THE RELATIONSHIP BETWEEN ARIMA-GARCH AND UNOBSERVED COMPONENT MODELS WITH GARCH DISTURBANCES[*]

Santiago Pellegrini[1], Esther Ruiz[2] and Antoni Espasa[3]

Abstract

The objective of this paper is to analyze the consequences of fitting ARIMA-GARCH models to series generated by conditionally heteroscedastic unobserved component models. Focusing on the local level model, we show that the heteroscedasticity is weaker in the ARIMA than in the local level disturbances. In certain cases, the IMA(1,1) model could even be wrongly seen as homoscedastic. Next, with regard to forecasting performance, we show that the prediction intervals based on the ARIMA model can be inappropriate as they incorporate the unit root while the intervals of the local level model can converge to the homoscedastic intervals when the heteroscedasticity appears only in the transitory noise. All the analytical results are illustrated with simulated and real time series.

---

[1] Departamento de Estadística, Universidad Carlos III de Madrid
[2] Departamento de Estadística, Universidad Carlos III de Madrid
[3] Departamento de Estadística, Universidad Carlos III de Madrid

The relationship between ARIMA-GARCH and unobserved component models with GARCH disturbances[*]

Santiago Pellegrini[†], Esther Ruiz[‡] and Antoni Espasa[§]

April 2007

Abstract

The objective of this paper is to analyze the consequences of fitting ARIMA-GARCH models to series generated by conditionally heteroscedastic unobserved component models. Focusing on the local level model, we show that the heteroscedasticity is weaker in the ARIMA than in the local level disturbances. In certain cases, the IMA(1,1) model could even be wrongly seen as homoscedastic. Next, with regard to forecasting performance, we show that the prediction intervals based on the ARIMA model can be inappropriate as they incorporate the unit root while the intervals of the local level model can converge to the homoscedastic intervals when the heteroscedasticity appears only in the transitory noise. All the analytical results are illustrated with simulated and real time series.

Keywords: State Space Models, Conditional Heteroscedasticity, Prediction Intervals.

1. Introduction

ARIMA and unobserved component models, also called structural models, are alternative specifications to represent the dynamic properties of series with stochastic components, such as trends and seasonals. It is well known that when the disturbances are i.i.d. and Gaussian, the reduced

[†]spellegr@est-econ.uc3m.es

[‡]ortega@est-econ.uc3m.es

[§]espasa@est-econ.uc3m.es

form of an unobserved component model is an ARIMA model with restrictions in the parameters; see, for example, Harvey (1989). The specification of ARIMA models includes one aggregated disturbance while unobserved component models incorporate component disturbances. Thus, provided that the structural formulation is correct, working with the unobserved components may lead to the discovery of 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 focus on the presence of conditional heteroscedasticity in the form of GARCH processes.

From an empirical point of view, the presence of conditional heteroscedasticity in both ARIMA and unobserved component models, have previously interested many authors. The literature that considers ARIMA models with GARCH disturbances is very extensive; 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, and Chadha and Sarno (2002) and Moore and Schaller (2002) for applications to price volatility and term structure of interest rates, respectively. More recently, Stock and Watson (2007) find that a simple unobserved component model with conditionally heteroscedastic noises well describe the dynamics of inflation.

Little is known about the properties of the reduced form ARIMA model when the unobserved disturbances are conditionally heteroscedastic. Our first objective is to analyze them. For simplicity, we focus on the Local Level (LL) model that assumes that the series of interest, y_t, is composed by an underlying stochastic level, μ_t, and a transitory component, ε_t, i.e.

$$y_t = \mu_t + \varepsilon_t, \tag{1}$$

$$\mu_t = \mu_{t-1} + \eta_t, \tag{2}$$

where ε_t and η_t are mutually independent and serially uncorrelated processes, with zero means and variances σ_ε^2 and σ_η^2, respectively. Taking first differences in (1) results in a stationary series given by

$$\Delta y_t = \eta_t + \Delta\varepsilon_t. \tag{3}$$

It is well known that the corresponding reduced form model is an ARIMA(0,1,1) model given by

$$\Delta y_t = a_t + \theta a_{t-1}, \tag{4}$$

where $\theta = \left[(q^2 + 4q)^{1/2} - 2 - q\right]/2$ and $q = \sigma_\eta^2/\sigma_\varepsilon^2$ is the signal-to-noise ratio. The parameter θ is restricted to be negative, i.e. $-1 < \theta < 0$. We derive general expressions of the kurtosis and autocorrelations of squares of Δy_t defined as in expression (3) and in the single-disturbance model (4). The comparison of both expressions, allows us to derive the properties of the reduced form disturbance, a_t, in terms of the unobserved disturbances, ε_t and η_t. We show that if ε_t and η_t are assumed to be GARCH processes, the conditional heteroscedasticity of 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 model.

Furthermore, one important objective when analyzing time series is to obtain prediction intervals of future values of the series. When the series are heteroscedastic, the amplitude of the intervals changes depending on whether the conditional variance at the moment of making the prediction is larger or smaller than the marginal variance. Our second objective is to analyze how the presence of conditional heteroscedasticity in the unobserved disturbances affect the prediction intervals. Denote 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. In this case, the prediction intervals obtained with the unobserved component model converge to the intervals of the corresponding homoscedastic model. On the other hand,

if the long-run component is heteroscedastic, then the intervals always depend on the excess volatility. However, due to the presence of the unit root in the model, the prediction intervals based on the ARIMA model for any prediction horizon always depend on the excess volatility. This could have important consequences when building prediction intervals in series in which only the transitory component is heteroscedastic. 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 statistical properties of the LL model when the disturbances are uncorrelated processes with symmetric distributions and finite fourth order moments. We focus on two particular cases of interest, namely, non-normal distributions and conditionally heteroscedastic disturbances. In Section 3, we derive the statistical properties of the reduced form ARIMA noise, a_t. We also analyze the performance of the ARIMA-GARCH model when fitted to represent the dynamic properties of the LL-GARCH model. We finish this section by illustrating the results with Monte Carlo experiments. Section 4 analyzes the forecasting performance of both models. Section 5 contains an empirical application which illustrates the results of previous sections. Finally, Section 6 concludes the paper.

2. Properties of the Local Level model

In this section we derive the statistical properties of Δy_t defined as in (3) 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 unobserved component models with non-Normal and/or conditionally heteroscedastic components, we focus on two moments often used to characterized these two features. In particular, we derive the kurtosis and the acf of squared observations. First, note that symmetric distributions of ε_t and η_t lead to all odd

moments of Δy_t equal to zero. On the other hand, the variance of Δy_t is given by

$$Var[\Delta y_t] = E[(\Delta y_t)^2] = \sigma_\varepsilon^2 (q + 2). \tag{5}$$

The fourth order unconditional moment can be derived as follows:

$$
\begin{aligned}
E[(\Delta y_t)^4] &= E[\eta_t^4] + 2E[\varepsilon_t^4] + 12E[\eta_t^2]E[\varepsilon_t^2] + 6E[\varepsilon_t^2 \varepsilon_{t-1}^2] \\
&= \sigma_\varepsilon^4 \left[q^2 \kappa_\eta + 12q + 2\kappa_\varepsilon + 6 \left(\rho_1^{\varepsilon^2} (\kappa_\varepsilon - 1) + 1 \right) \right],
\end{aligned} \tag{6}
$$

where κ_ε and κ_η are the kurtosis of ε_t and η_t, respectively, and $\rho_1^{\varepsilon^2}$ is the lag-one autocorrelation of ε_t^2. From (5) and (6) it is possible to obtain the following expression of the kurtosis of Δy_t:

$$\kappa_{\Delta y} = \frac{1}{(q+2)^2} \left[q^2 \kappa_\eta + 12q + 2\kappa_\varepsilon + 6 \left(\rho_1^{\varepsilon^2} (\kappa_\varepsilon - 1) + 1 \right) \right]. \tag{7}$$

Note that the signal-to-noise ratio, q, plays an important role in determining the relative influence of each noise on the kurtosis of the stationary transformation of y_t. In the limiting cases, when $q \to \infty$, $\Delta y_t = \eta_t$ (i.e. y_t is a pure random walk process) so that $\kappa_{\Delta y} = \kappa_\eta$. On the other hand, as $q \to 0$, $\Delta y_t = \Delta \varepsilon_t$ (i.e. y_t is a white noise process), and consequently Δy_t is a non-invertible MA(1) process whose kurtosis may be different from κ_ε depending on the value of $\rho_1^{\varepsilon^2}$.

The acf of squares has often been used to represent the dynamic dependence of heteroscedastic series. If y_t is generated by model (1), then the acf of $(\Delta y_t)^2$ is given by

$$\rho_\tau^{(\Delta y)^2} = \frac{q^2 \rho_\tau^{\eta^2} (\kappa_\eta - 1) + (\kappa_\varepsilon - 1)(\rho_{\tau-1}^{\varepsilon^2} + 2\rho_\tau^{\varepsilon^2} + \rho_{\tau+1}^{\varepsilon^2})}{q^2 (\kappa_\eta - 1) + 8q + 2(\kappa_\varepsilon - 1)(1 + 3\rho_1^{\varepsilon^2}) + 4}, \quad \tau \geq 1, \tag{8}$$

where ρ_τ^x stands for the lag-τ autocorrelation of x. The numerator of (8) is defined as a weighted sum of two factors that depend on τ. The first one, $\rho_\tau^{\eta^2}$, has a weight which is a function of q and κ_η, while the second, $\rho_{\tau-1}^{\varepsilon^2} + 2\rho_\tau^{\varepsilon^2} + \rho_{\tau+1}^{\varepsilon^2}$, has a weight depending only on κ_ε. As long as the acf of squares of both disturbances converge to zero, each of these factors disappears as τ increases, and therefore the acf of $(\Delta y_t)^2$ also converges to zero. This issue is studied later with more detail.

Next, we derive particular cases of the kurtosis of Δy_t and acf of $(\Delta y_t)^2$ depending on different specifications of ε_t and η_t. In particular, we will consider noises which are homoscedastic although non-Normal and GARCH(1,1) noises[1].

2.1 Non-Gaussian disturbances

Consider first that the disturbances are homoscedastic. If they are further assumed to be Gaussian, then, as expected, the kurtosis in (7) is given by $\kappa_{\Delta y} = 3$. However, if the disturbances are non-Normal, then[2]

$$\kappa_{\Delta y} = \frac{1}{(q+2)^2} \left[q^2 \kappa_\eta + (12q + 6) + 2\kappa_\varepsilon \right]. \tag{9}$$

From (9) we can see that if $q^2 > 2$, the contribution of κ_η to the kurtosis of Δy_t is greater than the contribution of κ_ε. The opposite happens when $q^2 < 2$. Furthermore, note that the slopes of $\kappa_{\Delta y}$ with respect to k_ε and k_η are different:

$$\frac{\partial \kappa_{\Delta y}}{\partial \kappa_\eta} = \frac{q^2}{(q+2)^2} \rightarrow 1 \text{ as } q \text{ tends to } +\infty,$$

$$\frac{\partial \kappa_{\Delta y}}{\partial \kappa_\varepsilon} = \frac{2}{(q+2)^2} \rightarrow 0.5 \text{ as } q \text{ tends to zero.}$$

Figure 1 plots $\kappa_{\Delta y}$ as a function of κ_ε and κ_η for different values of the signal-to-noise ratio, q. It can be observed that, as the derivatives above show, the relation between $\kappa_{\Delta y}$ and each kurtosis is linear, with the slope with respect to κ_η being steeper when $q > \sqrt{2}$. In any case, it is interesting to observe that the slope is always smaller than one. Consider, for example, that

---

[1]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 LL model with GQARCH disturbances to account for asymmetries in volatility.

[2]Note that if ε_t is a white noise process with $E[\varepsilon_t^4] > 0$ and $\kappa_\varepsilon \neq 1$, the assumption of conditional homoscedasticity implies that $\rho_j^{\varepsilon^2} = 0$, $j > 0$, regardless of its conditional distribution. To prove it, note that conditional homoscedasticity means $E[\varepsilon_t^2 | \mathcal{I}_{t-j}] = E[\varepsilon_t^2] = \sigma_\varepsilon^2$, $\forall j > 0$ where $\mathcal{I}_{t-j}$ is the information set available at time $t - j$. Taking into account that

$$\rho_j^{\varepsilon^2} = \frac{E[\varepsilon_t^2 \, \varepsilon_{t-j}^2]}{(\kappa_\varepsilon - 1)\sigma_\varepsilon^4} - \frac{1}{\kappa_\varepsilon - 1},$$

and knowing that

$$E[\varepsilon_t^2 \, \varepsilon_{t-j}^2] = E\{E[\varepsilon_t^2 \, \varepsilon_{t-j}^2 | \mathcal{I}_{t-j}]\} = E\{\varepsilon_{t-j}^2 \, E[\varepsilon_t^2 | \mathcal{I}_{t-j}]\} = E[\varepsilon_{t-j}^2] \, \sigma_\varepsilon^2 = \sigma_\varepsilon^4,$$

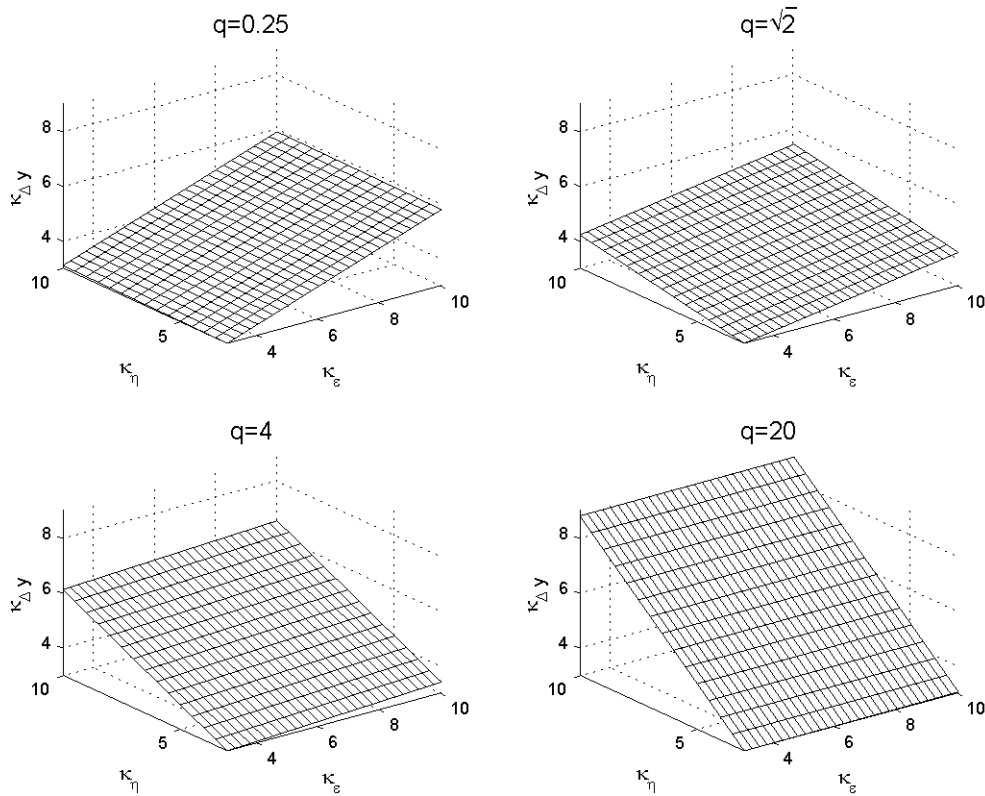the result $\rho_j^{\varepsilon^2} = 0$, $j > 0$ is straightforward.

Figure 1: Relationship between $\kappa_{\Delta y}$, κ_ε and κ_η in homoscedastic non-Gaussian LL models

$\kappa_\varepsilon = 6$ and $\kappa_\eta = 5$. In this case, if $q = 0.25$, $\kappa_{\Delta y} = 4.21$, while if $q = \sqrt{2}$, then $\kappa_{\Delta y} = 3.86$. Finally, if $q = 4$, then $\kappa_{\Delta y} = 4.06$. Note that in this example the excess kurtosis of the stationary transformation of y_t is smaller than the corresponding for each component. As a general result, we show in the Appendix that for all values of q, $\kappa_{\Delta y}$ is always smaller than or equal to the maximum of κ_ε and κ_η. Furthermore, as long as both disturbances have positive excess kurtosis, we can always find an interval for q within which $\kappa_{\Delta y} < \min(\kappa_\varepsilon, \kappa_\eta)$.

With respect to the acf of squares, if the disturbances are homoscedastic, expression (8) is reduced to

$$\rho_\tau^{(\Delta y)^2} = \begin{cases} \frac{(\kappa_\varepsilon - 1)}{(q+2)^2(\kappa_{\Delta y} - 1)}, & \tau = 1 \\ \\ 0, & \tau \geq 2 \end{cases} \tag{10}$$

Note that under normality, $\rho_1^{(\Delta y)^2} = (q + 2)^{-2}$, which turns out to be the squared lag-one autocorrelation of Δy_t; see Maravall (1983). However, for non-Gaussian processes the acf of squares is not necessarily the square of the acf . From (10) it can be seen that, if κ_ε is greater

(smaller) than $\kappa_{\Delta y}$, $\rho_1^{(\Delta y)^2}$ is greater (smaller) than in the Gaussian case. For example, if $\kappa_\varepsilon = 6$, $\kappa_\eta = 5$ and $q = \sqrt{2}$, so that $\kappa_{\Delta y} = 3.86$, then $\rho_1^{(\Delta y)^2} = 0.15 > 0.09 = (q+2)^{-2}$. On the other hand, when the excess kurtosis comes only from the permanent component (i.e. $\kappa_\varepsilon = 3$ and $\kappa_\eta > 3$), $\rho_1^{(\Delta y)^2}$ is always smaller than in the Gaussian case.

2.2 GARCH disturbances

Consider now that each noise is a conditionally Normal GARCH(1,1) process. Therefore, they 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 Gaussian white noise processes and

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

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

where the parameters α_0, α_1, α_2, γ_0, γ_1 and γ_2 are assumed to satisfy the usual positivity and stationarity conditions. Then, substituting κ_ε, κ_η and $\rho_1^{\varepsilon^2}$ in expression (7) by their expressions when the disturbances are GARCH[3], $\kappa_{\Delta y}$ is given by

$$\kappa_{\Delta y} = \frac{3}{(q+2)^2}\left[q^2 \frac{1-(\gamma_1+\gamma_2)^2}{1-3\gamma_1^2-2\gamma_1\gamma_2-\gamma_2^2} + 4q \right.$$
$$\left. + 4\frac{1-(\alpha_1+\alpha_2)^2+\alpha_1(1-\alpha_1-\alpha_1\alpha_2-\alpha_2^2)}{1-3\alpha_1^2-2\alpha_1\alpha_2-\alpha_2^2}\right]. \tag{13}$$

Figure 2 plots for different values of the signal-to-noise ratio, the relationship between the kurtosis of Δy_t and the persistence of the volatility of both noises, measured as the sum of the ARCH (α_1, γ_1) and GARCH (α_2, γ_2) coefficients. For ease of exposition we keep the GARCH coefficients equal to 0.85. Note that the slope with respect to the persistence of η_t is steeper as q increases, and also that varying q significantly affects $\kappa_{\Delta y}$.

We consider now the autocorrelations of squares. If ε_t and η_t are GARCH(1,1) processes,

---

[3]The properties of GARCH processes can be found in, for example, Tsay (2005).
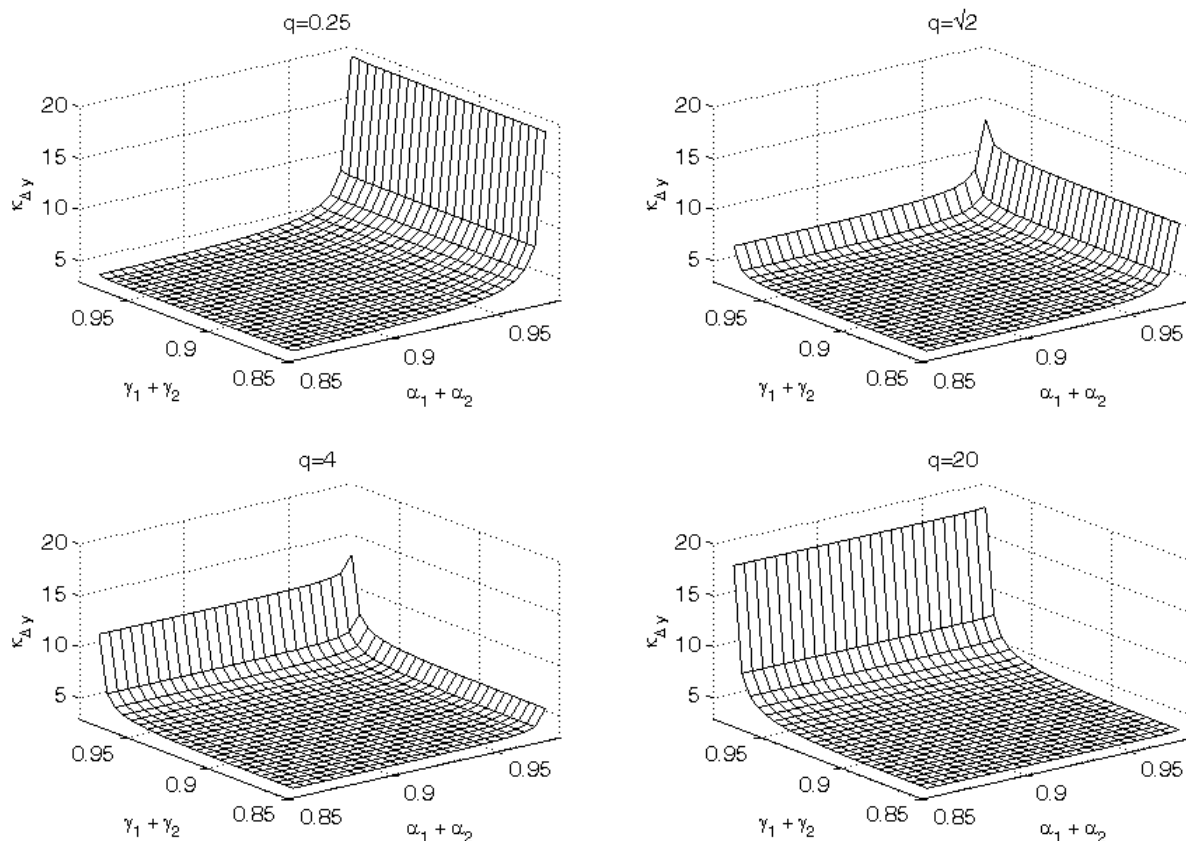
Figure 2: Relationship between $\kappa_{\Delta y}$ and persistence. The GARCH coefficients, α_2 and γ_2 are fixed to 0.85.

the acf of $(\Delta y_t)^2$ is given by

$$
\rho_\tau^{(\Delta y)^2} =
\begin{cases}
\frac{1}{(q+2)^2(\kappa_{\Delta y}-1)}\left[q^2\rho_1^{\eta^2}(\kappa_\eta - 1) + (\kappa_\varepsilon - 1)(1 + \rho_1^{\varepsilon^2}(2 + \alpha_1 + \alpha_2))\right], & \tau = 1 \\[2mm]
(\alpha_1 + \alpha_2)\rho_{\tau-1}^{(\Delta y)^2} + \frac{1}{(q+2)^2(\kappa_{\Delta y}-1)}(\gamma_1 + \gamma_2 - \alpha_1 - \alpha_2)q^2(\gamma_1 + \gamma_2)^{\tau-2}\rho_1^{\eta^2}(\kappa_\eta - 1), & \tau \geq 2
\end{cases}
$$

$$(14)$$

where $\rho_1^{\eta^2} = \gamma_1(1 - \gamma_1\gamma_2 - \gamma_2^2)/(1 - 2\gamma_1\gamma_2 - \gamma_2^2)$, with $\rho_1^{\varepsilon^2}$ defined analogously.

From (14) we can see that when both noises are heteroscedastic, and $\gamma_1 + \gamma_2 = \alpha_1 + \alpha_2$, the acf of squares has an exponential decay, as in the 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 Δy_t is not GARCH. Figure 3 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 the case in which both disturbances follow the same GARCH process, and the second shows a case where both noises

follow GARCH processes with different persistences. The last two rows consider the cases 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[4]. Consequently, exponential structures such as the ones implied by GARCH processes can be a good approximation for the acf of $(\Delta y_t)^2$.
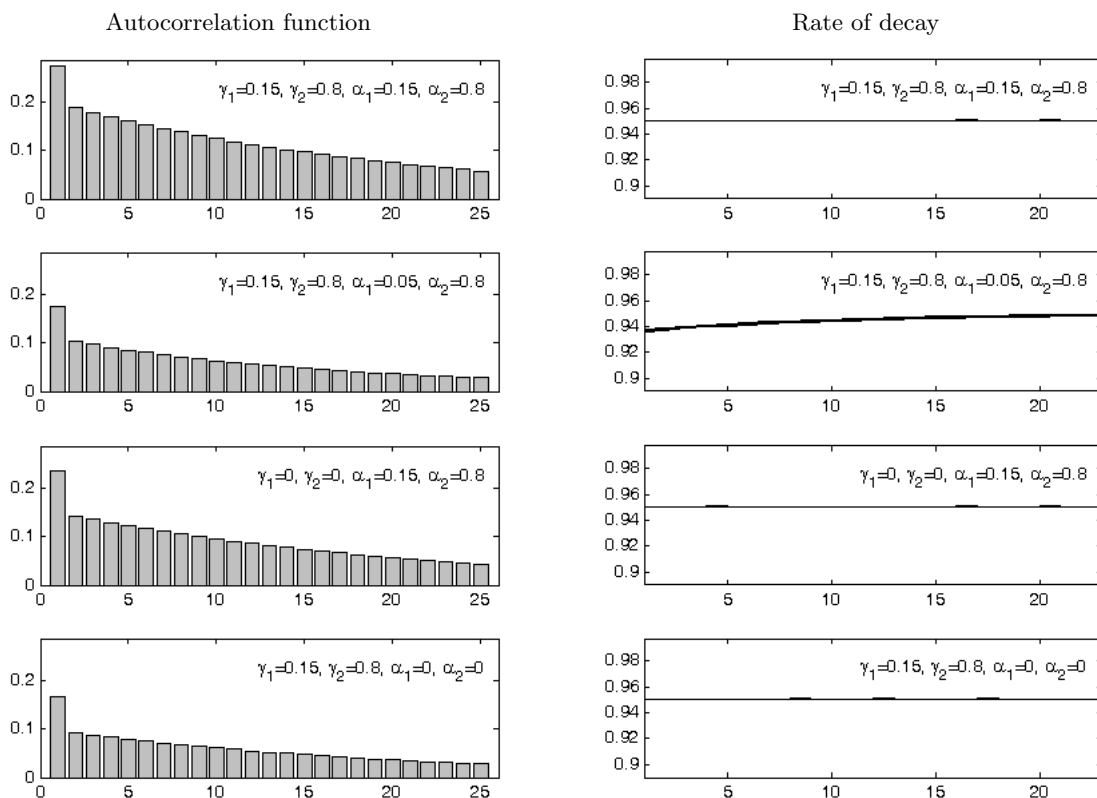

Figure 3: Autocorrelations of $(\Delta y_t)^2$ for several Local Level models with conditionally Normal GARCH(1,1) disturbances and $q = \sqrt{2}$. The rate of decay reported in the right column is defined by the ratio $\rho_\tau^{(\Delta y)^2} / \rho_{\tau-1}^{(\Delta y)^2}$.

---

[4]It can be proved that the rate of decay of $\rho_\tau^{(\Delta y)^2}$ given by (14) converges to the $max(\alpha_1 + \alpha_2; \gamma_1 + \gamma_2)$ as τ 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 τ.

3. Properties of the residuals of the reduced form ARIMA model

We have seen that ε_t and η_t being mutually and serially uncorrelated is sufficient to prove that the reduced form of the LL model is a restricted IMA(1,1), as stated in (4), with a_t being also serially uncorrelated. Furthermore, if both ε_t and η_t 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 show which are the parameters expected to be obtained if a_t is assumed to be a GARCH(1,1) model.

3.1 Kurtosis and acf of squares of a_t in terms of ε_t and η_t

From the model given by (4), it is possible to derive the following expression of the kurtosis of Δy_t,

$$\kappa_{\Delta y} = \frac{1}{(1+\theta^2)^2} \left[\kappa_a (1+\theta^4) + 6\theta^2 [\rho_1^{a^2} (\kappa_a - 1) + 1] \right], \tag{15}$$

where κ_a and $\rho_1^{a^2}$ are the kurtosis of a_t and the lag-one autocorrelation of a_t^2, respectively. Note from (15) that $\kappa_{\Delta y}$ is a function not only of κ_a but also of $\rho_1^{a^2}$. This result is consistent with (7), where $\kappa_{\Delta y}$ is affected also by $\rho_1^{\varepsilon^2}$. Finally, the acf of Δy_t^2 is given by

$$\rho_\tau^{(\Delta y)^2} = \frac{\kappa_a - 1}{(1+\theta^2)^2 (\kappa_{\Delta y} - 1)} \left[(1+\theta^4)\rho_\tau^{a^2} + \theta^2 (\rho_{\tau-1}^{a^2} + \rho_{\tau+1}^{a^2}) \right], \ \tau \geq 1. \tag{16}$$

Note that expressions (15) and (16) are defined for a general IMA(1,1) process, resulting from a LL model with ε_t and η_t being mutually independent and serially uncorrelated processes, with symmetric distributions around zero and finite fourth order moments.

In order to find an expression of κ_a and $\rho_\tau^{a^2}$, as functions of the parameters from the unobserved component model, consider the kurtoses of Δy_t given by (7) and (15). After equalling

these expressions we obtain[5]

$$(\kappa_a - 1)\left(1 + \theta^4 + 6\theta^2 \rho_1^{a^2}\right) \equiv (1+\theta)^4(\kappa_\eta - 1) - 8\theta(1+\theta)^2 + 2\theta^2(\kappa_\varepsilon - 1)\left(1 + 3\rho_1^{\varepsilon^2}\right). \quad (17)$$

On the other hand, after equalling the autocorrelation of order $\tau = 1, 2, \ldots$ in (8) and (16) we obtain

$$(\kappa_a - 1)\left[(1+\theta^4)\rho_\tau^{a^2} + \theta^2\left(\rho_{\tau-1}^{a^2} + \rho_{\tau+1}^{a^2}\right)\right] \equiv \theta^2(\kappa_\varepsilon - 1)\left(\rho_{\tau-1}^{\varepsilon^2} + 2\rho_\tau^{\varepsilon^2} + \rho_{\tau+1}^{\varepsilon^2}\right)$$
$$+ (1+\theta)^4(\kappa_\eta - 1)\rho_\tau^{\eta^2}. \quad (18)$$

With (17) and the identities that follow from (18) for different values of τ, we are able to fully characterize the kurtosis and acf of squares of a_t. In the simplest case, when ε_t and η_t are homoscedastic Gaussian processes, the above expressions reduce to

$$(\kappa_a - 1)\left(1 + \theta^4 + 6\theta^2 \rho_1^{a^2}\right) \equiv 2(1+\theta^4), \quad (19)$$

$$(\kappa_a - 1)\left[(1+\theta^4)\rho_1^{a^2} + \theta^2\left(1 + \rho_2^{a^2}\right)\right] \equiv 2\theta^2, \quad (20)$$

$$(\kappa_a - 1)\left[(1+\theta^4)\rho_\tau^{a^2} + \theta^2\left(\rho_{\tau-1}^{a^2} + \rho_{\tau+1}^{a^2}\right)\right] \equiv 0 \ , \ \forall \ \tau > 1 \quad (21)$$

The system of identities given by (19)-(21) yields $\rho_\tau^{a^2} = 0$, $\forall \tau > 0$ and $\kappa_a = 3$, as expected given that a_t is a linear combination of independent Gaussian noises.

When assuming that ε_t and η_t are homoscedastic but not necessarily normal, we have seen in the previous section that $\rho_1^{(\Delta y)^2}$ may differ from $(\rho_1^{\Delta y})^2$. 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 values of q, κ_ε and κ_η, coming from the resolution of the system given by (17) and (18). Observe

---

[5]To obtain (17), recall that θ can be defined in terms of q, so that the following expressions result:

$$1 + \theta^2 = -\theta(q+2),$$
$$1 + \theta^4 = \theta^2(q^2 + 4q + 2).$$

that non-normality in either or both noises may generate structure in the acf of squares (specially in the lag-one autocorrelation), thus reflecting the serial dependence. It must be pointed out that, although this structure does not reflect the presence of GARCH effects in the series (by just looking at the whole acf of squares we realize that it does not follow any specific pattern), it is possible to obtain values for some usual statistics, e.g. the McLeod-Li (1983) test, that may wrongly lead to fit conditionally heteroscedastic models to the series. Also note that this pattern in which only the lag-one autocorrelation of squares seem to be significant, can be confused with the effect of outliers; see Carnero et al. (2006).

q	κ_ε	κ_η	θ	$\kappa_{\Delta y}$	$\rho_1^{(\Delta y)^2}$	κ_a	$\rho_1^{a^2}$	$\rho_2^{a^2}$	$\rho_3^{a^2}$	$\rho_4^{a^2}$	$\rho_5^{a^2}$
0.5	3	6	-0.5	3.120	0.151	3.273	-0.030	0.008	-0.002	0.001	0.000
$\sqrt{2}$	3	6	-0.324	3.515	0.068	3.665	-0.026	0.003	0.000	0.000	0.000
0.5	6	6	-0.5	4.080	0.260	3.818	0.194	-0.048	0.012	-0.003	0.001
$\sqrt{2}$	6	6	-0.324	4.029	0.142	4.120	0.063	-0.007	0.001	0.000	0.000
0.5	6	3	-0.5	3.960	0.270	3.546	0.241	-0.060	0.015	0.004	0.001
$\sqrt{2}$	6	3	-0.324	3.515	0.171	3.456	0.109	-0.011	0.001	0.000	0.000

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

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 identities given by (17) and (18) to construct the following set of equations:

$$\left[(1 + \theta^4) - 6\theta^2 Q(1)\right] \rho_1^{a^2} + \theta^2 \rho_2^{a^2} = Q(1)(1 + \theta^4) - \theta^2, \tag{22}$$

$$\left[\theta^2 - 6\theta^2 Q(2)\right] \rho_1^{a^2} + (1 + \theta^4)\rho_2^{a^2} + \theta^2 \rho_3^{a^2} = Q(2)(1 + \theta^4), \tag{23}$$

$$-6\theta^2 Q(\tau)\rho_1^{a^2} + \theta^2 \rho_{\tau-1}^{a^2} + (1 + \theta^4)\rho_\tau^{a^2} + \theta^2 \rho_{\tau+1}^{a^2} = Q(\tau)(1 + \theta^4) , \ \tau > 2 \tag{24}$$

where $Q(\tau)$ depends on the structural noises in the following way:

$$Q(\tau) = \frac{(1+\theta)^4 \rho_\tau^{\eta^2}(\kappa_\eta - 1) + \theta^2(\kappa_\varepsilon - 1)\left(\rho_{\tau-1}^{\varepsilon^2} + 2\rho_\tau^{\varepsilon^2} + \rho_{\tau+1}^{\varepsilon^2}\right)}{(1+\theta)^4(\kappa_\eta - 1) - 8\theta(1+\theta)^2 + 2\theta^2(\kappa_\varepsilon - 1)\left(1 + 3\rho_1^{\varepsilon^2}\right)}. \tag{25}$$

When assuming that ε_t and η_t are stationary GARCH processes, $Q(\tau)$ converges to zero as τ increases. In other words, there exists a value of τ, say τ_{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 3. In general, the magnitude of the autocorrelations of a_t^2 is smaller than the corresponding ones of the disturbances of the LL 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. It might be the case for instance that if the permanent component, μ_t, presents a significant heteroscedastic structure but the transitory component, ε_t, is homoscedastic; the stationary transformation, Δy_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 IMA-GARCH model and the reduced form of Δy_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 model. However, the decay of the autocorrelations of squares could be approximated by such a model and this is the usual practice when analyzing real time series. Consequently, in this subsection, we derive the value of the GARCH parameters that would be obtained if one fits a GARCH(1,1) model to the disturbance of the IMA(1,1) 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

$$\sigma_t^2 = \delta_0 + \delta_1 a_{t-1}^2 + \delta_2 \sigma_{t-1}^2. \tag{26}$$
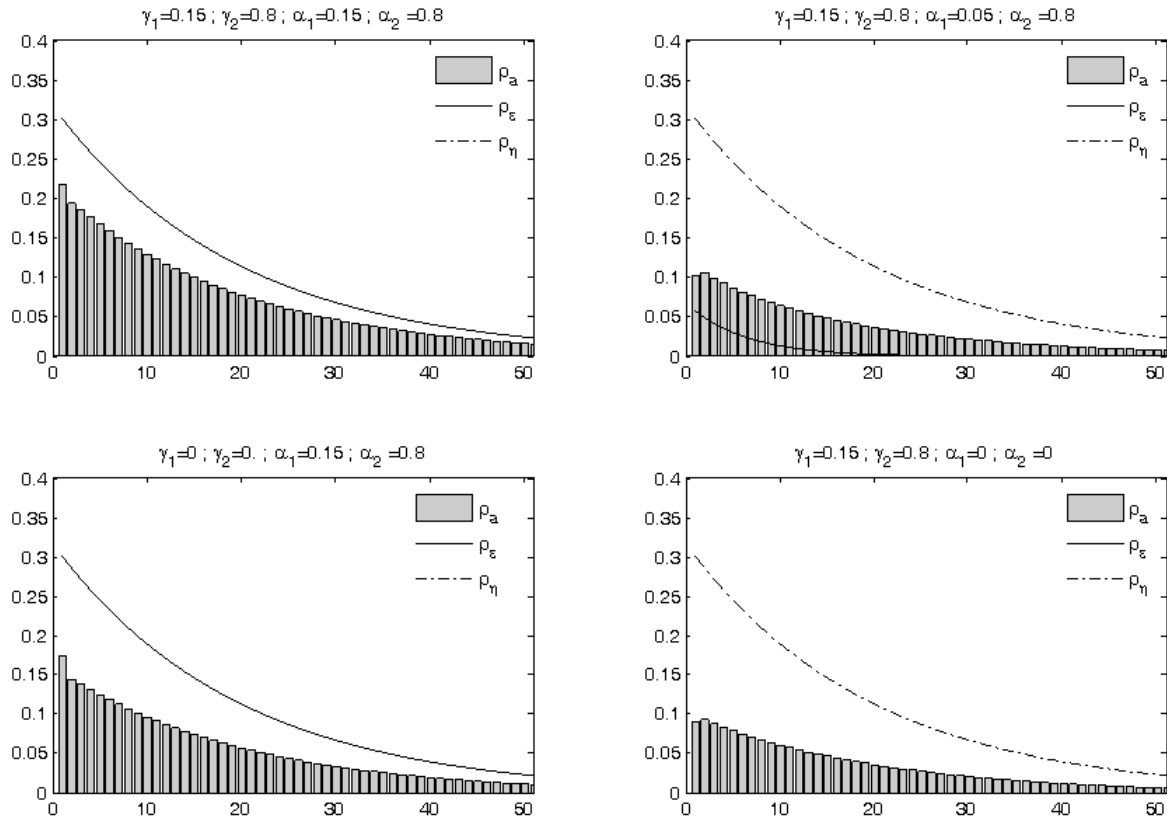
Figure 4: Autocorrelations of a_t^2, resulting from different conditionally Normal GARCH(1,1) disturbances with $q = \sqrt{2}$. The solid and dash-dotted lines draw the autocorrelations of ε_t^2 and η_t^2, respectively.

We want to obtain expressions of the parameters δ_1 and δ_2 as functions of κ_a and $\rho_\tau^{a^2}$ derived in the previous subsection. First, note that in the GARCH(1,1) model in (26) κ_a and $\rho_\tau^{a^2}$ are given by

$$\kappa_a = \frac{3\left[1 - (\delta_1 + \delta_2)^2\right]}{1 - 3\delta_1^2 - 2\delta_1\delta_2 - \delta_2^2}, \tag{27}$$

$$\rho_\tau^{a^2} = (\delta_1 + \delta_2)^{\tau-1}\frac{\delta_1(1 - \delta_1\delta_2 - \delta_2^2)}{1 - 2\delta_1\delta_2 - \delta_2^2}, \quad \tau \geq 1. \tag{28}$$

From (27) and (28) it is possible to derive the following expression of δ_1,

$$\delta_1 = \frac{3(\kappa_a - 1)\rho_\tau^{a^2} - (\delta_1 + \delta_2)^\tau(\kappa_a - 3)}{2\kappa_a(\delta_1 + \delta_2)^{\tau-1}}, \tag{29}$$

for any $\tau \geq 1$. On the other hand, given the parameters of the GARCH models of ε_t and η_t, the kurtosis, κ_a, and autocorrelations of squares, $\rho_\tau^{a^2}$, can be obtained by solving the system (22) to (24). Then $(\delta_1 + \delta_2)$ is given by the ratio $\rho_\tau^{a^2}/\rho_{\tau-1}^{a^2}$. Finally, substituting κ_a, $\rho_\tau^{a^2}$ and $(\delta_1 + \delta_2)$

in (29), we obtain the corresponding value of δ_1. Table 2 shows some examples. For instance, consider the model where both noises are heteroscedastic and $q = \sqrt{2}$. In this case, $\kappa_a = 4.45$, $\rho_2^{a^2} = 0.19$, $\rho_3^{a^2} = 0.18$ and $\delta_1 + \delta_2 = 0.95$, so that $\delta_1 = 0.083$ (the upper-left case of Figure 4). This value, which measures the level of heteroscedasticity in a series, is clearly smaller than the corresponding values for the structural noises ($\alpha_1 = \gamma_1 = 0.15$). If we now consider the model in which only the trend component is heteroscedastic, the values are given by $\kappa_a = 3.4$, $\rho_2^{a^2} = 0.094$, $\rho_3^{a^2} = 0.089$ and $\delta_1 + \delta_2 = 0.95$, thus obtaining $\delta_1 = 0.05$. In this case, the ARCH coefficient in the reduced form disturbance is one-third the value of γ_1 (0.15). Furthermore, if in this case, we change the value of q from 0.5 to $\sqrt{2}$ keeping the rest unchanged, we see that $\delta_1 = 0.014$, that is one-tenth of γ_1. Notice from these examples that it is possible to reject heteroscedasticity although at least one of the underlying noises are clearly heteroscedastic.

q	α_1	α_2	γ_1	γ_2	κ_ε	κ_η	θ	κ_a	$\rho_1^{a^2}$	$\rho_2^{a^2}$	$\rho_3^{a^2}$	$\rho_4^{a^2}$	δ_1	δ_2
0.5	0.15	0.80	0.15	0.80	5.57	5.57	-0.5	4.910	0.251	0.223	0.216	0.204	0.100	0.850
$\sqrt{2}$	0.15	0.80	0.15	0.80	5.57	5.57	-0.324	4.451	0.217	0.193	0.185	0.175	0.083	0.867
0.5	0	0	0.15	0.80	3	5.57	-0.5	3.083	0.023	0.026	0.024	0.023	0.014	0.936
$\sqrt{2}$	0	0	0.15	0.80	3	5.57	-0.324	3.396	0.092	0.094	0.089	0.084	0.049	0.901
0.5	0.15	0.80	0	0	5.57	3	-0.5	4.828	0.244	0.214	0.208	0.196	0.093	0.857
$\sqrt{2}$	0.15	0.80	0	0	5.57	3	-0.324	4.055	0.174	0.144	0.139	0.132	0.051	0.899

Table 2: Theoretical values of the GARCH parameters, δ_1 and δ_2, of the reduced form noise corresponding to different models for ε_t and η_t.

To illustrate how the heteroscedasticity may be hidden when fitting an IMA(1,1)-GARCH(1,1) model to series drawn from a LL-GARCH(1,1) model, we generate 1000 series with parameters $\alpha_1 = \alpha_2 = 0$, $\gamma_1 = 0.15$, $\gamma_2 = 0.8$ and $q = 1$[6]. The series are generated with four different sample sizes ($T = 200$, 500, 1000, and 5000). For each simulated series, we first fit an homoscedastic IMA(1,1) model and test for conditional heteroscedasticity in the residuals using the test

---

[6]Results for other alternative models are available upon request.

proposed by Rodriguez and Ruiz (2005):

$$Q_1(10) = T \sum_{k=1}^{9} \left[\tilde{r}(k) + \tilde{r}(k+1) \right]^2, \tag{30}$$

where $\tilde{r}(k) = \sqrt{(T+2)/(T-k)} r(k)$ is the standardized sample autocorrelation of order k. Table 3 reports the mean and standard deviations through all Monte Carlo replicates of the QML estimates of θ. We see that $\hat{\theta}$ is asymptotically unbiased in the sense that it converges to the value implied by q. The table also reports 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 ($T = 5000$) all the IMA residuals capture the conditional heteroscedasticity coming from the underlying unobserved components. However, for small or moderate samples, this statistic show a large proportion of cases where the homoscedasticity cannot be rejected. For instance, for $T = 200$, in more than 75% of the cases the residuals do not present evidence of conditional heteroscedasticity when in fact the permanent component is heteroscedastic.

Then, for each simulated series, the GARCH(1,1) model is fitted to the residuals. Table 3 reports the mean of the QML estimates, as well as the variance, kurtosis and autocorrelations of squares of Δy_t implied by these estimates. With respect to the estimates of the ARCH parameter, δ_1, their values are consistent with the analytical results found above. Given the parameters of the LL model used to simulate the series, the implied value of δ_1 is 0.05. Regarding the mean estimates of the GARCH coefficient, δ_2, we see that they increase with the sample size, so that the estimated sum $\delta_1 + \delta_2$ converges to the sum $\gamma_1 + \gamma_2$.

Finally, Table 3 reports the percentage of rejection of homoscedasticity when looking at the 5% significance of the ARCH parameter. For large samples we can see that there are no serious dangers in terms of heteroscedasticity rejection. However, in small or moderate samples, a higher sampling error may accentuate the problems of estimating the reduced form IMA-GARCH as the heteroscedasticity can be rejected in a very large proportion of cases. For instance, if $T = 200$, we find significant ARCH effects in only 9% of the series.

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)
$\sigma^2_{\Delta y} = 3$	$\hat{\sigma}^2_{\Delta y}$	2.923 (0.673)	3.016 (0.946)	2.973 (0.253)	2.967 (0.101)
$\kappa_{\Delta y} = 3.286$	$\hat{\kappa}_{\Delta y}$	3.275 (0.631)	3.243 (0.574)	3.192 (0.165)	3.177 (0.058)
$\rho_1^{(\Delta y)^2} = 0.164$	$\hat{\rho}_1^{(\Delta y)^2}$	0.191 (0.108)	0.174 (0.053)	0.166 (0.025)	0.163 (0.010)
$\rho_2^{(\Delta y)^2} = 0.063$	$\hat{\rho}_2^{(\Delta y)^2}$	0.061 (0.122)	0.056 (0.063)	0.055 (0.032)	0.057 (0.014)
Homosc. Rejection $(t_{(\delta_1)} > 1.64)$:		9%	39%	73%	100%

Table 3: Monte Carlo results on the QML estimator of the parameters of the IMA-GARCH when the series are generated by a LL-GARCH model.

4. Analysis of the forecasting performance

In the present section we analyze the forecasting performance of the IMA-GARCH model when implemented to construct prediction intervals of series generated by the LL-GARCH model. We compare the theoretical mean squared forecast errors (MSFEs) and prediction intervals obtained from both models. The results are illustrated using simulated series.

4.1 The MSFE of both models

Consider y_t generated by the LL model in (1). It is well known that if one wants to minimize the MSFE, then the conditional mean is the optimal point predictor of y_{T+k}. Assuming that the parameters are known, the Kalman Filter can be implemented to obtain estimates of the underlying state at time $t = 1, 1, \ldots, T$ denoted by m_t. Then, the optimal linear point predictor is given by

$$\hat{y}_{T+k} = \underset{T}{E}[y_{T+k}] \quad = \quad m_T, \ k = 1, 2, \ldots \tag{31}$$

where the T under the expectation means that it is conditional on the information available at time T, i.e. $\{y_1, \ldots, y_T\}$. From (31) it is easy to derive the MSFE, as function of both noises:

$$MSFE(\hat{y}_{T+k}) \;=\; P_T + \underset{T}{E}[\varepsilon_{T+k}^2] + \sum_{j=1}^{k} \underset{T}{E}[\eta_{T+j}^2], \;\; k = 1, 2, \ldots \tag{32}$$

where $P_T = \underset{T}{E}[(\mu_T - m_T)^2]$. When η_t and ε_t are GARCH processes, given in (11) and (12), it is easy to see that

$$\underset{T}{E}[\varepsilon_{T+k}^2] \;=\; \sigma_\varepsilon^2 + (\alpha_1 + \alpha_2)^{k-1}(h_{T+1} - \sigma_\varepsilon^2), \;\; k = 1, 2, \ldots \tag{33}$$

$$\underset{T}{E}[\eta_{T+k}^2] \;=\; \sigma_\eta^2 + (\gamma_1 + \gamma_2)^{k-1}(q_{T+1} - \sigma_\eta^2), \;\; k = 1, 2, \ldots \tag{34}$$

where h_{T+1} and q_{T+1} are the conditional variances of ε_{T+1} and η_{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. Since we are assuming stationarity in second moments, any disequilibrium (i.e. an excess volatility different from zero) vanishes as k increases and the long term forecast converges to the marginal variance. Plugging (33) and (34) into (32) we get the following expression for the $MSFE(\hat{y}_{T+k})$:

$$\begin{aligned} MSFE(\hat{y}_{T+k}) \;=\;& P_T + \sigma_\varepsilon^2 + k\,\sigma_\eta^2 + \frac{1 - (\gamma_1 + \gamma_2)^k}{1 - (\gamma_1 + \gamma_2)}\,(q_{T+1} - \sigma_\eta^2) \\ &+ (\alpha_1 + \alpha_2)^{k-1}\,(h_{T+1} - \sigma_\varepsilon^2), \;\; k = 1, 2, \ldots \end{aligned} \tag{35}$$

Note that the MSFE of the homoscedastic LL model is given by the first three terms of (35). Also note that the MSFE of the LL-GARCH becomes a linear function of k in the long run, with the same slope as its linear counterpart, but with a different intercept due to the contribution of the fourth term in (35). However, for short and medium horizons, the influence of the excess volatility in both noises leads to a MSFE smaller or greater than that of the homoscedastic LL model. It is also important to note that there is a significant distinction 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.

Now, consider that y_t is given by the reduced form IMA(1,1) model in (4). In this case the optimal linear predictor of y_{T+k} given the information available at time T is given by

$$\hat{y}_{T+k} = y_T + \theta a_T, \ k = 1, 2, ... \tag{36}$$

with MSFE

$$MSFE(\hat{y}_{T+k}) = \begin{cases} \underset{T}{E}[a_{T+1}^2], & k = 1 \\ \\ \underset{T}{E}[a_{T+k}^2] + (1+\theta)^2 \sum_{j=1}^{k-1} \underset{T}{E}[a_{T+j}^2], & k = 2, 3, ... \end{cases} \tag{37}$$

If a_t is a GARCH(1,1) model given by (26) then

$$\underset{T}{E}[a_{T+k}^2] = \sigma_a^2 + (\delta_1 + \delta_2)^{k-1}(\sigma_{T+1}^2 - \sigma_a^2), \ k = 1, 2, ... \tag{38}$$

where σ_{T+1}^2 is the conditional variance of a_{T+1}, and $(\sigma_{T+1}^2 - \sigma_a^2)$ is the measure of the excess volatility, analogous to those of the LL-GARCH disturbances. Again, by plugging (38) into (37) we find that

$$MSFE(\hat{y}_{T+k}) = \begin{cases} \sigma_a^2 + (\sigma_{T+1}^2 - \sigma_a^2), & k = 1 \\ \\ \left[(1+\theta)^2(k-1) + 1 \right] \sigma_a^2 + \\ \\ \left[\frac{(1+\theta)^2 - (\delta_1+\delta_2)^{k-1}(\theta(2+\theta)+\delta_1+\delta_2)}{1-(\delta_1+\delta_2)} \right] (\sigma_{T+1}^2 - \sigma_a^2), & k = 2, 3, ... \end{cases} \tag{39}$$

As in the LL-GARCH case, the MSFE of the IMA-GARCH model can be separated into a linear and a nonlinear part, defined by the first and second terms of (39), respectively. It is clear from this expression that the $MSFE(\hat{y}_{T+k})$ is also a linear function of the horizon as k increases. However, as long as the excess volatility is different from zero, the path of the IMA-GARCH MSFE never converges to that of the linear (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 linear IMA. In this sense, the behavior is similar to that of the LL model with heteroscedastic long-run disturbances.

4.2 Some illustrations using prediction intervals

In this subsection we illustrate with simulated data the differences between the MSFEs of the LL-GARCH and IMA-GARCH models and their consequences when constructing prediction intervals for y_{T+k}. We generate series from the unobserved component LL model with either or both disturbances being GARCH processes and, assuming that the parameters of both models are known, we find the MSFE of the LL-GARCH and IMA-GARCH models at a given time T. Then, we construct 95% Gaussian prediction intervals[7] 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.

Assume first that the series is generated from a LL-GARCH model where only the permanent component is heteroscedastic, with $\gamma_1 = 0.15$, $\gamma_2 = 0.8$, and $q = 1$. The selected time point corresponds to a volatile period since $q_{T+1} - \sigma_\eta^2 = 1.58 > 0$, and obviously all the excess volatility comes from the permanent component. Figure 5 shows that the resulting MSFEs of the IMA-GARCH model produce narrower prediction intervals compared to those of the LL-GARCH. In order to obtain the observed coverage, we simply count the number of observations lying outside each prediction interval and then divide it by B. Figure 6 shows the 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 observed coverage of the IMA-GARCH stays around 92%, when the nominal is 95%.

---

[7]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.
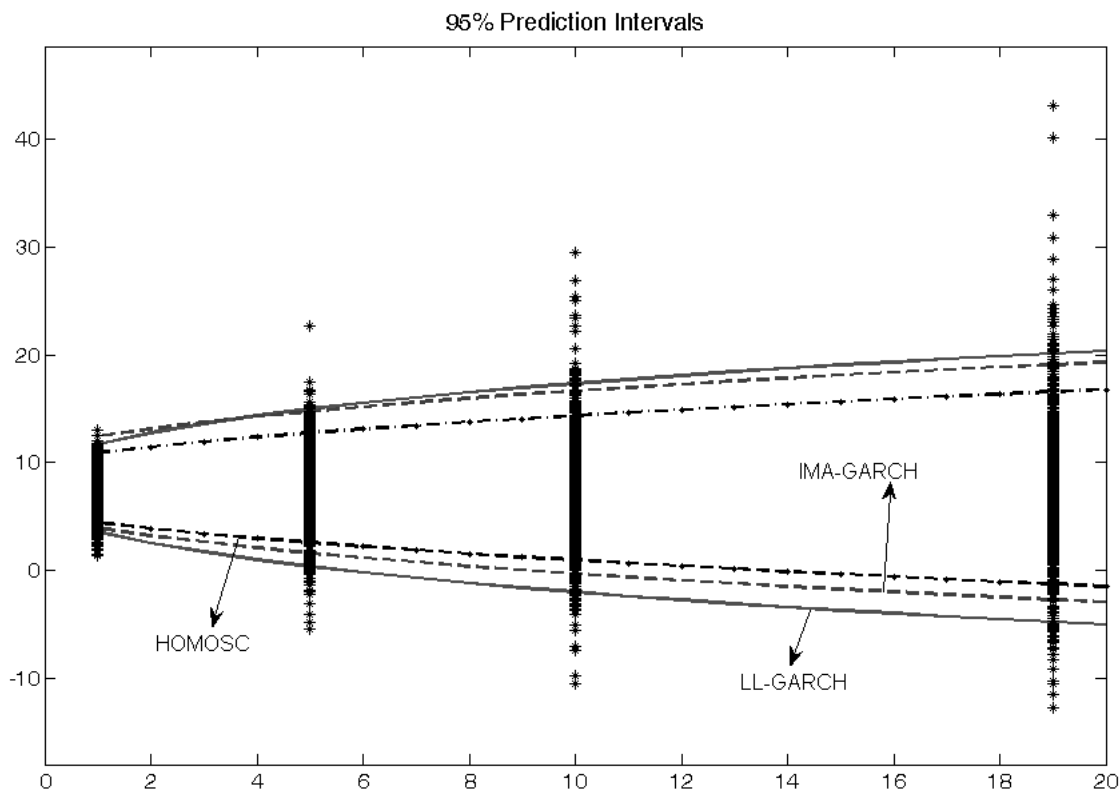
Figure 5: Prediction intervals of a LL-GARCH series where only the permanent component is heteroscedastic, with $\gamma_1 = 0.15$, $\gamma_2 = 0.8$, and $q = 1$. The time point is selected in a highly volatile period.

The homoscedastic prediction intervals generate the worst scenario because the observed coverage starts at 90% and then stays around 87%. In this case, given that a volatility shock in the permanent component does not vanish as the horizon increases, the homoscedastic prediction intervals do not converge to those of the LL-GARCH model.

Now we illustrate the case where only ε_t is heteroscedastic, with parameters $\alpha_1 = 0.15$, $\alpha_2 = 0.8$, and $q = 1$. We take two time points in this series, a highly volatile and a quiet period. The first case is reported in Figure 7. In this case, where $h_{t+1} - \sigma_\varepsilon^2 = 4.2$, we see that the IMA-GARCH MSFEs produce too wide prediction intervals. Indeed, Figure 8 shows that IMA-GARCH prediction intervals cover almost 100% of the observations in the medium term, whereas the coverage of the LL-GARCH is always around the nominal.

In the second case, a quiet period where $h_{t+1} - \sigma_\varepsilon^2 = -0.46$, the IMA-GARCH MSFEs
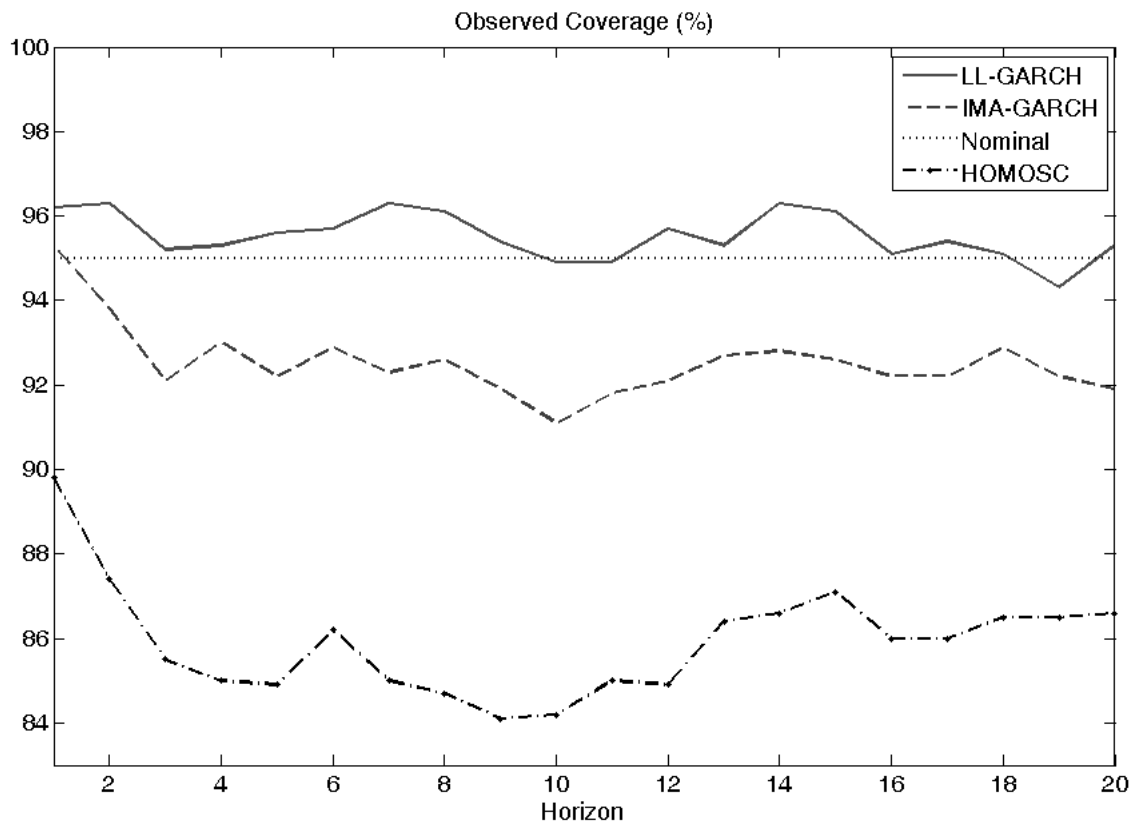
Figure 6: Observed coverage of the prediction intervals shown in Figure 5, computed as the percentage of trajectories within the prediction intervals.

produce too narrow prediction intervals, covering around 92% of the observations against a 95.8% of the LL-GARCH counterparts. Figure 10 plots the coverage in this last case.

In these two cases where ε_t is the only heteroscedastic component, it is worth to note that, as the shocks are purely transitory, the homoscedastic and the LL-GARCH prediction intervals stick to each other. However, the IMA-GARCH counterparts remain always above or below depending on the sign of the excess volatility. This leads to significant differences between the two prediction intervals, specially for medium and long term.

Summarizing, we have seen that the MSFE of the IMA-GARCH model may produce inaccurate prediction intervals when the series is generated by heteroscedastic unobserved component models. This is due to its incapacity of distinguishing whether the heteroscedasticity affects the long or the short run components. Therefore, the use of reduced form ARIMA models to construct prediction intervals may be inappropriate to capture the underlying uncertainty of the
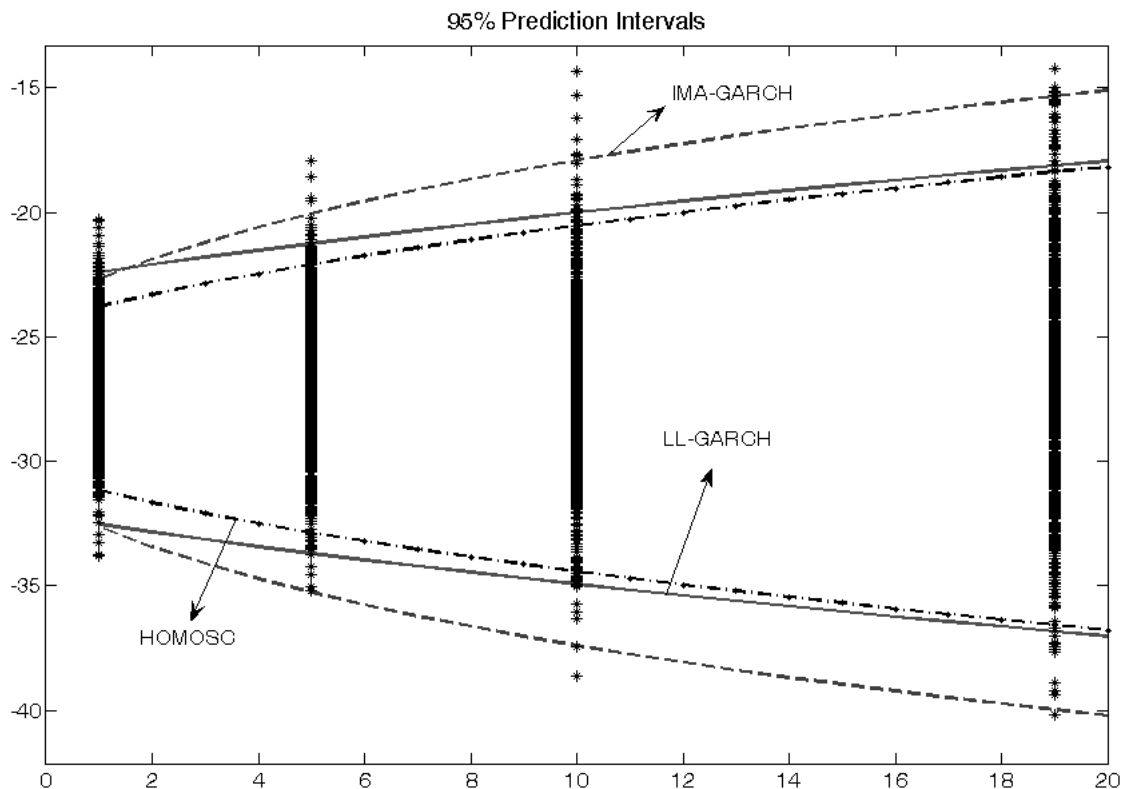
Figure 7: Prediction intervals of a LL-GARCH series where only the transitory component is heteroscedastic, with $\alpha_1 = 0.15$, $\alpha_2 = 0.8$, and $q = 1$. The time point is selected in a highly volatile period.

heteroscedastic components.

5. An empirical illustration

In this section we fit the LL-GARCH and IMA(1,1)-GARCH models to daily closing prices of the Pound/Euro exchange rate observed from January 3, 2000 to March 29, 2006 with T $=$ 1626[8]. The objective is to compare the conclusions about the conditional heteroscedasticity derived from the estimates of both models. We also compare the prediction intervals obtained when both models are used to forecast future values of returns. Figure 11 plots the observed prices, p_t, as well as the returns computed, as usual, as $y_t = 100 \times \Delta log(p_t)$. From the graphs we see that the exchange rates follow a non stationary pattern while returns are stationary with

---

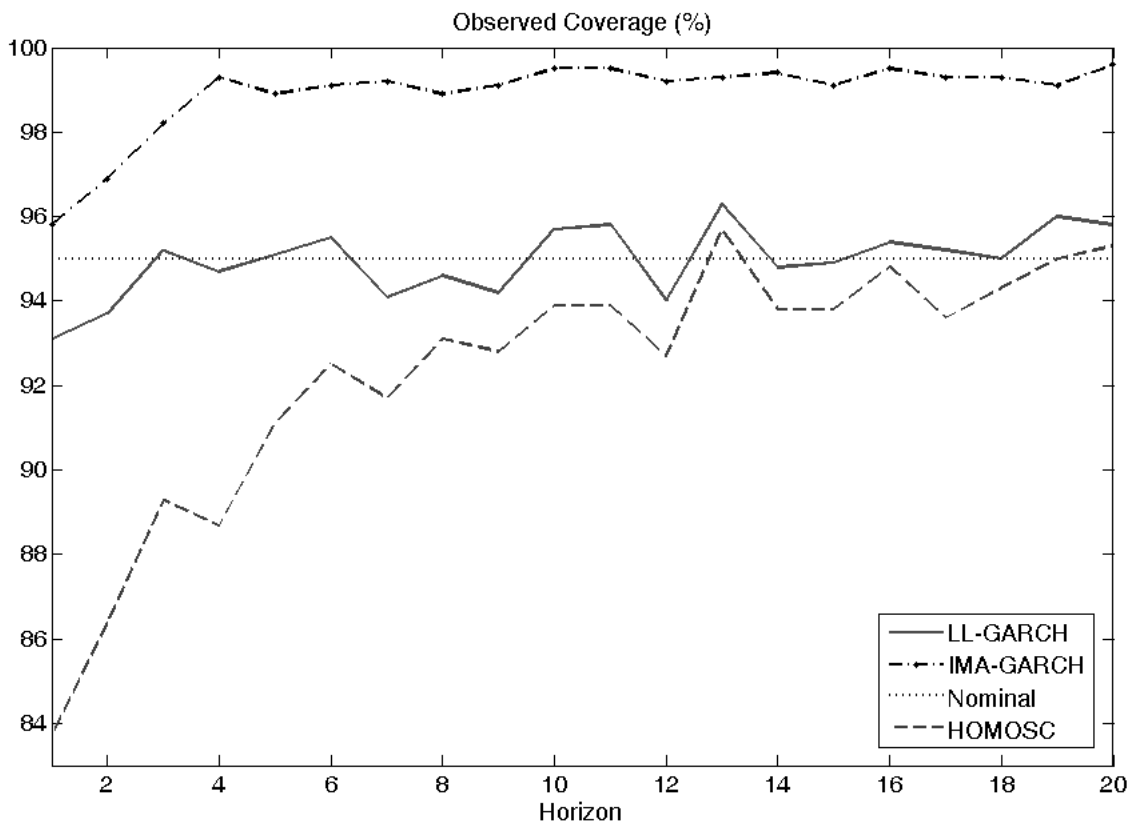[8]The series has been downloaded from the EcoWin database.

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

periods of clustered volatility. On the other hand, only the sample first order autocorrelation is significant (and negative), thus suggesting that the IMA(1,1) and/or the LL model may be appropriate models to the series. Furthermore, the sample kurtosis, 3.84, and the significant autocorrelations of squares may indicate the presence of conditional heteroscedasticity. In order to decide whether the heteroscedasticity may affect the transitory or the long-run component of the LL model, we first fit the homoscedastic model and then use the correlograms of the squared auxiliary residuals; see Broto and Ruiz (2006) for further details.[9]

Table 4 reports the estimates of the two homoscedastic models. We can see that the permanent component has more weight in the series as $\hat{q} > 1$. Furthermore, the estimate of θ (-0.247) is significant and almost identical to the one implied by $\hat{q}$ (-0.248). This confirms once more that

---

[9]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). The program found two outliers in the transitory component (ε_t) and three in the noise of the permanent component (η_t), representing just a 0.3% of the observations.
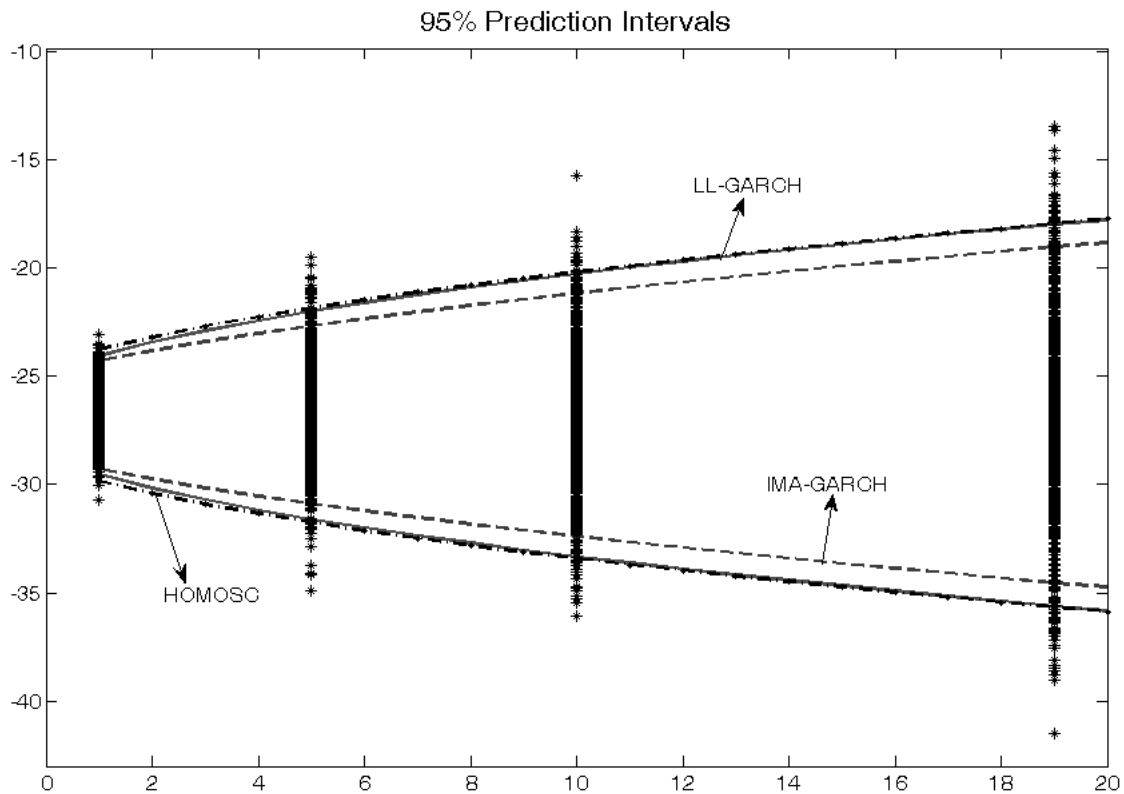
Figure 9: Prediction intervals of a LL-GARCH series where only the transitory component is heteroscedastic, with $\alpha_1 = 0.15$, $\alpha_2 = 0.8$, and $q = 1$. The time point is selected in a quiet period.

possible nonlinearities in the form of conditional heteroscedasticity do not affect the equivalence in the conditional mean of both models. Table 4 also reports the sample mean, skewness (SK), kurtosis (κ), and autocorrelations of the standardized one-step-ahead residuals, $\hat{v}_t$ in the LL model, and $\hat{a}_t$ in the IMA model. They are clearly uncorrelated suggesting that the two models seem to be appropriate to fit the conditional mean. However, when we look at the sample autocorrelations of the squares, it is clear the presence of a certain structure in the variance of the series that is not captured by the homoscedastic models. This fact, in conjunction with a significant excess kurtosis, leads to propose a GARCH process to account for this structure in the variance.

Table 4 also reports the sample moments and autocorrelations of the auxiliary residuals and their squares in the LL model. These values are useful tools to identify which of the components present evidence of conditional heteroscedasticity. Since by construction both auxiliary
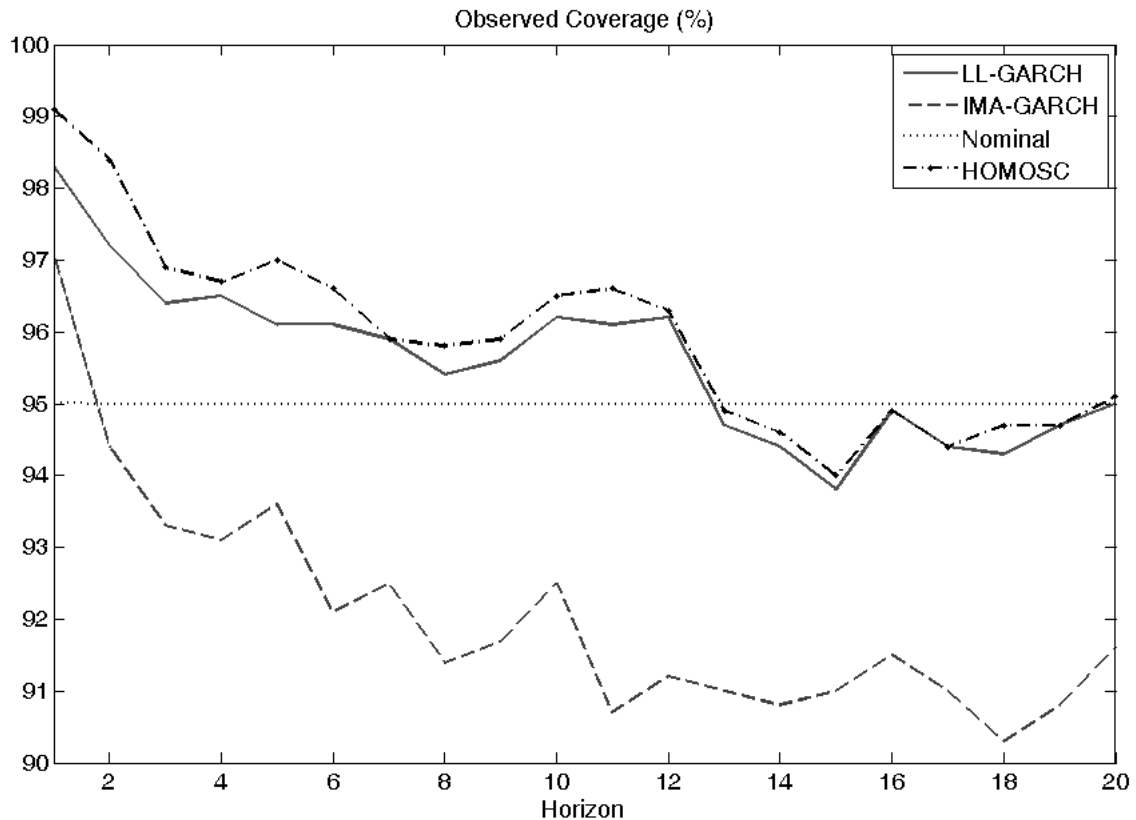
Figure 10: Actual coverage measured as the percentage of trajectories within the prediction intervals.

residuals are serially correlated, we have to find the sample autocorrelations of squares adjusted by the square of the sample acf to account for a significant nonlinear structure in the form of conditional heteroscedasticity. Figure 12 shows both adjusted correlograms. From them we can conclude that both components seem to have conditional heteroscedasticity, being the transitory component the one with the highest level of heteroscedasticity. Then, in the selection of the structural model that best captures the conditional heteroscedasticity, we should include a GARCH specification in both noises.

On the other hand, for the reduced form model selection, the sample autocorrelations given in Table 4 suggest that the IMA(1,1)-GARCH(1,1) could also be an adequate reduced form model to fit the series. Therefore, we fit the two models to the log of the daily £/€ exchange rate. Table 5 reports the estimation results.

The estimates of the LL-GARCH model given in Table 5 imply that both noises are condi-
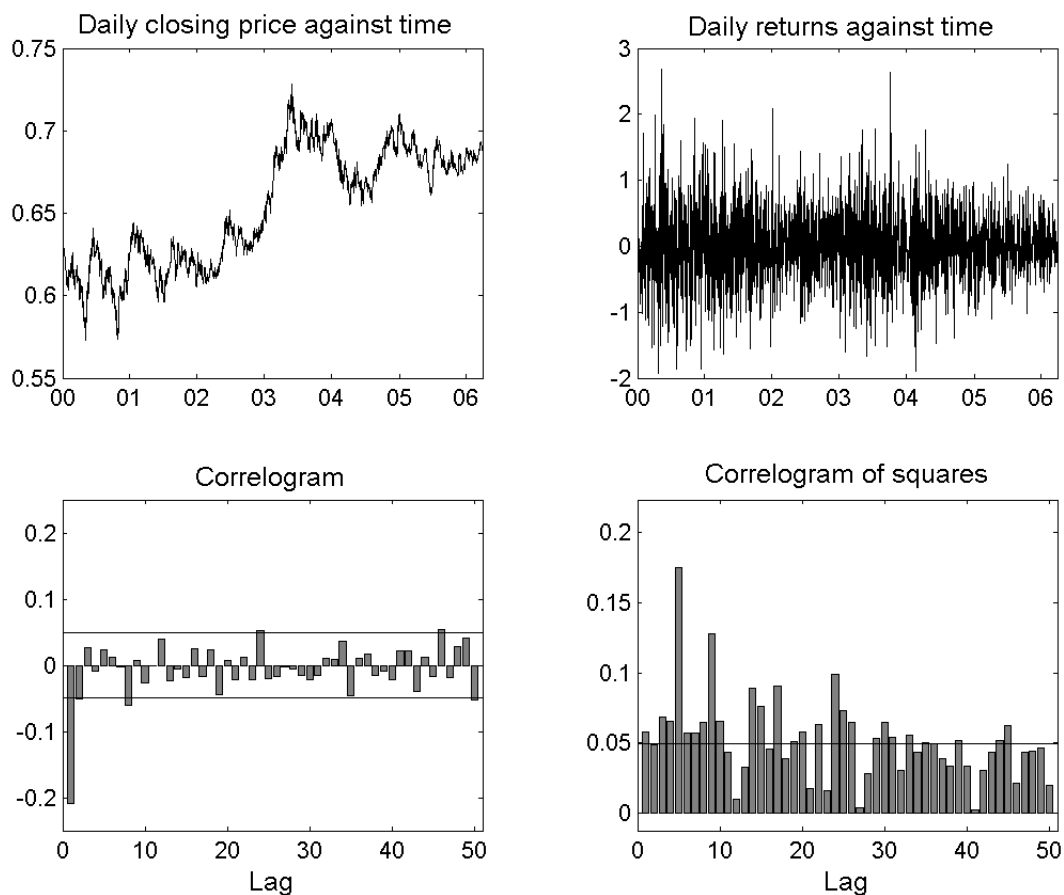
Figure 11: The daily £/€ exchange rate, from 01/2000 to 03/2006.

tionally heteroscedastic. Furthermore, as expected, the transitory component seems to be more heteroscedastic than the permanent one as $\hat{\alpha}_1 > \hat{\gamma}_1$, although both share almost the same persistence. On the other hand, compared to the homoscedastic specification, the introduction of a GARCH process in each noise increases the Log-Likelihood from -1439 to -1372. In addition to this, the residuals standardized by their estimated conditional variances not only are uncorrelated but also present almost no evidence of conditional heteroscedasticity (all the Q-statistics are insignificant at 1% and only one significant at 5%). We have to point out, however, that a small structure in the conditional variances of both noises still remains after fitting the GARCH process.

The fact that the structural LL model estimates not only each component but also its volatility, allows to decompose the total volatility of the series into the sum of the volatility

	Local Level model				IMA(1,1) model	
	$\hat{\sigma}_\varepsilon^2 =$	0.084**			$\hat{\sigma}_a^2 =$	0.344**
	$\hat{\sigma}_\eta^2 =$	0.196**			$\hat{\theta} =$	-0.247**
	$\hat{q} =$	2.338				
	$\hat{v}_t$	$\hat{\varepsilon}_t$	$\hat{\eta}_t$			$\hat{a}_t$
Mean	0.016	0.000	0.019	Mean		0.016
SK	0.205	-0.002	0.240	SK		0.201
κ	3.981	3.486	4.056	κ		3.962
ρ_1	0.012	-0.353**	0.244**	ρ_1		0.014
ρ_2	-0.044	-0.148**	0.022	ρ_2		-0.043
ρ_3	0.017	0.007	0.024	ρ_3		-0.017
ρ_4	0.002	-0.017	0.015	ρ_4		-0.002
ρ_5	0.028	0.023	0.032	ρ_5		0.028
ρ_{10}	-0.031	-0.021	-0.036	ρ_{10}		-0.031
$Q(10)$	15.434	247.2**	116.28**	$Q(10)$		15.489
	$\hat{v}_t^2$	$\hat{\varepsilon}_t^2$	$\hat{\eta}_t^2$			$\hat{a}_t^2$
ρ_1	0.038	0.189**	0.092**	ρ_1		0.038
ρ_2	0.029	0.078**	0.016	ρ_2		0.029
ρ_3	0.054*	-0.094**	0.035	ρ_3		-0.054*
ρ_4	0.052*	-0.080**	0.056*	ρ_4		-0.052*
ρ_5	0.170**	0.151**	0.154**	ρ_5		0.171**
ρ_{10}	0.065**	0.087**	0.079**	ρ_{10}		0.065**
$Q_1(10)$	365.44**	352.30**	600.90**	$Q_1(10)$		369.99**
$LogL$	-1439			$LogL$		-1437

Table 4: Estimates and sample moments of the residuals of the homoscedastic LL and IMA(1,1) models fitted to the £/€ exchange rate. *(**) Significant at 5% (1%) level.

of the transitory component (h_t) and the permanent component (q_t). Figure 13 plots these two volatilities against time, as well as the volatility of the reduced form disturbance, a_t. All these volatilities show a common pattern. First, at the beginning of the Euro as a common currency, the uncertainty about its behavior leads to a highly volatile period. Then it begins to
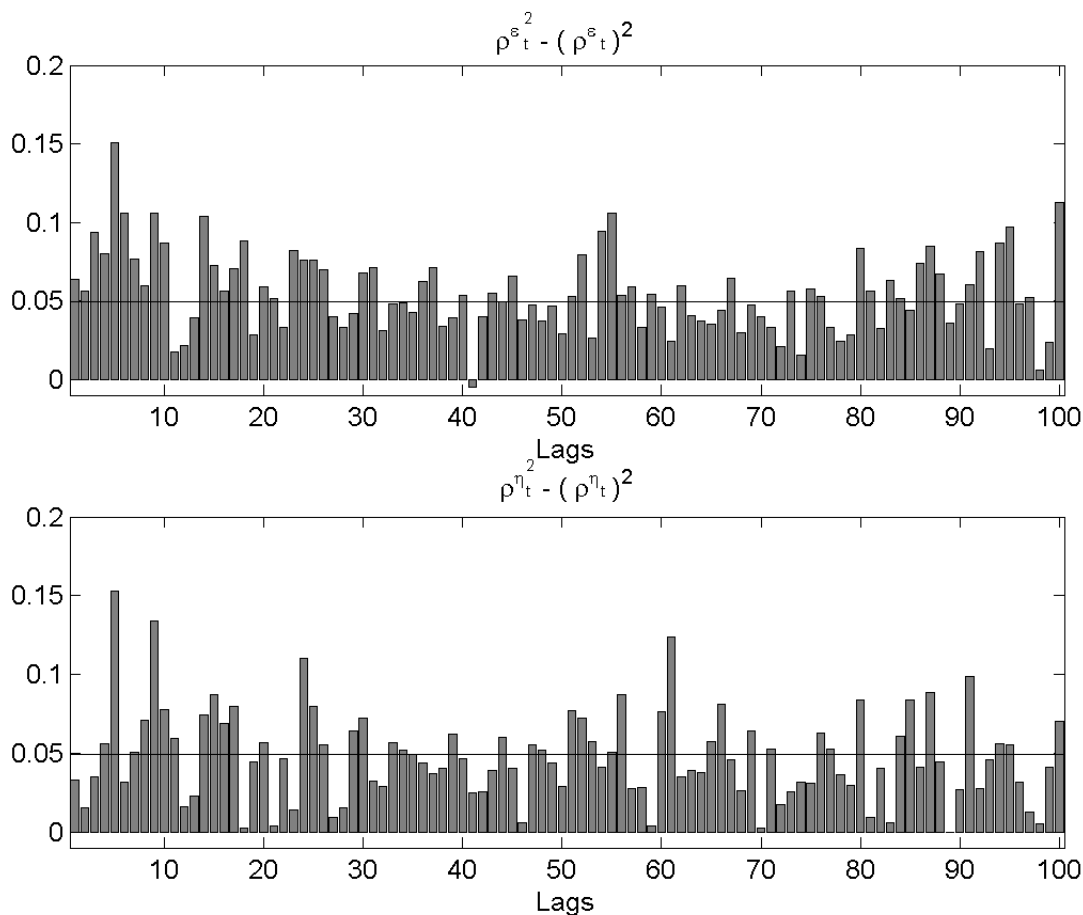
Figure 12: Correlogram of squares of the auxiliary residuals, corrected by the square of the sample autocorrelations.

decrease until the end of 2003, where there is a new increase of the uncertainty surrounding the exchange rate. Finally, the last two years show a smooth decreasing pattern. By construction, the information given by the reduced form disturbance, a_t, cannot provide any extra information about the sources of these highly volatile and quiet periods. However, from the volatility of both structural noises, we observe that the contribution of each component to the total volatility has been different. Thus, while the first highly volatile period is almost totally driven by the permanent component, the source of volatility in the second period is shared by the two components. However, we can see that the decreasing behavior of the last two years is accompanied by a gradual reduction in the contribution of the transitory component.

In reference to the IMA-GARCH model, we can see that the overall fit is almost identical to

Local Level - GARCH(1,1)				IMA(1,1) - GARCH(1,1)		
	Estimates	(t-stat)			Estimates	(t-stat)
$\hat{\alpha}_0 =$	1.00E-04	(10.2)		$\hat{\delta}_0 =$	0.001	(9.18)
$\hat{\alpha}_1 =$	0.072	(73.6)		$\hat{\delta}_1 =$	0.026	(4.27)
$\hat{\alpha}_2 =$	0.920	(934.1)		$\hat{\delta}_2 =$	0.971	(158.6)
$\hat{\gamma}_0 =$	3.00E-04	(20.0)				
$\hat{\gamma}_1 =$	0.036	(81.8)				
$\hat{\gamma}_2 =$	0.961	(2023.8)				
	$\hat{v}_t^{\dagger}$				$\hat{a}_t^{\dagger}$	
Mean	0.018			Mean	0.019	
SK	0.144			SK	0.129	
κ	3.350			κ	3.333	
ρ_1	0.003			ρ_1	0.009	
ρ_2	-0.033			ρ_2	-0.027	
ρ_3	0.006			ρ_3	0.005	
ρ_4	-0.012			ρ_4	-0.012	
ρ_5	0.031			ρ_5	0.035	
ρ_{10}	-0.025			ρ_{10}	-0.025	
$Q(10)$	11.160			$Q(10)$	10.930	
	$\hat{v}_t^{\dagger\,2}$				$\hat{a}_t^{\dagger\,2}$	
ρ_1	0.001			ρ_1	0.006	
ρ_2	0.011			ρ_2	0.006	
ρ_3	0.009			ρ_3	0.013	
ρ_4	0.004			ρ_4	0.002	
ρ_5	0.073**			ρ_5	0.069**	
ρ_{10}	0.001			ρ_{10}	0.001	
$Q_1(10)$	18.78			$Q_1(10)$	18.27	
$LogL$	-1372			$LogL$	-1374	

Table 5: Estimates of the LL-GARCH and IMA(1,1)-GARCH(1,1) models fitted to the £/€ exchange rate. The sample moments and correlograms reported refer to the residuals after being standardized by their estimated conditional variances. *(**) Significant at 5% (1%) level.
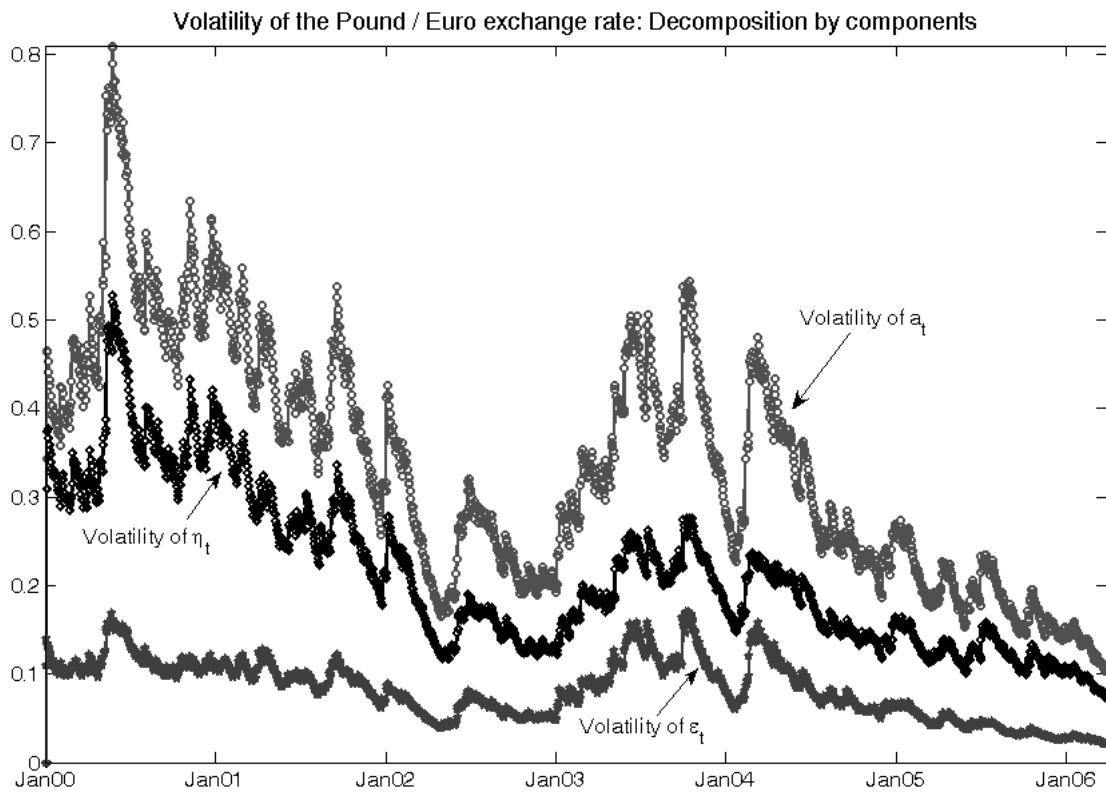
Figure 13: Estimated volatility of the daily £/€ exchange rate decomposed by its components.

the LL-GARCH model. Furthermore, it is also clear that the nonlinear model improves the fit in relation with the homoscedastic IMA model. However, as stated before, the estimated ARCH coefficient, $\hat{\delta}_1$ is smaller than both, $\hat{\alpha}_1$ and $\hat{\gamma}_1$, suggesting that the level of heteroscedasticity of the reduced form model is inferior compared to both components. Furthermore, the value of $\hat{\delta}_1$ is approximately equal to the implied δ_1 (0.029) obtained after plugging the estimates of the structural parameters into the equations considered in Section 3.

Finally, we construct prediction intervals at two particular time points of the series. From Figure 13, we choose them in a highly volatile and a quiet period. The former corresponds to the first part of our sample, while the latter corresponds to the end of the sample when the exchange rate volatility is lower. In both cases we set the prediction horizon to $k = 40$ which is approximately two months. In the first case we have re-estimated the parameters of both models with the information up to the selected time point to construct the prediction

intervals[10]. Figures 14 and 15 present the two cases. Note first that the prediction intervals resulting from the homoscedastic model significantly differ from those of the two nonlinear models. In the first case, the homoscedastic prediction intervals are too narrow because they cannot account for a conditional variance much higher than the marginal, as it is expected in highly volatile periods. The opposite happens in Figure 15, where the homoscedastic model produces too wide prediction intervals because they cannot capture the effect of a conditional variance being smaller than the marginal.
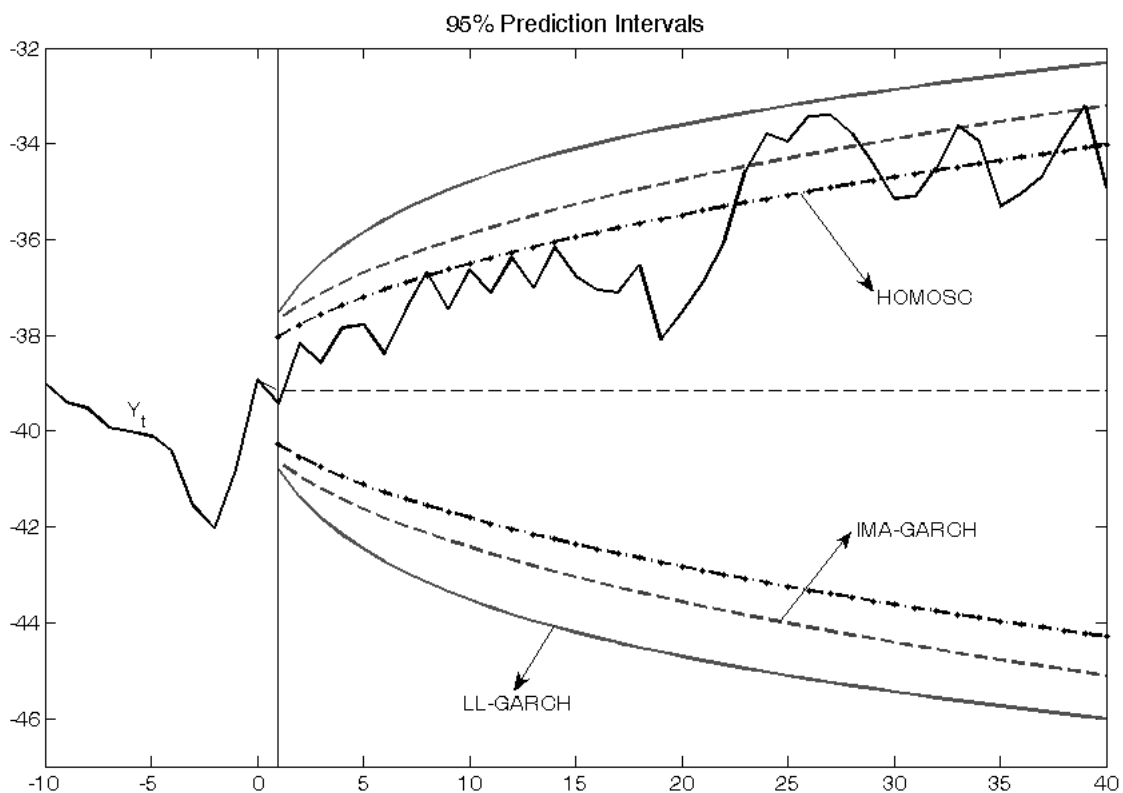

Figure 14: Prediction intervals in a highly volatile period (Jun-2000).

With regard to the IMA-GARCH and LL-GARCH prediction intervals, we observe that they are almost identical for $k = 1$, but then the LL-GARCH prediction intervals start to move away from those of the IMA-GARCH as the horizon increases. This is explained by the ability

---

[10]Since this case is at the beginning of the sample, we have added previous values of the exchange rate in order to have the same sample size as in the original sample. We have not observed significant changes in the estimations reported above.
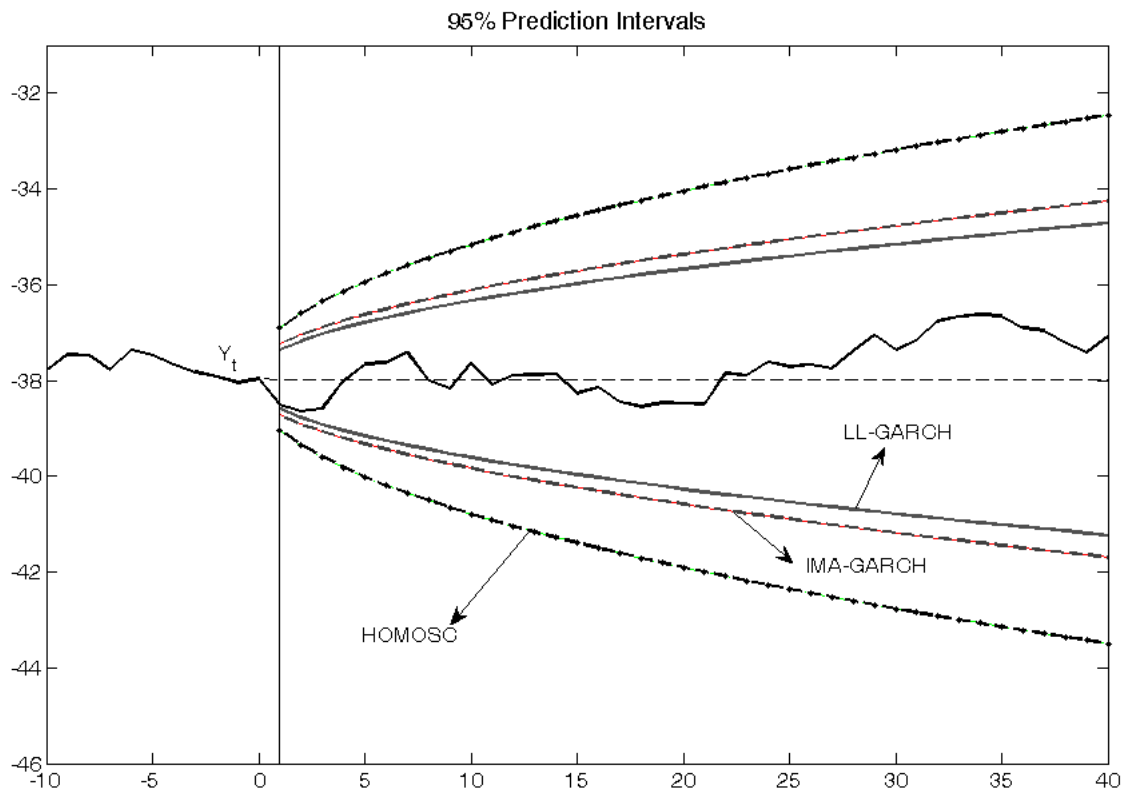
Figure 15: Prediction intervals in a quiet period (Feb-2006)

of the LL-GARCH model to distinguish whether the shocks come from the permanent or the transitory component. Thus, the highly volatile period of the first case is mainly due to shocks in the permanent component ($q_{t+1} - \sigma_\eta^2 = 0.4$ and $h_{t+1} - \sigma_\varepsilon^2 = 0.04$), so that a great proportion of the excess volatility does not vanish as k increases. Therefore, the MSFEs of the LL-GARCH take this into account and the resulting prediction intervals become wider than the corresponding intervals of the IMA-GARCH. We see a similar pattern in the quiet period. Provided that the negative excess volatility is mainly determined by the permanent component ($q_{t+1} - \sigma_\eta^2 = -0.16$ and $h_{t+1} - \sigma_\varepsilon^2 = -0.06$), the resulting LL-GARCH prediction intervals are the narrowest for all k.

6. Conclusions

In this paper we analyze the relationship between unobserved component models and their reduced form when the disturbances are GARCH(1,1) models. We show that although working with the ARIMA reduced form model is simpler because there is only one disturbance, its heteroscedasticity is weaker. Therefore, working with the unobserved component model may lead to the discovery of conditional heteroscedasticity that could not be apparent in the reduced form noise.

We also show that working with the unobserved component models generates more accurate prediction intervals when the heteroscedasticity only affects the short-run 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. In any case, the coverages obtained with the unobserved component models are closer to the nominal than those obtained in the ARIMA model.

Finally, the empirical application with the Pound-Euro exchange 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. With respect to the latter, it is likely that in other series, such as inflation rates, the divergences between the IMA-GARCH and the LL-GARCH prediction intervals may be important enough to be studied in depth.

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.

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, in press.*

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.

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

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. Wadhwani (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.

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.

Tsay, R. S. (2005). *Analysis of Financial Time Series* (2nd ed.). USA: Wiley.

Appendix

The relationship between $\kappa_{\Delta y}$, κ_ε and κ_η in non-Gaussian homoscedastic models

Within the framework of the LL model defined in (1)-(2), we derive the relationship between the kurtosis of Δy_t and the corresponding ones of the disturbances, ε_t and η_t. Let first consider that $\kappa_\varepsilon \geq 3$, $\kappa_\eta \geq 3$ and $\rho_1^{\varepsilon^2} = 0$. Define $\overline{\kappa} \equiv \max(\kappa_\varepsilon, \kappa_\eta)$. Then, we can show that

$$\kappa_{\Delta y} \;\; \leq \;\; \overline{\kappa}, \;\; \forall q \geq 0. \tag{A.1}$$

To prove (A.1), note from equation (9) that

$$
\begin{aligned}
\kappa_{\Delta y} - \overline{\kappa} \;\; &= \;\; \frac{1}{(q+2)^2} \left[q^2 \kappa_\eta + 12q + 6 + 2\kappa_\varepsilon \right] - \overline{\kappa} \\
&= \;\; \frac{1}{(q+2)^2} \left[q^2 \kappa_\eta + 12q + 6 + 2\kappa_\varepsilon - q^2 \overline{\kappa} - 4q\overline{\kappa} - 4\overline{\kappa} \right] \\
&= \;\; \frac{1}{(q+2)^2} \left[q^2 (\kappa_\eta - \overline{\kappa}) + 4q(3 - \overline{\kappa}) + 2(3 + \kappa_\varepsilon - 2\overline{\kappa}) \right].
\end{aligned}
\tag{A.2}
$$

Note that by assumption, $(\kappa_\eta - \overline{\kappa}) \leq 0$, $(3 - \overline{\kappa}) \leq 0$ and $(3 + \kappa_\varepsilon - 2\overline{\kappa}) \leq 0$[11]. Therefore, from (A.2) we obtain $\kappa_{\Delta y} - \overline{\kappa} \leq 0$, $\;\forall q \geq 0$. Note also in (A.2) that if at least one of the disturbances has excess kurtosis, i.e. $\overline{\kappa} > 3$, then one of these terms is strictly negative and therefore, $\kappa_{\Delta y} < \overline{\kappa}$ for all values of q except for the limiting cases, $q = 0$ and $q = \infty$.

On the other hand, if $\kappa_\varepsilon > 3$, $\kappa_\eta > 3$ and $\rho_1^{\varepsilon^2} = 0$, then

$$\kappa_{\Delta y} \;\; < \;\; \underline{\kappa}, \;\; \text{for some } q \geq 0, \tag{A.3}$$

where $\underline{\kappa} \equiv \min(\kappa_\varepsilon, \kappa_\eta)$. To prove statement (A.3), note that

$$\kappa_{\Delta y} - \underline{\kappa} \;\; = \;\; \frac{1}{(q+2)^2} \left[q^2 (\kappa_\eta - \underline{\kappa}) + 4q(3 - \underline{\kappa}) + 2(3 + \kappa_\varepsilon - 2\underline{\kappa}) \right]. \tag{A.4}$$

In this case, $(\kappa_\eta - \underline{\kappa}) \geq 0$ and $(3 - \underline{\kappa}) < 0$. However, the sign of $(3 + \kappa_\varepsilon - 2\underline{\kappa})$ depends on the values of κ_ε and κ_η. Thus, we can distinguish two cases.

Consider first that $\underline{\kappa} = \kappa_\eta$. Then, from (A.4) we find that

$$f(q) = \kappa_{\Delta y} - \kappa_\eta \;\; = \;\; \frac{1}{(q+2)^2} \left[4q(3 - \kappa_\eta) + 2(3 + \kappa_\varepsilon - 2\kappa_\eta) \right]. \tag{A.5}$$

---

[11]Note that $(3 + \kappa_\varepsilon - 2\overline{\kappa}) = (3 - \kappa_\varepsilon) + 2(\kappa_\varepsilon - \overline{\kappa}) \leq 0$

Note that $f(q)$ is nonlinear and continuous over $q \geq 0$, with $f(\infty) = 0$. Additionally, it is easy to prove that $f(q)$ has a global minimum at a certain value q^*, where $0 < q^* < \infty$. This implies that $f(q^*) < f(q)$ for all $q \geq 0$ and, in particular, $f(q^*) < f(\infty) = 0$. Consequently, there exists an interval for q within which $f(q) < 0$ or, in other words, $\kappa_{\Delta y} < \underline{\kappa}$.

If we now consider that $\underline{\kappa} = \kappa_\varepsilon$, we obtain from (A.4) that

$$f(q) = \kappa_{\Delta y} - \kappa_\varepsilon \quad = \quad \frac{1}{(q+2)^2} \left[q^2(\kappa_\eta - \kappa_\varepsilon) + 4q(3 - \kappa_\varepsilon) + 2(3 - \kappa_\varepsilon) \right], \quad (A.6)$$

where $f(q)$ is again continuous over $q \geq 0$. Moreover, we know that $f(0) = 0.5(3 - \kappa_\varepsilon) < 0$. Therefore, by property of continuous functions, there exists an interval for q within which $f(q) < 0$ and thus $\kappa_{\Delta y} < \underline{\kappa}$ for some $q \geq 0$.