Working Paper Series, N. 25, December 2007

Misspecification tests for Periodic Long Memory GARCH models

Massimiliano Caporin
Department of Economics
University of Padua
Italy

Francesco Lisi
Department of Statistical Sciences
University of Padua
Italy

Abstract: Distributional theory for Quasi-Maximum Likelihood estimators in long memory conditional heteroskedastic models is not formally defined, even asympotically. Because of that, this paper analyses the performance of the Likelihood Ratio and the Lagrange Multiplier misspecification tests for Periodic Long Memory GARCH models. The real size and power of these tests are studied by means of Monte Carlo simulations with respect to the class of Generalized Long Memory GARCH models. An application to the USD/JPY exchange rate is also provided.

Keywords: Long Memory, Generalized Long Memory GARCH models, PLM-GARCH models, misspecification tests.

Contents

Department of Statistical Sciences
Via Cesare Battisti, 241
35121 Padova
Italy

tel: +39 049 8274168
fax: +39 049 8274170
http://www.stat.unipd.it

Corresponding author:
Massimilano Caporin
tel: +39 049 827 4258
massimilano.caporin@unipd.it

Misspecification tests for Periodic Long Memory GARCH models

Massimilano Caporin
Department of Economics
University of Padua
Italy

Francesco Lisi
Department of Statistical Sciences
University of Padua
Italy

Abstract: Distributional theory for Quasi-Maximum Likelihood estimators in long memory conditional heteroskedastic models is not formally defined, even asympotically. Because of that, this paper analyses the performance of the Likelihood Ratio and the Lagrange Multiplier misspecification tests for Periodic Long Memory GARCH models. The real size and power of these tests are studied by means of Monte Carlo simulations with respect to the class of Generalized Long Memory GARCH models. An application to the USD/JPY exchange rate is also provided.

Keywords: Long Memory, Generalized Long Memory GARCH models, PLM-GARCH models, misspecification tests.

1 Introduction

Many works in the last years discussed the phenomenon of long memory in the volatility of fincial time series and findings of this are well documented in the literature. Several models were also proposed in the statistical and econometric literature to capture the observed persistence in the conditional variance; among these FIGARCH and FIEGARCH models (Baillie *et al.*, 1996; Bollerslev and Mikkelsen, 1996; Andersen and Bollerslev, 1997) and the Long Memory Stochastic Volatility model (Breidt *et al.*, 1998) are well known and very common.
In order to model the empirical evidences of periodic long memory behaviour in the volatility of intra-daily financial returns, more recently, Bordignon *et al.*(2005, 2007) introduced new GARCH-type models characterised by long memory behaviour of periodic type. These models, called Periodic Long–Memory GARCH (PLM-GARCH) and Generalised Long Memory GARCH (G–GARCH), generalise the FIGARCH and FIEGARCH models introducing suitable filters allowing to account also for periodic long memory patterns in conditional variance (associated to the zero frequency of the power spectrum). As a result, G-GARCH and PLM-GARCH also nest some

traditional long memory GARCH specifications.
The filter used for G–GARCH is the most general and allows the description of quite complex long memory behaviours. However, it also requires a richer and less parsimonious parametrization and is more difficult to estimate. In turn, PLM-GARCH is more complex than a simple short memory GARCH including seasonal lags.
It is thus important to be able to discriminate between short and long memory periodic dependence and, when periodic long memory occurs, to evaluate the suitability of the G–GARCH representation.
Since the estimation of PLM– and G–GARCH models is based on likelihood methods, classical misspecifiction tests, for example Likelihood Ratio (LR) and Lagrange Multiplier (LM), may be used to select the more appropriate model.
Bordignon *et al.* (2005, 2007) showed, by a simulation study, the practical applicability and the good performance of the Quasi-Maximum Likelihood (QML) procedure for parameters estimation. However, they also highlighted the lack of formal results concerning consistency or distributional theory, even asympotically, for estimators based on likelihood methods in long memory models.
For this reason, the aim of this study is to check that, despite the mentioned limitations, Likelihood Ratio and Lagrange Multiplier tests can be safely used as misspecification tests when generalised long memory patterns are involved in the conditional variance. This is done through Monte Carlo simulations, exploiting the nesting relations between G–GARCH and other GARCH-type models and taking advantage of the computation of the analytical expressions of the Gradient and of the Hessian of the G–GARCH model.
The paper is organized as follows: in section 2 periodic long memory filters and the frameworks of PLM- and G–GARCH models are briefly reviewed. The plan of the Monte Carlo simulations is described in section 3. Section 4 provides an example based on the time series of the two-hourly USD/JPY exchange rate. Conclusions are given in section 5, whereas technicalities, i.e. analytical derivatives, are given in the Appendix.

2 Periodic Long Memory filters and Generalised–GARCH models

According to Woodward *et al.* (1998), an $(h+1)-$factor Gegenbauer ARMA (GARMA) model allowing for long memory behaviour associated with $h+1$ frequencies in $[0,\pi]$ is defined by

$$\Phi(L)\prod_{j=0}^{h}\left(1-2\cos(\omega_j)L+L^2\right)^{d_j}(y_t-\mu)=\Theta(L)\varepsilon_t, \tag{1}$$

where h is an integer, ε_t is a white noise with variance σ^2_ε, μ is the mean of the process, ω_j $(j=0,...,h)$ are frequencies at which the long memory behaviour occurs, d_j $(j=0,...,h)$ are long memory parameters indicating how slowly the autocorrelations are damped and $\Phi(L)$ and $\Theta(L)$ are standard short memory autoregressive and moving average polynomials with roots satisfying the usual conditions for stationarity and invertibility. The main characteristic of model (1) is given by the presence

of the Gegenbabuer polynomial $P(L) = \prod_{j=0}^{h} \left(1 - 2\cos(\omega_j)L + L^2\right)^{d_j}$ that models the long memory periodic behaviour at frequencies ω_j through the parameters d_j. When we think of the ω_j as the driving frequencies of a cyclical pattern of length S, $\omega_j = \left(\frac{2\pi j}{S}\right)$ and $h+1 = [S/2]+1$, where $[\cdot]$ stands for the integer part.
To highlight the contributions at frequencies $\omega = 0$ and $\omega = \pi$, $P(L)$ can be also written as:

$$P(L) = (1-L)^{d_0}(1+L)^{d_h I(E)} \prod_{j=1}^{h-1} \left(1 - 2\cos(\omega_j)L + L^2\right)^{d_j}, \qquad (2)$$

where $I(E) = 1$ if S is even and zero otherwise and $h+1 = [S/2]+1-I(E)$.
Bordignon *et al.* (2007) proposed to include the generalized long memory filter $P(L)$ into a GARCH structure in order to describe periodic long memory patterns in the conditional variance of a time series. Such kind of patterns are observed, for example, in some intra-daily financial time series.The resulting class of models was called G–GARCH.
Due to the constraints needed for conditional variance positivity, G–GARCH models themeselves are not always feasible. For this reason, Bordignon *et al.* (2007) suggested to use the logarithmic specification (Log–G–GARCH), wich is easier to estimate.
The Log–G–GARCH model is given by

$$y_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t\, z_t \qquad\qquad \varepsilon_t | I_{t-1} \sim D(0, \sigma_t^2)$$

where μ_t is the conditional mean of y_t, z_t is an i.i.d. random variable with zero mean and unitary variance, and $\varepsilon_t | I_{t-1} \sim D(0,\sigma_t^2)$ with conditional variance σ_t^2, I_{t-1} being the information up to time $t-1$.
The dynamics of the log-conditional variance is given by

$$\begin{aligned} \ln(\sigma_t^2) &= \gamma + \beta(L)\ln\sigma_t^2 + \{1 - \beta(L) + \\ &- \left[(1-L)^{d_0}(1+L)^{d_h I(E)} \prod_{j=1}^{h-1}\left(1 - 2\cos(\omega_j)L + L^2\right)^{d_j}\right] \times \\ &\times \phi(L)\} \left[\ln\left(\varepsilon_t^2\right) - \tau\right], \end{aligned} \qquad (3)$$

where $\phi(L) = 1 - \sum_{i=1}^{q} \phi_i L^i$ and $\beta(L) = \sum_{i=1}^{p} \beta_i L^i$ are suitable polynomials in the lag operator L and $\tau = E\left[(ln(z_t^2)\right]$ (in the gaussian case $\tau = -1.27$). The d_j $(j = 0, ..., h)$ are (long) memory parameters associated to the frequencies ω_j indicating how slowly the autocorrelations are damped. In the G–GARCH model, thus, each periodic frequency is modelled by means of a specific long memory parameter d_i.
When $d_0 = d_1 = ... = d_h$ all the involved frequencies have the same degree of memory. Under the additional assumption that the remarkable frequencies are associated to a single periodic component, the specification of the conditional variance (3) corresponds to that of a logarithmic Periodic Long–Memory GARCH (Log-PLM–GARCH) model, introduced by Bordignon *et al.* (2005). Again, the logarithmic

specifiction is considered for obtain computational advantages. In this case, the expression (3) becomes

$$\ln\left(\sigma_t^2\right) = \gamma + \beta(L)\ln\left(\sigma_t^2\right) + \left[1 - \beta(L) - \left(1 - L^S\right)^d (1 - \phi(L))\right]\left[\ln\left(\varepsilon_t^2\right) - \tau\right] \tag{4}$$

which can be derived from model (3) under the restriction $d_0 = d_1 = ... = d_h$.
The main difference between PLM-GARCH and G–GARCH is that the former assumes equal degrees of memory for all interested frequencies and, thus, models the whole long memory behaviour with just a single parameter, leading to a very parsimonious description of the dynamics.
G–GARCH models nest, as particular cases, some of the existing GARCH models. For example, standard GARCH models – included Seasonal GARCH (Bollerslev and Hodrick, 1992; Bollerslev and Ghysel, 1996) – can be obtained by putting $d_j = 0$ $(j = 0, ..., h)$, while the FIGARCH model is equivalent to $S = 1$, $0 < d_0 < 1$ and $d_j = 0$ $(j = 1, ..., h)$. Also PLM-GARCH includes the same GARCH specifications. Whereas model (3) is clearly more flexible than the nested models, it is also evident that it is more complex and less parsimonious. It is, thus, particularly useful to have suitable tests for establishing when using G–GARCH models instead of PLM-GARCH models. Similarly, it is of interest to test the opportunity of fitting a long memory periodic model rather than a simpler short memory periodic one.
Since the G–GARCH class encompasses PLM-GARCH as well as other short memory GARCH specifications, it is possible to apply the standard LR test as misspecification test. For example, testing a PLM-GARCH form *versus* a possible G–GARCH specification implies verifying the hypothesis of equality of all memory coefficients d_i, while testing PLM-GARCH or G–GARCH forms *versus* a short memory GARCH with coefficients at periodic lags implies to test the not significance of all coefficients d_i because this induces a GARCH model where $\alpha(L) = \phi(L) - \beta(L)$.
It is well known that the application of the LM test requires the computation of the derivatives of the likelihood. To this purpose, the analytical gradient and Hessian of the G–GARCH model have been calculated and they are given in Appendix. The gradient and the hessian of the PLM model may be obtained with suitable simplifications.

3 Monte Carlo simulations

In order to proof the reliability of LR and LM tests, in this section nominal and real levels of the tests, as well as powers, are compared through Monte Carlo simulations. To evaluate levels and powers $M = 1000$ simulation trials were considered for series of length $n = 500$, 1000 and 2000. All the data generating models are expressed in the logarithmic form, with no mean component ($\mu_t = 0$) and with a periodic component of period $S = 7$ $(h + 1 = 4)$.
Since this work mainly focuses on the ability of distinguishing between short and long memory periodic behaviour and between PLM– and G–GARCH specifications, only Seasonal-GARCH, PLM–GARCH and G–GARCH models were considered and compared among them. In detail, data were generated from:

a. Log–G–GARCH models without and with short memory component.
For models with only the long memory component we set $\gamma = -0.05$, $\beta(L) = 0$, $\phi(L) = 0$, and the following memory parameters:
$d_1 = 0.1$, $d_2 = 0.2$, $d_3 = 0.25$ and $d_4 = 0.25$ (model M_1); $d_1 = 0.3$, $d_2 = 0.4$, $d_3 = 0.5$ and $d_4 = 0.6$ (model M_2); $d_1 = 0.45$, $d_2 = 0.6$, $d_3 = 0.7$ and $d_4 = 0.8$ (model M_3);
The models including also a short memory component are M_1, M_2 and M_3 with $\phi_1 = 0.1$, $\phi_7 = 0.5$, $\beta_1 = 0.2$ $\beta_7 = 0.3$. They will be referred, respectively, as M_4, M_5 and M_6 models.

b. Log–PLM–GARCH models without and with short memory component.
For models with only the long memory component we set $\beta(L) = 0$, $\phi(L) = 0$, and parameters $d_1 = d_2 = d_3 = d_4 = d$ with: $d = 0.1$ (model M_7); $d = 0.25$ (model M_8) and $d = 0.4$ (model M_9).
The same models were also considered with a short memory component defined by $\phi_1 = 0.1$, $\phi_7 = 0.5$, $\beta_1 = 0.2$ $\beta_7 = 0.3$. They will be referred, respectively, as M_{10}, M_{11} and M_{12} models.

c. Seasonal short memory GARCH models with coefficients at lags 1 and S both for $\alpha(L)$ and $\beta(L)$. Parameters were chosen in order to induce peristent periodic behaviour. In particular they were set to $\alpha_1 = 0.05$, $\alpha_7 = 0.05$, $\beta_1 = 0.05$ and $\beta_7 = 0.8$ (model M_{13}) and $\alpha_1 = 0.02$, $\alpha_7 = 0.02$, $\beta_1 = 0$ and $\beta_7 = 0.95$ (model M_{14}).

Within these cases, when two models are compared the simpler one is denoted by M_0 and the more complex by M_A. In all simulations the hypotesis system under study is

$$\begin{cases} H_0 & : & M_0 \\ H_1 & : & M_A \end{cases} \tag{5}$$

System (5) is verified with respect to q constraints on some model parameters. For example, when M_0 is an S–GARCH and M_A a PLM–GARCH, which means to test short memory *versus* long memory periodic behaviour, system (5) becomes $H_0 : d = 0$ against $H_1 : d \neq 0$. Instead, if M_0 is a PLM–GARCH and M_1 a G–GARCH, the implied null hypothesis is $H_0 : d_i = d$ $(i = 1, ..., h)$.
When PLM–GARCH and G–GARCH models have only the long memory component, they do not nest S-GARCH models; in this case the LR test was applied only for discriminating between PLM- and G–GARCH models. Also, when the data generating process (DGP) is short memory the hypothesis PLM–GARCH *versus* G–GARCH has not been considered because not interesting.
Real levels are studied considering Log–PLM–GARCH and Log-S–GARCH as data generating processes and results on this point are contained in Tables 3, 4 and 5. For $n = 500$, results indicate that the tests perform poorly, but for $n \geq 1000$ real levels are globally satisfactory and more consistent with the nominal ones. Furthermore, when the DGP is short memory (Table 5) the LM test tends to be conservative and thus to under-reject the hypothesis of short memory DGP.

Results concerning the power of tests are dispalyed in Tables 1, 2, 3 and 4.
Again, for $n = 500$ the tests perform sometimes poorly but for $n \geq 1000$ powers are generally high. In particular, we note that when the DGP is long memory the S-GARCH is definitely rejected and both tests, and particularly the LM test, show a very high power in discriminating between PLM- and G–GARCH generators. An exception is when the DGP is model M_4: in this case the test is too much conservative with respect the hypotesis of PLM-type generating process. This is not strange because model M_4 has parameters d_i quite similar to those of a PLM-type model with $d_i = 0.2$ or $d_i = 0.25$. We note, however, that in general the inclusion of short memory components makes more difficult to discriminate between different forms of periodic long memory. Finally, it seems that the LM test is more powerful than the LR test.
In Table 4, the effects of the inclusion of short memory components appear as an underrejection of the null hypothesis, particularly evident with small sample size ($n = 500$), and for model M_{10}. In this last case, the convergence to the appropriate frequencies is slower than for the other DGPs, and still unsatisfactory for lenght $n = 2000$. This depends on the particular values of parameters of the DGP, a PLM model with memory coefficient set to 0.1. In fact, the limited memory of the process combined with the occurence of a short memory dynamics introduces a larger uncertainty which can be reduced only increasing the sample size.
In this study, the analyses were limited to just one combination of short memory parameters. In fact, the estimation of periodic long memory models requires a higher CPU time when models include also short memory coefficients. Despite the overall estimating time is not elevate, around 15 minutes for a series of length 2000, their inclusion within a Monte Carlo experiment greatly increment the time required to run the simulations. Note also that PLM– and G–GARCH models without short memory coefficients require estimation times considerably lower, in the order of 1 to 5 minutes, depending on the starting values used in association with the true memory degree and the series length. Furthermore, the specific chosen short memory parameters induce a mildly persistent short memory pattern. Under the hypothesis of no long memory, these short memory coefficients induce a Short memory GARCH with parameters equal to $\alpha_1 = -0.1$, $\alpha_7 = 0.2$, $\beta_1 = 0.2$, $\beta_7 = 0.3$; the negative coefficient is not a problem given that the model is expressed in the logs.

4 An application: the USD/JPY exchange rate

As an empirical application of the previous testing framework, the intra-daily series of the exchange rate US Dollar *versus* Japanese Yen was analysed. The covered period is March 1, 2000 - February 28, 2005, for a total of 1304 working days. Data were provided by Olsen & Associates at a frequency of 5–minutes, but in the application two-hourly data were considered. Within every week, data range from the 22.00 of Sunday to 22.00 of Friday. The length of the resulting series is $n = 15648$. The return time series (r_t) is uncorrelated with a not significant mean, thus no model for the conditional mean is required. Squared and log–squared returns, instead, are significantly correlated and show a periodic behaviour. Since in the fol-

	LM test						LR test	
level	0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05
DGP: M_1	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G	
500	0.673	0.883	0.932	0.992	0.933	0.994	0.855	0.955
1000	0.964	0.994	0.997	1.000	0.998	1.000	0.997	0.999
2000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DGP: M_2	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G	
500	0.999	1.000	0.987	0.996	0.990	0.998	1.000	1.000
1000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
2000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
DGP: M_3	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G	
500	1.000	1.000	0.983	0.990	0.986	0.995	1.000	1.000
1000	1.000	1.000	0.947	0.989	1.000	1.000	1.000	1.000
2000	1.000	1.000	0.911	0.982	1.000	1.000	1.000	1.000

Table 1: Data generating process: G–GARCH without short memory component. Parameters: Model M_1 (0.1 - 0.2 - 0.25 - 0.25); Model M_2 (0.3 - 0.4 - 0.5 - 0.6), Model M_3 (0.45 - 0.6 - 0.7 - 0.8)

lowing only log–GARCH–type models will be considered, hereafter we concentrate on log–squared residuals, $\ln(r_t^2)$. Figure 1 shows the correlogram and periodogram of $\ln(r_t^2)$.

The very slow periodic decaying of the autocorrelation function and the pronounced peaks of the periodogram at the origin and at seasonal frequencies may indicate a cyclical long memory behaviour. The four main peaks are located at the frequencies $\omega_0 = 0$, $\omega_1 = 0.083$, $\omega_2 = 0.166$ and $\omega_3 = 0.25$, which correspond, respectively, to a traditional long memory component and to three possibly long memory periodic components of daily ($S = 12$), semi-daily ($S = 6$) and 4–hourly ($S = 2$) periods. Besides these periodic components, in the spectrum some minor peaks, for example that at $\omega = 0.333$, are present.

The peaks in the periodogram may also suggest a deterministic periodic behaviour which could be accounted using seasonal dummy variables. A visual inspection of the periodogram of the standardized squared residuals of a regression on dummy variables revealed that this approach is not satisfactory and that a stochastic modelling seems more appropriate.

In order to specify the form of a suitable stochastic model for the observed periodic pattern, three periodic GARCH-type models were selected and estimated for the log–squared returns: a Seasonal-GARCH (S-GARCH) a PLM-GARCH and a G–GARCH. The first is short memory whereas the last two are long memory. Results, displayed in Table 6, suggest considering the G–GARCH specification.

The estimation results are listed in Table 7 and show that the memory parameters for the G–GARCH model are all significant and that, in terms of loglikelihood, the G–GARCH specification is, in actual fact, the best one.

Finally, the analysis of the autocorrelation function and of the periodogram of the standardized squared residuals of model G–GARCH confirms that there is not sig-

	LM test						LR test					
level	0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05
DGP: M_4	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G		S *vs* G		S *vs* PLM	
500	0.024	0.071	0.978	0.998	0.999	1.000	0.099	0.235	0.376	0.558	0.367	0.536
1000	0.021	0.073	0.996	0.999	1.000	1.000	0.145	0.320	0.354	0.556	0.259	0.505
2000	0.057	0.155	0.997	0.1000	1.000	1.000	0.286	0.542	0.587	0.817	0.416	0.683
DGP: M_5	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G		S *vs* G		S *vs* PLM	
500	0.985	0.997	0.988	0.997	1.000	1.000	0.997	0.998	0.999	1.000	0.914	0.962
1000	1.000	1.000	0.999	1.000	1.000	1.000	0.999	0.999	1.000	1.000	0.986	0.997
500	1.000	1.000	1.000	1.000	1.000	1.000	0.999	0.999	1.000	1.000	1.000	1.000
DGP: M_6	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G		S *vs* G		S *vs* PLM	
500	0.983	0.992	0.993	0.998	1.000	1.000	0.998	1.000	1.000	1.000	0.876	0.927
1000	0.999	0.999	0.978	0.996	0.998	0.998	0.999	0.999	1.000	1.000	0.979	0.992
2000	1.000	1.000	0.999	1.000	1.000	1.000	1.000	1.000	1.000	1.000	0.998	1.000

Table 2: Data generating process: G–GARCH including a short memory component. Parameters: Models M_4,M_5 and M_6 are defined as models M_1, M_2 and M_3 with also $\phi_1 = 0.1$, $\phi_7 = 0.5$, $\beta_1 = 0.2$ $\beta_7 = 0.3$

	LM test						LR test	
level	0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05
DGP: M_7	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G	
500	0.076	0.201	0.934	0.993	0.945	0.998	0.010	0.046
1000	0.014	0.067	0.991	0.997	0.999	1.000	0.013	0.052
2000	0.012	0.068	1.000	1.000	1.000	1.000	0.007	0.059
DGP: M_8	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G	
500	0.017	0.072	0.917	0.988	0.928	0.995	0.013	0.065
1000	0.020	0.062	0.993	1.000	0.999	1.000	0.014	0.063
2000	0.014	0.057	1.000	1.000	1.000	1.000	0.012	0.047
DGP: M_9	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G	
500	0.015	0.076	0.913	0.977	0.932	0.987	0.012	0.062
1000	0.018	0.072	0.988	0.996	0.994	0.999	0.021	0.073
2000	0.010	0.045	0.999	1.000	1.000	1.000	0.011	0.051

Table 3: Data generating process: PLM–GARCH without short memory component. Parameters: Model M_7 $d = 0.1$; M_8 $d = 0.25$, Model M_9 $d = 0.4$.

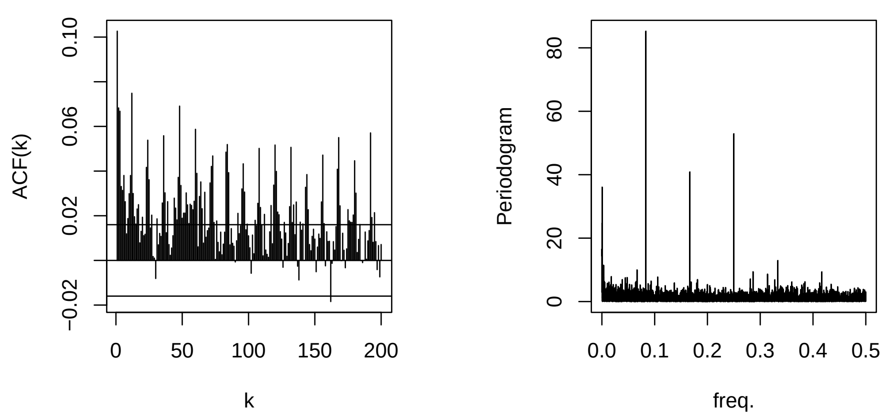

Figure 1: UDS/JPY exchange rate: autocorrelation function and periodogram of $\ln(r_t^2)$.

nificant residual correlation nor dominant peaks at the seasonal frequencies or at some neighbourhood in the periodogram.

5 Conclusions

In this paper it was shown that LM and LR tests can be safely used as model selection tools when the underlying data generating process may include long memory periodic components. By mean of a Monte Carlo analysis the real size and power of these tests were derived evidencing their reliability apart from some special and limited

	LM test						LR test					
level	0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05	0.01	0.05
DGP: M_{10}	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G		S *vs* G		S *vs* PLM	
500	0.004	0.010	0.923	0.993	0.999	1.000	0.011	0.054	0.086	0.210	0.164	0.315
1000	0.002	0.017	0.992	0.999	0.999	0.999	0.010	0.058	0.202	0.400	0.370	0.574
2000	0.004	0.020	1.000	1.000	1.000	1.000	0.016	0.062	0.509	0.724	0.719	0.862
DGP: M_{11}	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G		S *vs* G		S *vs* PLM	
500	0.008	0.025	0.985	1.000	1.000	1.000	0.025	0.099	0.505	0.701	0.662	0.849
1000	0.003	0.020	0.994	1.000	1.000	1.000	0.018	0.087	0.856	0.952	0.949	0.980
2000	0.004	0.035	1.000	1.000	1.000	1.000	0.013	0.057	0.947	0.989	0.996	1.000
DGP: M_{12}	PLM *vs* G		S *vs* G		S *vs* PLM		PLM *vs* G		S *vs* G		S *vs* PLM	
500	0.003	0.027	0.943	0.998	1.000	1.000	0.022	0.097	0.603	0.805	0.794	0.899
1000	0.005	0.018	1.000	1.000	1.000	1.000	0.025	0.080	0.959	0.991	0.988	0.998
2000	0.003	0.040	1.000	1.000	1.000	1.000	0.014	0.055	0.989	0.999	0.999	1.000

Table 4: Data generating process: PLM–GARCH including short memory component. Parameters: Models M_{10},M_{11} and M_{12} are defined as models M_7, M_8 and M_9 with also $\phi_1 = 0.1$, $\phi_S = 0.5$, $\beta_1 = 0.2$ $\beta_S = 0.3$

	LM test			
level	0.01	0.05	0.01	0.05
DGP: M_{13}	S *vs* PLM		S *vs* G	
500	0.010	0.037	0.013	0.035
1000	0.008	0.028	0.010	0.031
2000	0.005	0.026	0.007	0.029
DGP: M_{14}	S *vs* PLM		S *vs* G	
500	0.012	0.039	0.023	0.045
1000	0.011	0.035	0.022	0.041
2000	0.008	0.034	0.019	0.040

Table 5: Data generating process: S–GARCH. Parameters: $\alpha_1 = 0.05$, $\alpha_7 = 0.05$, $\beta_1 = 0.05$ and $\beta_7 = 0.8$ for model M_{13}; $\alpha_1 = 0.02$, $\alpha_7 = 0.02$, $\beta_1 = 0$ and $\beta_7 = 0.95$ for model M_{14}.

	LR–test	n. restrictions	p-value
Log–GARCH *vs* Log-PLM-GARCH	348.830	1	< 0.001
Log–GARCH *vs* Log–G–GARCH	446.902	7	< 0.001
Log–PLM–GARCH *vs* Log–G–GARCH	98.072	6	< 0.001
	LM–test	n. restrictions	p-value
Log–GARCH *vs* Log–PLM–GARCH	203.395	1	< 0.001
Log–GARCH *vs* Log–G–GARCH	204.275	7	< 0.001
Log–PLM–GARCH *vs* Log–G–GARCH	941.317	6	< 0.001

Table 6: UDS/JPY exchange rate: LM and LR tests for $\ln(r_t^2)$.

cases.

The test performances are however influenced by the sample length and results show that about a thousand observations are required in order to obtain reliable conclusions. This is not an unexpected result given the presence of a long memory of periodic type.

Finally, an application showing how our testing approach can be used in the model identification has been provided for an intra-daily series of the USD/JPY exchange rate.

References

Andersen T.G., Bollerslev T. (1997), Intraday Periodicity and Volatility Persistence in Financial Markets, *Journal of Empirical Finance*, 4, 115-158.

Baillie R.T., Bollerslev T., Mikkelsen H.O. (1996), Fractionally integrated generalized autoregressive conditional heteroskedasticity, *Journal of Econometrics*, 74, 3-30.

Bollerslev, T., Ghysel, E. (1996), Periodic autoregressive conditional heteroscedas-

Model	Estimate	Std. error	$t-$stat	LogLik
LoG–GARCH				19910.299
ω	-0.297	0.349	-0.852	
α_1	0.016	0.032	0.504	
α_{12}	0.052	0.009	5.617	
β_1	0.012	0.074	0.168	
β_{12}	0.824	0.072	11.535	
Log-PLM-GARCH				20084.714
ω	-0.385	0.130	-2.970	
d	0.163	0.037	4.436	
α_1	-0.018	0.008	-2.404	
α_{12}	0.642	0.051	12.598	
β_1	0.017	0.008	2.212	
β_{12}	0.529	0.053	10.009	
Log–G–GARCH				20133.750
ω	-0.133	0.041	-3.285	
d_0	0.123	0.005	24.347	
d_{12}	0.206	0.011	18.126	
d_6	0.201	0.012	16.299	
d_4	0.194	0.014	13.767	
d_3	0.217	0.017	12.995	
$d_{2.4}$	0.214	0.017	12.357	
d_2	0.100	0.010	10.202	
α_1	-0.034	0.026	-1.312	
α_{12}	0.709	0.041	17.451	
β_1	-0.022	0.057	-0.391	
β_{12}	0.553	0.044	12.620	

Table 7: UDS/JPY exchange rate: estimated parameters of conditional variance of models.

ticity, *Journal of Business and Economic Statistics*,14, 139-151.

Bollerslev, T., Mikkelsen, H.O. (1996), Modeling and pricing long memory in stock market volatility, *Journal of Econometrics*, 73, 151-184

Bordignon S., Caporin M., Lisi F. (2005), Periodic Long Memory GARCH models, *Econometric Reviews*, forthcoming. Working Paper n.19-2005, Department of Statistical Sciences, University of Padua, Italy.

Bordignon S., Caporin M., Lisi F. (2007), Generalized long memory GARCH for intra-daily volatility modelling, *Computational Statistics & Data Analysis*, 51, 5900-5912.

Breidt F.J., N. Crato, de Lima P. (1998), The detection and estimation of long memory in stochastic volatility, *Journal of Econometrics* 83, 325-348.

Lombardi J.M., Gallo G.M. (2002), Analytic Hessian matrices and the computation of FIGARCH estimates, *Statistical Methods and Applications*, 11, 139-264.

Woodard W.A., Cheng Q.C., Gray H. (1998), A k-factor GARMA long memory model, *Journal of Time Series Analysis* , 19, 485-504.

Appendix: derivatives

This appendix reports the analytical derivatives of the Log–G–GARCH model. The equations for the Log-PLM-GARCH model, nested in the Log–G–GARCH model, can be obtained by exploiting the nesting constraints.
Let us recall the model and some known results. The mean equation is:

$$y_t - \mu_t = z_t \sigma_t = \varepsilon_t.$$

with ε_t assumed to follow a GARCH structure. In particular, defined $h_t = \ln\left(\sigma_t^2\right)$ and $e_t = \ln\left(\varepsilon_t^2\right) - k(\theta)$, the log-conditional variance of ε_t has the following specification:

$$h_t = \omega + \beta(L) h_t + \left[1 - \beta(L) - \left[\prod_{i=0}^{h}\left(1 - 2\eta_i L + L^2\right)^{d_i}\right]\phi(L)\right] e_t.$$

where $\beta(L) = \sum_{i=1}^{p} \beta_i L^i$ and $\phi(L) = 1 - \sum_{i=1}^{q} \phi_i L^i$.
The the long memory polynomial can be expanded by using the expression

$$\prod_{i=0}^{h}\left(1 - 2\eta_i L + L^2\right)^{d_i} = \prod_{i=0}^{h}\sum_{j=0}^{\infty} c_{i,j}(d_i, \eta_i) L^j.$$

The coefficients of the expansion depend on the memory coefficients d_i and on the η_i frequencies at which the long memory operates. For simplicity, the dependence of the $c_{i,j}(d_i, \eta_i)$ coefficients on the memory levels and frequencies will be suppressed and the simpler notation $c_{i,j}$ will be used, where the first subscript identifies the memory and frequency coefficients while the second subscript identifies the lag. The $c_{i,j}$ coefficients have the following recursive structure

$$\begin{aligned} c_{l,0} &= 1 \\ c_{l,1} &= -2d_l\eta_l \\ c_{l,y} &= 2\eta_l\left(\frac{-d_l - 1}{y} + 1\right)c_{l,y-1} - \left(2\frac{-d_l - 1}{y} + 1\right)c_{l,y-2} \end{aligned}$$

Analytical gradient of Log–G–GARCH models

In order to define the gradient the model likelihood is defined as $L = \sum_{t=1}^{T} L_t$, with L_t the likelihood for time t

$$L_t = -\frac{1}{2}h_t - \frac{1}{2}\frac{\varepsilon_t^2}{\exp(h_t)}$$

We also collect the constant of the conditional variance, ω, the short memory coefficients included in $\beta(L)$ and $\phi(L)$, and the long memory coefficients in a single parameter set denoted by $\psi = (\omega, \beta_1, ..., \beta_p, \phi_1, ..., \phi)$. The gradient of the log-likelihood L is then

$$\frac{\partial L}{\partial \psi'} = \sum_{t=1}^{T}\frac{\partial L_t}{\partial \psi'}$$

and the gradient for time t is

$$\frac{\partial L_t}{\partial \psi'} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{\exp(h_t)} - 1\right]\frac{\partial h_t}{\partial \psi'}.$$

The expressions for the single coefficient derivatives are reported grouping them according to the coefficients. Notice that we report directly the expression for the derivative with respect to the log-variances; the gradient for the entire likelihood can be obtained by substitution.
The derivative with respect to the variance constant is

$$\frac{\partial h_t}{\partial \omega} = 1 + \beta(L)\frac{\partial h_t}{\partial \omega}.$$

The derivatives with respect to the short memory coefficients in $\beta(L)$ are

$$\frac{\partial h_t}{\partial \beta_j} = \beta(L)\frac{\partial h_t}{\partial \beta_j} + h_{t-j} - e_{t-j}.$$

The derivatives with respect to the short memory coefficients in $\phi(L)$ are

$$\frac{\partial h_t}{\partial \phi_l} = \beta(L)\frac{\partial h_t}{\partial \phi_l} - \left[\prod_{i=0}^{h}\left(1 - 2\eta_i L + L^2\right)^{d_i}\right] e_{t-l}.$$

Note that all derivatives have a recursive structure and are very similar to those reported in Lombardi and Gallo (2002).
The most complex set of derivatives is that with respect to the memory coefficients. In this case the derivatives with respect to the long memory coefficients require the computation of the derivatives for the coefficients in the long memory polynomial expansion

$$\frac{\partial}{\partial d_l}\prod_{i=0}^{h}\sum_{j=0}^{\infty} c_{i,j}(d_i,\eta_i)L^j = \left[\prod_{i=0 i\neq l}^{h}\sum_{j=0}^{\infty} c_{i,j}(d_i,\eta_i)L^j\right]\left(\frac{\partial}{\partial d_l}\sum_{y=0}^{\infty} c_{l,y}(d_l,\eta_l)L^y\right).$$

The derivative with respect to a single long memory coefficient can be written as

$$\frac{\partial}{\partial d_l}\sum_{y=0}^{\infty} c_{l,y}(d_l,\eta_l)L^y = \sum_{y=0}^{\infty}\frac{\partial c_{l,y}(d_l,\eta_l)}{\partial d_l}L^y,$$

where the derivatives of the long memory expansion coefficients are

$$\begin{aligned}
\frac{\partial c_{l,y}}{\partial d_l} &= 2\eta_l\left(\frac{-d_l-1}{y}+1\right)\frac{\partial c_{l,y-1}}{\partial d_l} - \left(2\frac{-d_l-1}{y}+1\right)\frac{\partial c_{l,y-2}}{\partial d_l} - 2\eta_l\frac{1}{y}c_{l,y-1} + \frac{2}{y}c_{l,y-2},\\
\frac{\partial c_{l,0}}{\partial d_l} &= 0, \qquad \frac{\partial c_{l,1}}{\partial d_l} = -2\eta_l.
\end{aligned}$$

Summarizing, the derivative with respect to a single memory coefficient is thus

$$\frac{\partial h_t}{\partial d_l} = \beta(L)\frac{\partial h_{t-1}}{\partial d_l} - \phi(L)\left[\prod_{\substack{i=0\\ i\neq l}}^{h}\sum_{j=0}^{\infty} c_{i,j}L^j\right]\left(\sum_{y=0}^{\infty}\frac{\partial c_{l,y}}{\partial d_l}L^y\right)e_t.$$

Note that also the derivatives with respect to the memory coefficients have a recursive structure in the main derivative and in the polynomial coefficients.

Analytical Hessian of Log–G–GARCH models

Using the same notation previously introduced, the Hessian at time t is

$$\frac{\partial L_t}{\partial \psi \partial \psi'} = \frac{1}{2}\left[\frac{\varepsilon_t^2}{\exp(h_t)} - 1\right]\frac{\partial^2 h_t}{\partial \psi \partial \psi'} - \frac{1}{2}\frac{\varepsilon_t^2}{\exp(h_t)}\frac{\partial h_t}{\partial \psi}\frac{\partial h_t}{\partial \psi'}.$$

The elements entering the second part of the Hessian can be obtained by using the results achieved for the Gradient. Here only the elements characterizing the second order derivatives of the log-variances are reported.
The second order derivative with respect to the constant is

$$\frac{\partial^2 h_t}{\partial \omega^2} = \beta(L)\frac{\partial^2 h_t}{\partial \omega^2}.$$

The second order derivatives with respect to a couple of coefficients of the polynomial $\beta(L)$ are

$$\frac{\partial^2 h_t}{\partial \beta_j \partial \beta_i} = \beta(L)\frac{\partial^2 h_t}{\partial \beta_j \partial \beta_i} + \frac{\partial h_{t-i}}{\partial \beta_j} + \frac{\partial h_{t-j}}{\partial \beta_i} \qquad i, j \quad = \quad 1, 2, ...p.$$

The second order derivatives with respect to a couple of coefficients of the polynomial $\phi(L)$ are

$$\frac{\partial^2 h_t}{\partial \phi_j \partial \phi_l} = \beta(L)\frac{\partial^2 h_t}{\partial \phi_j \partial \phi_l} j, l \quad = \quad 1, 2, ...q.$$

The cross-second order derivatives between the constant ω and the coefficients ϕ_i and β_j are

$$\begin{aligned}
\frac{\partial^2 h_t}{\partial \omega \partial \beta_j} &= \beta(L)\frac{\partial^2 h_t}{\partial \omega \partial \beta_j} + \frac{\partial h_{t-j}}{\partial \omega} && j = 1, 2, ...p; \\
\frac{\partial^2 h_t}{\partial \omega \partial \phi_l} &= \beta(L)\frac{\partial^2 h_t}{\partial \omega \partial \phi_l} && l = 1, 2, ...q; \\
\frac{\partial^2 h_t}{\partial \beta_j \partial \phi_l} &= \beta(L)\frac{\partial^2 h_t}{\partial \beta_j \partial \phi_l} + \frac{\partial h_{t-i}}{\partial \omega \phi_l} && j = 1, 2, ...p; \quad l = 1, 2, ...q.
\end{aligned}$$

The second order derivatives with respect to memory coefficients are

$$\begin{aligned}
\frac{\partial^2 h_t}{\partial d_l \partial d_k} &= \beta(L)\frac{\partial^2 h_{t-1}}{\partial d_l \partial d_k} - \phi(L)\left[\prod_{\substack{i=1\\ i\neq l,k}}^{k}\sum_{j=0}^{\infty} c_{i,j}L^j\right]\left(\sum_{y=0}^{\infty}\frac{\partial c_{l,y}}{\partial d_l}L^y\right)\left(\sum_{m=0}^{\infty}\frac{\partial c_{k,m}}{\partial d_k}L^m\right)e_t \\
& l, k = 0, 1, 2, ...h \quad l \neq k \\
\frac{\partial^2 h_t}{\partial d_l^2} &= \beta(L)\frac{\partial^2 h_{t-1}}{\partial d_l^2} - \phi(L)\left[\prod_{\substack{i=1\\ i\neq l}}^{k}\sum_{j=0}^{\infty} c_{i,j}L^j\right]\left(\sum_{y=0}^{\infty}\frac{\partial^2 c_{l,y}}{\partial d_l^2}L^y\right)e_t \\
& l = 0, 1, 2, ...h
\end{aligned}$$

where the derivatives of the memory expansion coefficients are

$$\begin{aligned}
\frac{\partial^2 c_{l,y}}{\partial d_l^2} &= 2\eta_l\left(\frac{-d_l-1}{y}+1\right)\frac{\partial^2 c_{l,y-1}}{\partial d_l^2}-\left(2\frac{-d_l-1}{y}+1\right)\frac{\partial^2 c_{l,y-2}}{\partial d_l^2} \\
&\quad -\frac{4}{y}\left(\eta_l\frac{\partial c_{l,y-1}}{\partial d_l}-\frac{\partial c_{l,y-2}}{\partial d_l}\right) \\
\frac{\partial^2 c_{l,0}}{\partial d_l^2} &= 0 \qquad \frac{\partial^2 c_{l,1}}{\partial d_l^2}=0.
\end{aligned}$$

Finally, the cross-derivatives involving memory coefficients are

$$\begin{aligned}
\frac{\partial^2 h_t}{\partial\omega\partial d_l} &= \beta(L)\frac{\partial^2 h_t}{\partial\omega\partial d_l} \\
\frac{\partial^2 h_t}{\partial\beta_j\partial d_l} &= \beta(L)\frac{\partial^2 h_t}{\partial\beta_j\partial d_l}+\frac{\partial h_{t-j}}{\partial d_l} \\
\frac{\partial^2 h_t}{\partial\phi_l\partial d_m} &= \beta(L)\frac{\partial^2 h_t}{\partial\phi_l\partial d_m}-\left[\prod_{\substack{i=1\\ i\neq m}}^{k}\sum_{j=0}^{\infty}c_{i,j}L^j\right]\left(\sum_{y=0}^{\infty}\frac{\partial c_{m,y}}{\partial d_m}L^y\right)e_{t-l}.
\end{aligned}$$

Working Paper Series

Department of Statistical Sciences, University of Padua

You may order paper copies of the working papers by emailing wp@stat.unipd.it

Most of the working papers can also be found at the following url: http://wp.stat.unipd.it

UNIVERSITÀ DEGLI STUDI DI PADOVA
DIPARTIMENTO DI SCIENZE STATISTICHE

Department of Statistical Sciences
University of Padua
Italy