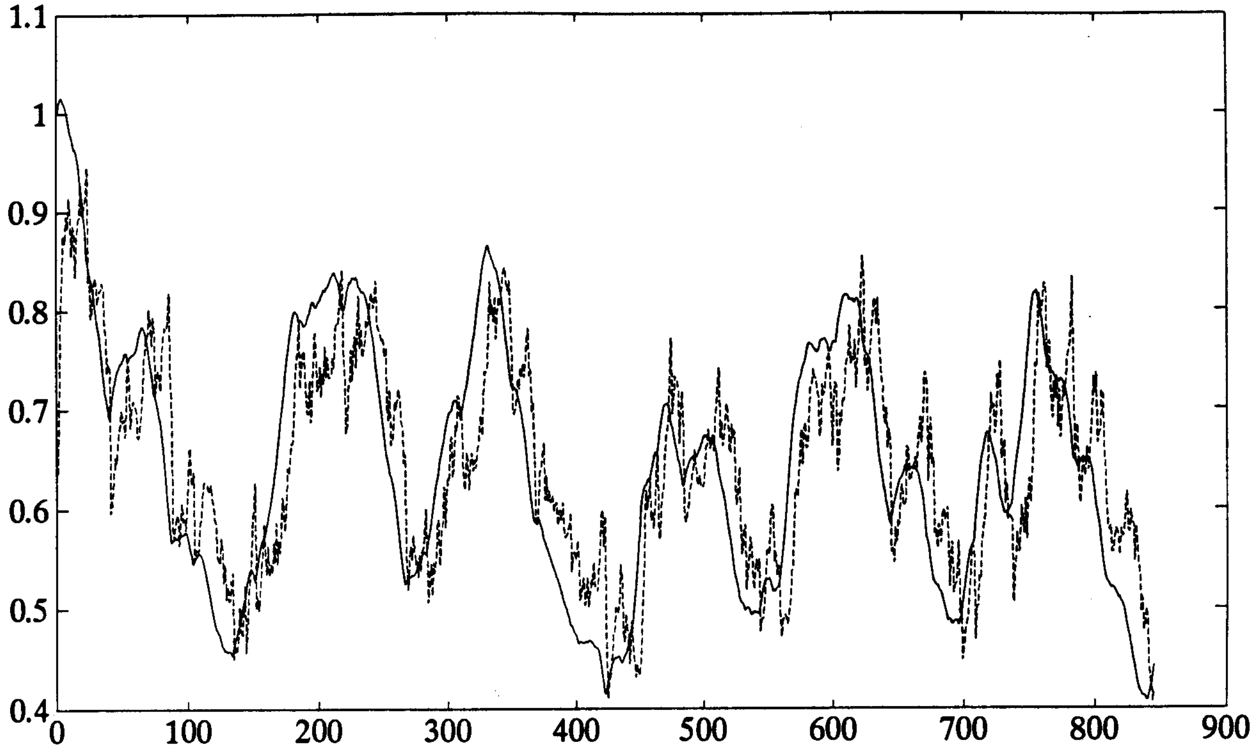


Deustchmark



EGARCH-ARV

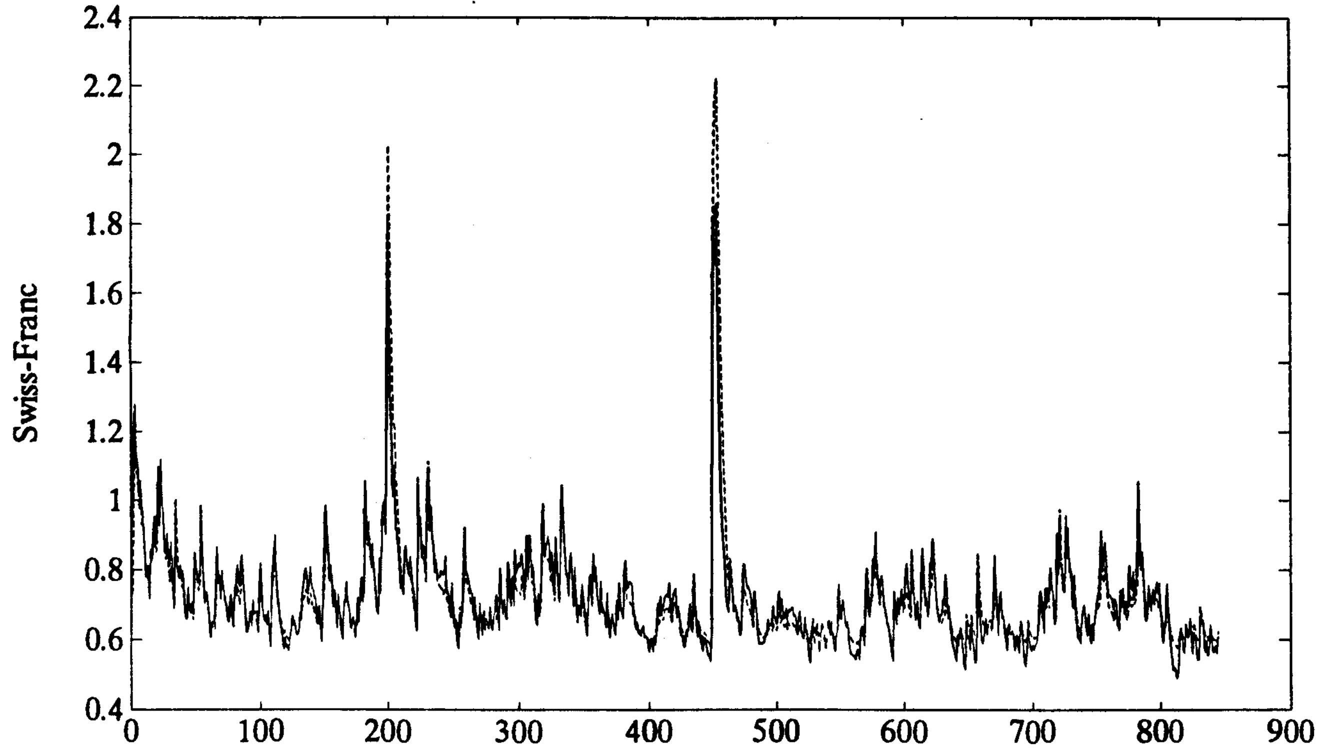
ARV One-step-ahead and Smoothed Volatility Estimates



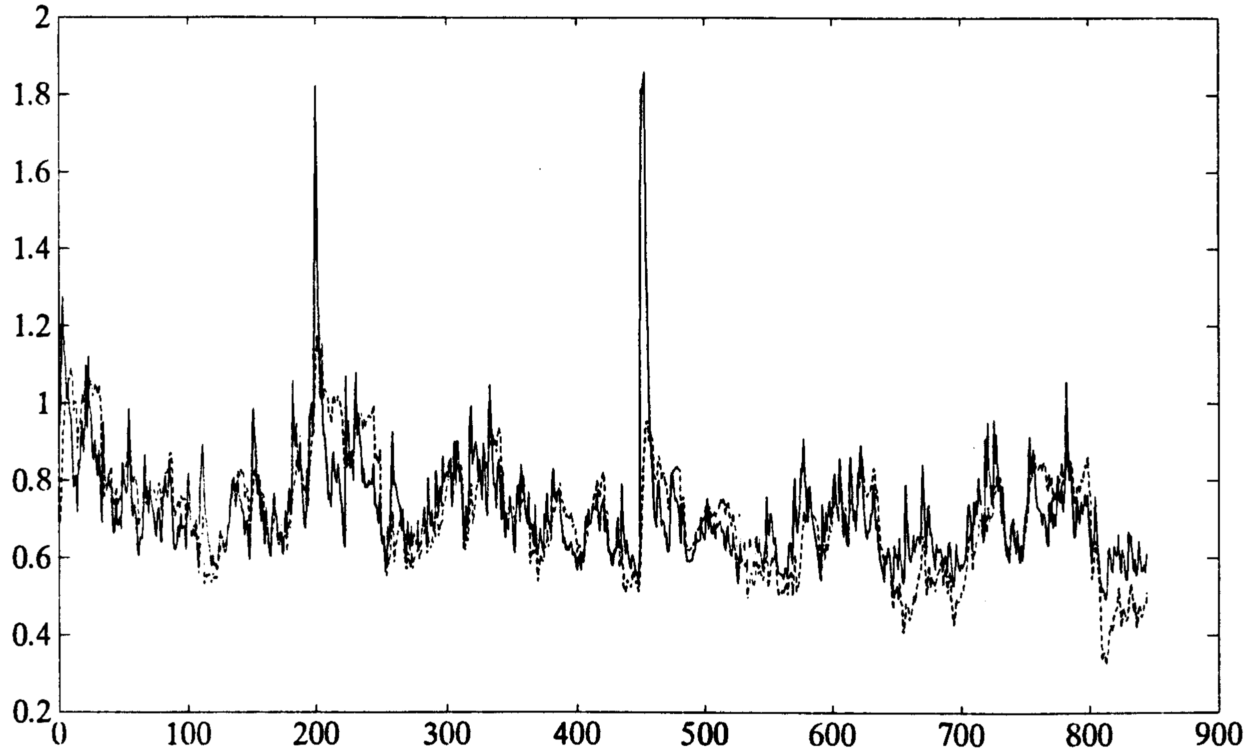
Deustschemark

c) Swiss - Franc

GARCH(1,1)-EGARCH(1,0)

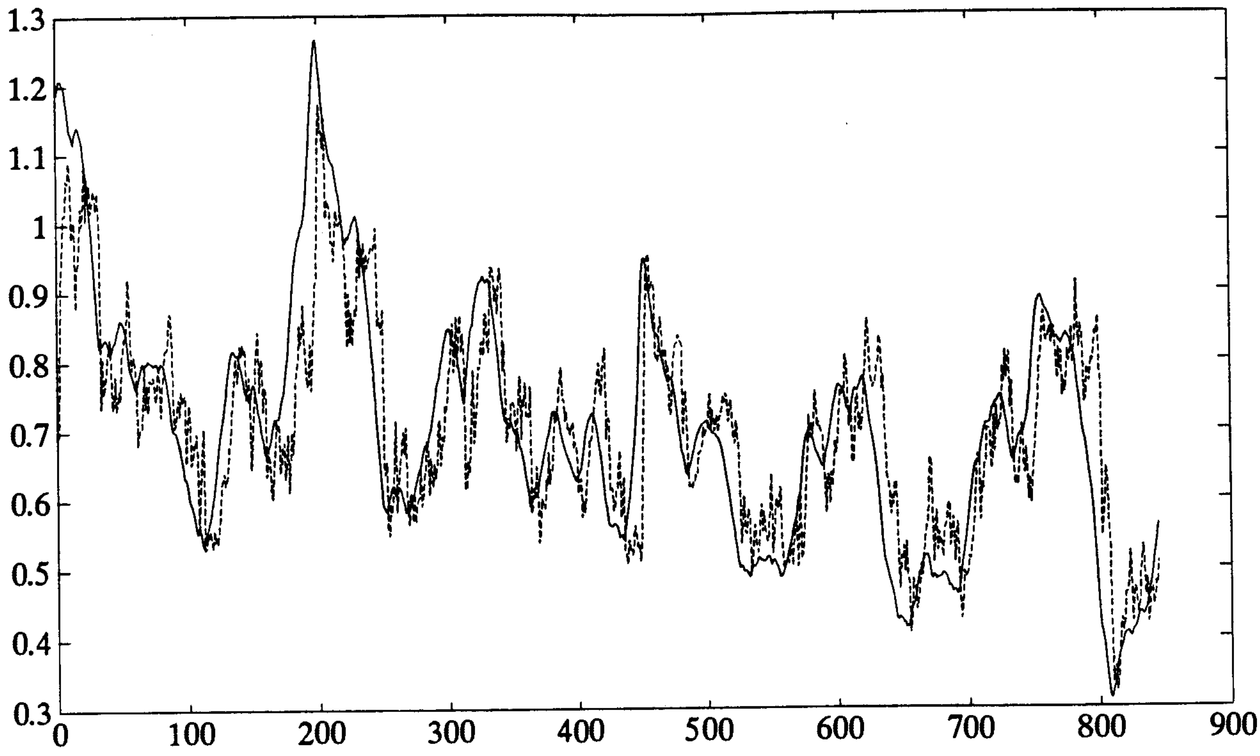


Swiss-Franc



EGARCH-ARV

ARV One-step-ahead and Smoothed Volatility Estimates

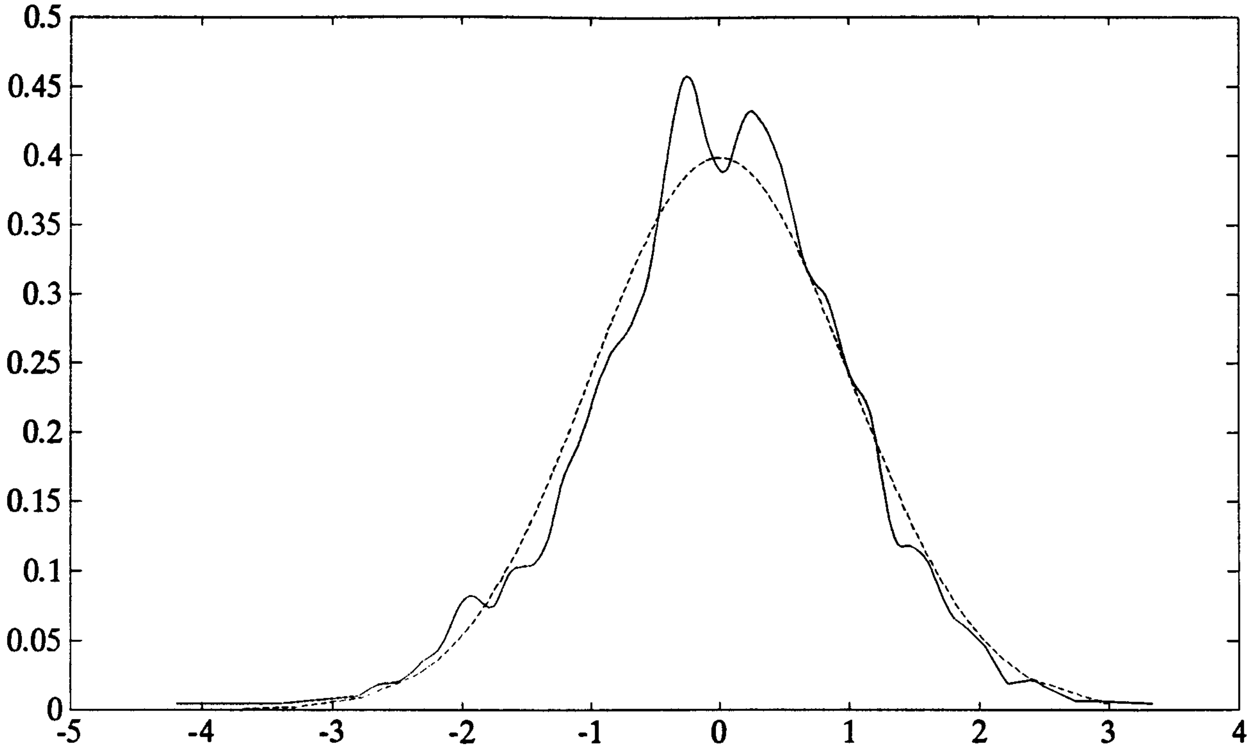


Swiss-Franc

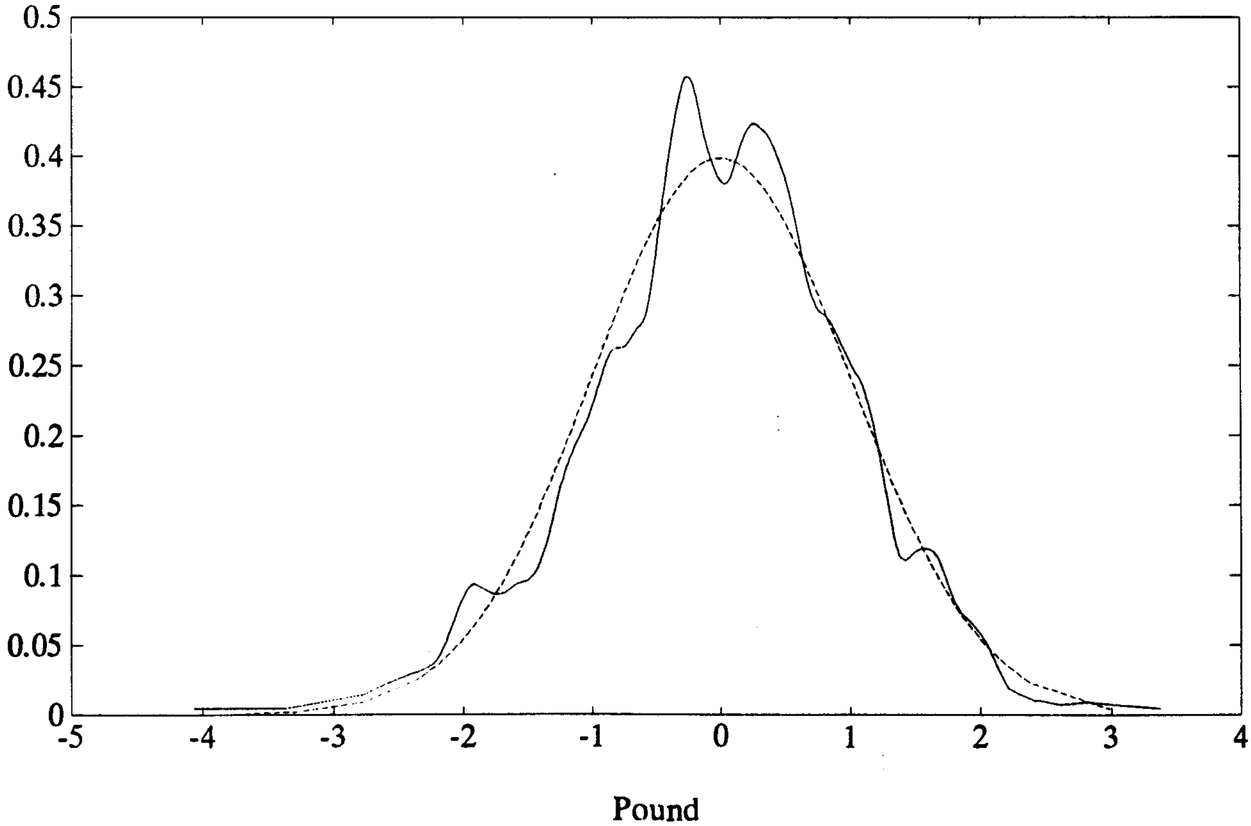
FIGURE 3. ESTIMATED DENSITIES OF STANDARDIZED OBSERVATIONS AND NORMAL DENSITY.

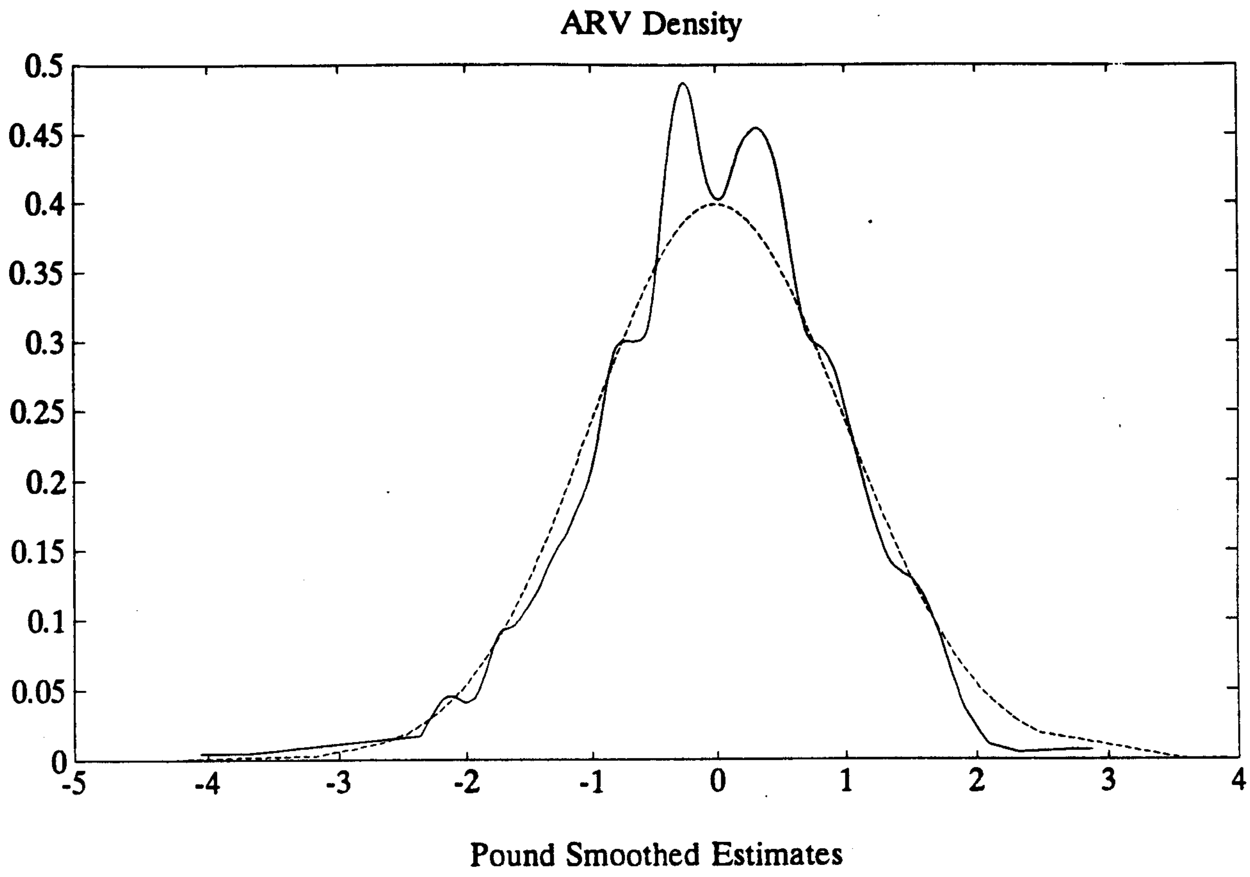
a) POUND

GARCH Density



EGARCH Density

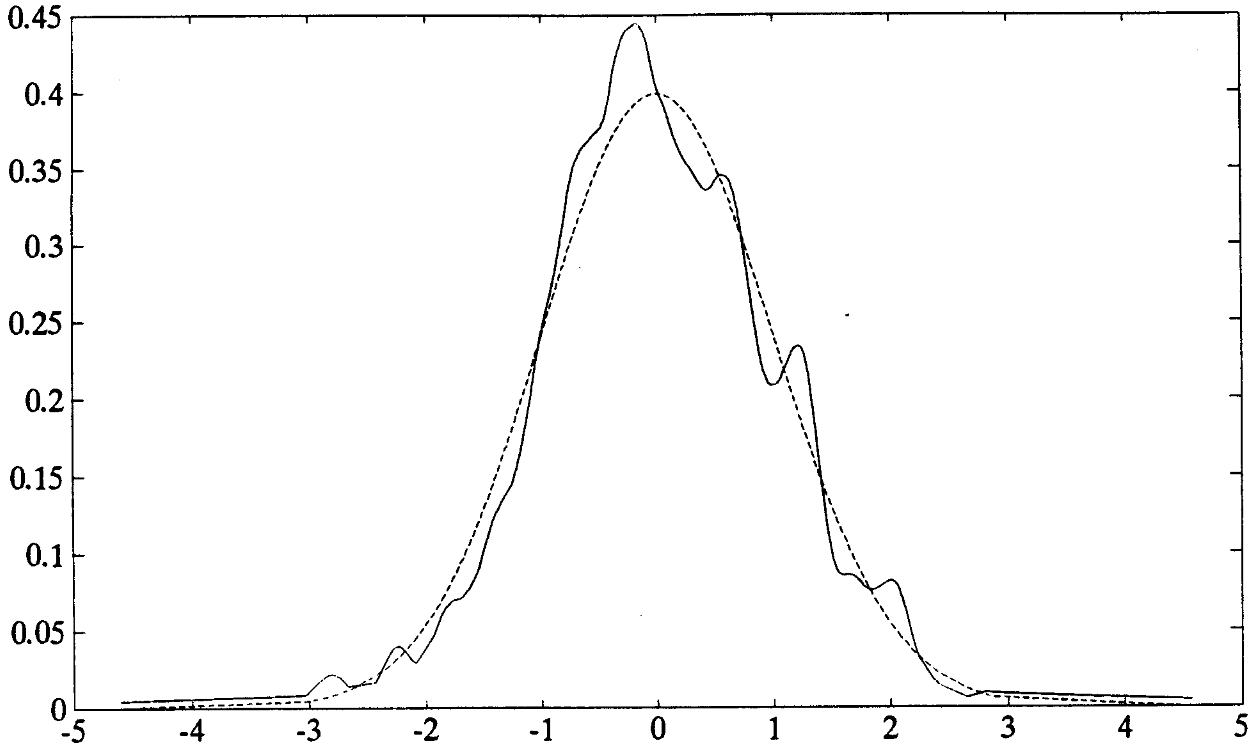




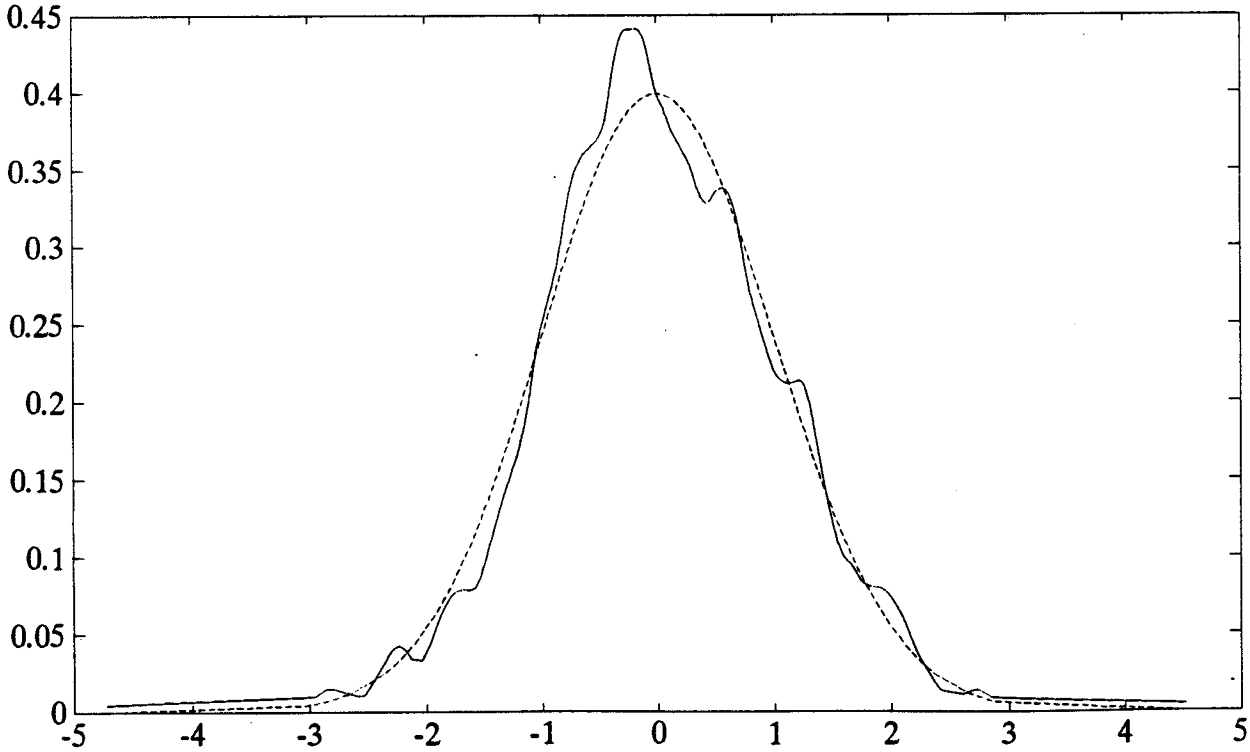
ARV Density
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-5
-4
-3
-2
-1
0
1
2
3
4
Pound Smoothed Estimates

5) DEUSTSCHEMARK

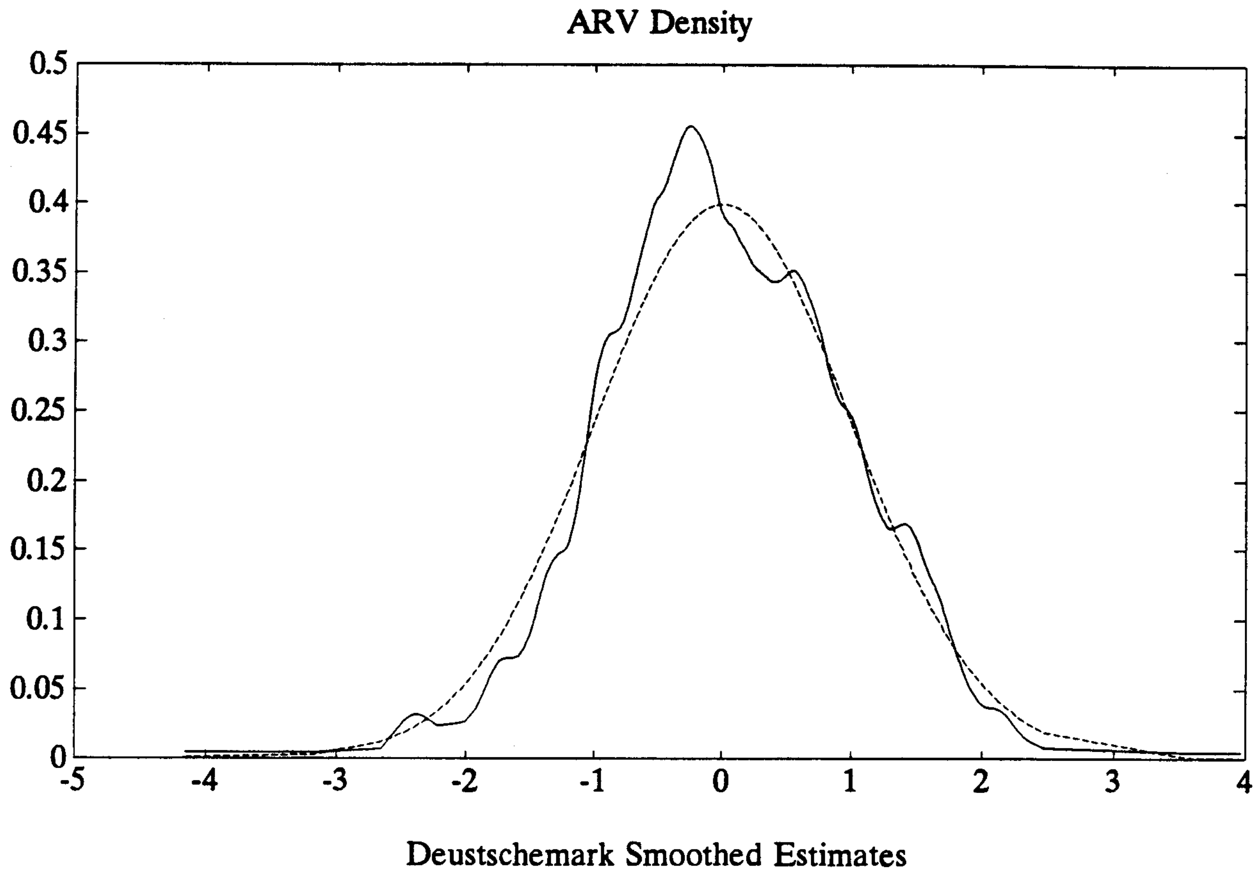
GARCH Density



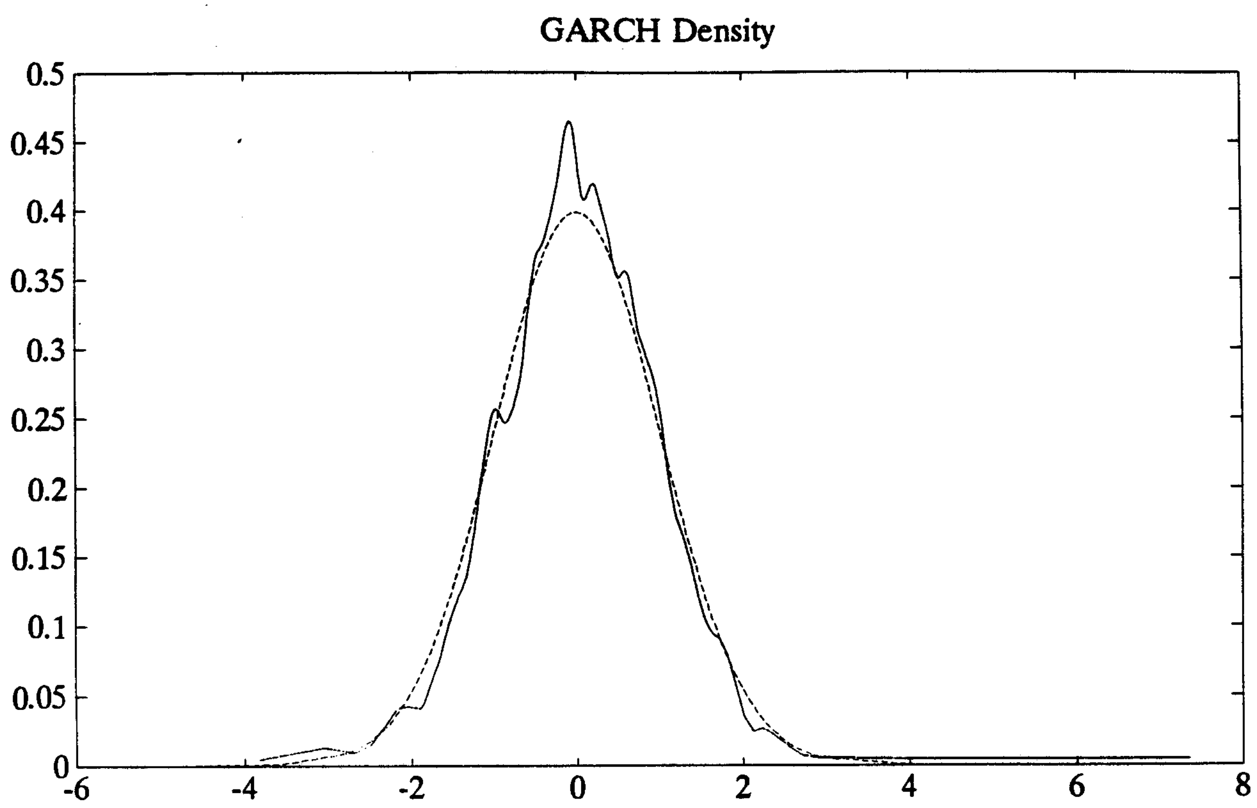
EGARCH Density

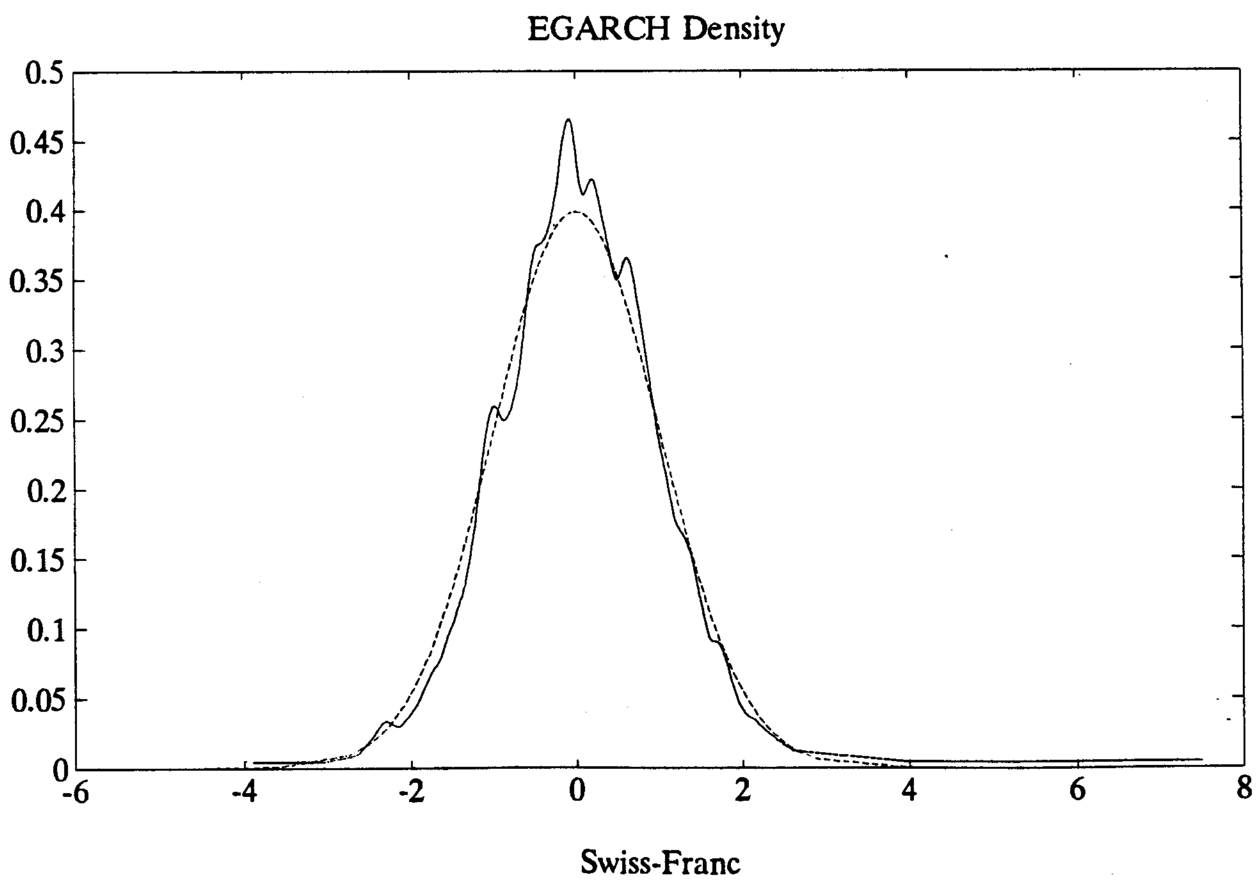


Deustchemark


ARV Density
0.5
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-5
-4
-3
-2
-1
0
1
2
3
4
Deustschemark Smoothed Estimates

c) SWISS - FRANC





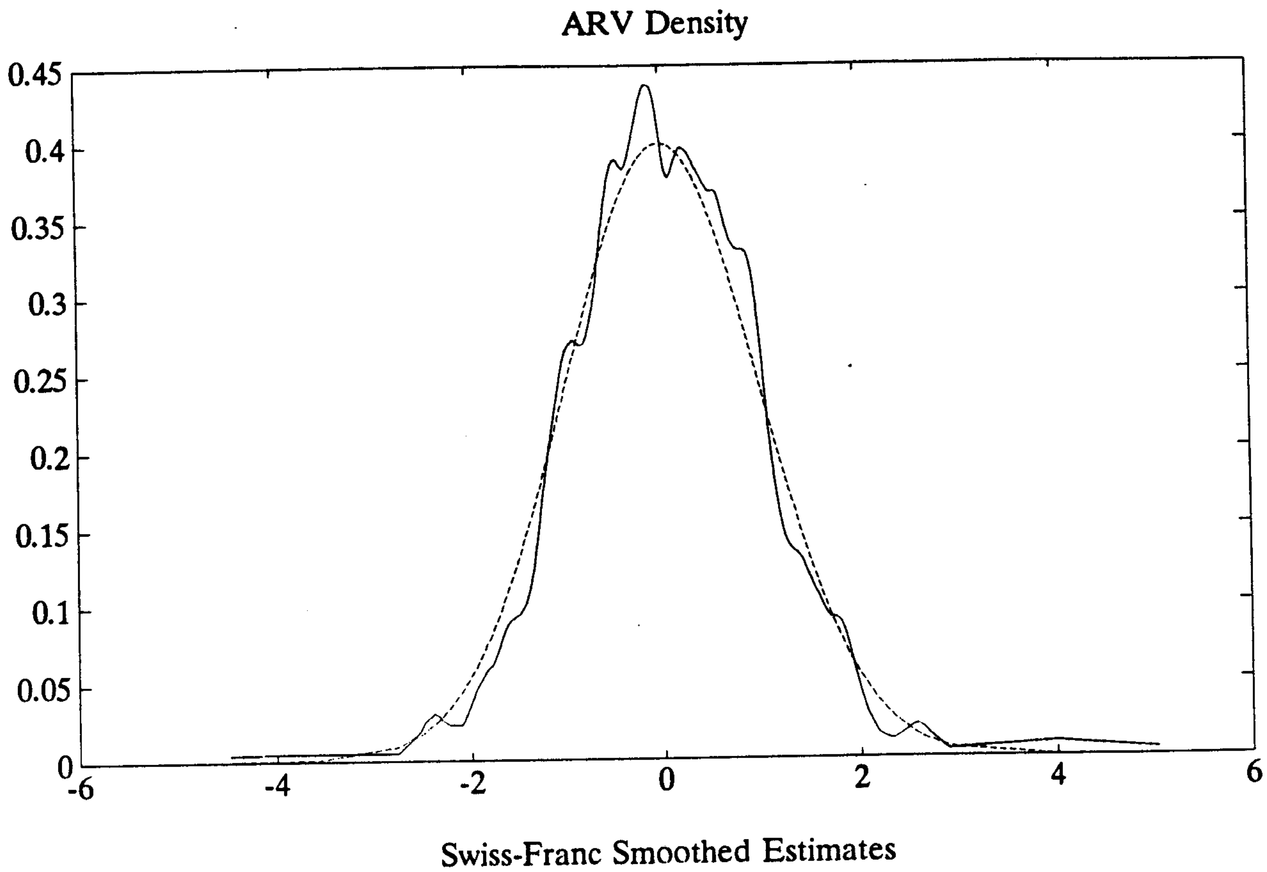
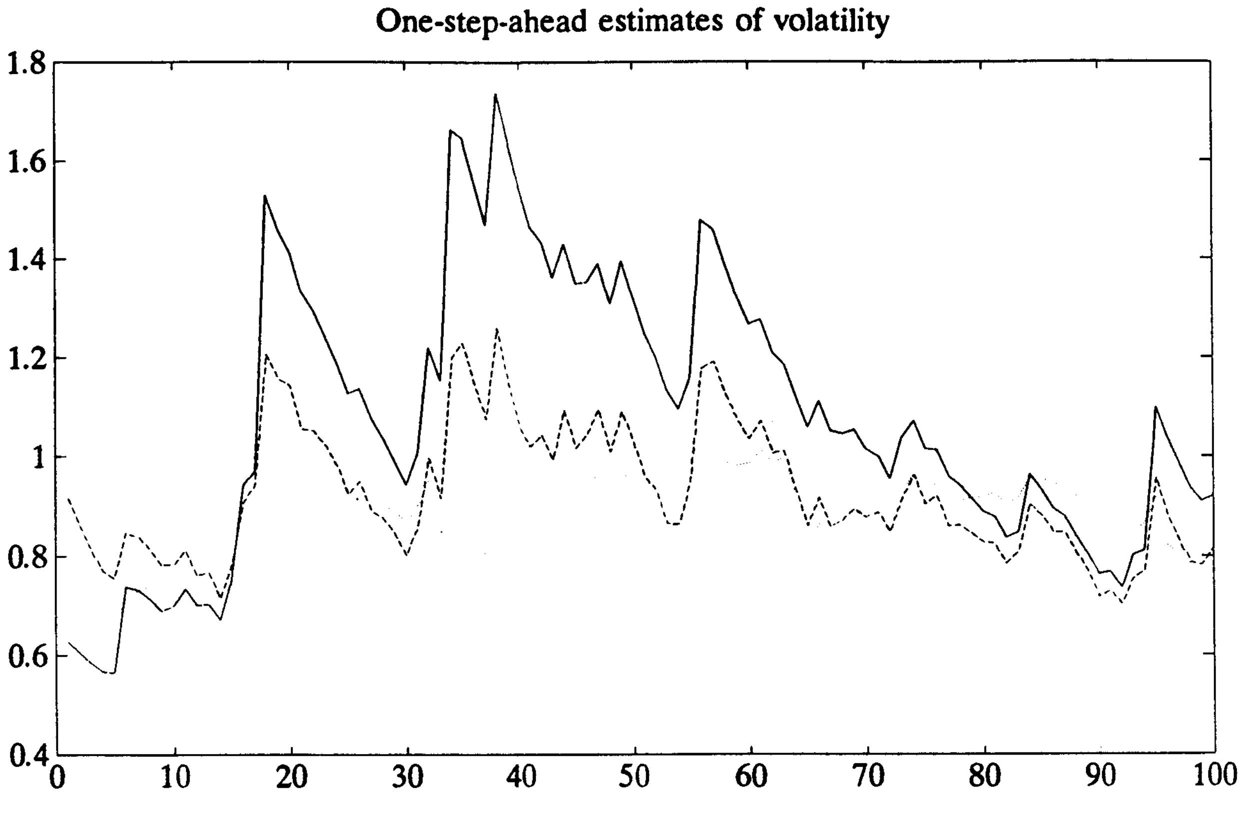
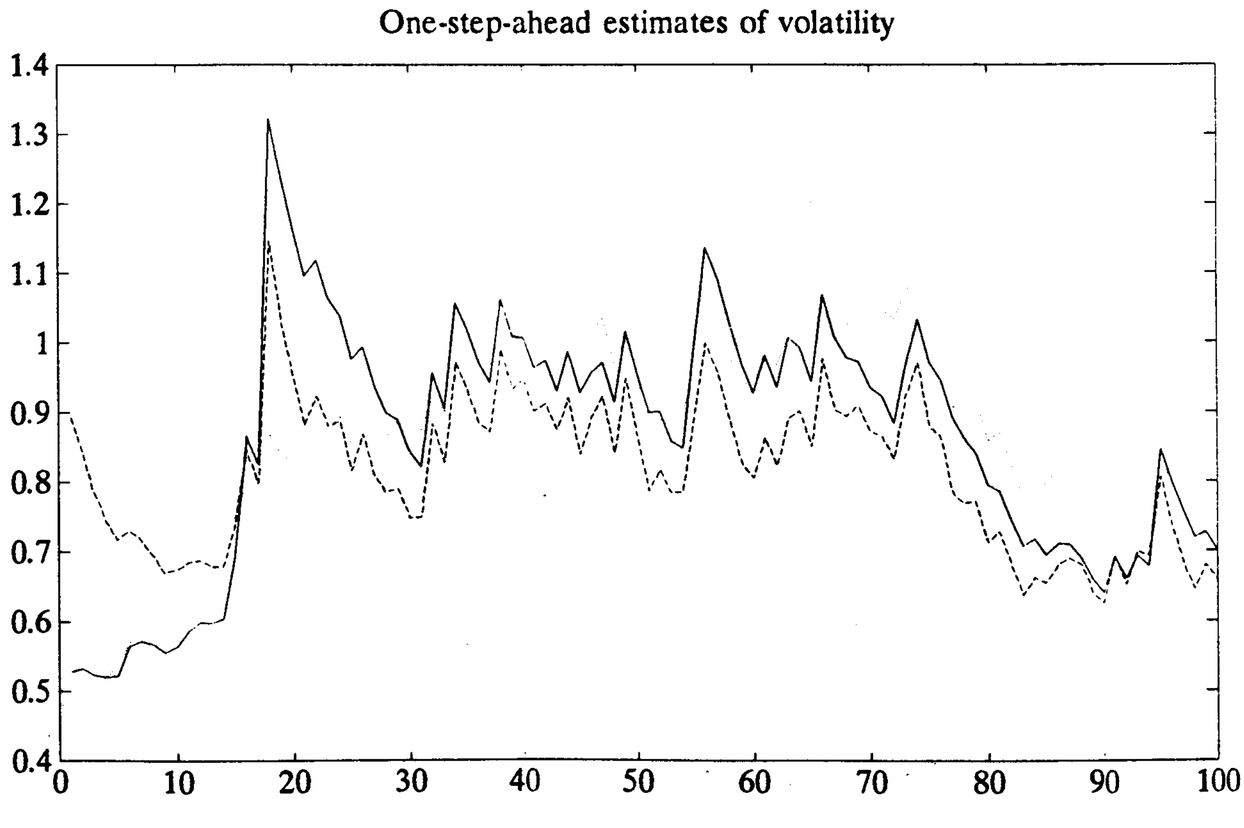

ARV Density
0.45
0.4
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
-6
-4
-2
0
2
4
6
Swiss-Franc Smoothed Estimates

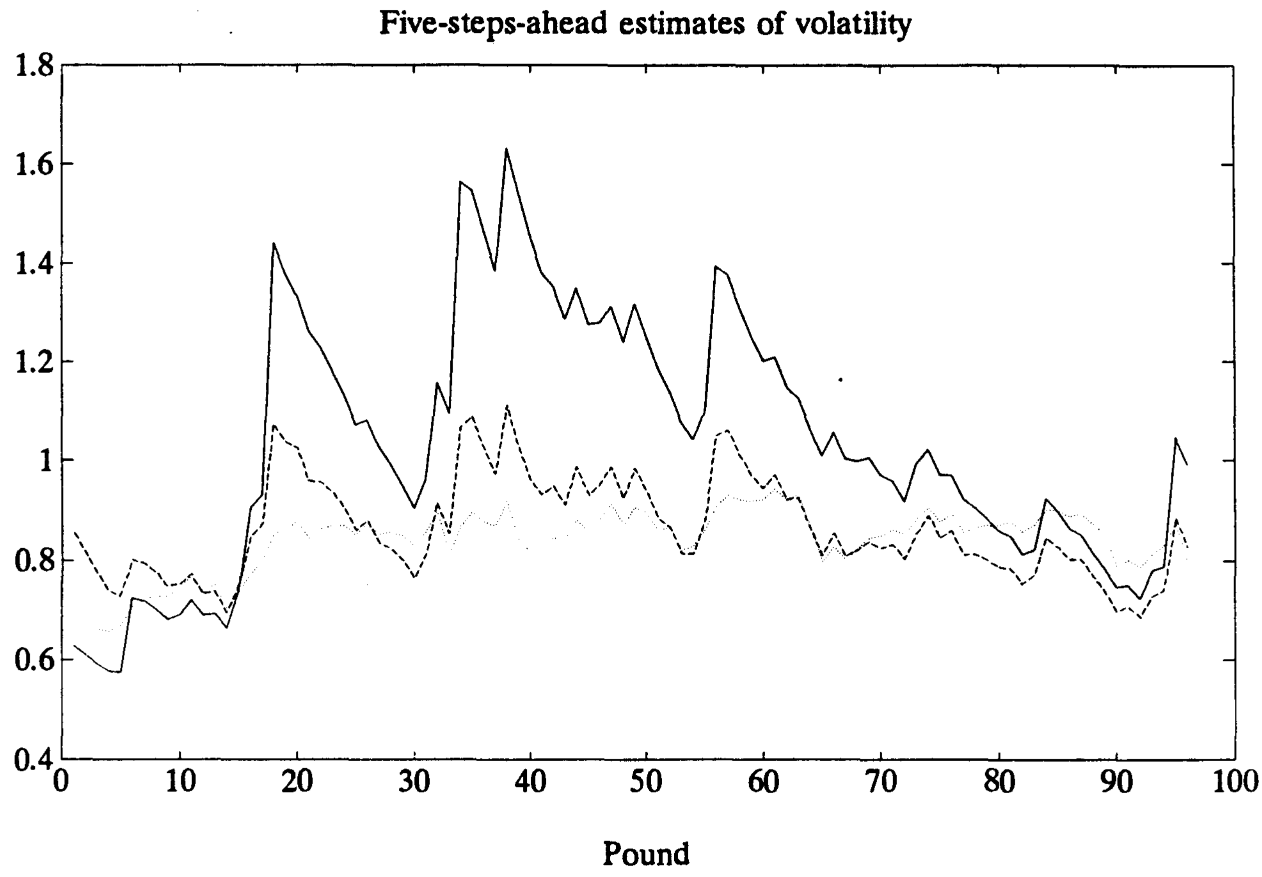
FIGURE 4. FORECASTED VOLATILITIES



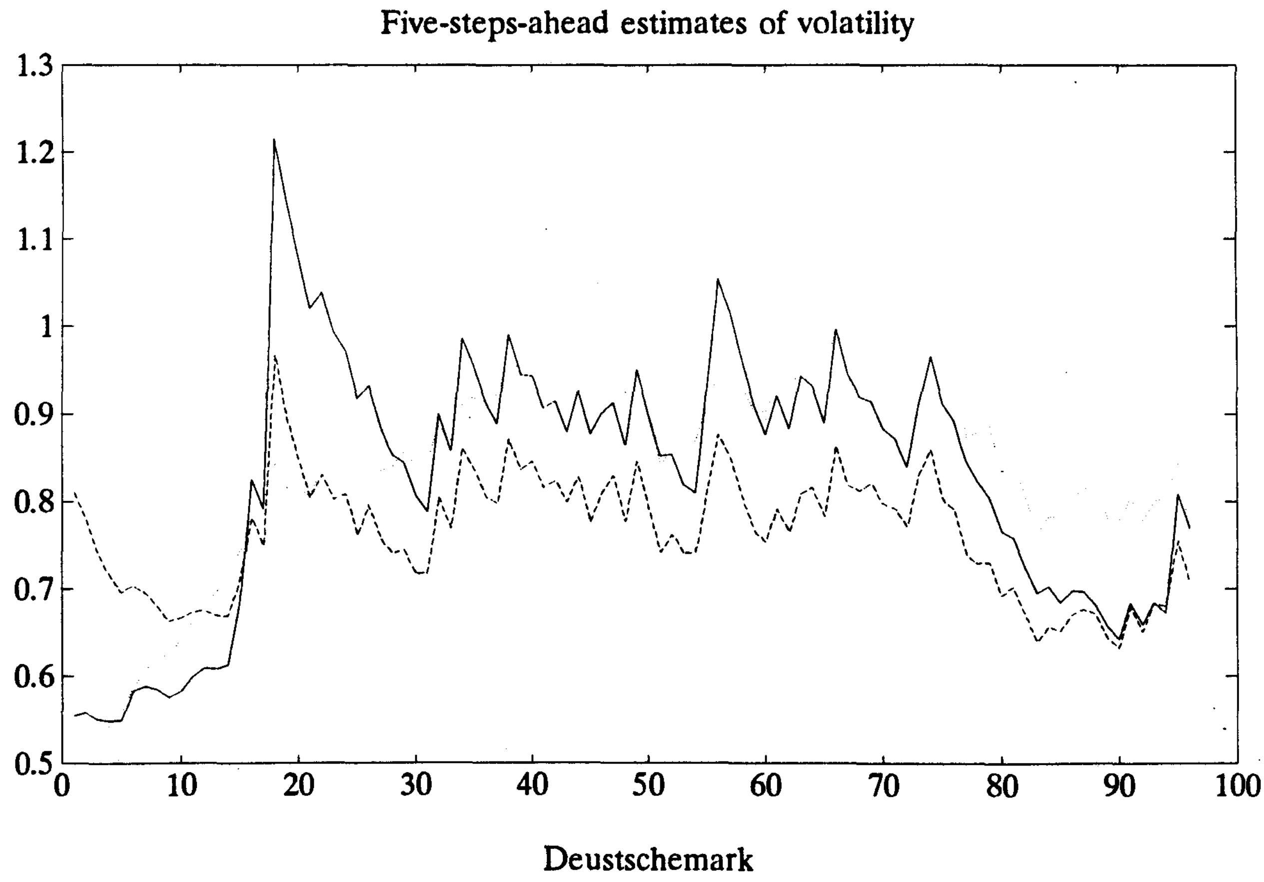
Pound



Deutschemark

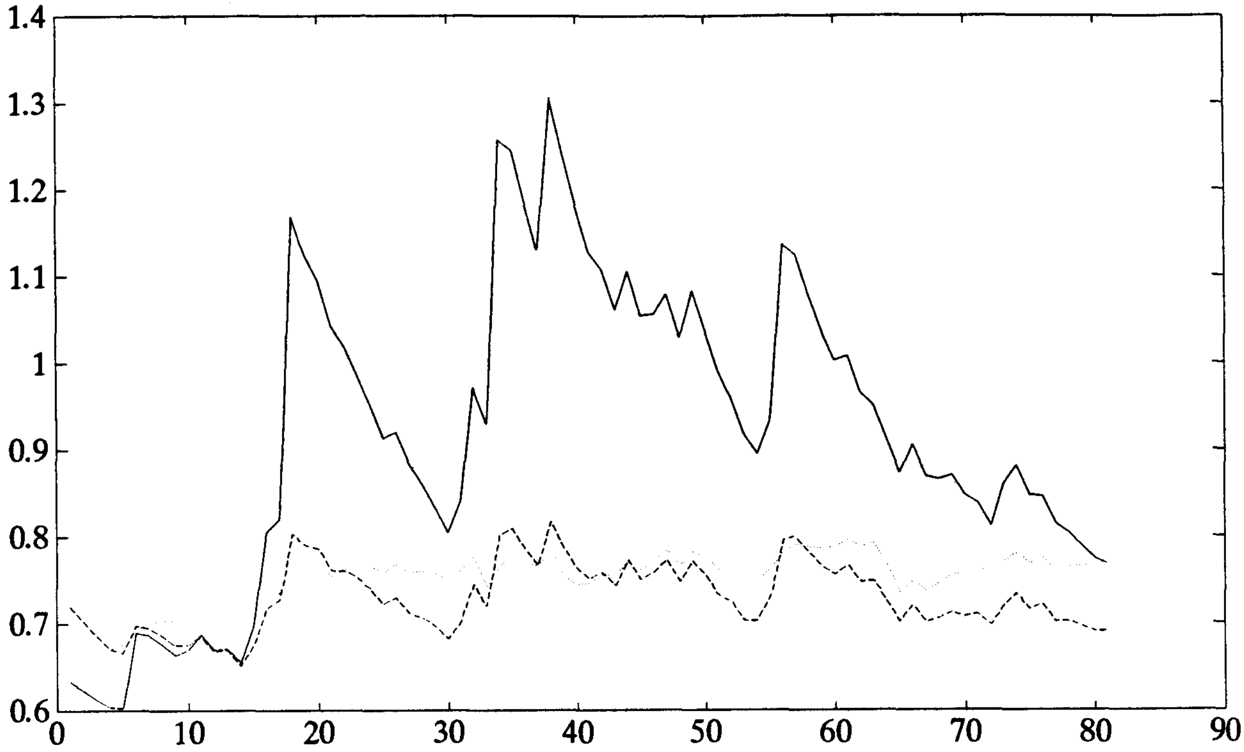


Five-steps-ahead estimates of volatility
1.8
1.6
1.4
1.2
1
0.8
0.6
0.4
0
10
20
30
40
50
60
70
80
90
100
Pound



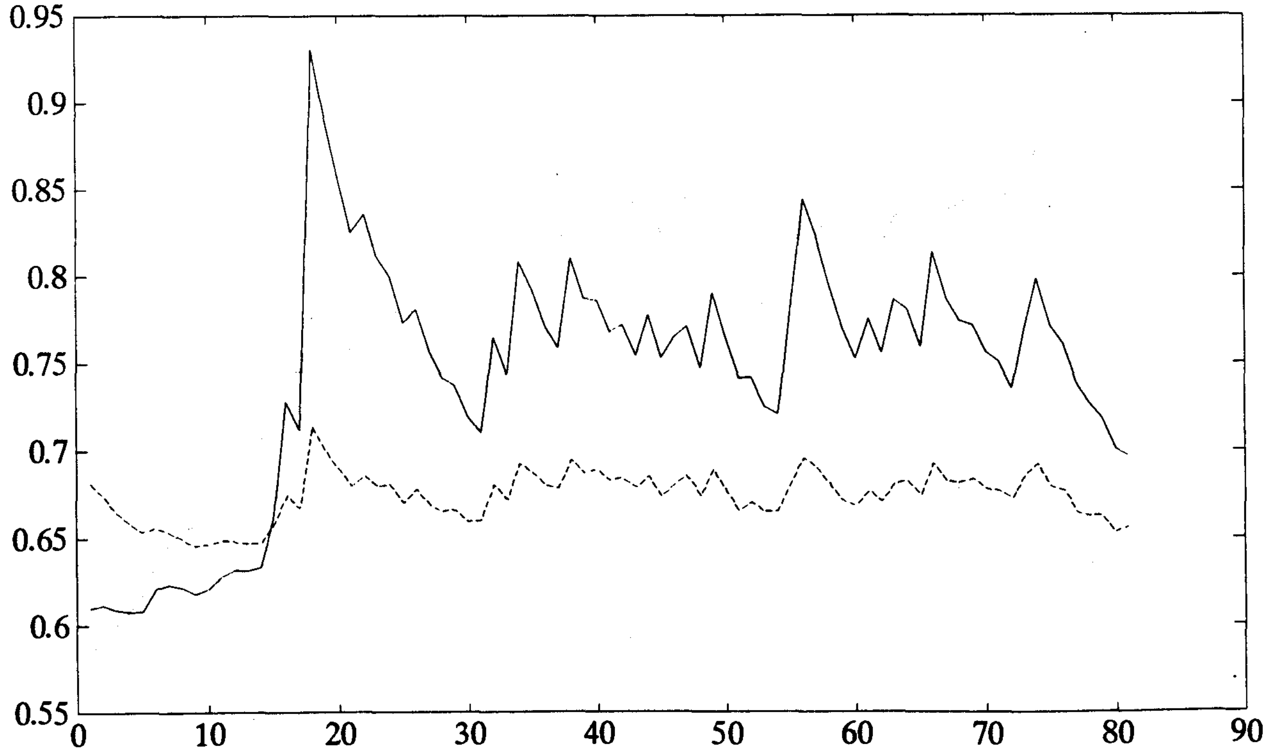
Five-steps-ahead estimates of volatility
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0
10
20
30
40
50
60
70
80
90
100
Deustschemark

Twenty-steps-ahead estimates of volatility



Pound

Twenty-steps-ahead estimates of volatility



Deustschemark

Appendix 1. Moments of $\log(y^2_t)$ when y_t is an EGARCH(1,0) process with normal conditional density

Consider the EGARCH(1,0) process given by

$$y_t = \varepsilon_t \sigma_t \qquad (1.a)$$

$$\log(\sigma^2_t) = \omega + \gamma \varepsilon_{t-1} + \alpha[|\varepsilon_{t-1}| - \sqrt{2/\pi}] + \beta \log(\sigma^2_{t-1}) \qquad (1.b)$$

where $\varepsilon_t \sim$ NID(0,1).

The first two moments of $\log(y^2_t)$ are given by

$$E(\log(y^2_t)) = -1.27 + \omega/(1-\beta) \qquad (2.a)$$

$$Var(\log(y^2_t)) = \frac{\pi^2}{2} + \frac{\gamma^2+\alpha^2(1-\frac{2}{\pi})}{(1-\beta^2)} \qquad (2.b)$$

Proof: The result is straighforward using the moments $E(\log(\varepsilon^2_t)) \approx -1.27$, $Var(\log(\varepsilon^2_t)) = \pi^2/2$ and $E\{\varepsilon_t (|\varepsilon_t| - (2/\pi)^{1/2})\} = 0$; see Harvey *et al.* (1994).

The expression of the acf of $\log(y^2_t)$ is given by

$$\rho_h = \beta^h \frac{\gamma^2+\alpha^2(1-\frac{2}{\pi}) + 1.1058\frac{\alpha}{\beta}}{\frac{\pi^2}{2} + \frac{\gamma^2+\alpha^2(1-\frac{2}{\pi})}{1-\beta^2}}, \qquad h \geq 1 \qquad (3)$$

Proof: First, the autocovariance of order h is given by

$$Cov(\log(y^2_t), \log(y^2_{t-h})) = E[(\log(\sigma^2_t)-\omega/(1-\beta))(\log(\sigma^2_{t-h})-\omega/(1-\beta)) + (\log(\sigma^2_t)-\omega/(1-\beta))(\log(\varepsilon^2_{t-h}) + 1.27)] \qquad (4)$$

To get an expression for the above expected value, the covariance between $|\varepsilon_t|$ and $\log(\varepsilon^2_t)$ is required.

$$E(|\epsilon|\log\epsilon^2)=\int_{-\infty}^{\infty}|\epsilon|\log(\epsilon^2)\frac{1}{(2\pi)^{1/2}}\exp(-\frac{1}{2}\epsilon^2)\,d\epsilon=$$
$$=\frac{1}{(2\pi)^{1/2}}\int_0^{\infty}\log(x)\exp(-\frac{1}{2}x)\,dx=$$
$$=-\frac{2}{(2\pi)^{1/2}}(C+\log(\frac{1}{2})$$

where C is the Euler's constant; see Gradshteyn and Ryzhik (1980) for the solution of the integral.

Then,

$$\mathrm{Cov}(|\varepsilon|, \log(\varepsilon^2)) = (2/\pi)\{-(0.5772 + \log(1/2)) + 1.27\} = 1.1058 \qquad (5)$$

Finally, using (4) and (5) and after some straighforward algebra,

$$\mathrm{Cov}(\log(y^2_t), \log(y^2_{t-h})) = (\gamma^2 + \alpha^2(1 - 2/\pi))\,\beta^h/(1-\beta^2) + 1.1058\,\alpha\,\beta^{h-1} \qquad (6)$$

From (2.b) and (6), it is easy to obtain the expression for the acf of $\log(y^2_t)$.

REFERENCES

Andersen, T.G. (1992), Volatility, manuscript Northwestern University.

Bera, A.K. and M.L. Higgins (1993), A Survey of ARCH Models: Properties, Estimation and Testing, Journal of Economic Surveys, forthcoming.

Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedastic Process, Journal of Econometrics, 31, 307-327.

Bollerslev, T. (1987), A Conditional Heteroskedastic Time Series Model for Speculative Prices and Rates of Return, 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, 5-59.

Bollerslev, T., R.F. Engle and D.B. Nelson (1993), ARCH, The Handbook of Econometrics, Vol. 4, forthcoming.

Danielsson, J. (1992), Stochastic Volatility in Asset Prices: Estimation with Simulated Maximum Likelihood, Journal of Econometrics, forthcoming.

Dassios, A. (1992), Asymptotic Approximations to Stochastic Variance Models, mimeo, London School of Economics.

Engle, R. (1982), Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of U.K. Inflation, Econometrica, 50, 987-1008.

Engle, R. and T. Bollerslev (1986), Modelling the Persistence of Conditional Variances, Econometrics Review, 5, 1, 1-50.

Gradshteyn, I.S. and I.M. Ryzhik (1980), Table of Integrals, Series and Products, Academic Press, Orlando.

Harvey, A.C. (1989), Forecasting, Structural Time Series Models and the Kalman Filter, Cambridge University Press, Cambridge.

Harvey, A.C., E. Ruiz and N.G. Shephard (1994), Multivariate Stochastic Variance Model, Review of Economic Studies, forthcoming.

Jacquier, E., N.G. Polson, P.E. Rossi (1994), Bayesian Analysis of Stochastic Volatility Models, Journal of Business and Economic Statistics, forthcoming.

Nelson, D.B. (1988), The Time Series Behavior of Stock Market Volatility and Returns, PhD dissertation, Massachusetts Institute of Technology.

Nelson, D.B. (1991), Conditional Heteroskedasticity in Asset Returns: A New Approach, Econometrica, 59, 2, 347-370.

Newey, W.K. and K.D. West (1987), A Simple Positive-definite, Heteroskedasticity Autocorrelation Consistent Covariance Matrix, Econometrica, Vol. 55, No. 3, 703-708.

Melino, A. and S. Turnbull (1990), Pricing Foreign Currency Options with Stochastic Volatility, Journal of Econometrics, 45, 7-39.

Pagan, A.R. and G.W. Schwert (1990), Alternative Models for Conditional Stock Volatility, Journal of Econometrics, 45, 267-290.

Ruiz, E. (1994), Quasi-Maximum Likelihood Estimation of Stochastic Variance Models, Journal of Econometrics, forthcoming.

Shephard, N.G. (1994), Fitting non-linear time series models, Journal of Applied Econometrics, forthcoming.

Taylor, S.J. (1986), Modelling Financial Time Series, John Wiley, Chichester, U.K.

Taylor, S.J. (1993) Modeling Stochastic Volatility, Mathematical Finance, forthcoming.