

# ARIMA-GARCH and unobserved component models with GARCH disturbances: Are their prediction intervals different?\*

Santiago Pellegrini<sup>†</sup> Esther Ruiz<sup>†</sup> and Antoni Espasa<sup>†</sup>

Job Market Paper

## Abstract

We analyze the effects on prediction intervals of fitting ARIMA models to series with stochastic trends, when the underlying components are heteroscedastic. We show that ARIMA prediction intervals may be inadequate when only the transitory component is heteroscedastic. In this case, prediction intervals based on the unobserved component models tend to the homoscedastic intervals as the prediction horizon increases. However, prediction intervals based on the ARIMA model incorporate the unit root, so they diverge for ever from the homoscedastic intervals. We focus on the local level and smooth trend models. All the results are illustrated with simulated and real time series.

**Keywords:** State Space Models, Conditional Heteroscedasticity, Non-Gaussian Distributions, Excess Kurtosis, Autocorrelations of Squares.

---

\*Financial support from Project SEJ2006-03919 by the Spanish government is gratefully acknowledged. We are grateful to participants at the 22<sup>nd</sup> International Symposium of Forecasting in Santander 2006, at the Econometrics Meeting in Budapest 2007, and at the 2<sup>nd</sup> International Workshop on Computational and Financial Econometrics in Neuchatel, 2008. The usual disclaims apply.

<sup>†</sup>Statistics Department, Universidad Carlos III, Madrid, Spain.

## 1. Introduction

Economic and financial time series often have stochastic trends. In this case, it is common to take differences in order to obtain a stationary transformation. Then, an ARMA model is fitted to this transformation to represent the transitory dependence. Alternatively, the dynamic properties of series with stochastic trends may be represented by unobserved component models. It is well known that both models are equivalent when the disturbances are i.i.d. and Gaussian given that, in this case, the reduced form of an unobserved component model is an ARIMA model with restrictions in the parameters; see, for example, Harvey (1989). The main difference between both specifications is that while the ARIMA model includes only one disturbance, the corresponding unobserved component model incorporates several disturbances. Consequently, working with the ARIMA specification is usually simpler. However, using the unobserved components model may lead to discover features of the series that are not apparent in the reduced form model. In this paper, we consider one of these features. In particular, we analyze the effects of working with the reduced form ARIMA model on prediction intervals for future values of the series when the underlying components are heteroscedastic.

From an empirical point of view, the presence of conditional heteroscedasticity in both ARIMA and unobserved component models, have previously interested many authors. There is a large literature that considers ARIMA models with GARCH disturbances; see Bollerslev et al. (1992), Bollerslev et al. (1994), Diebold and Lopez (1995), and Diebold (2004) for detailed surveys. On the other hand, unobserved component models with GARCH disturbances have been receiving a lot of attention as they allow to distinguish which components are heteroscedastic. One of the earliest implementations of these models is Harvey et al. (1992), which consider latent factor models; see also King

et al. (1994), Sentana and Fiorentini (2001), Chang and Kim (2004) and Sentana (2004) for other applications related with latent factor models. Chadha and Sarno (2002) and Moore and Schaller (2002) fit unobserved component models with GARCH disturbances to price volatility and term structure of interest rates, respectively. With respect to inflation rates, the presence of heteroscedasticity in the transitory and/or permanent components is a broadly debated issue in the recent literature; see Broto and Ruiz (2009) for a detailed survey. For example, Stock and Watson (2007) find that a simple unobserved component model with conditionally heteroscedastic noises describe well the dynamics of the US inflation.

To simplify the presentation, consider that the series of interest,  $y_t$ , is composed by a transitory component,  $\varepsilon_t$ , and a stochastic trend,  $\mu_t$ , with a stochastic slope  $\beta_t$ . Consequently, the stochastic trend model is given by

$$y_t = \mu_t + \varepsilon_t, \quad (1a)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} + \eta_t, \quad (1b)$$

$$\beta_t = \beta_{t-1} + \xi_t, \quad (1c)$$

where  $\varepsilon_t$ ,  $\eta_t$  and  $\xi_t$  are mutually independent and serially uncorrelated processes, with zero means and variances  $\sigma_\varepsilon^2$ ,  $\sigma_\eta^2$  and  $\sigma_\xi^2$ , respectively. In this paper, we focus on two particular cases of model (1), which are of interest from an empirical point of view. The first one is obtained when  $\sigma_\xi^2 = 0$ . Under this set up, the slope is fixed and the trend reduces to a random walk with a drift given by  $\beta_0$ . In this case, taking first differences in (1) results in a stationary series given by

$$\Delta y_t = \beta_0 + \eta_t + \Delta \varepsilon_t. \quad (2)$$

Without loss of generality, we will also assume that  $\beta_0 = 0$ , so that  $y_t$  follows a local level model; see, for example, Durbin and Koopman (2001) for a detailed description and

applications of this model. With this assumption, it is well known that model (2) can be represented by the following IMA(1,1) model

$$\Delta y_t = a_t + \theta a_{t-1}, \quad (3)$$

where, if  $\Delta y_t$  is invertible, then  $\theta = [(q_\eta^2 + 4q_\eta)^{1/2} - 2 - q_\eta] / 2$ , with  $q_\eta = \sigma_\eta^2 / \sigma_\varepsilon^2$  being the signal-to-noise ratio. Note that the parameter  $\theta$  is restricted to be negative, i.e.  $-1 < \theta < 0$ . Finally, the reduced form disturbance  $a_t$  is an uncorrelated process with zero mean and positive variance equal to  $\sigma_a^2 = -\frac{\sigma_\varepsilon^2}{\theta}$ .

On the other hand, when  $\sigma_\xi^2 > 0$  but  $\sigma_\eta^2 = 0$ , the smooth trend model is obtained; see Harvey and Jaeger (1993) and Nyblom and Harvey (2001) for description and applications. In this case, the trend is an integrated random walk and therefore two differences are necessary to obtain the stationary transformation, that is

$$\Delta^2 y_t = \xi_{t-1} + \Delta^2 \varepsilon_t. \quad (4)$$

The corresponding reduced form is a restricted IMA(2,2) given by

$$\Delta^2 y_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2}, \quad (5)$$

where the parameters  $\theta_1$ ,  $\theta_2$  and  $\sigma_a^2$  are the solutions of the following system

$$\sigma_a^2(1 + \theta_1^2 + \theta_2^2) = \sigma_\varepsilon^2(6 + q_\xi), \quad (6a)$$

$$\frac{\theta_1(1 + \theta_2)}{1 + \theta_1^2 + \theta_2^2} = -\frac{4}{6 + q_\xi}, \quad (6b)$$

$$\frac{\theta_2}{1 + \theta_1^2 + \theta_2^2} = \frac{1}{6 + q_\xi}, \quad (6c)$$

with  $q_\xi = \sigma_\xi^2 / \sigma_\varepsilon^2$ . There are four solutions of system (6) but only one of them contains a pair of real values for  $\theta_1$  and  $\theta_2$  that falls inside the invertibility region.

Given that we are interested in analyzing the reduced form model in the presence of GARCH unobserved noises and that GARCH processes are characterized by having

---

excess kurtosis and positive autocorrelations of squares, in this paper, we work out these moments for the stationary transformation of the local level and the smooth trend models when the underlying disturbances are GARCH(1,1) processes. For the local level model, we consider the case in which either or both disturbances,  $\varepsilon_t$  and  $\eta_t$ , are conditionally Normal GARCH processes, while for the smooth trend model we consider the case where only the transitory component,  $\varepsilon_t$ , is heteroscedastic. Then, we derive the properties of the reduced form ARIMA disturbance,  $a_t$ , in terms of the unobserved disturbances,  $\varepsilon_t$ ,  $\eta_t$  and  $\xi_t$ . We will show that there is no a simple GARCH model for  $a_t$  with exactly the same kurtosis and autocorrelations of squares as those generated in the reduced form noise when the unobserved components are GARCH. In any case, fitting ARIMA-GARCH models is a popular methodology, and, consequently, we also find by simulation, the parameters of an ARIMA-GARCH with similar theoretical kurtosis and autocorrelations of squares to those of conditionally heteroscedastic unobserved component models.

As a general conclusion, we show that if  $\varepsilon_t$  and  $\eta_t$  are assumed to be GARCH processes, the conditional heteroscedasticity of the ARIMA noise  $a_t$  is weaker than the one present in the unobserved disturbances. In some cases,  $a_t$  could even be seen as homoscedastic. Therefore, the heteroscedasticity is more evident in the unobserved component model and can be overlooked when working with the reduced form ARIMA model. This result could be expected as the heteroscedasticity weakens under contemporaneous aggregation of GARCH processes; see, for example, Zaffaroni (2007).

More interestingly, we also analyze the performance of prediction intervals obtained when both, the unobserved component and the ARIMA model are fitted. In both cases, if the series is heteroscedastic, the amplitudes of the intervals change depending on whether the conditional variance at the moment of making the prediction is larger or smaller than the marginal variance. Denoting by excess volatility, the difference between the conditional

and the marginal variances at the moment when the prediction is made, we show that, if only the transitory component is heteroscedastic, then the excess volatility disappears as the prediction horizon increases when the intervals are obtained using the unobserved component model. That is, the prediction intervals obtained with the unobserved component model converge to the intervals of the corresponding homoscedastic model as the prediction horizon increases. However, given that the prediction intervals obtained using the reduced form ARIMA model always contain at least one unit root, they depend on the size and sign of the shock at the time the prediction is made, for any prediction horizon. In this case, depending on whether the excess volatility is positive or negative, the multi-step prediction intervals based on the ARIMA model can be too wide or too narrow respectively, when compared with the intervals based on the corresponding unobserved component model.

The rest of this paper is structured as follows. In Section 2, we derive the kurtosis and autocorrelations of squares of the stationary transformation of the local level and smooth trend models when the disturbances are serially uncorrelated with symmetric distributions and finite fourth order moments. Then, we particularize these results for conditionally Normal GARCH(1,1) disturbances. In Section 3, we derive the statistical properties of the reduced form ARIMA noise,  $a_t$  in the two models considered. We also derive the parameters of the ARIMA-GARCH model that give comparable kurtosis and autocorrelations of squares with those of the unobserved component model with GARCH disturbances. The results in this section are illustrated with Monte Carlo experiments. Section 4 derives prediction intervals for both models. Section 5 contains an empirical application. Finally, Section 6 concludes the paper.

## 2. Properties of the local level and smooth trend models

In this section we derive the statistical properties of the stationary transformation of  $y_t$  in the local level and smooth trend models when the disturbances are assumed to be uncorrelated processes with symmetric densities and finite fourth order moments. As the interest in this paper is to analyze the properties of conditionally heteroscedastic unobserved component models, we focus on the excess kurtosis and the autocorrelation function (acf) of squared observations. We then particularize these results for the case in which the disturbances are GARCH(1,1) processes.

### 2.1 The local level model

Consider the local level model given by

$$y_t = \mu_t + \varepsilon_t \quad (7a)$$

$$\mu_t = \mu_{t-1} + \eta_t. \quad (7b)$$

It is easy to show that the variance of the stationary transformation,  $\Delta y_t$ , in (2) is given by

$$\text{Var}[\Delta y_t] = \sigma_\varepsilon^2(q_\eta + 2). \quad (8)$$

Additionally, the excess kurtosis is given by

$$\bar{\kappa}_{\Delta y} = \frac{q_\eta^2 \bar{\kappa}_\eta + 2\bar{\kappa}_\varepsilon + 6(\bar{\kappa}_\varepsilon + 2)\rho_1^{\varepsilon^2}}{(q_\eta + 2)^2}, \quad (9)$$

where  $\bar{\kappa}_\varepsilon$  and  $\bar{\kappa}_\eta$  are the excess kurtosis of  $\varepsilon_t$  and  $\eta_t$ , respectively, and  $\rho_1^{\varepsilon^2}$  is the lag-one autocorrelation of  $\varepsilon_t^2$ . Note that the signal-to-noise ratio,  $q_\eta$ , plays an important role in determining the relative influence of each noise on  $\bar{\kappa}_{\Delta y}$ . In the limiting cases, when  $q_\eta \rightarrow \infty$ ,  $\Delta y_t = \eta_t$  (i.e.  $y_t$  is a pure random walk process) so that  $\bar{\kappa}_{\Delta y} = \bar{\kappa}_\eta$ . On the other hand, as  $q_\eta \rightarrow 0$ ,  $\Delta y_t = \Delta \varepsilon_t$  (i.e.  $y_t$  is a white noise process), and consequently  $\Delta y_t$  is a

non-invertible MA(1) process whose excess kurtosis may be different from  $\bar{\kappa}_\varepsilon$  depending on the value of  $\rho_1^{\varepsilon^2}$ .

Finally, after some tedious although straightforward algebra, it is possible to show that the acf of  $(\Delta y_t)^2$  is given by

$$\rho_\tau^{(\Delta y)^2} = \frac{q_\eta^2(\bar{\kappa}_\eta + 2)\rho_\tau^{\eta^2} + (\bar{\kappa}_\varepsilon + 2)(\rho_{\tau-1}^{\varepsilon^2} + 2\rho_\tau^{\varepsilon^2} + \rho_{\tau+1}^{\varepsilon^2})}{(\bar{\kappa}_{\Delta y} + 2)(q_\eta + 2)^2}, \quad \tau \geq 1. \quad (10)$$

Note that for Gaussian homoscedastic noises,  $\rho_1^{(\Delta y)^2} = (q_\eta + 2)^{-2}$ , which turns out to be the squared lag-one autocorrelation of  $\Delta y_t$ ; see Maravall (1983). However, when assuming that  $\varepsilon_t$  and  $\eta_t$  are homoscedastic but not necessarily Gaussian, it is possible to see from (10) that  $\rho_1^{(\Delta y)^2}$  may differ from  $(\rho_1^{\Delta y})^2$ , depending on the values of the excess kurtosis of each noise. In the general case, the numerator of (10) is defined as a weighted sum of two factors that depend on  $\tau$ . The first one,  $\rho_\tau^{\eta^2}$ , has a weight that is a function of  $q_\eta$  and  $\bar{\kappa}_\eta$ , while the second,  $\rho_{\tau-1}^{\varepsilon^2} + 2\rho_\tau^{\varepsilon^2} + \rho_{\tau+1}^{\varepsilon^2}$ , has a weight depending only on  $\bar{\kappa}_\varepsilon$ . As long as the acf of squares of both disturbances converges to zero, each of these factors disappears as  $\tau$  increases, and therefore the acf of  $(\Delta y_t)^2$  also converges to zero.

Next, we derive the excess kurtosis of  $\Delta y_t$  and acf of  $(\Delta y_t)^2$  when  $\varepsilon_t$  and  $\eta_t$  are conditionally Normal GARCH(1,1) processes<sup>1</sup>. In this case, the noises in model (7) are given by  $\varepsilon_t = \varepsilon_t^\dagger h_t^{1/2}$  and  $\eta_t = \eta_t^\dagger q_t^{1/2}$ , where  $\varepsilon_t^\dagger$  and  $\eta_t^\dagger$  are mutually and serially independent Normal processes with zero mean and unit variance, and

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1}, \quad (11)$$

$$q_t = \gamma_0 + \gamma_1 \eta_{t-1}^2 + \gamma_2 q_{t-1}, \quad (12)$$

where the parameters  $\alpha_0$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\gamma_0$ ,  $\gamma_1$  and  $\gamma_2$  are assumed to satisfy the usual positivity and stationarity conditions. Substituting  $\bar{\kappa}_\varepsilon$ ,  $\bar{\kappa}_\eta$  and  $\rho_1^{\varepsilon^2}$  in expression (9) by their expres-

<sup>1</sup>The general expression for the kurtosis and acf of  $(\Delta y_t)^2$  can be utilized in other specifications of the noises. For instance, Broto and Ruiz (2006) derive these quantities for the particular case of a local level model with GQARCH disturbances to account for asymmetries in volatility.

sions for GARCH(1,1) processes given by  $\bar{\kappa}_\varepsilon = \frac{2\alpha_1^2}{1-3\alpha_1^2-2\alpha_1\alpha_2-\alpha_2^2}$ ,  $\bar{\kappa}_\eta = \frac{2\gamma_1^2}{1-3\gamma_1^2-2\gamma_1\gamma_2-\gamma_2^2}$  and  $\rho_1^{\varepsilon^2} = \frac{\alpha_1(1-\gamma_1\gamma_2-\gamma_2^2)}{1-2\gamma_1\gamma_2-\gamma_2^2}$ , respectively,  $\bar{\kappa}_{\Delta y}$  is given by

$$\bar{\kappa}_{\Delta y} = \frac{3}{(q_\eta + 2)^2} \left[ q_\eta^2 \frac{2\gamma_1^2}{1-3\gamma_1^2-2\gamma_1\gamma_2-\gamma_2^2} + 4 \frac{\alpha_1(1+\alpha_1-\alpha_1\alpha_2-\alpha_2^2)}{1-3\alpha_1^2-2\alpha_1\alpha_2-\alpha_2^2} \right]. \quad (13)$$

As an illustration, Figure 1 plots for different values of the signal-to-noise ratio, the relationship between the kurtosis of  $\Delta y_t$  and the persistence of the volatility of both noises, measured by  $\alpha_1 + \alpha_2$  and  $\gamma_1 + \gamma_2$ . We set  $\alpha_2 = \gamma_2 = 0.85$  in order to illustrate the situations with high persistence and small values of  $\alpha_1$  and  $\gamma_1$ , as commonly seen in empirical applications. Note that the slope with respect to the persistence of  $\eta_t$  is steeper as  $q_\eta$  increases, and also that varying  $q_\eta$  significantly affects  $\kappa_{\Delta y}$ .

The autocorrelations of squares when  $\varepsilon_t$  and  $\eta_t$  are GARCH(1,1) processes given in (10), become

$$\rho_\tau^{(\Delta y)^2} = \begin{cases} \frac{q_\eta^2 \rho_1^{\eta^2} (\bar{\kappa}_\eta + 2) + (\bar{\kappa}_\varepsilon + 2)(1 + \rho_1^{\varepsilon^2}(2 + \alpha_1 + \alpha_2))}{(q_\eta + 2)^2 (\bar{\kappa}_{\Delta y} + 2)}, & \tau = 1 \\ (\alpha_1 + \alpha_2) \rho_{\tau-1}^{(\Delta y)^2} + \frac{(\gamma_1 + \gamma_2 - \alpha_1 - \alpha_2) q_\eta^2 (\gamma_1 + \gamma_2)^{\tau-2} \rho_1^{\eta^2} (\bar{\kappa}_\eta + 2)}{(q_\eta + 2)^2 (\bar{\kappa}_{\Delta y} + 2)}, & \tau \geq 2, \end{cases} \quad (14)$$

where  $\rho_1^{\eta^2} = \frac{\gamma_1(1-\gamma_1\gamma_2-\gamma_2^2)}{1-2\gamma_1\gamma_2-\gamma_2^2}$ .

From (14) we can see that when the persistence of both noises is the same, i.e.  $\gamma_1 + \gamma_2 = \alpha_1 + \alpha_2$ , the acf of squares has an exponential decay, as in a GARCH(p,q) process. We can also observe an exponential decay when only one noise is heteroscedastic. However, in general, the decay of the autocorrelations in (14) is not exponential. Consequently, the behavior of  $\Delta y_t$  is not GARCH. The shape of the acf of  $(\Delta y_t)^2$  is illustrated in Figure 2 that plots the acf of squares for different specifications of the disturbances and the corresponding rates of decay from the second lag. The first row shows a model in which both disturbances follow the same GARCH process, while in the second model both noises follow GARCH processes with a different persistence. The last two rows consider models

in which only one noise is heteroscedastic. Note first that the cases in the first and the last two rows illustrate the situations mentioned above where we obtain an exponential decay in the acf of  $(\Delta y_t)^2$ . Moreover, in the case where  $\gamma_1 + \gamma_2 \neq \alpha_1 + \alpha_2$ , although the rate is slightly increasing, it can be approximated by a constant<sup>2</sup>. Consequently, exponential structures such as the ones implied by GARCH processes can be a good approximation for the acf of squares<sup>3</sup>.

## 2.2 The smooth trend model

The smooth trend model is given by

$$y_t = \mu_t + \varepsilon_t \quad (15a)$$

$$\mu_t = \mu_{t-1} + \beta_{t-1} \quad (15b)$$

$$\beta_t = \beta_{t-1} + \xi_t. \quad (15c)$$

In this case, the variance and excess kurtosis of the stationary transformation in (4) are given by

$$Var[\Delta^2 y_t] = \sigma_\varepsilon^2(q_\xi + 6), \quad (16)$$

$$\bar{\kappa}_{\Delta^2 y} = \frac{q_\xi^2 \bar{\kappa}_\xi + 18\bar{\kappa}_\varepsilon + 6(\bar{\kappa}_\varepsilon + 2)(8\rho_1^{\varepsilon^2} + \rho_2^{\varepsilon^2})}{(q_\xi + 6)^2}, \quad (17)$$

respectively.

Furthermore, the acf of  $(\Delta^2 y)^2$  is given by

$$\rho_\tau^{(\Delta^2 y)^2} = \begin{cases} \frac{q_\xi^2(\bar{\kappa}_\xi + 2)\rho_1^{\varepsilon^2} + (\bar{\kappa}_\varepsilon + 2)(8 + 35\rho_1^{\varepsilon^2} + 8\rho_2^{\varepsilon^2} + \rho_3^{\varepsilon^2})}{(\bar{\kappa}_{\Delta^2 y} + 2)(q_\xi + 6)^2}, & \tau = 1, \\ \frac{q_\xi^2(\bar{\kappa}_\xi + 2)\rho_\tau^{\varepsilon^2} + (\bar{\kappa}_\varepsilon + 2)(\rho_{\tau-2}^{\varepsilon^2} + 8\rho_{\tau-1}^{\varepsilon^2} + 18\rho_\tau^{\varepsilon^2} + 8\rho_{\tau+1}^{\varepsilon^2} + \rho_{\tau+2}^{\varepsilon^2})}{(\bar{\kappa}_{\Delta^2 y} + 2)(q_\xi + 6)^2}, & \tau \geq 2. \end{cases} \quad (18)$$

<sup>2</sup>It can be proved that the rate of decay of  $\rho_\tau^{(\Delta y)^2}$  implicit in (14) converges to the  $\max(\alpha_1 + \alpha_2; \gamma_1 + \gamma_2)$  as  $\tau$  increases. Therefore, in the cases where the persistence of the GARCH processes are close to each other, the rate of decay of  $\rho_\tau^{(\Delta y)^2}$  will be approximately constant for almost all values of  $\tau$ .

<sup>3</sup>It can be shown that this rate of decay is similar to that of some stochastic volatility models, so that current research is devoted to study this possibility.

The smooth trend model assumes that the slope of the trend evolves smoothly. Therefore, it seems sensible to assume that its noise is homoscedastic. Consequently, we only consider the possibility of the transitory noise being a GARCH(1,1) process as defined in (11). In this case, the excess kurtosis and acf of  $(\Delta^2 y_t)^2$  are given by

$$\bar{\kappa}_{\Delta^2 y} = \frac{18\bar{\kappa}_\varepsilon + 6(\bar{\kappa}_\varepsilon + 2)(8\rho_1^{\varepsilon^2} + \rho_2^{\varepsilon^2})}{(6 + q_\xi)^2}, \quad (19)$$

$$\rho_\tau^{(\Delta^2 y)^2} = \begin{cases} \frac{(\bar{\kappa}_\varepsilon + 2)(8 + 35\rho_1^{\varepsilon^2} + 8\rho_2^{\varepsilon^2} + \rho_3^{\varepsilon^2})}{(\bar{\kappa}_{\Delta^2 y} + 2)(q_\xi + 6)^2}, & \tau = 1 \\ \frac{(\bar{\kappa}_\varepsilon + 2)(\rho_{\tau-2}^{\varepsilon^2} + 8\rho_{\tau-1}^{\varepsilon^2} + 18\rho_\tau^{\varepsilon^2} + 8\rho_{\tau+1}^{\varepsilon^2} + \rho_{\tau+2}^{\varepsilon^2})}{(\bar{\kappa}_{\Delta^2 y} + 2)(q_\xi + 6)^2}, & \tau = 2, 3 \\ (\alpha_1 + \alpha_2)\rho_{\tau-1}^{(\Delta^2 y)^2}, & \tau \geq 4. \end{cases} \quad (20)$$

Expression (20) implies that models like the GARCH(p,q), which produce exponential structures in the acf squares, could be suitable for the stationary transformation of the smooth trend model with a GARCH(1,1) transitory component. As an illustration, Figure 3 plots the acf of squares and its rate of decay for different GARCH parameters and different  $q_\xi$ . Note that the rates are all constant from the third lag, as implied in (20).

### 3. Properties of the noise $a_t$ in the ARIMA model

It is well known that  $\varepsilon_t$ ,  $\eta_t$  and  $\xi_t$  being mutually and serially uncorrelated is sufficient to prove that the reduced form of the local linear trend model in (1) is a restricted IMA model. Consequently, the reduced form noise,  $a_t$ , is also serially uncorrelated. Furthermore, if the three disturbances are fourth-moment stationary and have symmetric distributions, then these properties are also shared by  $a_t$ . Taking this into account, the objective of this section is to derive the moments of  $a_t$  as functions of the moments of the disturbances of the unobserved component model. We will also derive the parameters expected to be obtained if  $a_t$  is assumed to be a GARCH(1,1) model.

### 3.1 The local level model

Consider the reduced form IMA(1,1) model given in (3). In this case, the excess kurtosis of  $\Delta y_t$ , is given by

$$\bar{\kappa}_{\Delta y} = \frac{\bar{\kappa}_a(1 + \theta^4) + 6\theta^2\rho_1^{a^2}(\bar{\kappa}_a + 2)}{(1 + \theta^2)^2}, \quad (21)$$

where  $\bar{\kappa}_a$  and  $\rho_1^{a^2}$  are the excess kurtosis of  $a_t$  and the lag-one autocorrelation of  $a_t^2$ , respectively. On the other hand, it is easy to show that the acf of  $\Delta y_t^2$  is given by

$$\rho_\tau^{(\Delta y)^2} = \frac{\bar{\kappa}_a + 2}{(1 + \theta^2)^2(\bar{\kappa}_{\Delta y} + 2)} \left[ (1 + \theta^4)\rho_\tau^{a^2} + \theta^2(\rho_{\tau-1}^{a^2} + \rho_{\tau+1}^{a^2}) \right], \quad \tau \geq 1. \quad (22)$$

In order to find expressions of  $\bar{\kappa}_a$  and  $\rho_\tau^{a^2}$ , as functions of the parameters of the unobserved component model, we equal the excess kurtosis of  $\Delta y_t$  given by (9) and (21), and the autocorrelations of order  $\tau = 1, 2, \dots$  in (10) and (22). The following system is then obtained<sup>4</sup>

$$\begin{aligned} (\bar{\kappa}_a + 2) \left( 1 + \theta^4 + 6\theta^2\rho_1^{a^2} \right) &\equiv (1 + \theta)^4(\bar{\kappa}_\eta + 2) - 8\theta(1 + \theta)^2 \\ &\quad + 2\theta^2(\bar{\kappa}_\varepsilon + 2) \left( 1 + 3\rho_1^{\varepsilon^2} \right), \end{aligned} \quad (23a)$$

$$\begin{aligned} (\bar{\kappa}_a + 2) \left[ (1 + \theta^4)\rho_\tau^{a^2} + \theta^2 \left( \rho_{\tau-1}^{a^2} + \rho_{\tau+1}^{a^2} \right) \right] &\equiv \theta^2(\bar{\kappa}_\varepsilon + 2) \left( \rho_{\tau-1}^{\varepsilon^2} + 2\rho_\tau^{\varepsilon^2} + \rho_{\tau+1}^{\varepsilon^2} \right) \\ &\quad + (1 + \theta)^4(\bar{\kappa}_\eta + 2)\rho_\tau^{\eta^2}, \quad \tau \geq 1. \end{aligned} \quad (23b)$$

As previously stated,  $\rho_1^{(\Delta y)^2}$  may differ from  $(\rho_1^{\Delta y})^2$  when  $\varepsilon_t$  and  $\eta_t$  are homoscedastic but not necessarily normal. Given that  $y_t$  follows an IMA(1,1) process,  $a_t$  must incorporate a nonlinear behavior that explains this difference. In other words, though still uncorrelated,  $a_t$  is not independent; see Breidt and Davis (1992). To illustrate the behavior of  $a_t$  in this particular set up, Table 1 shows the theoretical acf of squares for several

<sup>4</sup>To obtain (23), recall that  $\theta$  can be defined in terms of  $q_\eta$ , so that the following expressions result:

$$\begin{aligned} 1 + \theta^2 &= -\theta(q_\eta + 2), \\ 1 + \theta^4 &= \theta^2(q_\eta^2 + 4q_\eta + 2). \end{aligned}$$

values of  $q_\eta$ ,  $\bar{\kappa}_\varepsilon$  and  $\bar{\kappa}_\eta$ , coming from the resolution of the system given by (23). Observe that non-normality in either or both noises may generate non-zero autocorrelations of squares (specially in the lag-one autocorrelation). However, these autocorrelations do not follow any specific pattern and, consequently, they do not reflect the presence of GARCH effects in the series. However, it is possible to obtain values for some usual conditional homoscedasticity statistics, e.g. the McLeod-Li (1983) test, that may wrongly lead to reject the null of conditionally homoscedasticity. Note also that the pattern in which only the lag-one autocorrelation of squares is different from zero may be confused with the effect of outliers; see Carnero et al. (2006).

When either or both noises are heteroscedastic, the kurtosis and acf of squares of  $a_t$  are not easily derived. However, we can redefine the system given by (23) to construct the following set of equations:

$$[(1 + \theta^4) - 6\theta^2 Q(1)] \rho_1^{a^2} + \theta^2 \rho_2^{a^2} = Q(1)(1 + \theta^4) - \theta^2, \quad (24)$$

$$[\theta^2 - 6\theta^2 Q(2)] \rho_1^{a^2} + (1 + \theta^4) \rho_2^{a^2} + \theta^2 \rho_3^{a^2} = Q(2)(1 + \theta^4), \quad (25)$$

$$-6\theta^2 Q(\tau) \rho_1^{a^2} + \theta^2 \rho_{\tau-1}^{a^2} + (1 + \theta^4) \rho_\tau^{a^2} + \theta^2 \rho_{\tau+1}^{a^2} = Q(\tau)(1 + \theta^4), \quad \tau > 2 \quad (26)$$

where  $Q(\tau)$  depends on the moments of  $\varepsilon_t$  and  $\eta_t$  in the following way:

$$Q(\tau) = \frac{(1 + \theta)^4 \rho_\tau^{\eta^2} (\bar{\kappa}_\eta + 2) + \theta^2 (\bar{\kappa}_\varepsilon + 2) (\rho_{\tau-1}^{\varepsilon^2} + 2\rho_\tau^{\varepsilon^2} + \rho_{\tau+1}^{\varepsilon^2})}{(1 + \theta)^4 (\bar{\kappa}_\eta + 2) - 8\theta(1 + \theta)^2 + 2\theta^2 (\bar{\kappa}_\varepsilon + 2) (1 + 3\rho_1^{\varepsilon^2})}.$$

When assuming that  $\varepsilon_t$  and  $\eta_t$  are stationary GARCH processes,  $Q(\tau)$  converges to zero as  $\tau$  increases. In other words, there exists a value of  $\tau$ , say  $\tau_{max}$ , large enough to make  $Q(\tau) \approx 0$  and thus also make  $\rho_\tau^{a^2} \approx 0$  for  $\tau > \tau_{max}$ . Taking this into account we can find the kurtosis and acf of squares of  $a_t$  for a given set of parameters. Figure 4 plots the acf of squares of  $a_t$  for the same models considered in Figure 2. In general, the magnitude of the autocorrelations of  $a_t^2$  is smaller than the corresponding ones of the disturbances

of the local level model. This suggests that working with the reduced form of an unobserved component model may hide part of the heteroscedasticity of each component, by producing a reduced form disturbance,  $a_t$ , with less structure in its acf of squares. For instance, it might be the case that if the permanent component,  $\mu_t$ , presents a significant heteroscedastic structure but the transitory component,  $\varepsilon_t$ , is homoscedastic; the stationary transformation,  $\Delta y_t$ , may not provide significant evidence of heteroscedasticity at all.

It is worthy to note that, as found with the acf of  $(\Delta y_t)^2$ , the autocorrelations of squares of  $a_t$  may not show an exponential decay. Therefore,  $a_t$  may not follow a GARCH process as well.

### 3.2 The smooth trend model

The reduced form of the smooth trend is an IMA(2,2) as given in (4) with the MA parameters being close to the non-invertibility frontier. The excess kurtosis and autocovariances of squares of the  $\Delta^2 y_t$  are given by

$$\bar{\kappa}_{\Delta^2 y} = \frac{\bar{\kappa}_a (1 + \theta_1^4 + \theta_2^4) + 6(\bar{\kappa}_a + 2) \left[ \theta_1^2 (1 + \theta_2^2) \rho_1^{a^2} + \theta_2^2 \rho_2^{a^2} \right]}{\sigma_a^4 (1 + \theta_1^2 + \theta_2^2)^2} \quad (27)$$

$$\rho_{\tau}^{(\Delta^2 y)^2} = \begin{cases} \frac{\bar{\kappa}_a + 2}{(1 + \theta_1^2 + \theta_2^2)^2 (\bar{\kappa}_{\Delta^2 y} + 2)} \left[ \theta_1^2 (1 + \theta_2^2) + (1 + \theta_1^4 + \theta_2^4 + \theta_2^2 + 4\theta_1^2 \theta_2) \rho_1^{a^2} + \theta_1^2 (1 + \theta_2^2) \rho_2^{a^2} + \theta_2^2 \rho_3^{a^2} + \frac{4\theta_1^2 \theta_2}{\bar{\kappa}_a + 2} \right], & \tau = 1 \\ \frac{\bar{\kappa}_a + 2}{(1 + \theta_1^2 + \theta_2^2)^2 (\bar{\kappa}_{\Delta^2 y} + 2)} \left[ \theta_2^2 \rho_{\tau-2}^{a^2} + \theta_1^2 (1 + \theta_2^2) \rho_{\tau-1}^{a^2} + (1 + \theta_1^4 + \theta_2^4) \rho_{\tau}^{a^2} + \theta_1^2 (1 + \theta_2^2) \rho_{\tau+1}^{a^2} + \theta_2^2 \rho_{\tau+2}^{a^2} \right], & \tau \geq 2, \end{cases} \quad (28)$$

respectively. If  $\varepsilon_t$  is assumed to be GARCH(1,1), then the conditional heteroscedasticity should be also present in the resulting  $a_t$ . In this case, finding the implied acf of  $a_t^2$  in terms of  $q_{\xi}$  by equating the expressions of the kurtosis in (17) and (27), and the expressions of

the autocorrelations of squares in (18) and (20), is very complicate. Consequently, Figure 5 shows the mean estimates of the sample acf of simulated series. Looking at the plots of this Figure we find the same patterns as the ones derived analytically for the local level model. The autocorrelation structure of the squared innovations is markedly weaker than that of the transitory component. As expected, the difference between these two autocorrelation functions is higher as  $q_\xi$  increases ( $\sigma_\varepsilon^2$  decreases relative to  $\sigma_\xi^2$ ). However, it seems to be invariant for different GARCH specifications.

Overall, we find the same results as in the local level model. That is, the ARCH effects in the resulting ARIMA disturbance are less evident than in the unobserved components.

### 3.3 The IMA-GARCH model and the reduced form of $y_t$

In the previous subsection, we have seen that if the disturbances of an unobserved component model are GARCH, the noise of the corresponding reduced form model does not follow exactly a GARCH(1,1) model. However, the decay of the autocorrelations of squares could be approximated by such a model. On the other hand, it is usual when analyzing real time series to fit ARIMA-GARCH models. Consequently, in this subsection, we carry out Monte Carlo experiments to know which values of the GARCH parameters would be obtained if one fits a GARCH(1,1) model to the disturbance of the IMA(d,q) model for  $y_t$ . In particular, if  $a_t$  is assumed to be a conditionally Normal GARCH(1,1) model, then  $a_t = a_t^\dagger \sigma_t$ , where  $a_t^\dagger$  is a white noise Gaussian process and

$$\sigma_t^2 = \delta_0 + \delta_1 a_{t-1}^2 + \delta_2 \sigma_{t-1}^2. \quad (29)$$

For each of the two unobserved component models considered in this paper, 1000 series are generated with four different sample sizes ( $T = 200, 500, 1000, \text{ and } 5000$ )<sup>5</sup>. In the local level model, we fix  $\alpha_1 = \alpha_2 = 0$ ,  $\gamma_1 = 0.15$ ,  $\gamma_2 = 0.8$  and  $q_\eta = 1$ , while in the smooth

---

<sup>5</sup>For reason of space we only show one exercise per model, but the results can be generalized to different

trend model we set  $\alpha_1 = 0.1$ ,  $\alpha_2 = 0.85$  and  $q_\xi = 1$ . For each simulated series, we first fit an homoscedastic MA model to the reduced form and test for conditional heteroscedasticity in the residuals using the test proposed by Rodriguez and Ruiz (2005):

$$Q_1(10) = T \sum_{k=1}^9 [\tilde{r}(k) + \tilde{r}(k+1)]^2,$$

where  $\tilde{r}(k) = \sqrt{(T+2)/(T-k)}r(k)$  is the standardized sample autocorrelation of order  $k$ .

Tables 2 and 3 report the mean and standard deviations through all Monte Carlo replicates of the QML estimates of the MA parameters for each model respectively. We see that in the local level case  $\theta$  is estimated quite accurately with respect to the value implied by  $q_\eta$ . In the smooth trend model, however, some downward bias still remains even for large samples. These Tables also report the percentage of times when the null of homoscedasticity is rejected by the  $Q_1(10)$  test. We can see that for relatively large samples ( $T = 5000$ ) the IMA residuals in all models capture the conditional heteroscedasticity coming from the underlying unobserved components. However, for small or moderate samples, this statistic shows a large proportion of cases where the homoscedasticity cannot be rejected. For instance, for  $T = 200$ , the exercise with the smooth trend shows that 43% of the cases the residuals do not present evidence of conditional heteroscedasticity, rising to 77% for the local level model.

The next step in the exercise consists on fitting a GARCH(1,1) model to the residuals of each series. The second block of Tables 2 and 3 reports the Monte Carlo means and standard deviations of the QML estimates, as well as of the plug-in kurtosis and autocorrelations of squares of  $a_t$  obtained by substituting the estimated parameters in the analytical expressions of these moments. With respect to the estimates of the ARCH parameter,  $\delta_1$ , their values are consistent with the results found above, since they imply a

---

parameter specifications.

significant reduction in the ARCH effects. For example, in the local level model  $\gamma_1 = 0.15$  and the average estimate of  $\delta_1$  is around 0.05. With respect to the GARCH coefficient,  $\delta_2$ , we see that the average estimates increase with the sample size, so that the estimated sum  $\delta_1 + \delta_2$  converges to the sum  $\gamma_1 + \gamma_2$ . In the smooth trend model we observe pretty similar results. Finally, the last row of Tables 2 and 3 reports the percentage of series in which the ARCH parameter,  $\delta_1$ , is significant. Observe that only a 9% of the cases in the local level and a 17% in the smooth trend show significant ARCH effects for  $T = 200$ . Of course, for large samples these figures rapidly increase to reach the 100%.

In this section, we have found that the disturbance  $a_t$  of a reduced form ARIMA model has weaker ARCH effects compared to  $\varepsilon_t$  and/or  $\eta_t$ . This is reflected mainly in a smaller structure of the acf of squares, that leads to a value for  $\delta_1$  smaller than the corresponding to the unobserved components. Furthermore, when dealing with small samples, a higher sampling error may lead to the lack of rejection of the homoscedasticity hypothesis in series that are indeed generated as a sum of conditionally heteroscedastic processes.

#### 4. Analysis of the forecasting performance

In this section we derive expressions of the prediction intervals obtained when an unobserved component model with GARCH disturbances is considered to represent the dynamic evolution of series with stochastic trends. We will see how these intervals adapt their length depending on whether the conditional variance at the moment when the prediction is made is over or under the marginal variance. On the other hand, we also derive the intervals that are obtained when fitting the corresponding ARIMA model. We will see that there may be crucial differences between both prediction intervals mainly when the conditional heteroscedasticity affects only the transitory component.

#### 4.1 Prediction intervals in series with stochastic levels

Consider that  $y_t$  is given by the local level model in (7) with GARCH errors (LL-GARCH). It is well known that if one wants to minimize the mean squared forecast error (MSFE), then the conditional mean is the optimal point predictor of  $y_{T+k}$ . Assuming that the parameters are known, the Kalman Filter derived in Harvey et al. (1992) may be implemented to obtain estimates of the underlying state at time  $t = 1, 2, \dots, T$  denoted by  $m_t$ . Then, the optimal linear point predictor is given by

$$\hat{y}_{T+k} = m_T, \quad k = 1, 2, \dots \quad (30)$$

Taking into account that the T under the expectation means that it is conditional on the information available at time T, the corresponding MSFE is given by

$$\begin{aligned} MSFE(\hat{y}_{T+k}) &= E_T [(y_{T+k} - \hat{y}_{T+k})^2] \\ &= E_T [(\mu_T + \eta_{T+1} + \dots + \eta_{T+k} + \varepsilon_{T+k} - m_T)^2] \\ &= P_T^\mu + E_T[\varepsilon_{T+k}^2] + \sum_{j=1}^k E_T[\eta_{T+j}^2], \quad k = 1, 2, \dots \\ &= P_T^\mu + \sigma_\varepsilon^2 + k \sigma_\eta^2 + \frac{1 - (\gamma_1 + \gamma_2)^k}{1 - (\gamma_1 + \gamma_2)} (q_{T+1} - \sigma_\eta^2) \\ &\quad + (\alpha_1 + \alpha_2)^{k-1} (h_{T+1} - \sigma_\varepsilon^2), \quad k = 1, 2, \dots \end{aligned} \quad (31)$$

where  $P_T^\mu = E_T[(\mu_T - m_T)^2]$ , and  $h_{T+1}$  and  $q_{T+1}$  are the conditional variances of  $\varepsilon_{T+1}$  and  $\eta_{T+1}$ , respectively (see Harvey et al., 1992, for further details). Expressions  $(h_{T+1} - \sigma_\varepsilon^2)$  and  $(q_{T+1} - \sigma_\eta^2)$  may be interpreted as measures of the excess volatility at the time the prediction is made with respect to the marginal variance in both noises. Note that the MSFE of the homoscedastic local level model is given by the first three terms of (31). Furthermore, given that  $\alpha_1 + \alpha_2 < 1$ , the MSFE of the LL-GARCH becomes a linear function of  $k$  in the long run, with the same slope as its homoscedastic counterpart, but with a different intercept due to the contribution of the fourth term in (31). However,

for short and medium horizons, depending on whether the excess kurtosis is negative or positive, the influence of the excess volatility in both noises leads to a MSFE smaller or greater than that of the homoscedastic local level model. It is also important to note that there is a significant difference in the behavior of the MSFE depending on whether the conditional heteroscedasticity affects the long or the short-run components. An excess volatility in the permanent component affects the MSFE for all horizons while the effect of an excess volatility in the transitory component vanishes in the long run. Therefore, when the heteroscedasticity only affects the transitory noise, the MSFE converges to the one obtained in the homoscedastic model. However, when the long run component is heteroscedastic, depending on the sign of the excess volatility, the MSFE is over or under the MSFE obtained in the homoscedastic model for all prediction horizons.

Consider now that the predictions are obtained using the IMA-GARCH representation of  $y_t$  given by (3) and (29). In this case, the optimal linear predictor of  $y_{T+k}$  given the information available at time  $T$  is given by

$$\hat{y}_{T+k} = y_T + \theta a_T, \quad k = 1, 2, \dots \quad (32)$$

with MSFE

$$MSFE(\hat{y}_{T+k}) = \begin{cases} E_T[a_{T+1}^2], & k = 1 \\ E_T[a_{T+k}^2] + (1 + \theta)^2 \sum_{j=1}^{k-1} E_T[a_{T+j}^2], & k = 2, 3, \dots \end{cases} \quad (33)$$

where

$$E_T[a_{T+k}^2] = \sigma_a^2 + (\delta_1 + \delta_2)^{k-1} (\sigma_{T+1}^2 - \sigma_a^2), \quad k = 1, 2, \dots \quad (34)$$

In this case,  $\sigma_{T+1}^2$  is the conditional variance of  $a_{T+1}$ , and  $(\sigma_{T+1}^2 - \sigma_a^2)$  is the corresponding

measure of the excess volatility. By plugging (34) into (33) we find that

$$MSFE(\hat{y}_{T+k}) = \begin{cases} \sigma_a^2 + (\sigma_{T+1}^2 - \sigma_a^2), & k = 1 \\ [(1 + \theta)^2(k - 1) + 1] \sigma_a^2 + \\ \left[ \frac{(1 + \theta)^2 - (\delta_1 + \delta_2)^{k-1}(\theta(2 + \theta) + \delta_1 + \delta_2)}{1 - (\delta_1 + \delta_2)} \right] (\sigma_{T+1}^2 - \sigma_a^2), & k = 2, 3, \dots \end{cases} \quad (35)$$

Note that the MSFE in (35) can also be separated into a linear and a nonlinear part, defined by the first and second terms of the expression, respectively. It is clear from this expression that as  $k$  increases, the  $MSFE(\hat{y}_{T+k})$  is also a linear function of the horizon. However, as long as the excess volatility is different from zero, the path of the IMA-GARCH MSFE never converges to that of the homoscedastic IMA model. Moreover, the sign of the excess volatility at time  $T$  determines if the IMA-GARCH prediction variance will be smaller or greater for all  $k$  than the prediction variance of the homoscedastic IMA. In this sense, the behavior is similar to that of the local level model with heteroscedastic long-run disturbances.

In order to illustrate the behavior of the predictions intervals constructed from the two alternative ways of dealing with stochastic levels, we generate series with their component disturbances being GARCH processes and, assuming that the parameters of both models are known<sup>6</sup>, we find the MSFEs at a given time  $T$ . Then, we construct 95% Gaussian prediction intervals<sup>7</sup> and calculate their observed coverage by generating  $B = 1000$  trajectories of  $y_{T+k}$  conditional on the information at time  $T$ . The time points are arbitrarily chosen to illustrate the behavior of the MSFEs in highly volatile and more quiet periods.

The first case is reported in Figure 6 and assumes that we are in a highly volatile

<sup>6</sup>In the reduced form ARIMA case,  $\theta$  and  $\sigma_a^2$  are directly derived from  $\sigma_\varepsilon^2$  and  $\sigma_\eta^2$ , while for the parameters of the GARCH(1,1) model,  $\delta_1$  and  $\delta_2$ , we use the mean estimates given in the Monte Carlo simulations of the previous section.

<sup>7</sup>Although we know that if  $k > 1$  the forecasts distribution is not Gaussian, the results of Pascual et al. (2006) suggest that it may be a good approximation.

period of a series generated from a LL-GARCH model where only  $\varepsilon_t$  is heteroscedastic, with parameters  $\alpha_1 = 0.15$ ,  $\alpha_2 = 0.8$ , and  $q_\eta = 1$ . The 95% prediction intervals for the LL-GARCH and the IMA-GARCH models are in solid and dashed lines, respectively. The prediction intervals for the homoscedastic model is in dash-dotted lines. Finally, the vertical clouds of points (marked with stars) are the 1000 generated observations for some prediction horizons. Thus, according to the nominal coverage, 50 of these points should lay outside the intervals for each  $k$ .

At the selected time point  $T$ , where  $h_{T+1} - \sigma_\varepsilon^2 = 4.2$ , we see that the IMA-GARCH MSFEs produce too wide prediction intervals. In order to obtain the empirical coverage, we simply count the number of observations lying inside each prediction interval and then divide it by  $B$ . Figure 7 shows the empirical coverage of both models for each horizon  $k$ . We can see that at  $k = 1$  both intervals are pretty close to the nominal but as  $k$  increases, the empirical coverage of the IMA-GARCH is almost 100%. As expected, the homoscedastic model significantly underestimate the coverage for short horizons, but then it works better because the effect of the large shock vanishes with  $k$ .

In the second example, we consider a quiet period where  $h_{T+1} - \sigma_\varepsilon^2 = -0.46$ . Then, the IMA-GARCH MSFEs produce too narrow prediction intervals, covering around 92% of the observations against a 95.8% of the LL-GARCH counterparts. Figure 8 plots the coverage in this last case.

## 4.2 Prediction intervals for series with smooth stochastic trends

Consider now that  $y_t$  is given by the smooth trend model in (15), with  $\varepsilon_t$  being a GARCH(1,1) process (ST-GARCH). In this case, the optimal linear point predictor is given by

$$\hat{y}_{T+k} = m_T + k b_T, \quad k = 1, 2, \dots \quad (36)$$

where  $m_T$  and  $b_T$  are the Kalman Filter estimates of  $\mu_t$  and  $\beta_t$  at time  $T$ . From (36) we find the general expression of the MSFE:

$$\begin{aligned}
MSFE(\hat{y}_{T+k}) &= P_T^\mu + k^2 P_T^\beta + 2k P_T^{\mu,\beta} + \frac{k(k-1)(2k-1)}{6} \sigma_\xi^2 + E[\varepsilon_{T+k}^2] \\
&= P_T^\mu + k^2 P_T^\beta + 2k P_T^{\mu,\beta} + \frac{k(k-1)(2k-1)}{6} \sigma_\xi^2 + \sigma_\varepsilon^2 \\
&\quad + (\alpha_1 + \alpha_2)^{k-1} (h_{T+1} - \sigma_\varepsilon^2), \quad k = 1, 2, \dots
\end{aligned} \tag{37}$$

where  $P_T^\beta = E[(\beta_T - b_T)^2]$  and  $P_T^{\mu,\beta} = E[(\mu_T - m_T)(\beta_T - b_T)]$ . Note that the MSFE of the homoscedastic smooth trend model is given by the first five terms of (37), and the last term depends on the excess volatility. However, the long run predictions of the homoscedastic and GARCH smooth trend model will have the same MSFEs since the heteroscedasticity comes only from the transitory component.

Consider now the IMA(2,2)-GARCH(1,1) model in (5) and (29). In this case, the optimal linear predictor is given by

$$\hat{y}_{T+k} = \begin{cases} 2y_T - y_{T-1} + \theta_1 a_T + \theta_2 a_{T-1}, & k = 1 \\ k \hat{y}_{T+1} - (k-1)(y_T - \theta_2 a_T), & k = 2, 3, \dots \end{cases} \tag{38}$$

with MSFE

$$MSFE(\hat{y}_{T+k}) = \begin{cases} E[a_{T+1}^2], & k = 1 \\ \sum_{j=0}^{k-1} B_j^2 E[a_{T+k-j}^2], & k = 2, 3, \dots \end{cases} \tag{39}$$

where  $B_0 = 1$  and  $B_j = j(1 + \theta_1 + \theta_2) + (1 - \theta_2)$ ,  $j = 1, 2, \dots, k-1$ . After plugging expression (34) in (39) we obtain

$$MSFE(\hat{y}_{T+k}) = \begin{cases} \sigma_a^2 + (\sigma_{T+1}^2 - \sigma_a^2), & k = 1 \\ (k-1) \left[ (1 - \theta_2)^2 + \frac{k(2k-1)+6(1-\theta_2)k}{6} (1 + \theta_1 + \theta_2)^2 \right] \sigma_a^2 \\ \quad + \sum_{j=0}^{k-1} B_j^2 (\delta_1 + \delta_2)^{k-1-j} (\sigma_{T+1}^2 - \sigma_a^2), & k = 2, 3, \dots \end{cases} \tag{40}$$

In this case, the first term is the linear part of the MSFE and the second the nonlinear one, defined as a function of the excess volatility and the prediction horizon,  $k$ . When compared to the MSFE of the smooth trend model, however, the nonlinear term does not vanish when  $k$  increases. Therefore, as in the IMA(1,1)-GARCH case, any shock that leads to an excess volatility different from zero is permanent and thus the MSFE of the heteroscedastic model will differ from that of the homoscedastic one for all prediction horizons. Figure 9 presents a case in which the prediction intervals of simulated series are calculated at a highly volatile period. For the simulations, we set  $\sigma_\varepsilon^2 = 1$ ,  $\alpha_1 = 0.15$ ,  $\alpha_2 = 0.8$ , and  $\sigma_\xi^2 = 0.5$  (i.e  $q_\xi = 0.5$ ).

We see from Figure 9 that again, the IMA-GARCH MSFEs produce too wide prediction intervals. Indeed, they contain almost all observations for each prediction horizon. The empirical coverages have been plotted in Figure 10, where it can be observed that when the prediction horizon is small, IMA-GARCH prediction intervals underestimate the coverage but for horizons larger than 4, these intervals become wide enough to cover almost 100% of the observations. On the other hand, the coverages of the homoscedastic model are smaller than the nominal for short horizons but since the effects of volatility shocks are transitory, they converge to the nominal of 95% for horizons larger than approximately 8.

Consider now the same exercise in a quiet period. The observed coverage is shown in Figure 11. This is the opposite case since the IMA-GARCH prediction intervals underestimate the nominal coverage. In fact, since the reduced form model takes the low variance scenario as permanent, the IMA-GARCH MSFEs are smaller than those of the Smooth Trend-GARCH for all prediction horizons, thus producing too narrow intervals.

Summarizing, in the models where the transitory component,  $\varepsilon_t$ , is the only heteroscedastic component, the shocks to the variance are purely transitory. Consequently, the homoscedastic and the unobserved component-GARCH prediction intervals stick to

each other for large prediction horizons. However, depending on the sign of the excess volatility, the ARIMA-GARCH counterparts may be wider or thinner than the intervals obtained with the corresponding unobserved component model. This is due to its incapacity of distinguishing whether the heteroscedasticity affects the long or the short run components, and it may lead to significant differences between the two prediction intervals, specially for medium and long term. Therefore, the use of reduced form ARIMA models to construct prediction intervals may be inappropriate to capture the underlying uncertainty of the heteroscedastic components.

## 5. An empirical illustration

In this section we fit unobserved component models with GARCH disturbances to real time series with stochastic levels and/or trends. We also find their stationary transformations and fit ARMA-GARCH models to them. In particular, we analyze a series of daily Pound/Euro exchange rate and another of monthly UK inflation rate<sup>8</sup>.

With respect to the inflation rate, the period analyzed spans from July 1961 to August 2007, thus containing  $T = 554$  observations, while the exchange rates have been observed from January 3, 2000 to March 29, 2007 with  $T = 1626$  observations. Figure 12 plots both series together with their corresponding sample autocorrelations of their first differences. The correlogram of the exchange rate returns suggests that the first differenced series can be well represented by an MA(1) model with  $\theta < 0$  and, consequently, the dynamic dependence of the series of exchange rates,  $y_t$ , can be explained by the local level model.

With respect to the correlogram of the differenced inflation, it also suggests an MA(1)

---

<sup>8</sup>More specifically, the exchange rate corresponds to the the daily closing price of the Pound (£) per unit of Euro (€), while the inflation rate is defined as the log-difference of the monthly seasonally adjusted consumer price index (CPI), multiplied by 100 to have percentage rates. Both series have been downloaded from the EcoWin database. An intervention analysis of the series using auxiliary residuals (see Harvey and Koopman, 1992) was carried out with the program STAMP 6.20 of Koopman et al. (2000).

dependence but with more structure in its autocorrelations. Therefore, we have decided to include a cyclical component to capture this structure.

On the other hand, it is possible to observe that the sample autocorrelations of squared observations are in general greater than the squared sample autocorrelations of the original observations. This indicates the presence of a conditionally heteroscedastic structure. Following the procedure of Broto and Ruiz (2009), we first fit the homoscedastic unobserved component model to the series to obtain information about which component is conditionally heteroscedastic by analyzing the auxiliary residuals. We also fit an ARMA model to the corresponding first differences. According to the Akaike and Schwarz information criteria, the best fit for the Pound/Euro exchange rate is obtained with the IMA(1,1) model, and the best fit for the UK inflation rate is obtained with an ARIMA(1,1,1) model<sup>9</sup>. Table 4 reports the results. First note that the signal-to-noise ratios,  $q_\eta$ , for both series are different. For the exchange rate series,  $\hat{q}_\eta = 2.349$ , which implies an MA parameter equal to -0.244 that is rather close to the estimated MA parameter of -0.242. On the other hand, for the inflation rate, the variance of the transitory component is much larger than that of the permanent component, yielding  $\hat{q}_\eta = 0.098$ . In this case, the MA parameter implied by  $q_\eta$  is equal to -0.732, which is smaller than the estimated by the ARMA(1,1) model (-0.871). In order to evaluate the fit in each series, Table 4 also reports the sample mean, skewness (SK), excess kurtosis ( $\bar{\kappa}$ ), and the Ljung-Box Q-statistics of the one-step-ahead residuals of both models,  $\hat{v}_t$  and  $\hat{a}_t$ . In general, it seems that the models selected fit relatively well the conditional mean of the series, although in the UK case, we see that the ARMA(1,1) has a little bit worse fit than the local level plus cycle model. In any case, a positive

---

<sup>9</sup>Within the ARIMA framework, in order to capture the cyclical component of the UK inflation rate, at least an ARMA(2,1) is needed. However, the ARMA(1,1) model not only has a better fit according to the criteria selected, but also is the one chosen by TRAMO (Time Series Regression with ARIMA Noise, Missing Observations and Outliers), which is a program developed by Victor Gómez and Agustín Maravall for automatic modelling, estimation and forecasting of ARIMA models; see Gómez and Maravall (1996) for further details.

excess kurtosis and a significantly large Q-statistic of the squared residuals suggest that the models fail to capture the heteroscedasticity of the series. This leads to propose a GARCH process in all models to account for the dependence of the variance.

In order to decide whether to impose a GARCH process in the transitory and/or permanent components of the unobserved component models, we analyze the difference between the squared autocorrelations of the auxiliary residuals and the sample autocorrelations of the squared auxiliary residuals; see Broto and Ruiz (2009) for further details. With respect to the exchange rate series, we found that both components may be conditionally heteroscedastic<sup>10</sup>, so that the model is specified with both,  $\varepsilon_t$  and  $\eta_t$  being GARCH(1,1) processes. For the UK inflation rate, given that  $q_\eta$  is very small, we decided to include GARCH(1,1) effects only in the transitory component,  $\varepsilon_t$ , so that the permanent component noise,  $\eta_t$ , is homoscedastic.

Table 5 shows the estimation results of the described unobserved component models with GARCH disturbances and the corresponding ARIMA-GARCH models. As a general conclusion, we see that these results are in concordance with the analytical findings of previous sections. First, after comparing the estimates of the ARCH parameters,  $\alpha_1$ ,  $\gamma_1$ , of the local level model with  $\delta_1$  of the reduced form IMA model, we see that  $\delta_1$  is clearly smaller. This result is more evident in the UK inflation rate, where  $\hat{\alpha}_1 \approx 2\hat{\delta}_1$ . Second, we observe that in both series  $\hat{\alpha}_1 + \hat{\alpha}_2 \approx \hat{\delta}_1 + \hat{\delta}_2$ , as expected because in one case we have assumed only one heteroscedastic component and in the other case, the transitory and permanent components have the same persistence. Finally, the estimates of the conditional mean and marginal variances do not change significantly after including the ARCH effects, supporting the fact that the introduction of nonlinearities in the form of GARCH process does not affect the fit of the conditional mean. With respect to the

---

<sup>10</sup>Results of this exploratory analysis are available upon request.

---

diagnostics, observe that in both series of residuals the excess kurtosis is smaller than in the corresponding homoscedastic models. Furthermore, the autocorrelations of squares are not any longer significant.

Finally, conditional on the estimated models for each series, we find the prediction intervals obtained by using the theoretical mean squared forecast errors. They are plotted in Figure 13. For the exchange rate series, the time point selected corresponds to a period, in June of 2000<sup>11</sup>. Note that the LL-GARCH prediction intervals have the same width as those of the IMA-GARCH for the one-step-ahead predictions. However, the prediction intervals of the LL-GARCH are wider for larger prediction horizons. Finally, it is worth noting that the homoscedastic model is not able to incorporate the information about the volatility shock and therefore it produces prediction intervals that are really worse in terms of coverage.

With respect to the UK inflation rate, the period considered to make predictions is December, 2001. Compared to the high volatility period of the 70's and the beginning of the 80's, the time point selected is within a quiet period. Consequently, the excess volatility is negative and the theoretical MSFEs of the heteroscedastic models are smaller than those of the homoscedastic model. Thus, this model produces too wide prediction intervals. On the other hand, the ARIMA-GARCH model assumes that this period of low volatility is permanent, so that the resulting prediction intervals are always narrower than those of the LL-GARCH. However, as the volatility shocks are assumed to be transitory by the LL-GARCH model, the prediction intervals are narrow at first but then become wider to follow the path of the homoscedastic prediction intervals.

---

<sup>11</sup>For this series, we have re-estimated the models with a sample of the same size but taking into account data prior to 2000. The results have not changed significantly compared to those reported previously.

## 6. Conclusions

In this paper we analyze the effects of taking differences and fit ARMA models to conditionally heteroscedastic series with stochastic trends. In particular, we consider the local level and smooth trend models with GARCH(1,1) disturbances. We show that although working with the ARIMA reduced form model is simpler because there is only one disturbance, working with the unobserved component model may lead to discover conditionally heteroscedastic structures that could not be apparent in the reduced form noise.

More interestingly, we also show that working with the unobserved component models generates more accurate prediction intervals when the heteroscedasticity only affects the transitory component. In this case, the effects of the heteroscedasticity disappear as the prediction horizon increases. Therefore, the prediction intervals produced by the unobserved component models converge to the homoscedastic intervals. However, the presence of a unit root in the reduced form model, lead to generate prediction intervals that never converge to the homoscedastic intervals.

Finally, the empirical application with the Pound-Euro exchange rate and the UK inflation rate illustrates our main findings. Namely, a weaker heteroscedasticity of the reduced form noise compared to that of the unobserved components, and a differential behavior of the IMA-GARCH and LL-GARCH prediction intervals depending on the source of the volatility shocks.

## References

Bollerslev, T., R. Y. Chou, and K. F. Kroner (1992). ARCH modelling in finance: A selective review of the theory and empirical evidence. *Journal of Econometrics* 52, 5–59.

- 
- Bollerslev, T., R. F. Engle, and D. B. Nelson (1994). ARCH models. *The Handbook of Econometrics* 2, 2959–3038.
- Breidt, F. J. and R. A. Davis (1992). Time-reversibility, identifiability and independence of innovations for stationary time series. *Journal of Time Series Analysis* 13, 377–390.
- Broto, C. and E. Ruiz (2006). Unobserved component models with asymmetric conditional variances. *Computational Statistics and Data Analysis* 50, 2146–2166.
- Broto, C. and E. Ruiz (2009). Testing for conditional heteroscedasticity in the components of inflation. *Studies in Nonlinear Dynamics and Econometrics*, forthcoming.
- Carnero, M. A., D. Peña, and E. Ruiz (2006). Effects of outliers on the identification and estimation of GARCH models. *Journal of Time Series Analysis*, 28, 471–497.
- Chadha, J. S. and L. Sarno (2002). Short- and long-run price level uncertainty on investment under different monetary policy regimes: an international comparison. *Oxford Bulletin of Economics and Statistics* 64, 183–212.
- Chang, K. H. and M. J. Kim (2004). Jumps and time-varying correlations in daily foreign exchange rates. *Journal of International Money and Finance* 119, 257–289.
- Diebold, F. X. (2004). *Measuring and Forecasting Financial Market Volatilities and Correlations*. New York: Norton.
- Diebold, F. X. and J. Lopez (1995). *Modelling Volatility Dynamics*, pp. 427–472. Boston: Kluwer Academic Press.
- Durbin, J. and S. J. Koopman (2001). *Time Series Analysis by State Space Methods*. Oxford, UK: Oxford University Press.

- Gómez, V. and A. Maravall (1996). Programs TRAMO (Time series Regression with Arima noise, Missing observations, and Outliers) and SEATS (Signal Extraction in Arima Time Series). Instructions for the User. Working Paper 9628, Banco de España.
- Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge: Cambridge University Press.
- Harvey, A. C. and A. Jaeger (1993). Detrending, stylized facts and the business cycle. *Journal of Applied Econometrics* 8, 231–247.
- Harvey, A. C. and S. Koopman (1992). Diagnostic checking of unobserved-components time series models. *Journal of Business and Economic Statistics* 10, 377–389.
- Harvey, A. C., E. Ruiz, and E. Sentana (1992). Unobserved component time series models with ARCH disturbances. *Journal of Econometrics* 52, 129–157.
- King, M. A., E. Sentana, and E. Wadhvani (1994). Volatility and links between national stock markets. *Econometrica* 62, 901–933.
- Koopman, S. J., A. C. Harvey, J. A. Doornik, and N. Shephard (2000). *STAMP: Structural Time Series Analyser, Modeller and Predictor*. London: Timberlake Consultants Press.
- Maravall, A. (1983). An application of nonlinear time series forecasting. *Journal of Business and Economic Statistics* 1, 66–74.
- Moore, B. and L. Schaller (2002). Persistent and transitory shocks, learning and investment dynamics. *Journal of Money, Credit and Banking* 34, 650–677.
- Nyblom, J. and A. Harvey (2001). Testing against smooth stochastic trends. *Journal of Applied Econometrics* 3, 415–429.

- Pascual, L., J. Romo, and E. Ruiz (2006). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics and Data Analysis*, 50, 2293–2312.
- Rodriguez, J. and E. Ruiz (2005). A powerful test for conditional heteroscedasticity for financial time series with highly persistent volatilities. *Statistica Sinica* 15, 505–525.
- Sentana, E. (2004). Factor representing portfolios in large asset markets. *Journal of Econometrics* 119, 257–289.
- Sentana, E. and G. Fiorentini (2001). Identification, estimation and testing of conditionally heteroscedastic factor models. *Journal of Econometrics* 102, 143–164.
- Stock, J. H. and M. W. Watson (2007). Why has U.S. inflation become harder to forecast? *Journal of Money, Credit and Banking* 39, 3.
- Zaffaroni, P. (2007). Contemporaneous aggregation of GARCH processes. *Journal of Time Series Analysis* 28, 521–544.

## 7. Figures and Tables

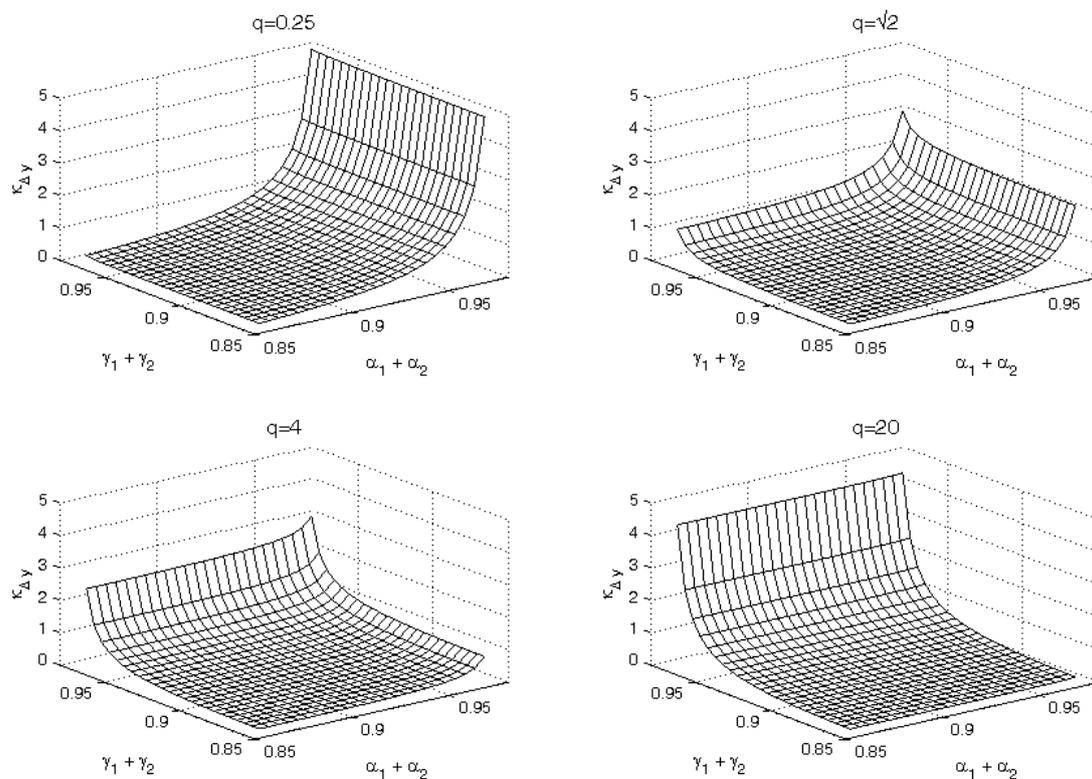


Figure 1: Relationship between  $\bar{\kappa}_{\Delta y}$  and persistence. The GARCH coefficients,  $\alpha_2$  and  $\gamma_2$  are fixed to 0.85.

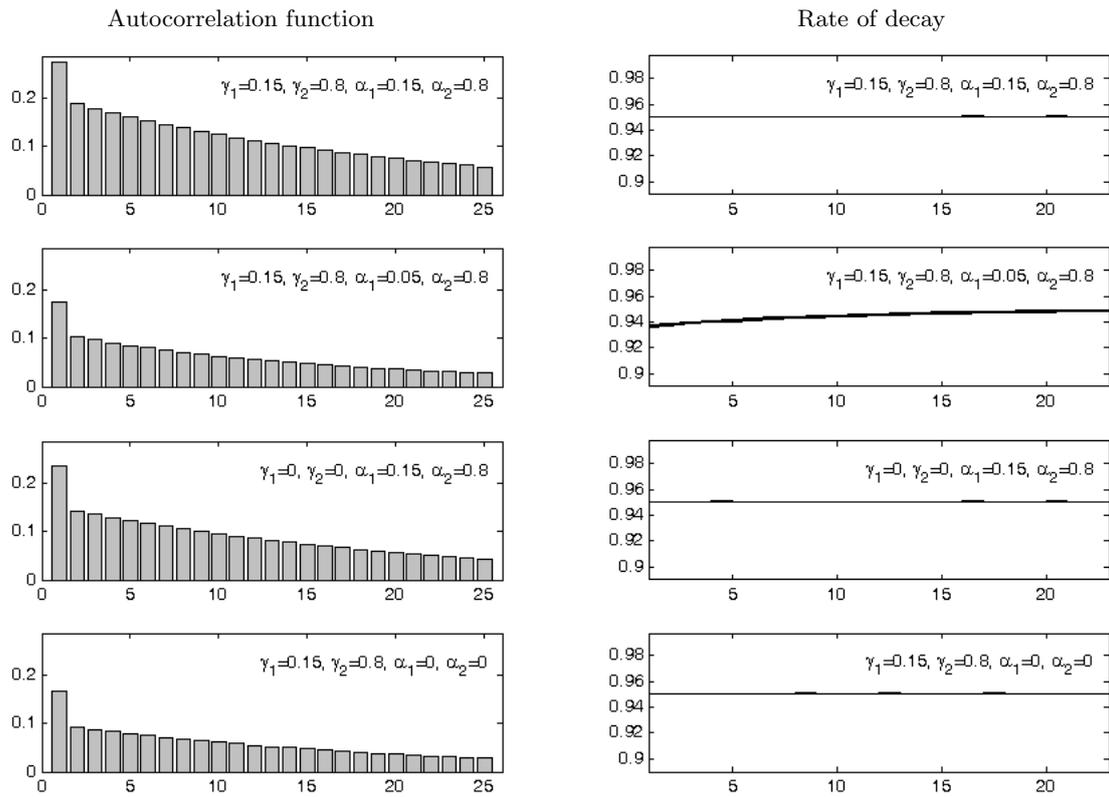


Figure 2: Autocorrelations of  $(\Delta y_t)^2$  for several local level models with conditionally Normal GARCH(1,1) disturbances and  $q_\eta = \sqrt{2}$ . The rate of decay reported in the right column is defined as the ratio  $\rho_\tau^{(\Delta y)^2} / \rho_{\tau-1}^{(\Delta y)^2}$ .

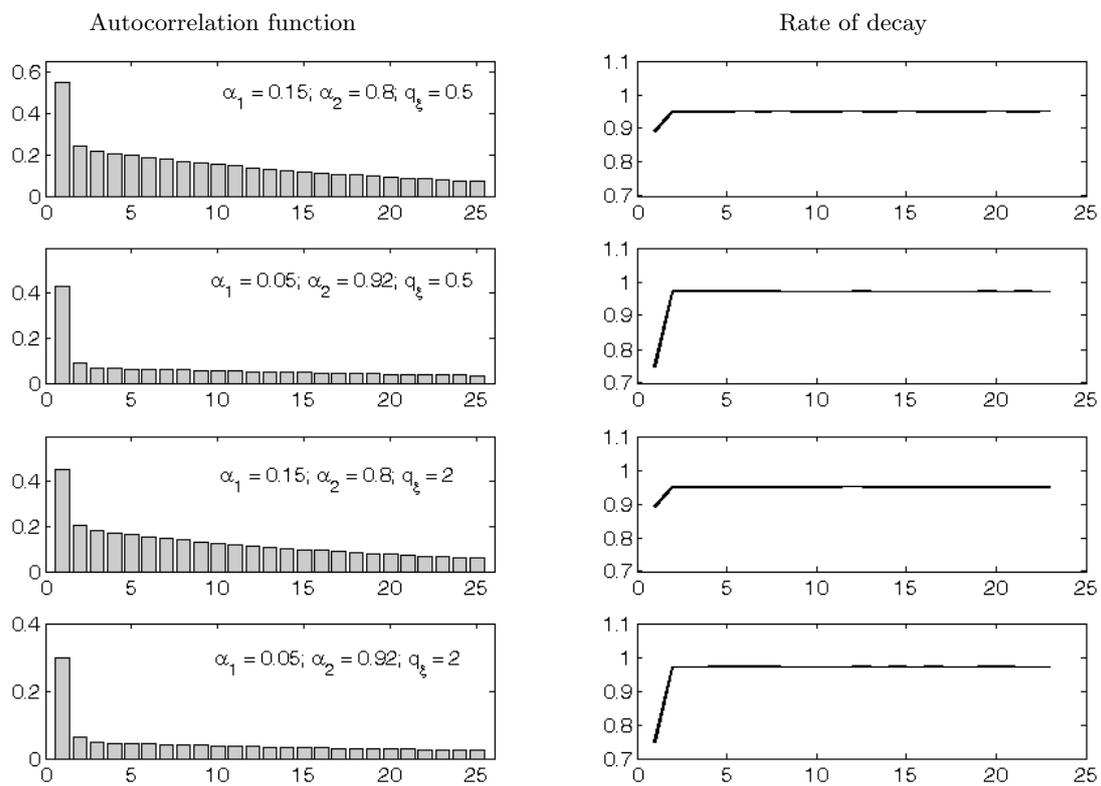


Figure 3: Autocorrelations of  $(\Delta^2 y_t)^2$  for several smooth trend models with conditionally Normal GARCH(1,1) disturbances. The rate of decay reported in the right column is defined as the ratio  $\rho_\tau^{(\Delta y)^2} / \rho_{\tau-1}^{(\Delta y)^2}$ .

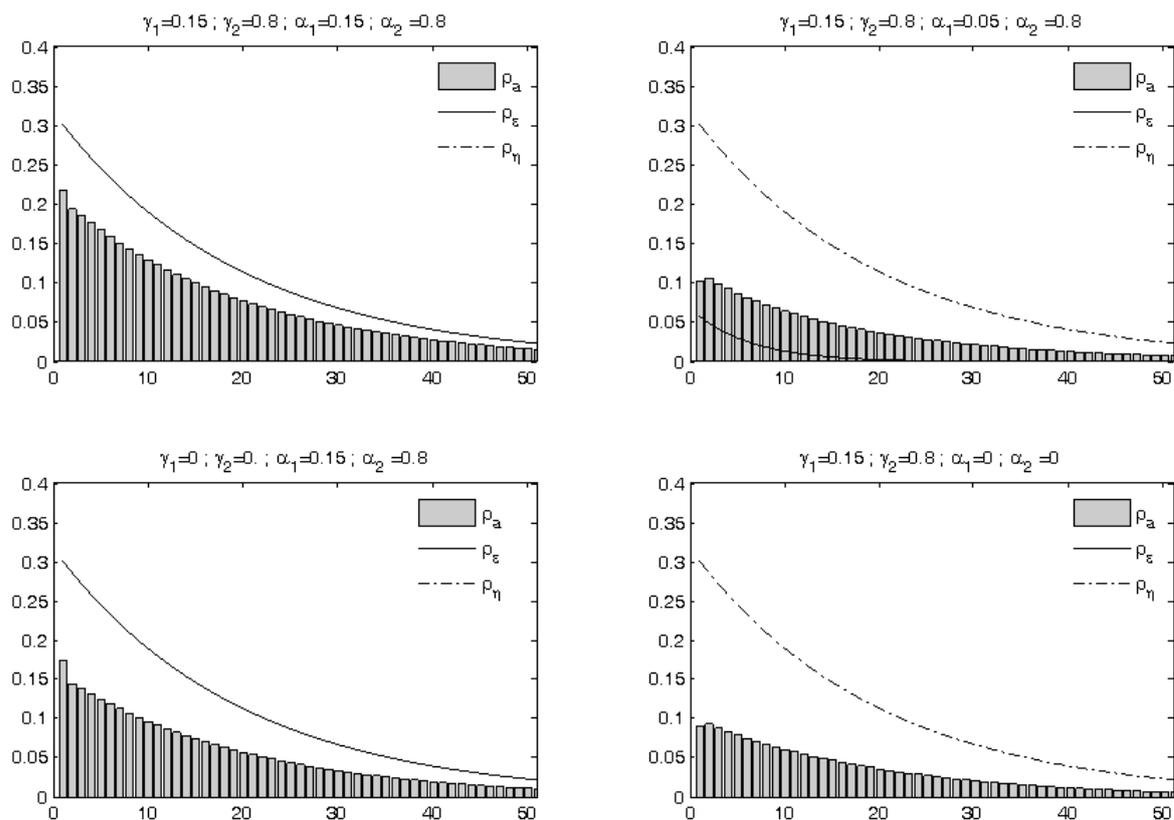


Figure 4: Autocorrelations of  $a_t^2$ , resulting from different conditionally Normal GARCH(1,1) disturbances with  $q_\eta = \sqrt{2}$ . The solid and dash-dotted lines draw the autocorrelations of  $\varepsilon_t^2$  and  $\eta_t^2$ , respectively.

q	$\bar{\kappa}_\varepsilon$	$\bar{\kappa}_\eta$	$\theta$	$\bar{\kappa}_{\Delta y}$	$\rho_1^{(\Delta y)^2}$	$\bar{\kappa}_a$	$\rho_1^{a^2}$	$\rho_2^{a^2}$	$\rho_3^{a^2}$	$\rho_4^{a^2}$	$\rho_5^{a^2}$
0.5	0	3	-0.5	0.120	0.151	0.273	-0.030	0.008	-0.002	0.001	0.000
$\sqrt{2}$	0	3	-0.324	0.515	0.068	0.665	-0.026	0.003	0.000	0.000	0.000
0.5	3	3	-0.5	1.080	0.260	0.818	0.194	-0.048	0.012	-0.003	0.001
$\sqrt{2}$	3	3	-0.324	1.029	0.142	1.120	0.063	-0.007	0.001	0.000	0.000
0.5	3	0	-0.5	0.960	0.270	0.546	0.241	-0.060	0.015	0.004	0.001
$\sqrt{2}$	3	0	-0.324	0.515	0.171	0.456	0.109	-0.011	0.001	0.000	0.000

Table 1: Theoretical moments of  $a_t$  resulting from local level models with either or both non-Gaussian homoscedastic noises.

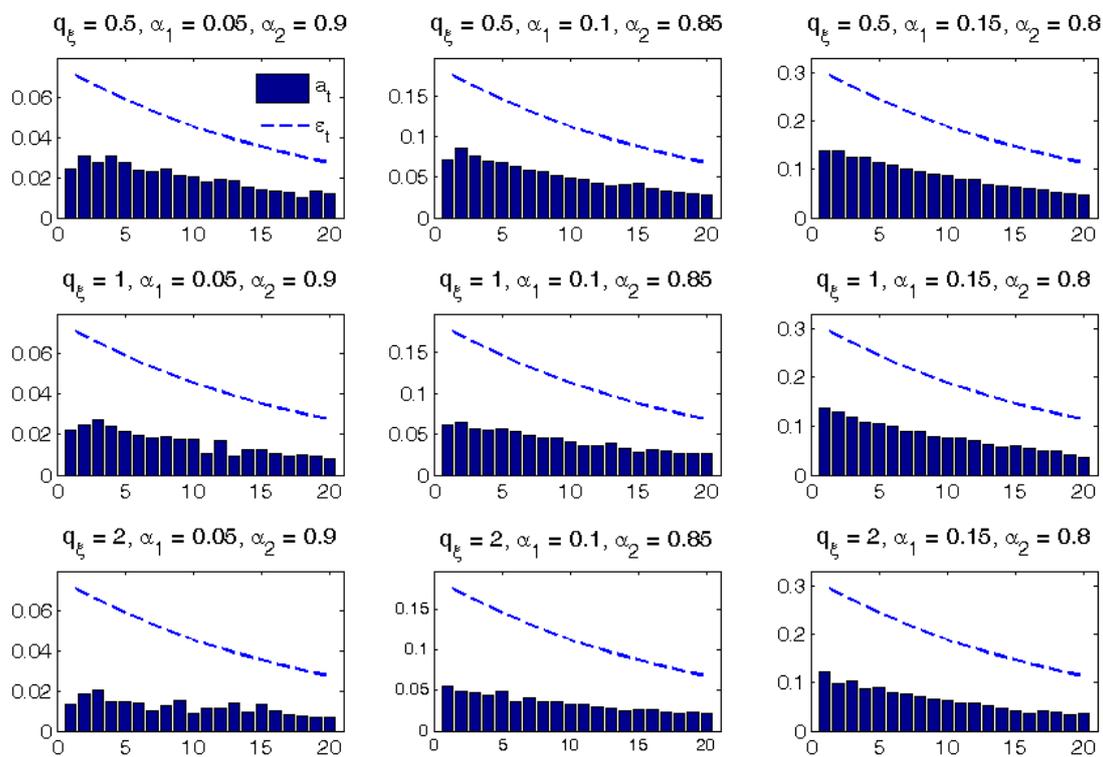


Figure 5: Simulated autocorrelations of  $a_t^2$  compared to the theoretical acf of  $\varepsilon_t$ , for different values of  $q_\xi$  and GARCH parameters,  $\alpha_1$  and  $\alpha_2$ .

LL-GARCH(1,1)	Estimated IMA(1,1)-GARCH(1,1)				
Parameters	Estimates	$T = 200$	$T = 500$	$T = 1000$	$T = 5000$
$q = 1$					
$\theta = -0.382$	$\hat{\theta}$	-0.425 (0.099)	-0.398 (0.067)	-0.391 (0.050)	-0.384 (0.023)
	$Q_1(10)$	23%	47%	76%	100%
$\gamma_1 = 0.15 ; \gamma_2 = 0.8$	$\hat{\delta}_1$	0.055 (0.058)	0.053 (0.039)	0.050 (0.025)	0.048 (0.011)
$\alpha_1 = 0 ; \alpha_2 = 0$	$\hat{\delta}_2$	0.668 (0.300)	0.748 (0.243)	0.821 (0.175)	0.880 (0.034)
$\kappa_\eta = 5.57$	$\hat{\kappa}_a$	3.137 (0.434)	3.277 (0.562)	3.367 (0.609)	3.466 (0.640)
$\rho_1^{\eta^2} = 0.3$	$\hat{\rho}_1^{a^2}$	0.031 (0.082)	0.057 (0.071)	0.072 (0.068)	0.094 (0.054)
$\rho_2^{\eta^2} = 0.285$	$\hat{\rho}_2^{a^2}$	0.033 (0.084)	0.051 (0.068)	0.064 (0.063)	0.081 (0.047)
ARCH effects (at 5%)		9%	39%	73%	100%

Table 2: Monte Carlo averages and standard deviations (in parenthesis) of the QML estimates of the IMA-GARCH parameters when the series are generated by a local level model with GARCH noises.

Smooth Trend - GARCH(1,1)	Estimated IMA(2,2)-GARCH(1,1)				
Parameters	Estimates	$T = 200$	$T = 500$	$T = 1000$	$T = 5000$
$q_\xi = 1$					
$\theta_1 = -0.750$	$\hat{\theta}_1$	-0.733 (0.093)	-0.729 (0.056)	-0.729 (0.041)	-0.727 (0.019)
$\theta_2 = 0.231$	$\hat{\theta}_2$	0.218 (0.080)	0.218 (0.048)	0.223 (0.034)	0.221 (0.016)
	$Q_1(10)$	57%	79%	94%	100%
$\alpha_1 = 0.1$	$\hat{\delta}_1$	0.054 (0.057)	0.048 (0.032)	0.048 (0.022)	0.045 (0.009)
$\alpha_2 = 0.85$	$\hat{\delta}_2$	0.719 (0.273)	0.813 (0.185)	0.853 (0.127)	0.900 (0.023)
$\kappa_\varepsilon = 3.774$	$\hat{\kappa}_a$	3.161 (0.471)	3.291 (0.448)	3.35 (0.424)	3.445 (0.234)
$\rho_1^{\varepsilon^2} = 0.179$	$\hat{\rho}_1^{a^2}$	0.032 (0.084)	0.051 (0.070)	0.063 (0.057)	0.069 (0.031)
$\rho_2^{\varepsilon^2} = 0.170$	$\hat{\rho}_2^{a^2}$	0.033 (0.085)	0.054 (0.067)	0.064 (0.054)	0.075 (0.033)
ARCH effects (at 5%):		17%	39%	64%	100%

Table 3: Monte Carlo averages and standard deviations (in parenthesis) of the QML estimates of the IMA-GARCH parameters when the series are generated by the smooth trend model with a GARCH transitory component.

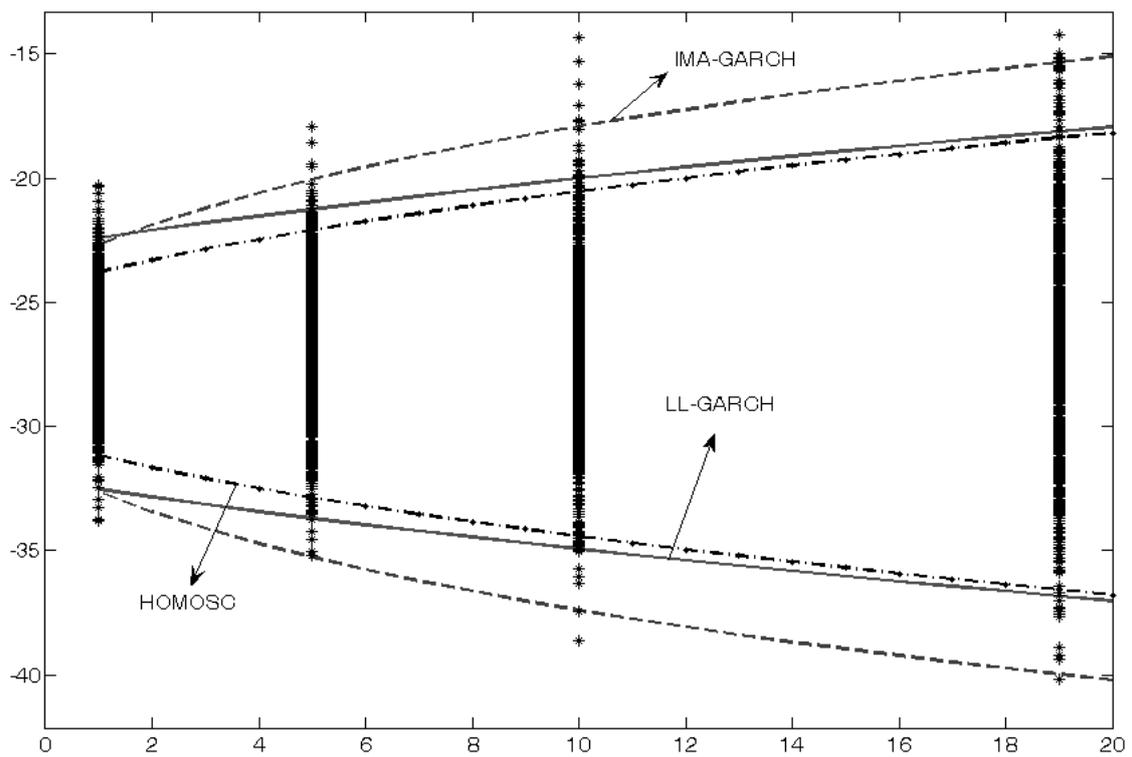


Figure 6: Prediction intervals of a LL-GARCH series where only the transitory component is heteroscedastic, with  $\alpha_1 = 0.15$ ,  $\alpha_2 = 0.8$ , and  $q_\eta = 1$ . The time point is selected in a highly volatile period.

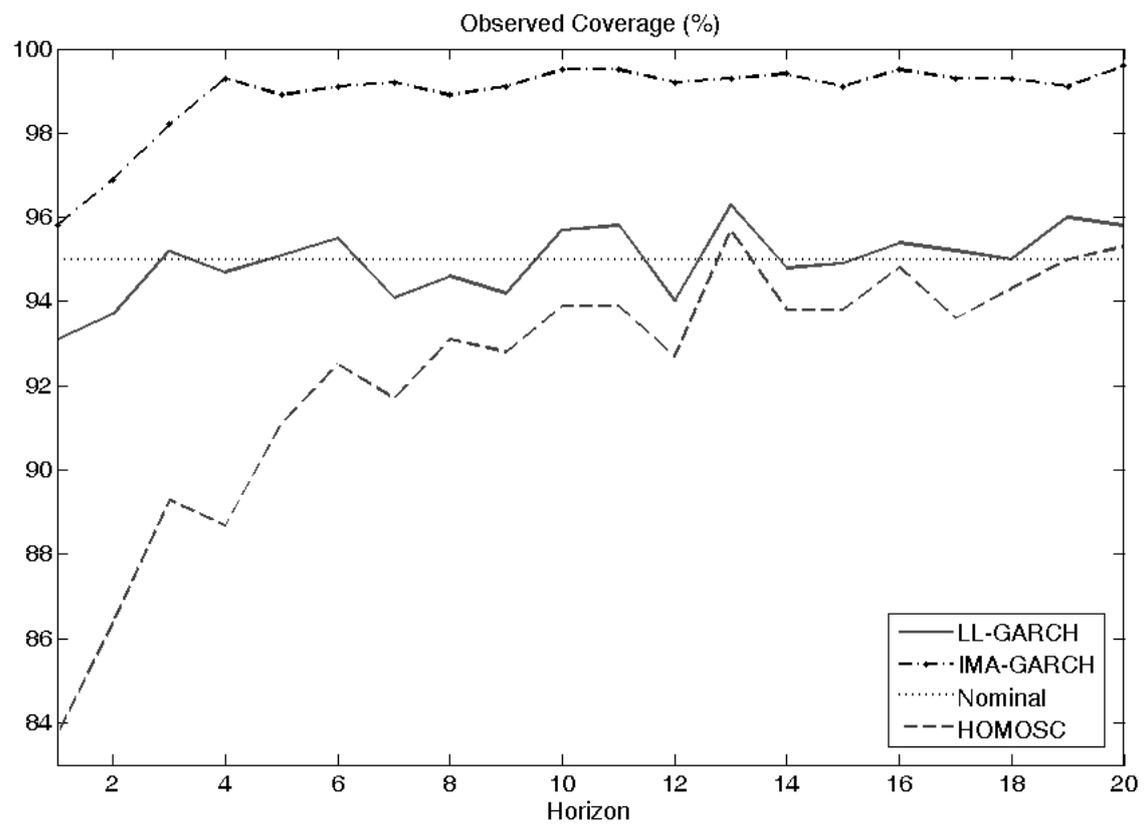


Figure 7: Observed coverage measured as the percentage of trajectories within the prediction intervals.

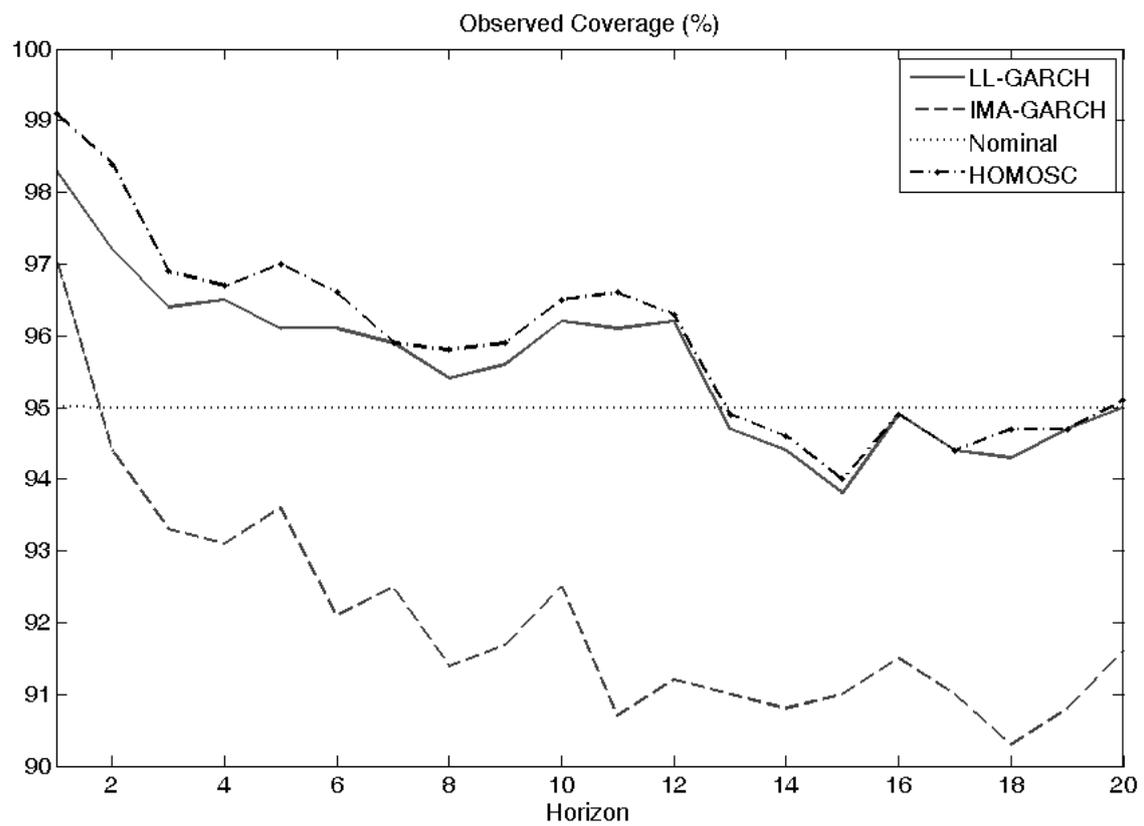


Figure 8: Observed coverage measured as the percentage of trajectories within the prediction intervals.

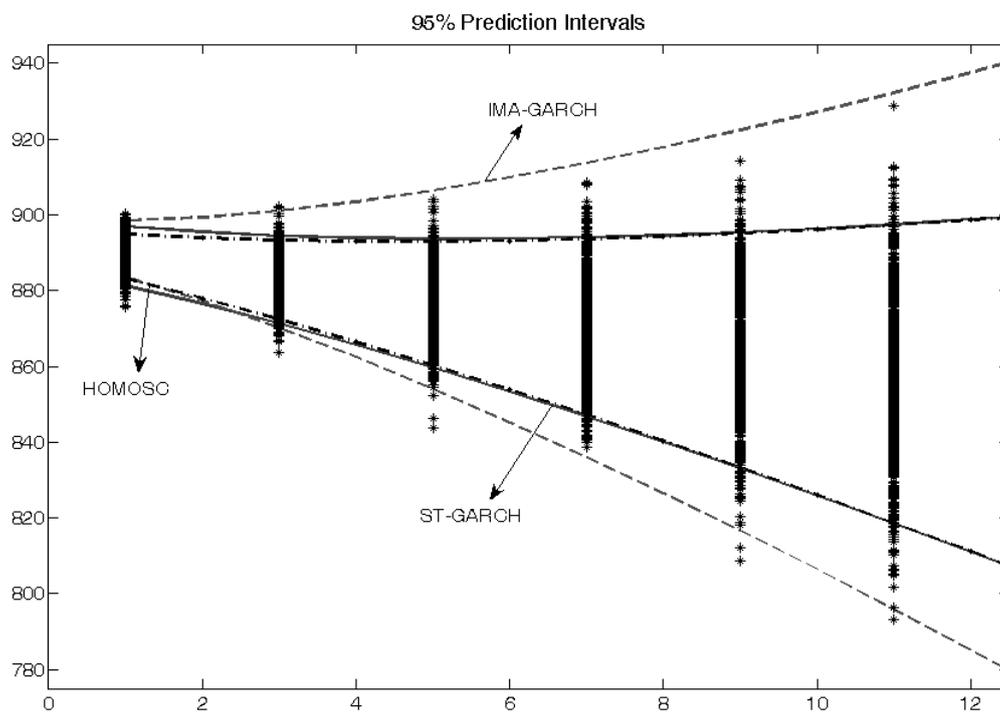


Figure 9: Prediction intervals of an ST-GARCH series, with  $\alpha_1 = 0.15$ ,  $\alpha_2 = 0.8$ , and  $q_\xi = 0.5$ . The time point is selected in a highly volatile period.

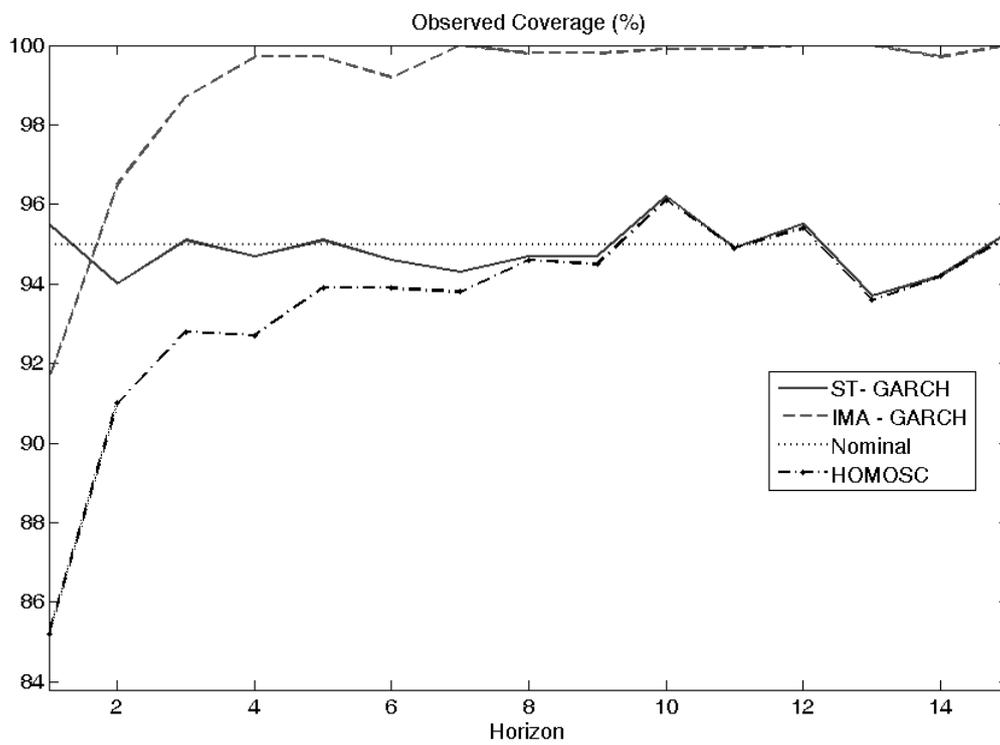


Figure 10: Observed coverage measured as the percentage of trajectories within the prediction intervals.

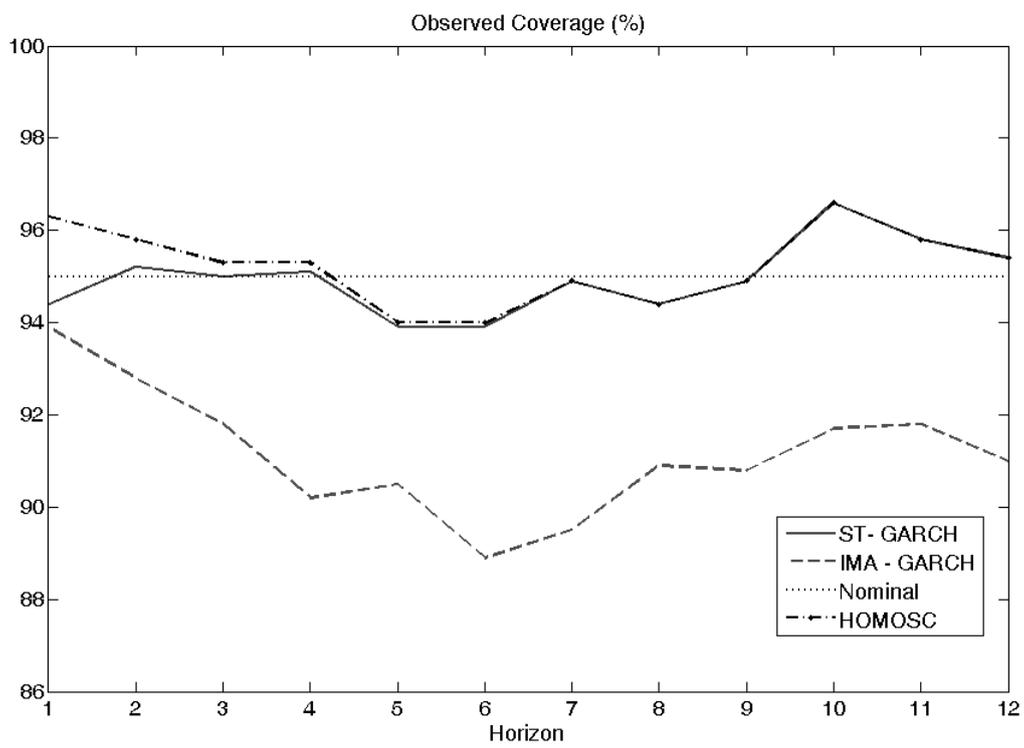


Figure 11: Observed coverage measured as the percentage of trajectories within the prediction intervals.

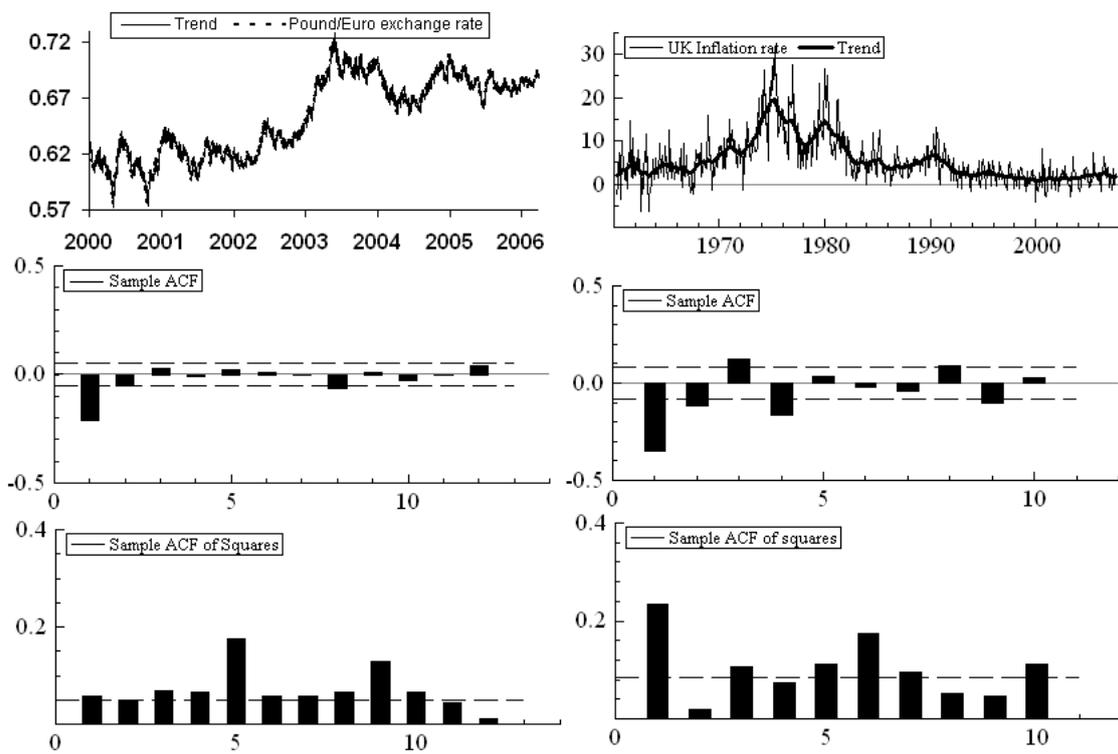


Figure 12: UK inflation rate and the Pound/Euro (£/€) exchange rate: Evolution against time and sample autocorrelations of their first differences.

Pound/Euro exchange rate				UK inflation rate			
Local level		IMA(1,1)		Local level + cycle		ARIMA(1,1,1)	
$\hat{\sigma}_\varepsilon^2 =$	0.083**	$\hat{\sigma}_a^2 =$	0.344**	$\hat{\sigma}_\varepsilon^2 =$	0.037**	$\hat{\sigma}_a^2 =$	0.08**
$\hat{\sigma}_\eta^2 =$	0.195**	$\hat{\theta} =$	-0.242**	$\hat{\sigma}_\eta^2 =$	0.004**	$\hat{\theta} =$	-0.871**
				$\hat{\sigma}_\kappa^2 =$	0.015**	$\hat{\phi} =$	0.357**
	$\hat{v}_t$		$\hat{a}_t$		$\hat{v}_t$		$\hat{a}_t$
Mean	0.009		0.007	Mean	-0.034		-0.001
SK	0.209		0.209	SK	0.571		0.558
$\bar{\kappa}$	0.997		0.994	$\bar{\kappa}$	1.964		1.834
$Q(10)$	15.44		15.42	$Q(10)$	15.30		18.95*
$Q_2(10)$	109.8**		111.1**	$Q_2(10)$	69.44**		82.05**

Table 4: Estimates and sample moments of the residuals of the homoscedastic unobserved component and ARIMA models fitted to the Pound/Euro exchange rate and UK inflation rate.  $\hat{\sigma}_\kappa^2$  is the estimate of the cycle component variance and  $\hat{\phi}$  is the estimate of the AR parameter,  $\hat{v}_t$  and  $\hat{a}_t$  are the estimates of the one-step-ahead innovations, SK is the skewness,  $\bar{\kappa}$  is the excess kurtosis and  $Q(10)$  and  $Q_2(10)$  are the classic lag-10 Ljung-Box statistics applied to the level and squared residuals, respectively. (\*\*) Significant at 5% (1%) level.

Pound/Euro exchange rate				UK inflation rate			
LL-GARCH		IMA(1,1)-GARCH		LLCycle-GARCH		ARIMA(1,1,1)-GARCH	
$\hat{\alpha}_0 =$	1.00E-04**	$\hat{\delta}_0 =$	6.00E-04**	$\hat{\alpha}_0 =$	1.05E-04**	$\hat{\delta}_0 =$	4.44E-04**
$\hat{\alpha}_1 =$	0.069**	$\hat{\delta}_1 =$	0.026**	$\hat{\alpha}_1 =$	0.079**	$\hat{\delta}_1 =$	0.041**
$\hat{\alpha}_2 =$	0.930**	$\hat{\delta}_2 =$	0.971**	$\hat{\alpha}_2 =$	0.919**	$\hat{\delta}_2 =$	0.951**
$\hat{\gamma}_0 =$	2.00E-04**	$\hat{\theta} =$	-0.244**	$\hat{\sigma}_\eta^2 =$	0.003**	$\hat{\theta} =$	-0.897**
$\hat{\gamma}_1 =$	0.034**			$\hat{\sigma}_\kappa^2 =$	0.011**	$\hat{\phi} =$	0.344**
$\hat{\gamma}_2 =$	0.965**						
	$\hat{v}_t^\dagger$		$\hat{a}_t^\dagger$		$\hat{v}_t^\dagger$		$\hat{a}_t^\dagger$
Mean	0.018		0.019	Mean	-0.013		-0.011
SK	0.159		0.129	SK	0.553		0.618
$\bar{\kappa}$	0.412		0.333	$\bar{\kappa}$	1.106		1.217
$Q(10)$	11.37		10.93	$Q(10)$	8.133		9.548
$Q_2(10)$	11.88		12.10	$Q_2(10)$	14.03		15.20

Table 5: Estimates and sample moments of the residuals of the unobserved component and ARIMA models with GARCH noises fitted to the Pound/Euro exchange rate and UK inflation rate.  $\hat{v}_t^\dagger$  and  $\hat{a}_t^\dagger$  are the estimates of the one-step-ahead innovations, divided by their conditional standard deviations. See Table 4 for further explanation.

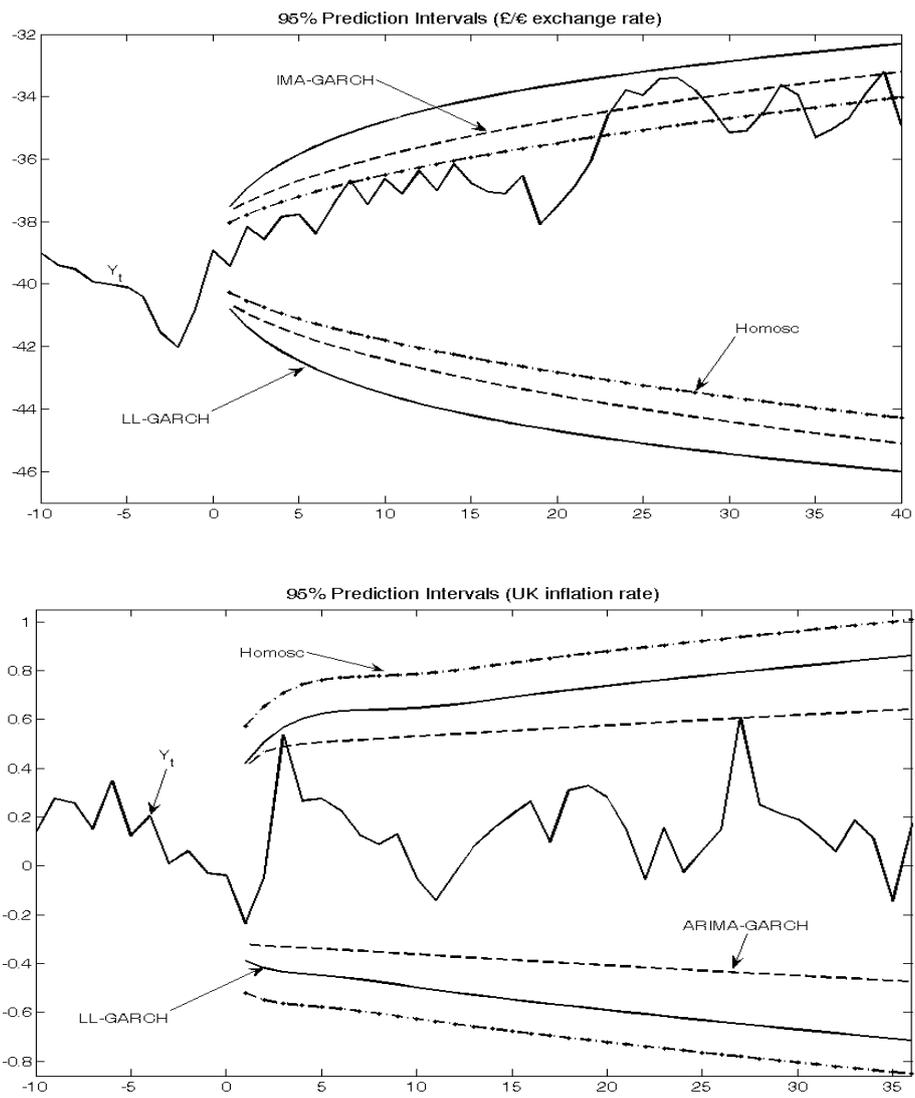


Figure 13: Prediction intervals of the Pound/Euro exchange rate and UK inflation rate.