Prediction intervals in conditionally heteroscedastic time series with stochastic components

Santiago Pellegrini[a,1], Esther Ruiz[b,c,*], Antoni Espasa[b,c,2]

[a] *Middle Office, Dirección de Mercados, Repsol YPF, Madrid, Spain*
[b] *Department of Statistics, Universidad Carlos III, Getafe 28903, Madrid, Spain*
[c] *Instituto Flores de Lemus, Universidad Carlos III, Getafe 28903, Madrid, Spain*

Abstract

Differencing is a very popular stationary transformation for series with stochastic trends. Moreover, when the differenced series is heteroscedastic, authors commonly model it using an ARMA-GARCH model. The corresponding ARIMA-GARCH model is then used to forecast future values of the original series. However, the heteroscedasticity observed in the stationary transformation should be generated by the transitory and/or the long-run component of the original data. In the former case, the shocks to the variance are transitory and the prediction intervals should converge to homoscedastic intervals with the prediction horizon. We show that, in this case, the prediction intervals constructed from the ARIMA-GARCH models could be inadequate because they never converge to homoscedastic intervals. All of the results are illustrated using simulated and real time series with stochastic levels.

1. Introduction

Economic and financial time series often have stochastic trends. In this case, following Box and Jenkins (1976), it is common to obtain a stationary transformation by taking differences, and then fit an ARMA model to the differenced series. Another characteristic which is often observed in this transformation is the evolution of its uncertainty, which is usually modelled by assuming GARCH errors. The corresponding ARIMA-GARCH model is then used to construct prediction intervals for future values of the original series. For example, Doornik and Ooms (2008) model UK prices and Soares and Medeiros (2008) model hourly electricity loads by estimating a GARCH model, after removing the trend

---

* Corresponding author at: Department of Statistics, Universidad Carlos III, c/ Madrid 126, Getafe 28903, Madrid, Spain. Tel.: +34 91 6249851; fax: 34 91 6249849.

E-mail addresses: santiago.pellegrini@repsol.com (S. Pellegrini), ortega@est-econ.uc3m.es (E. Ruiz), espasa@est-econ.uc3m.es (A. Espasa).

[1] Tel.: +34 91 2044307.
[2] Tel.: +34 91 6249803; fax: +34 91 6249849.

by taking differences. Even more recently, Bowden and Payne (2008) and Payne (2009) considered similar models for electricity prices and Thailand inflation, respectively.

In this paper, we analyze the effects of differencing conditionally heteroscedastic time series with stochastic trends on the performance of prediction intervals for the original observations. As an illustration, we consider that the series of interest, y_t, has a stochastic level, μ_t, characterized by a random walk and a transitory component, ε_t, which is a white noise.[3] Then, y_t is given by the local level model as follows

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

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

where ε_t and η_t are mutually independent and serially uncorrelated processes, with zero means and marginal variances σ_ε^2 and σ_η^2, respectively; see, for example, Durbin and Koopman (2001) for a detailed description and applications of this model. As we mentioned above, it is very popular to differentiate the non-stationary series y_t in order to obtain the stationary transformation

$$\Delta y_t = \eta_t + \Delta \varepsilon_t. \tag{2}$$

It is well known that the marginal variance and autocorrelations of Δy_t are the same as those of the following IMA(1,1) model:

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

where, if Δy_t is invertible, then $\theta = [(q^2 + 4q)^{1/2} - 2 - q]/2$, with $q = \sigma_\eta^2/\sigma_\varepsilon^2$ being known as the signal-noise ratio. Note that the parameter θ 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 $\sigma_a^2 = -\frac{\sigma_\varepsilon^2}{\theta}$. When Δy_t is conditionally heteroscedastic, the evolution of its uncertainty over time can be incorporated in the IMA(1,1) model in Eq. (3) by assuming that a_t is a GARCH(1,1) process. In this case, $a_t = a_t^\dagger \sigma_t$, where $a_t^\dagger$ is a white noise Gaussian process and

$$\sigma_t^2 = \delta_0 + \delta_1 a_{t-1}^2 + \delta_2 \sigma_{t-1}^2. \tag{4}$$

---

[3] The results for more complex models are similar to those reported in this paper.

However, note that it is obvious that the heteroscedasticity observed in Δy_t should come from the transitory component, ε_t, and/or from the long-run noise, η_t. When the long-run noise is heteroscedastic, the shocks to the variance of y_t are permanent, and consequently, the prediction intervals never converge to the prediction intervals based on the corresponding homoscedastic model. However, the effects of shocks to the transitory component are not permanent. Therefore, when the long-run noise is homoscedastic and only the transitory component is heteroscedastic, the prediction intervals for y_t should converge to the corresponding homoscedastic intervals with the prediction horizon. However, the prediction intervals constructed from the ARIMA-GARCH model, Eqs. (3)–(4), incorporate the unit root and never converge to the homoscedastic intervals. Therefore, they can be inadequate when the heteroscedasticity observed in Δy_t emerges only from the transitory component of y_t, i.e. when only ε_t is heteroscedastic.

Furthermore, it is important to note that when the heteroscedasticity only affects the transitory component of the series, the ARIMA-GARCH model fitted to y_t will not show any signs of misspecification, in the sense that the correlations between the residuals and the squared standardized residuals will not be significantly different from zero. Therefore, after fitting this model one may think that it is appropriate for obtaining prediction intervals for future values of y_t.

In this paper, we analyze the effects of differencing heteroscedastic time series with stochastic trends on the performance of prediction intervals for the original series.

The rest of this paper is structured as follows. Section 2 derives the prediction intervals for the local level model with GARCH disturbances and for the IMA-GARCH model. Section 3 reports the results of several Monte Carlo experiments which were carried out to analyze the performances of both prediction intervals. Section 4 contains an empirical application. Finally, Section 5 concludes the paper.

2. Prediction intervals

In this section, we derive expressions for the prediction intervals for future values of the series of interest when it has a stochastic level and is conditionally heteroscedastic. We consider the local

level model in Eq. (1), and, in order to incorporate conditional heteroscedasticity, assume that the errors are represented by GARCH(1,1) processes. Then, $\varepsilon_t = \varepsilon_t^\dagger h_t^{1/2}$ 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

$$h_t = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 h_{t-1}, \tag{5a}$$

$$q_t = \gamma_0 + \gamma_1 \eta_{t-1}^2 + \gamma_2 q_{t-1}, \tag{5b}$$

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; see Pellegrini, Ruiz, and Espasa (2010) for the relationship between the GARCH parameters in Eq. (5) and the GARCH parameters of the moving-average innovation in Eq. (4). Obviously, the lengths of the prediction intervals will be different, depending on whether the conditional variances of ε_t and η_t at the moment of forecasting are larger or smaller than the corresponding marginal variances.

The objective is to obtain prediction intervals for y_{T+k} given $\{y_1, y_2, \ldots, y_T\}$. The optimal point predictor of y_{T+k} that minimizes the mean squared forecast error (MSFE), denoted by $\hat{y}_{T+k}$, is its conditional mean, i.e. $\hat{y}_{T+k} = \mathrm{E}_T(y_{T+k})$, where the T under the expectation means that it is conditional on the information available at time T. Note that from Eq. (1), it is easy to see that $\mathrm{E}_T(y_{T+k}) = \mathrm{E}_T(\mu_{T+k}) = \mathrm{E}_T(\mu_T) \equiv \hat{\mu}_T$. Furthermore, from Eq. (5) it is possible to derive the following expressions for the variances of ε_{T+k} and η_{T+k} conditional on $\{y_1, y_2, \ldots, y_T\}$

$$
\begin{aligned}
\mathrm{E}_T(\varepsilon_{T+k}^2) &= \mathrm{E}_T[(\varepsilon_{T+k}^\dagger)^2]\mathrm{E}_T(h_{T+k}) \\
&= \alpha_0 + (\alpha_1 + \alpha_2)\mathrm{E}_T(h_{T+k-1}) \\
&= \alpha_0 \left[\frac{1 - (\alpha_1 + \alpha_2)^{k-1}}{1 - \alpha_1 - \alpha_2} \right] \\
&\quad + (\alpha_1 + \alpha_2)^{k-1}\mathrm{E}_T(h_{T+1}) \\
&= \sigma_\varepsilon^2 + (\alpha_1 + \alpha_2)^{k-1}(\hat{h}_{T+1} - \sigma_\varepsilon^2), \quad k \geq 1, \quad \text{(6a)}
\end{aligned}
$$

where $\hat{h}_{T+1} = \mathrm{E}_T(h_{T+1})$ and $\sigma_\varepsilon^2 = \frac{\alpha_0}{1-\alpha_1-\alpha_2}$. By analogy, it is straightforward to show that

$$
\begin{aligned}
\mathrm{E}_T(\eta_{T+k}^2) &= \sigma_\eta^2 + (\gamma_1 + \gamma_2)^{k-1}(\hat{q}_{T+1}^2 - \sigma_\eta^2), \\
& \qquad k \geq 1, \tag{6b}
\end{aligned}
$$

where $\hat{q}_{T+1} = \mathrm{E}_T(q_{T+1})$ and $\sigma_\eta^2 = \frac{\gamma_0}{1-\gamma_1-\gamma_2}$. Using the expressions given in Eq. (6), it is easy to show that

the MSFE of $\hat{y}_{T+k}$ is given by

$$
\begin{aligned}
\mathrm{MSFE}(\hat{y}_{T+k}) &= \mathrm{E}_T\left[(y_{T+k} - \hat{y}_{T+k})^2\right] \\
&= \mathrm{E}_T[(\mu_T + \eta_{T+1} + \cdots + \eta_{T+k} \\
&\quad + \varepsilon_{T+k} - \hat{\mu}_T)^2] \\
&= P_T^\mu + \sigma_\varepsilon^2 + k\,\sigma_\eta^2 \\
&\quad + \frac{1 - (\gamma_1 + \gamma_2)^k}{1 - (\gamma_1 + \gamma_2)}\,(\hat{q}_{T+1} - \sigma_\eta^2) \\
&\quad + (\alpha_1 + \alpha_2)^{k-1}\left(\hat{h}_{T+1} - \sigma_\varepsilon^2\right), \\
& \qquad k = 1, 2, \ldots, \tag{7}
\end{aligned}
$$

where $P_T^\mu = \mathrm{E}_T[(\mu_T - \hat{\mu}_T)^2]$. Note that the local level with GARCH disturbances (LL-GARCH henceforth) is not conditionally Gaussian, even when the standardized disturbances, $\varepsilon_t^\dagger$ and $\eta_t^\dagger$, are Gaussian. Therefore, it is not straightforward to obtain the conditional expectations involved in the MSFE in Eq. (7). Harvey, Ruiz, and Sentana (1992) show that, in practice, it is possible to obtain approximations of the quantities $\hat{\mu}_T$, P_T^μ, $\hat{h}_{T+1}$ and $\hat{q}_{T+1}$ in Eq. (7) by using an augmented version of the Kalman filter. This is the approximation adopted in this paper in order to obtain the MSFE of $\hat{y}_{T+k}$.

The expressions $(\hat{h}_{T+1} - \sigma_\varepsilon^2)$ and $(\hat{q}_{T+1} - \sigma_\eta^2)$ in Eq. (7) 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 Eq. (7). Furthermore, given that $\alpha_1 + \alpha_2 < 1$, the MSFE of $\hat{y}_{T+k}$ becomes a linear function of k for large forecasting horizons, with the same slope as its homoscedastic counterpart σ_η^2, but with a different intercept, due to the contribution of the fourth term in Eq. (7), i.e. the contribution of the long-run excess volatility. However, for short and medium horizons, the influence of the excess volatility in both noises leads to a MSFE either smaller or greater than that of the homoscedastic local level model, depending on the sign of the excess volatility.

Once $\hat{y}_{T+k}$ and its MSFE are available, one can obtain prediction intervals for y_{T+k} by assuming that the distribution of the k-step-ahead prediction errors is normal.[4] Therefore, approximated $(1-\alpha)\%$ prediction

---

[4] As we commented before, the prediction error distribution in unobserved component models with GARCH disturbances is

intervals for y_{T+k} are given by

$$\hat{\mu}_T \pm z_{\alpha/2} \sqrt{\text{MSFE}(\hat{y}_{T+k})}, \tag{8}$$

where $\text{MSFE}(\hat{y}_{T+k})$ is given by Eq. (7) and $z_{\alpha/2}$ is the $\alpha/2$ quantile of the standard normal density. On the other hand, when assuming homoscedasticity, $\alpha_1 = \alpha_2 = \gamma_1 = \gamma_2 = 0$, and the prediction intervals are given by

$$\hat{\mu}_T \pm z_{\alpha/2} \sqrt{P_T^\mu + \sigma_\varepsilon^2 + k\,\sigma_\eta^2}. \tag{9}$$

It is important to note that there is a significant difference in the behavior of the prediction intervals in Eq. (8), depending on whether the conditional heteroscedasticity affects the long- or short-run components. Excess volatility in the permanent component affects the MSFE for all horizons, while the effect of excess volatility in the transitory component vanishes in the long run. Therefore, when the heteroscedasticity only affects the transitory noise, i.e. $\gamma_1 = \gamma_2 = 0$, the prediction intervals in Eq. (8) converge to those of the homoscedastic model in Eq. (9). However, when the long-run component is heteroscedastic, depending on the sign of the excess volatility, the prediction intervals of the heteroscedastic local level model are wider or narrower than those obtained from the homoscedastic model for all prediction horizons. As an illustration, Fig. 1 plots the prediction intervals obtained for a series simulated by the local level model with parameters $\alpha_0 = 0.05$, $\alpha_1 = 0.10$, $\alpha_2 = 0.85$, $\gamma_1 = \gamma_2 = 0$ and $q = 1$; i.e., only ε_t is heteroscedastic. The point in time at which the prediction is made is selected in such a way that the excess volatility is positive. Assuming that the parameters are known and using the Kalman filter proposed by Harvey et al. (1992) to approximate the MSFE, we construct the prediction intervals for the LL-GARCH (in solid lines) as in Eq. (8), and for the homoscedastic model (in dash-dotted lines) as in Eq. (9). Note that the LL-GARCH model produces wider intervals for short horizons than the homoscedastic model, because the conditional variance is higher than the marginal. However, since the shock producing the positive excess volatility is transitory, the prediction intervals of the LL-GARCH stick to

---

not Gaussian. However, the results of Pascual, Romo, and Ruiz (2006) suggest that it could be approximated well by a Gaussian distribution.
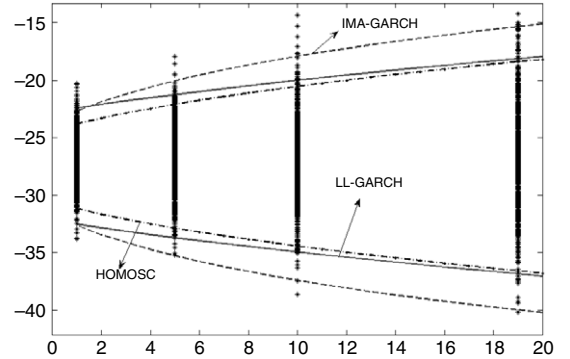

Fig. 1. 95% prediction intervals for a simulated series with a stochastic level and heteroscedasticity in the transitory component, with $\alpha_0 = 0.05$, $\alpha_1 = 0.10$, $\alpha_2 = 0.85$, and $q = 1$. The point in time at which the forecast is made is selected to be in a highly volatile period.

those of the homoscedastic model as k increases. In order to have a visual insight into the coverage of the prediction intervals for this particular series, Fig. 1 also plots possible trajectories for y_{T+k}, represented by the vertical clouds of points.

As we mentioned in the introduction, it is a very common practice to deal with stochastic trends by differencing the original series in order to obtain its stationary transformation. Then, after fitting an ARMA model to the differenced series, the conditional heteroscedasticity is modelled by fitting a GARCH model to the residuals. In the case of the LL-GARCH model, the resulting model is the IMA(1,1)-GARCH model (Eqs. (3)–(4)). Next, we derive the prediction intervals for y_t obtained when this is the methodology chosen to deal with the presence of unit roots in the data. In the IMA-GARCH model, assuming as usual that the within-sample innovations are observable, the optimal predictor of y_{T+k}, given the information available at time T, is given by

$$\hat{y}_{T+k} = y_T + \theta a_T, \quad k = 1, 2, \ldots, \tag{10}$$

with the MSFE given by the equation in Box I.

Once more, $(\sigma_{T+1}^2 - \sigma_a^2)$ is a measure of the excess volatility of a_t at the moment when the forecast is made. Note that the MSFE in Eq. (11) can also be separated into a linear and a nonlinear part, defined by the first and second terms, respectively. It is clear from Eq. (11) that as k increases, the $\text{MSFE}(\hat{y}_{T+k})$ is also a linear function of the horizon. Note that, as long as the excess volatility is different from zero,

$$
\begin{aligned}
&\mathrm{MSFE}(\hat{y}_{T+k}) \\
&= \begin{cases}
\sigma_a^2 + (\sigma_{T+1}^2 - \sigma_a^2), & k = 1 \\
\left[(1+\theta)^2(k-1) + 1\right]\sigma_a^2 & \\
\quad + \left[\dfrac{(1+\theta)^2 - (\delta_1 + \delta_2)^{k-1}\,(\theta(2+\theta) + \delta_1 + \delta_2)}{1 - (\delta_1 + \delta_2)}\right](\sigma_{T+1}^2 - \sigma_a^2), & k = 2, 3, \ldots.
\end{cases}
\end{aligned}
\tag{11}
$$

<div align="center">Box I.</div>

the path of MSFE in the IMA-GARCH model is always either above or below the path of the MSFE in the corresponding homoscedastic IMA model. This implies that the sign of the excess volatility at time T determines whether the IMA-GARCH prediction variance will be smaller or larger than the prediction variance of the homoscedastic IMA, for all prediction horizons. In this sense, the behavior is similar to that of the local level model with heteroscedastic long-run disturbances. However, when only the transitory component is heteroscedastic, the MSFEs still depend on the excess volatility at the moment when the prediction is made and do not converge to the corresponding homoscedastic intervals as they should, given that the heteroscedasticity is transitory.

As in the LL-GARCH model, k-step-ahead intervals based on the IMA-GARCH model can be obtained by assuming that the forecast errors are normally distributed for all values of k. In this case, the approximated $(1-\alpha)\%$ prediction intervals for the IMA-GARCH model are given by

$$
y_T + \theta\,a_T \pm z_{\alpha/2}\,\sqrt{\mathrm{MSFE}(\hat{y}_{T+k})}, \tag{12}
$$

where the $\mathrm{MSFE}(\hat{y}_{T+k})$ is given by Eq. (11). Finally, we can construct the prediction intervals for future values of y_{T+k} by assuming a homoscedastic IMA model. We do not consider these intervals further, as they are identical to those obtained using the homoscedastic local level model in Eq. (9).

Given that the reduced form IMA(1,1) model contains one unit root, the corresponding prediction intervals in Eq. (12) depend on the excess volatility at the moment when the prediction is made, for all prediction horizons, regardless of whether the conditional heteroscedasticity of Δy_t is due to the transitory or the long-run component of y_t (or both). When the long-run component is heteroscedastic, they are similar to those of the unobserved component

model in Eq. (8). However, when only the transitory component is heteroscedastic and the excess volatility is positive (negative), the multi-step prediction intervals based on the IMA(1,1) model will be too wide (narrow).

Going back to the illustration in Fig. 1, we also plot the IMA-GARCH prediction intervals (dashed lines) from Eq. (12). The values of θ and δ_0 in the IMA-GARCH model have been obtained using the functions that relate them to the signal-to-noise ratio, q, and the marginal variance, σ_ε^2, while the parameters of the GARCH process, δ_1 and δ_2, have been recovered from α_1 and α_2 following the procedure given by Pellegrini et al. (2010). By looking at the resulting intervals, we observe that they have almost the same length as those of the LL-GARCH for very short horizons, but become wider as k increases. This behavior is a consequence of taking the transitory shock as permanent, which means that the positive excess volatility leads to a higher MSFE and wider prediction intervals. In the next section we use simulated data to analyze whether the different patterns of the prediction intervals lead to significant differences in their coverages.

3. Forecasting performance

In order to analyze the performances of the prediction intervals constructed using the two alternative models considered in Section 2 for dealing with stochastic levels, we generate 1000 series of size $T = 1000$ with stochastic levels and with their component disturbances being GARCH processes, and construct 90% and 95% prediction intervals using the LL-GARCH, homoscedastic LL and IMA-GARCH models, as given in Eqs. (8), (9) and (12), respectively. We calculate the empirical coverages of these intervals by generating $B = 1000$ trajectories of y_{T+k}

Table 1

Mean Absolute Deviation (MAD) of the differences between the observed and nominal coverages for horizons $k = 1$, 6, 12 and 24, and for two confidence levels, 90% and 95%. The MAD is calculated in percentages. The series are simulated from the local level model with a GARCH(1,1) process in the transitory component.

	90%				95%			
	$k = 1$	$k = 6$	$k = 12$	$k = 24$	$k = 1$	$k = 6$	$k = 12$	$k = 24$
Parameters: $\alpha_1 = 0.10, \alpha_2 = 0.85, q = 1$								
Homoscedastic (true param)	3.164	1.210	0.837	0.784	2.186	0.867	0.633	0.560
LL-GARCH (true param)	1.575	0.885	0.742	0.773	1.104	0.630	0.571	0.549
IMA-GARCH (true param)	1.753	1.823	1.877	1.684	1.220	1.249	1.297	1.173
Homoscedastic (QML estim)	3.241	1.593	1.463	1.578	2.260	1.134	1.022	1.064
LL-GARCH (QML estim)	2.011	1.442	1.515	1.531	1.402	1.000	1.077	1.043
IMA-GARCH (QML estim)	2.179	2.079	2.176	2.104	1.483	1.387	1.491	1.409
Parameters: $\alpha_1 = 0.10, \alpha_2 = 0.85, q = 0.5$								
Homoscedastic (true param)	3.906	1.685	1.098	0.825	2.705	1.164	0.774	0.583
LL-GARCH (true param)	1.792	0.959	0.806	0.764	1.223	0.703	0.583	0.546
IMA-GARCH (true param)	1.930	1.933	2.019	1.934	1.306	1.332	1.392	1.280
Homoscedastic (QML estim)	3.940	2.034	1.735	1.683	2.698	1.395	1.210	1.179
LL-GARCH (QML estim)	2.064	1.505	1.541	1.635	1.442	1.067	1.088	1.134
IMA-GARCH (QML estim)	2.206	2.110	2.288	2.338	1.534	1.422	1.536	1.550

conditional on $\{y_1, y_2, \ldots, y_T, \mu_1, \mu_2, \ldots, \mu_T\}$, following the dynamics given by Eq. (1). That is, we fix the value of μ_T, and generate 1000 values of η_{T+1} from the GARCH model described above in order to find the trajectories of μ_{T+1}. Then, we obtain the trajectories of y_{T+1} by adding a draw of ε_{T+1} to each value of μ_{T+1}. This procedure is repeated in order to find the trajectories of y_{T+k} with $k > 1$.

We consider four designs, depending on whether the transitory or permanent components are heteroscedastic, and on the value of the signal-to-noise ratio. In the first two models, the transitory component is heteroscedastic while the long-run component is homoscedastic. Their parameters are $\alpha_0 = 0.05$, $\alpha_1 = 0.10$ and $\alpha_2 = 0.85$, with $q = 0.5$ in the first model and $q = 1$ in the second. In the last two models, the long-run component is heteroscedastic, with $\gamma_0 = 0.05$, $\gamma_1 = 0.10$, $\gamma_2 = 0.85$ and $q = 1$ in the third model and $q = 2$ in the last one. The prediction horizons considered are $k = 1$, 6, 12 and 24, and the nominal coverages are 90% and 95%.

Note that, in practice, the parameters needed to construct the prediction intervals in Eqs. (8), (9) and (12) should be estimated. Consequently, in order to analyze the effects of estimation on the performance of prediction intervals, we construct them, first, by assuming known parameters, and second, by

substituting the parameters with their QML estimates. To estimate the parameters of the heteroscedastic local level model, we follow the estimation approach proposed by Harvey et al. (1992). Tables 1 and 2 report the mean absolute deviation (MAD) of the observed coverages with respect to the nominal, for the two models with heteroscedasticity in the transitory component and the two models with a heteroscedastic long-run component, respectively. Consider first the results when the parameters are known. The first conclusion from both tables is that, regardless of the model and coverage considered, the deviations between the empirical and nominal coverages become smaller, in general, as the prediction horizon increases. Furthermore, note that for short horizons, $k = 1$, the coverages of the LL-GARCH and IMA-GARCH models have similar deviations from the nominal coverages, which are smaller than that observed in the homoscedastic case, in any case. One might expect that this deviation should be zero. However, remember that the future values of y_{T+k} were generated by an approximation of its conditional distribution, not by the conditional distribution itself, which is unknown. Furthermore, the Kalman filter for the LL-GARCH model only yields an approximated MSFE.

We now compare the results reported in Table 1 with those in Table 2. Table 1 reports the results for

Table 2

Mean Absolute Deviation (MAD) of the differences between the observed and nominal coverages for horizons $k = 1, 6, 12$ and 24, and for two confidence levels, 90% and 95%. The MAD is calculated in percentages. The series are simulated from the local level model with a GARCH(1,1) process in the permanent component.

	90%				95%			
	$k = 1$	$k = 6$	$k = 12$	$k = 24$	$k = 1$	$k = 6$	$k = 12$	$k = 24$
Parameters: $\gamma_1 = 0.10, \gamma_2 = 0.85, q = 1$								
Homoscedastic (true param)	2.325	3.606	3.411	2.804	1.607	2.421	2.310	1.850
LL-GARCH (true param)	1.624	2.631	2.497	2.087	1.132	1.775	1.711	1.400
IMA-GARCH (true param)	1.715	2.933	2.813	2.384	1.193	1.946	1.905	1.552
Homoscedastic (QML estim)	2.387	3.723	3.656	3.236	1.659	2.570	2.5210	2.216
LL-GARCH (QML estim)	1.858	3.000	2.985	2.727	1.299	2.095	2.098	1.945
IMA-GARCH (QML estim)	1.976	3.144	3.139	2.804	1.384	2.195	2.220	2.013
Parameters: $\gamma_1 = 0.10, \gamma_2 = 0.85, q = 2$								
Homoscedastic (true param)	2.995	3.755	3.471	2.808	2.060	2.527	2.262	1.868
LL-GARCH (true param)	1.969	2.483	2.299	1.904	1.355	1.698	1.524	1.286
IMA-GARCH (true param)	2.016	2.715	2.542	2.119	1.384	1.833	1.668	1.421
Homoscedastic (QML estim)	3.077	3.972	3.668	3.159	2.124	2.689	2.489	2.142
LL-GARCH (QML estim)	2.241	2.933	2.811	2.612	1.551	2.040	1.976	1.833
IMA-GARCH (QML estim)	2.320	3.077	2.974	2.648	1.600	2.165	2.091	1.918

the two models in which the long-run component is homoscedastic and the heteroscedasticity only appears in the transitory component. In this case, regardless of whether the parameters are known or estimated, we can observe that the homoscedastic models only have the largest deviations between the empirical and nominal coverages when predicting in the short run for $k = 1$. However, when the prediction horizon is longer, constructing the intervals using the IMA model leads to the worst coverages. This is due to the inability of the IMA-GARCH model to represent the dynamics of series in which only the transitory component is heteroscedastic. Also note that when the forecast horizon is large ($k = 24$), the deviations between the empirical and nominal coverages of the homoscedastic and LL-GARCH models are similar. Obviously, when the parameters are estimated, we observe larger deviations. However, the conclusions about the comparisons among intervals are the same as those obtained assuming known parameters.

We now consider the results reported in Table 2 for the models in which the short-run component is homoscedastic and the long-run component is heteroscedastic. In this case the homoscedastic prediction intervals are worse than any of the intervals constructed using the heteroscedastic models for all prediction horizons, regardless of whether the

parameters are known or estimated. Furthermore, comparing the heteroscedastic intervals between them, we observe that their MADs are very similar, in particular when the parameters are estimated. Therefore, when the long-run component is heteroscedastic, constructing the prediction intervals using the LL-GARCH model leads to only slight improvements in the performance of the prediction intervals with respect to constructing them using the simpler IMA-GARCH models. Once more, the deviations between the empirical and nominal coverages are only slightly larger when the parameters are estimated.

The information in Tables 1 and 2 allows us to compare the alternative models in terms of which generates the smallest deviations between the nominal and empirical coverages. However, it does not contain any information on the sign of these deviations. Therefore, we cannot conclude whether or not we are obtaining intervals that cover more or less than the nominal coverage. Obviously, the sign of the deviations depends on the sign of the excess volatility at the time the prediction is made. Consequently, we also compute the mean coverage of each model and prediction horizon, conditional on whether the excess volatility is positive or negative. Fig. 2 plots the average empirical coverages against the horizon
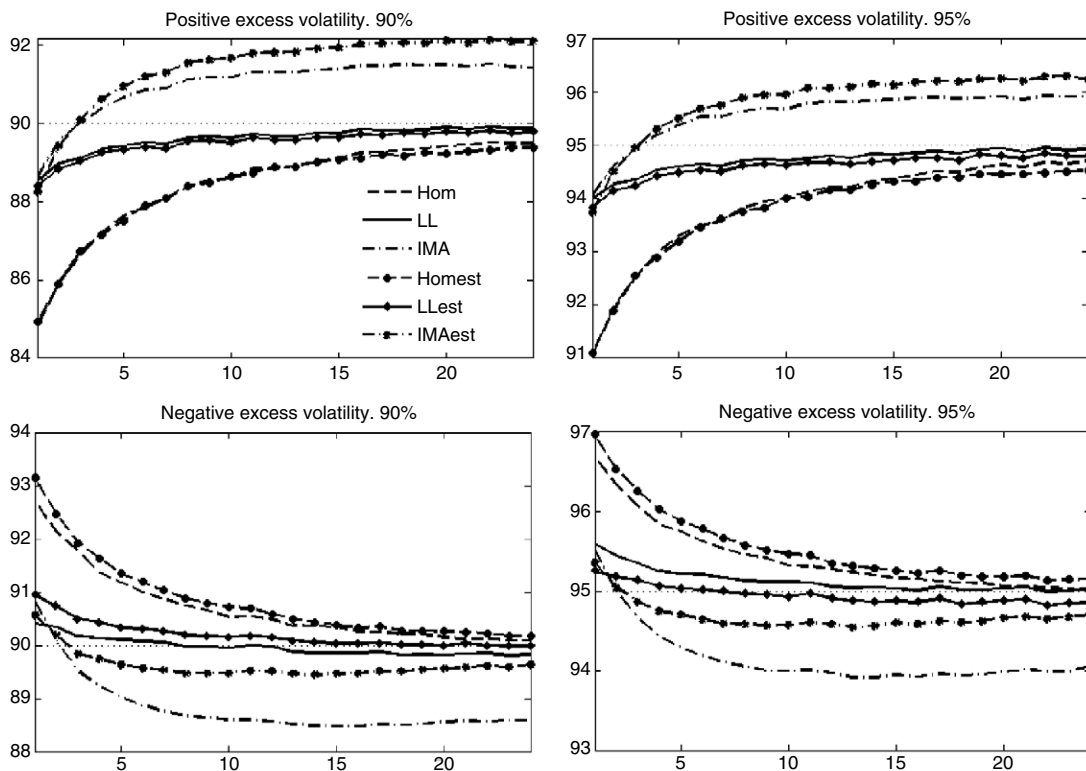
Fig. 2. Mean observed coverage in volatile (positive excess volatility) and quiet (negative excess volatility) periods. The series are generated from the local level model with a GARCH(1,1) noise in the transitory component only. The parameters are given by $\alpha_1 = 0.10$, $\alpha_2 = 0.85$ and $q = 0.5$.

k, computed both in a volatile period, when the marginal variance is smaller than the conditional (top row), and in a quiet period in which the marginal is larger than the conditional (bottom row) for the model in which the short-run component is heteroscedastic and $q = 0.5$. Fig. 2 shows that in both periods, the coverages of the homoscedastic model tend toward the nominal when the prediction horizon increases. However, in the short-run, the coverages of the homoscedastic prediction intervals are smaller (larger) than the nominal when the excess volatility is positive (negative). The short-run coverages of the two heteroscedastic intervals are closer to the nominal than those of the homoscedastic intervals, although there still exists a gap between the empirical and nominal coverages. On the other hand, the long-run coverages of the IMA-GARCH intervals are well above (below) the nominal when the excess volatility is positive (negative). As we mentioned above, this

model incorporates the unit root, and it cannot cope with the fact that only the transitory component is heteroscedastic. Finally, the average coverages of the LL-GARCH intervals are close to the nominal for all of the prediction horizons considered.

Finally, Fig. 2 illustrates that, for the large sample size considered in this paper, obtaining prediction intervals using estimated parameters implies slightly larger deviations of the empirical with respect to the nominal coverages. This is an interesting result, given the increased number of parameters that need to be estimated in this model. It seems that when the heteroscedastic component is identified correctly, the parameter estimation uncertainty does not have any impact on the prediction intervals. A further point which is worth making in relation to estimation is that estimating the parameters produces IMA-GARCH intervals with coverages which are closer to (further from) the nominal when the period of forecasting is quiet (volatile) than those obtained
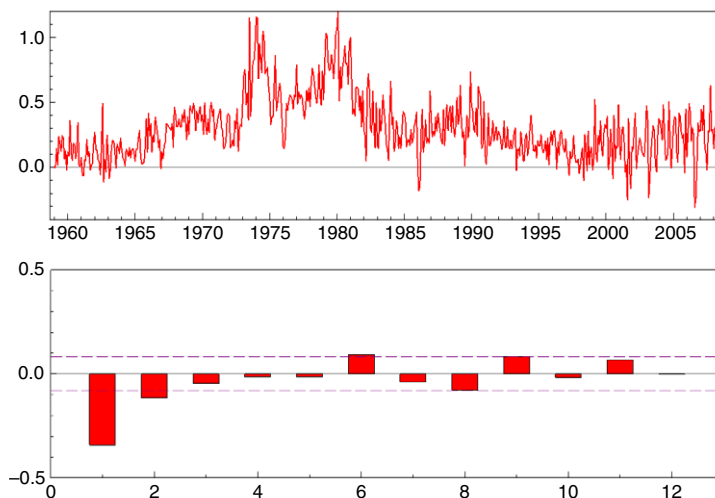
Fig. 3. Seasonally adjusted monthly US inflation rate observed from February 1959 to May 2008 (top panel), and sample autocorrelations of its first differences (bottom panel).

when the parameters are known. This is due to the fact that when the prediction is made in a quiet period, the IMA-GARCH intervals tend to be narrower than the nominal coverage. However, estimating the parameters generates wider intervals, which obviously have coverages which are closer to the nominal. On the other hand, when forecasting in a volatile period, estimating the parameters yields still wider intervals, which produce coverages which are even further from the nominal than when the parameters are known.

In this section we have generated the series using the LL-GARCH model, and the prediction intervals have then been constructed by: (i) using the augmented Kalman filter to approximate the conditional distribution of future observations; (ii) fitting the IMA-GARCH model; and (iii) fitting the corresponding homoscedastic model. One might think that we should also generate the series using the IMA-GARCH model and then obtain the three prediction intervals just described. However, given the results already reported, it is clear that the IMA-GARCH model implies a conditional heteroscedasticity with permanent effects. Therefore, the LL-GARCH model with conditionally heteroscedastic long-run noise will generate prediction intervals with coverages close to the nominal. Furthermore, when we differentiate a time series, it is because the original observations have stochastic trends. Therefore, it seems rather distant from the situation faced in reality to generate the dif-

ferenced series and then obtain the original observations from it.

4. An empirical illustration

In this section, we construct prediction intervals of the seasonally adjusted monthly US inflation rate.[5] The period analyzed spans 49 years, from February 1959 to May 2008, and thus contains $T = 592$ observations. We leave the last $R = 90$ observations (from January 2001 to May 2008) for the out-of-sample forecasting exercise. Fig. 3 plots the inflation series and the correlogram of its first differences, where the first order autocorrelation is negative and significant. Therefore, the inflation series can be represented well by an IMA(1,1) model with $\theta < 0$. Alternatively, the dynamic dependence of the monthly inflation can be explained by the local level model. In consequence, we fit both models. The estimation results are given in Table 3, together with the results of testing for homoscedasticity in the residuals. In the case of the IMA(1,1) model, we

---

[5] More specifically, the inflation rate is defined as the log-difference of the monthly personal consumption expenditure (PCE) deflator, multiplied by 100 in order to obtain percentage rates. The series was downloaded from the EcoWin database. An intervention analysis of the series using auxiliary residuals (see Harvey & Koopman, 1992) was carried out using the program STAMP 6.20 from Koopman, Harvey, Doornik, and Shephard (2000).

Table 3

Estimation results for the US inflation rate from February 1959 to December 2000. The top rows report the results for the homoscedastic local level and IMA(1,1) models, while the bottom rows report the estimation results for the LL-GARCH and IMA-GARCH models. The statistics $Q_\varepsilon(8)$ and $Q_\eta(8)$ are the lag-8 Box-Pierce statistics for joint significance of the differences between the autocorrelations of squares and the squared autocorrelations of the auxiliary residuals, while $Q_a(8)$ is the same statistic applied to the squared IMA residuals.

LL model	IMA(1,1) model
$\hat{\sigma}_\varepsilon^2 = 16.07 \times 10^{-3}$ **	$\hat{\sigma}_a^2 = 21.74 \times 10^{-3}$ **
$\hat{\sigma}_\eta^2 = 1.49 \times 10^{-3}$ **	$\hat{\theta} = -0.738$ **
$Q_\varepsilon(8) = 52.63$ **	$Q_a(8) = 44.84$ **
$Q_\eta(8) = 70.82$ **	
LL-GARCH(1,1) model	IMA(1,1)-GARCH(1,1) model
$\hat{\alpha}_0 = 1.30 \times 10^{-3}$ **	$\hat{\delta}_0 = 1.32 \times 10^{-3}$ **
$\hat{\alpha}_1 = 0.193$ **	$\hat{\delta}_1 = 0.142$ **
$\hat{\alpha}_2 = 0.738$ **	$\hat{\delta}_2 = 0.804$ **
$\hat{\sigma}_\eta^2 = 1.06 \times 10^{-3}$ **	$\theta = -0.775$ **

** indicates significance at the 1% level.

test for the joint significance of the autocorrelations of squared residuals by using the McLeod and Li (1983) test. Given that these autocorrelations are significant, we fit an IMA(1,1)-GARCH(1,1) model to the inflation series; see Doornik and Ooms (2008) and Payne (2009) for other authors who fit ARIMA-GARCH models for modelling inflation. In order to identify which component in the local level model is heteroscedastic, we compute the differences between the autocorrelations of squares and the squared autocorrelations of the auxiliary residuals and test for the joint significance of these differences; see Broto and Ruiz (2009). According to this analysis, and given that the variance of the permanent component is very small relative to that of the transitory component ($\hat{q} = 0.092$), we assume that the permanent component noise, η_t, is homoscedastic, and include GARCH(1,1) effects only in the transitory component, ε_t; see Stock and Watson (2007) for a local level model with heteroscedastic noises for US inflation. The estimation results of the heteroscedastic models appear in the bottom row of Table 3. Note that the estimates of q in the LL models and θ in the IMA models are very similar, regardless of whether we fit homoscedastic or heteroscedastic errors. These two parameters only depend on the marginal variances.
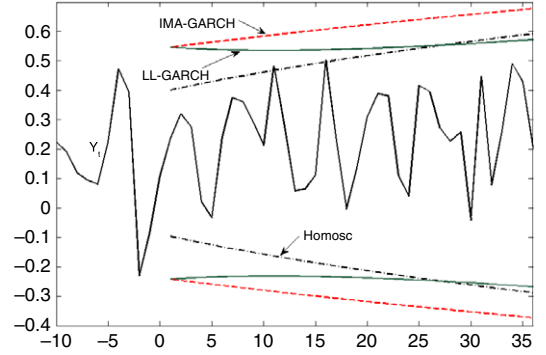

Fig. 4. 90% prediction intervals of the US inflation rate, obtained in June 2003.

Conditional on the estimated parameters of each model, we find the approximated $(1 - \alpha)\%$ prediction intervals of the two heteroscedastic and the homoscedastic models given in Eqs. (8), (9) and (12) for horizons running from $k = 1$ to $k = 36$. We consider $\alpha = 5\%$ and 10%. Finally, we re-estimate the three models for the whole out-of-sample period, from January 2001 to May 2008, by adding one new observation at a time and computing the prediction intervals again. Therefore, the results below are based on $90 - k + 1$ predictions of y_{T+k}. As an illustration, Fig. 4 plots the approximated 90% prediction intervals obtained in a period of high volatility, in June 2003. Note that the homoscedastic model produces narrower prediction intervals for small values of k, because it cannot capture the positive volatility shock. However, as the horizon increases, the intervals have almost the same length as those generated by the LL-GARCH model. On the other hand, the IMA-GARCH prediction intervals have the same length as those of the LL-GARCH in the short-term, but widen as the horizon increases because the model treats the shock as permanent.

To measure the accuracies of the coverages of each of the three prediction intervals considered, Table 4 reports the percentages of observations lying within the intervals, for horizons $k = 1$, 3, 6, 12 and 24. These prediction horizons have been selected because they are relevant, in the sense that they represent a month, a quarter, half a year, and one and two years ahead. We observe that when the nominal coverage is 90% and the prediction horizon is $k = 1$, 3, or 6, the coverage of the homoscedastic intervals is smaller than the nominal, while the two

Table 4
Empirical coverage of the US inflation rate, measured as the percentage of observations lying within the 90% and 95% prediction intervals.

Horizon	Homoscedastic		LL-GARCH(1,1)		IMA(1,1)-GARCH(1,1)	
	90%	95%	90%	95%	90%	95%
$k = 1$	81.11	88.89	88.89	91.11	90.00	91.11
$k = 3$	82.95	87.50	88.64	92.05	87.50	92.05
$k = 6$	84.71	91.76	89.41	94.12	90.59	94.12
$k = 12$	91.14	92.41	91.14	94.94	92.41	96.20
$k = 24$	95.52	97.01	94.03	95.52	94.03	98.51
$k = 36$	98.18	100.00	98.18	98.18	98.18	100.00

heteroscedastic models have similar coverages, close to 90%. However, for longer horizons, the three models generate intervals with very similar coverages, which are clearly larger than the nominal 90%. On the other hand, the conclusions are similar when the nominal coverage is 95%, although the coverages of the LL-GARCH intervals are slightly closer to the nominal, even when $k = 12$ or 24. Therefore, estimating the conditionally heteroscedastic models improves the coverage, especially for short and medium horizons. As the estimated value of the signal-to-noise ratio, q, is very small (around 0.1), we cannot expect big differences between the LL and IMA models (in the limiting case where $q = 0$, the local level component model collapses into a white noise process). However, estimating the conditionally heteroscedastic component directly seems to work better for medium to long horizons.

5. Conclusions

In this paper, we analyze the effects of differencing conditionally heteroscedastic time series with stochastic trends on the performances of prediction intervals for the original observations. In particular, we consider a conditionally heteroscedastic model with a stochastic level. We show that, in this model, if the long-run component is heteroscedastic, the corresponding ARIMA-GARCH prediction intervals have properties which are similar to those of the conditionally heteroscedastic unobserved component model for all prediction horizons. On the other hand, when the long-run disturbance is homoscedastic and the transitory component is the only heteroscedastic component, since the shocks to the variance are purely transitory. Consequently, the prediction intervalsbased on the corresponding homoscedastic and heteroscedastic models stick to each other for long prediction horizons. However, depending on the sign of the excess volatility, the ARIMA-GARCH counterparts may be wider or narrower than the intervals obtained using the corresponding unobserved component model for long prediction horizons. This is due to its incapacity to distinguish whether the heteroscedasticity affects the long or short run components, and may lead to significant differences between the two prediction intervals, especially for the medium to long term. When the prediction horizon is small, the simpler ARIMA-GARCH model generates prediction intervals with coverages similar to those of the conditionally heteroscedastic unobserved component intervals.

It is also important to mention that when the conditional heteroscedasticity is transitory, the incorrect prediction intervals from the ARIMA-GARCH models are the result of a model which does not appear to be misspecified when looking at the usual residual diagnostics. Therefore, it seems as though one goal for further research could be to analyze the way in which the properties of the prediction intervals can be used to identify the source of the heteroscedasticity. In this way, it may be possible to overcome some of the problems pointed out by Broto and Ruiz (2009) which are faced when trying to identify the source of the heteroscedasticity by using auxiliary residuals. Furthermore, our results also suggest that it is worth putting more effort into the identification and estimation of conditionally heteroscedastic unobserved component models. In any case, we have seen that, for moderate sample sizes, if the source of the heteroscedasticity is identified correctly, parameter estimation uncertainty does not have an impact on the prediction intervals. These results are illustrated using simulated data and a real time series of the US monthly inflation rate.

References

Bowden, N., & Payne, J. E. (2008). Short term forecasting of electricity prices for MISO hubs: evidence from ARIMA-EGARCH models. *Energy Economics*, *30*, 3186–3197.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: forecasting and control*. San Francisco: Holden-Day.

Broto, C., & Ruiz, E. (2009). Testing for conditional heteroscedasticity in the components of inflation. *Studies in Nonlinear Dynamics and Econometrics*, *13*, Article 4.

Doornik, J. A., & Ooms, M. (2008). Multimodality in the GARCH regression model. *International Journal of Forecasting*, *24*, 432–448.

Durbin, J., & Koopman, S. J. (2001). *Time series analysis by state space methods*. Oxford, UK: Oxford University Press.

Harvey, A. C., & Koopman, S. (1992). Diagnostic checking of unobserved components time series models. *Journal of Business and Economic Statistics*, *10*, 377–389.

Harvey, A. C., Ruiz, E., & Sentana, E. (1992). Unobserved component time series models with ARCH disturbances. *Journal of Econometrics*, *52*, 129–157.

Koopman, S. J., Harvey, A. C., Doornik, J. A., & Shephard, N. (2000). *STAMP: structural time series analyser, modeller and predictor*. London: Timberlake Consultants Press.

McLeod, A. I., & Li, W. K. (1983). Diagnostic checking ARMA time series models using squared-residual autocorrelations. *Journal of Time Series Analysis*, *4*, 269–273.

Pascual, L., Romo, J., & Ruiz, E. (2006). Bootstrap prediction for returns and volatilities in GARCH models. *Computational Statistics and Data Analysis*, *50*, 2293–2312.

Payne, J. E. (2009). Inflation targeting and the inflation-inflation uncertainty relationship: evidence from Thailand. *Applied Economics Letters*, *16*, 233–238.

Pellegrini, S., Ruiz, E., & Espasa, A. (2010). Conditionally heteroscedastic unobserved component models and their reduced form. *Economics Letters*, *107*, 88–90.

Soares, L. J., & Medeiros, M. C. (2008). Modeling and forecasting short-term electricity load: a comparison of methods with an application to Brazilian data. *International Journal of Forecasting*, *24*, 630–644.

Stock, J. H., & Watson, M. W. (2007). Why has US inflation become harder to forecast? *Journal of Money, Credit and Banking*, *39*, 3–34.