

Robust detection of nonlinearity in financial time series

Non linearità e robustezza nelle serie storiche finanziarie

Luigi Grossi and Fabrizio Laurini

Dipartimento di Economia, Sezione di Statistica ed Informatica, Università di Parma
e-mail: luigi.grossi@unipr.it

Riassunto: I test generalmente utilizzati per l'individuazione di effetti non-lineari nelle serie storiche si basano sull'utilizzo di modelli di regressione multipla sul quadrato dei residui. Esempio classico è il test ARCH di Engle (1982). La diffusione dei test basati sui moltiplicatori di Lagrange deriva dalla possibilità di ricavare in modo semplice la distribuzione del test sotto l'ipotesi nulla. Empiricamente i test di Lagrange hanno evidenziato scarsa potenza, non necessariamente causata dall'inclusione di valori anomali. Pertanto, alcuni ricercatori hanno suggerito l'utilizzo di opportuni pesi per ogni unità statistica considerata. In questo lavoro si vuole generalizzare il meccanismo di ponderazione suggerito nel contesto delle serie finanziarie. A tale scopo si suggerisce l'introduzione di un sistema di pesi derivanti dall'applicazione di metodi diagnostici della *forward search*, introdotta da Atkinson e Riani (2000).

Keywords: ARCH test, forward search, non-linearity, weighted regression.

1. Introduction

Modeling of daily observed prices of financial time series P_t , $t = 1, \dots, T$ has become increasingly popular in most financial econometric literature. A wide class of models deal with financial logarithmic returns $r_t = \ln(P_t) - \ln(P_{t-1})$ rather than original prices. On top of everything else the reason for modeling r_t processes is related to the mean-stationarity of such processes, no matter whether stock prices, market indices or equities are of concern.

However, there is a number of *stylized facts* which characterize financial returns r_t as a process whose memory has distinctive non-linear features. The combination of stationarity and non-linearity of r_t has led to the introduction of autoregressive conditionally heteroscedastic (ARCH) of Engle (1982) and their generalization (GARCH) due to Bollerslev (1986).

A general form of GARCH processes, that takes into account the linear term for modeling the conditional mean of r_t , is given by the ARMA(s, v)-GARCH(p, q) models, defined as

$$\begin{aligned} r_t &= \kappa + \sum_{h=1}^s \phi_h r_{t-h} + \sum_{l=1}^v \vartheta_l \eta_{t-l} + \eta_t \\ \eta_t &= z_t \sqrt{h_t} \\ h_t &= \alpha_0 + \sum_{i=1}^q \lambda_i \eta_{t-i}^2 + \sum_{j=1}^p \gamma_j h_{t-j} \end{aligned}$$

where $\{z_t\}$ is a zero-mean unit-variance independent and identically distributed (*iid*) sequence. ARMA parameters κ , ϕ_k , $k = 1, \dots, s$ and ϑ_l , $l = 1, \dots, v$ are assumed to satisfy all constraints for stationarity. GARCH parameters $\alpha_0 > 0$, $\lambda_i \geq 0$ and $\gamma_j \geq 0$ ($i = 1, \dots, q$, $j = 1, \dots, p$) must satisfy constraints for stationarity, which we assume. For any fixed t , h_t and z_t are independent and, if not stated otherwise, z_t will be assumed standard Gaussian. Special cases are represented by $s = v = 0$, which gives GARCH(p, q) or $\{s = v = p = 0, q = 1\}$, leading to ARCH(1).

The fit of heteroscedastic components requires the earlier stage of their identification, possibly through a statistical test. Engle (1982) derived a test for identification of ARCH components which is based on Lagrange multipliers (LM). The test proposed by Engle for the presence of ARCH components is based on the auxiliary regression of residuals of a conditional mean fit $\hat{\eta}_t = r_t - E_r$, where E_r is a fit for the conditional mean of the r_t process, like ARMA as defined above. Formally, the auxiliary regression is

$$\hat{\eta}_t^2 = \alpha_0 + \alpha_1 \hat{\eta}_{t-1}^2 + \dots + \alpha_q \hat{\eta}_{t-q}^2 + \varepsilon_t, \quad t = q + 1, \dots, T \quad (1)$$

where ε_t is the innovation term. The volatility process h_t is constant if the parameters corresponding to the lagged squared shocks $\hat{\eta}_{t-i}^2$, $i = 1, \dots, q$ are equal to zero. Therefore, the null hypothesis of conditional homoscedasticity can be formulated as $H_0 : \alpha_1 = \alpha_2 = \dots = \alpha_q = 0$. This way of testing for ARCH components can be computed using $\lambda = TR^2$, where R^2 is the multiple correlation coefficient of model (1). Under the null hypothesis, the LM test statistic has an asymptotic $\chi^2(q)$ distribution. Lee (1991) showed that the LM test against a GARCH(p, q) alternative is the same as the LM test against the alternative of a ARCH(q), therefore the method adopted here is quite general.

An empirical finding, is that for multiple regression models a few influential observations might yield to a wrong decision about, for instance, significance of t -statistics, of parameters coefficients, or accuracy of R^2 evaluation (see, Atkinson and Riani, 2000). For the ARCH test, the rejection of the null hypothesis occurs less often than it should and this motivated the introduction of several “weighted” methods for assessing the presence of non-linear components. Example of robust weighted tests are in van Dijk et al. (1999).

Testing for non-linearity in financial time series also involves the assessment for symmetry in modeling the conditional variance. Examples of models which account for asymmetric effects (or leverage) are the generalization of GARCH as Exponential GARCH (EGARCH) or Threshold GARCH (TGARCH). However, despite the different way of modeling asymmetric effects, the basic idea that justifies the introduction of these models is the assessment of asymmetry, which led to develop the class of statistical “Bias Test” (sign, negative size and positive size) of Engle and Ng (1993). The indicator function of negative residuals can be written as

$$S_t^- = \begin{cases} 1 & \text{if } r_t < 0 \\ 0 & \text{if } r_t \geq 0 \end{cases}$$

while $S_t^+ = 1 - S_t^-$ denotes positive residuals. Positive and negative residuals can be used in the regression model

$$\hat{\eta}_t = a + b_0 S_{t-1}^- + c_0 S_{t-1}^- \hat{\eta}_{t-1} + c_1 S_{t-1}^+ \hat{\eta}_{t-1} + \varepsilon_t$$

testing jointly the non-leverage effect under the null that parameter of the above auxiliary regression $H_0 : b_0 = c_0 = c_1 = 0$, where we have not introduced the term $b_1 S_{t-1}^-$ to avoid collinearity. As for the LM test of Engle, under the null hypothesis the statistic

$$\xi^* = TR^2$$

has an asymptotic χ^2 distribution with 3 degrees of freedom. Similarly to the LM test for ARCH components, the bias tests all suffer from the presence of influential observations. In this paper we only study LM test for detection of ARCH components, but results might be extended to any test based on multiple regression, such as the bias tests introduced above.

The paper is organized as follows. Section 2 reviews the forward search steps and introduces a method for weighting observations in time series. In Section 3 we present some results on simulated data, while in Section 4 we study real time series. We close with some final remarks and discussion.

2. Robust LM test through the forward search

The low power of the LM test for ARCH components could be given by the presence of extreme returns which strongly bias the regression parameter estimates and the coefficient of determination of the linear models on the squared residuals. A robust version of the test has been introduced by Van Dijk et al. (1999). The main advantage of Van Dijk et al. method is the allowance of a weight for outliers. Observations are multiplied by a factor $\psi \in [0, 1]$, with the most outlying units getting smaller weight. The weights, which are derived from the generalized moment method, reduce the impact of outliers in the estimation step. The procedure applies a backward method of outlier detection and, like all backward methods, is affected by the problem of masking which prevents from right identification of genuine outliers. A forward analysis of the LM test for GARCH components has been proposed by Grossi and Laurini (2004). In this paper we suggest to use the forward search in order to compute weights which are affected by the masking effect. In the next section the extension of the forward search to the LM tests will be introduced, then a procedure to obtain robust weights will be discussed.

2.1. Extension of the forward search to LM tests

The forward search approach (for more details, see Atkinson and Riani, 2000) is a procedure which combines robust and efficient estimators. It is essentially based on three steps:

- 1) use of very robust methods to sort the initial data set into a clean part (a subset free from outliers from now on called Clean Data Set, CDS) and potential outliers;
- 2) definition of criteria which indicate how to introduce new observations into the initial subset in order to progress the forward search until all observations are included;

- 3) monitoring of some relevant statistics during the procedure in order to built some plots which help to visualize abnormal influential units on parameters estimate, test statistics, etc.

2.2. Identification of the Clean Data Set

Usually the identification of the CDS is done by applying the Least Median of Squares (LMS) estimator introduced by Rousseeuw (1984), but other robust estimators could be used. When the alternative hypothesis of the LM test is the presence of q ARCH components, the initial subset is chosen among a set of blocks of observations of dimension $p = q + 1$. Let $z'_t = (1, \hat{\eta}_t^2, \dots, \hat{\eta}_{t-q}^2)$ with $t = p, \dots, T$ the $t - th$ row of the matrix Z of dimension $(T - p) \times (p)$. If T is not too large and $p \ll T$ we can find the best subset free from outliers through an exhaustive search of all possible $\binom{T-p}{p}$ distinct p tuples $S_{t_1, \dots, t_p}^{(p)} \equiv \{z_{t_1}, \dots, z_{t_p}\}$ where z_{t_1} is the $t_1 - th$ row of Z , for $p \leq t_1, \dots, t_p \leq T$ and $t_j \neq t_{j'}$. More specifically, let $\tau' = [t_1, \dots, t_p]$ and let $e_{t, S_{\tau'}^{(p)}}$ be the least squares residual estimated from model (1) for unit t given observations in $S_{\tau'}^{(p)}$. We choose as the CDS the one which satisfies the least median of squares (LMS) criterion. Among all considered blocks of size p we select the subset of p observations $S_*^{(p)}$ which minimizes the median of the squared OLS residuals, that is:

$$e_{[med], S_*^{(p)}}^2 = \min_{\tau} [e_{[med], S_{\tau'}^{(p)}}^2]$$

where $e_{[k], S_{\tau'}^{(p)}}^2$ is the $k - th$ ordered squared residual among $e_{i, S_{\tau'}^{(p)}}^2, i = p, \dots, T$ and med is the sample median, i.e. the integer part of $(T + 1)/2$. If $\binom{T-p}{p}$ is too large, we use instead a large number of samples such as 1000 or 3000. We remark that results are not sensitive to the choice of the number of the initial CDS (see, for example, Grossi and Riani, 2002).

2.3. How to progress along the search

Given the best subset $S_*^{(m)}$ of dimension $m \geq p$ detected at the first step of the search, we move from step m to $m + 1$ sorting all observations non-decreasingly with respect to the corresponding squared OLS residuals $e_{t, S_*^{(m)}}^2, t = p, \dots, T$. Then, we form a new subset given by the $m + 1$ observations with the smallest squared OLS residuals. The procedure is repeated until $m = T$, that is until all observations are included in the CDS. In the first steps, the forward search avoids the inclusion of outliers and immediately highlights units far away from the bulk of the data.

The forward search test statistic for ARCH components λ_{FS} (Grossi and Laurini, 2004), can be defined as a collection of test statistics derived from least squares estimators in each step of the forward search, that is

$$\lambda_{FS} = (\lambda_{(p)}, \dots, \lambda_{(T)})^T,$$

where $\lambda_{(m)}$ is the test statistic computed on m observations. In most moves from m to $m + 1$ just one new unit joins the subset. It may also happen that two or more observations

join the subset as one or more leave it. This behavior is called “interchange” and it is frequent in the first steps of the forward search when the initial CDS is unstable. Note that in the forward search approach a highly robust method is used and, at the same time, least squares estimators which are fully efficient.

2.4. Diagnostic monitoring

A major advantage of the forward approach is to provide the user with a number of informative pictures displaying all the diagnostics computed along the search. For instance, the anomalous behavior of outliers or level shifts is clearly revealed by the individual trajectories in the forward plot of squared residuals e_t^2 , even when standard deletion diagnostics suffer from masking. An additional bonus of the forward search algorithm is that it is often possible to rank the observations following their entrance step into the CDS. This gives an ordering of the data according to their degree of agreement with the underlying model, with observations furthest from it joining the CDS in the last steps of the procedure. During the forward search we can graphically monitor, for example, $\lambda_{(m)}$, coefficients $\hat{\alpha}_{i(m)}$ and t -statistics of coefficients; sharp changes of the curves point out the introduction of influential observations into the CDS.

2.5. Weighting observations

Classical outlier detection methods in linear regression are based on deletion diagnostics which evaluate how the deletion of one or more observations influences the parameter estimates, the residual sum of squares and the residuals. The algebra of deletion is important in deriving some quantities we monitor during the forward search and compare with given confidence regions in order to robustly quantify their degree of outlyingness. Let $y_t = e_t^2$ and $x'_t = (1, e_{t-1}^2, \dots, e_{t-q}^2)$, with $t = p, \dots, T$ and $\alpha' = (\alpha_0, \alpha_1, \dots, \alpha_q)$ be a p vector containing the parameter estimates coming from model (1). The test for the agreement of the observed and predicted value is

$$\tilde{e}_t = \frac{y_t - x'_t \hat{\beta}_m^*}{s_{S_*^{(m)}} \sqrt{[1 + x'_t (X'_{S_*^{(m)}} X_{S_*^{(m)}})^{-1} x_t]}} \quad (2)$$

where $\hat{\beta}_m^*$ is the least squares estimator based on the m observations contained in the subset, $s_{S_*^{(m)}}^2 = \sum_{i=1}^m e_{[i]}^2 / (m)$ is the residual mean square estimate of the variance σ^2 , $X_{S_*^{(m)}}$ is the $m \times p$ matrix that contains the m rows of the matrix X associated with the units which belong to the subset. During the forward search the residuals (2) can be used to test the effect of adding observations. When the t – th observation comes from the same population as the other observations $\tilde{e}_t$ has a Student- t distribution on $(T - 2p - 1)$ degrees of freedom. The strategy to obtain the weights works as follows:

- compare trajectories of $\tilde{e}_t$ during the forward search with a confidence region based on the t distribution; trajectories included in the region
- measure the degree of outlyingness of an observation as the area between the corresponding trajectory of (2) during the forward search and the bound of the

confidence region, that is:

$$\begin{aligned} &\text{when } \tilde{e}_{t,S_*^{(m)}} \in [-z_{\lambda/2}, +z_{\lambda/2}] \quad \forall m = p, \dots, T \\ &\quad \gamma_t = 0 \\ &\text{when } \tilde{e}_{t,S_*^{(m)}} > +z_{\lambda/2} \text{ for at least one } m = p, \dots, T \\ &\quad \gamma_t = \int_{(x: \tilde{e}_{t,S_*^{(x)}} > z_{\lambda/2}; p \leq x \leq T)} (\tilde{e}_{t,S_*^{(x)}} - z_{\lambda/2})^2 dH \end{aligned}$$

which, being H a discrete measure, can be written as

$$\gamma_t = \sum_{(x: \tilde{e}_{t,S_*^{(x)}} > z_{\lambda/2}; p \leq x \leq T)} (\tilde{e}_{t,S_*^{(x)}} - z_{\lambda/2})^2.$$

Similarly, when

$$\begin{aligned} &\tilde{e}_{t,S_*^{(m)}} < -z_{\lambda/2} \text{ for at least one } m = p, \dots, T \\ &\quad \gamma_t = \int_{(x: \tilde{e}_{t,S_*^{(x)}} < -z_{\lambda/2}; p \leq x \leq T)} (\tilde{e}_{t,S_*^{(x)}} - z_{\lambda/2})^2 dH \end{aligned}$$

that can be computed as

$$\gamma_t = \sum_{(x: \tilde{e}_{t,S_*^{(x)}} < -z_{\lambda/2}; p \leq x \leq T)} (\tilde{e}_{t,S_*^{(x)}} - z_{\lambda/2})^2$$

where $z_{\lambda/2}$ is the $\lambda/2 - th$ quantile of the t distribution on $(T - 2p - 1)$ degrees of freedom.

- The degree of outlyingness is finally used to obtain a weight w_t for each observation, that is:

$$w_t = \exp(-\gamma_t)$$

which gives low weight to influential observations, i.e. observations with high squared residuals.

The last step of the procedure is based on a weighted version of (1) where the vector of weighted parameter estimates α^* is given by

$$\alpha^* = (X'WX)^{-1}X'Wy. \quad (3)$$

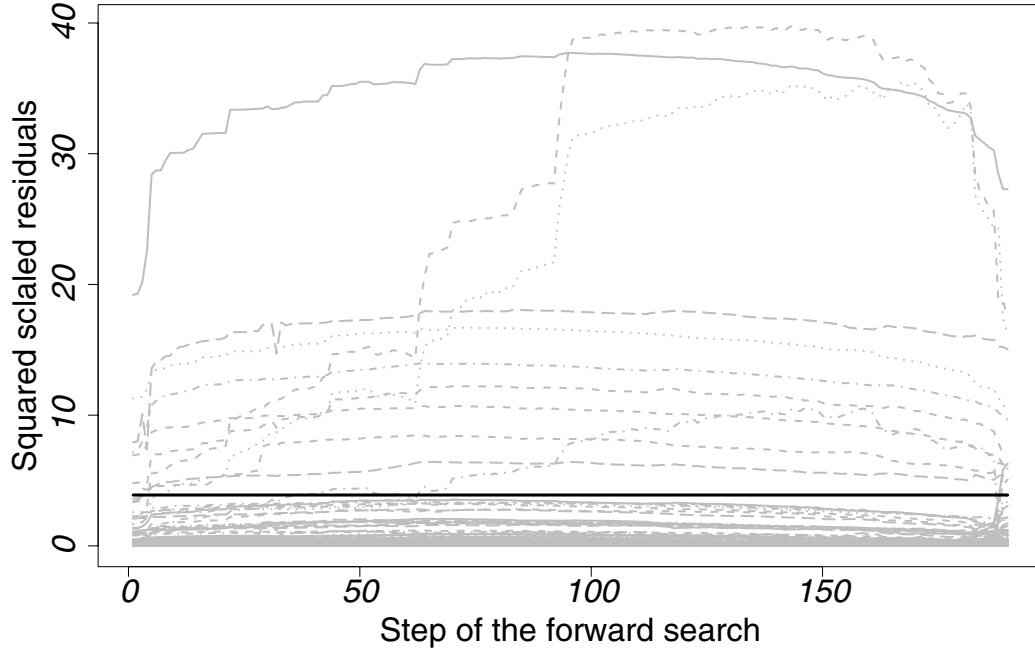
W is a $(T - p) \times (T - p)$ diagonal matrix containing the weights on the main diagonal.

The robust version of the LM test is be given by

$$\lambda^* = (T - p)(R^2)^* \quad (4)$$

where $(R^2)^*$ is the coefficient of determination of the linear model with parameter estimates obtained from expression (3).

Figure 1: Example for the evaluation of weights adopting the forward search trajectories in simulated data



3. Some results on simulated series

In this section we show that the robust test suggested in section 2 has higher power than the traditional LM test, even if the return series does not contains outliers. We have generated many series of 200 observations from a ARCH(1) model with parameters $\alpha_0 = 0.1$, $\alpha_1 = 0.5$. The necessary starting values for both y_t and η_t are set equal to zero, while the starting value for σ_t^2 is set equal to the unconditional variance. The first 100 observations of each series are discarded in order to avoid possible dependence of results on the initial conditions. Figure 1 reports trajectories of squared studentized residuals computed as in (2) at each step of the forward search. Each trajectory refers to an observation, while the straight bold line indicates the 95% quantile of a F distribution with 1 and $T - 2p + 1$ degrees of freedom which is the reference distribution of squared studentized residuals. According to what has been stated in section 2, the degree of outlyingness of the units beyond the critical value are proportional to the area delimited by the F -quantile and the corresponding trajectory. The most extreme observations along the entire procedure will be downweighted with respect to the other. Note that the weighting method applied in this paper takes into account the history of each unit during the search avoiding the effects of masking and swamping. Classical methods of outlier detection are instead based only on the last step highlighted in the plot where the ranking of the units is badly biased by the influence of extreme observations on the parameter estimates. The classical LM Engle's test does not reject the null hypothesis with p -value equal to 0.118, while the

robust version of the test rightly rejects the null hypothesis of residuals homoscedasticity with p -value equal to zero.

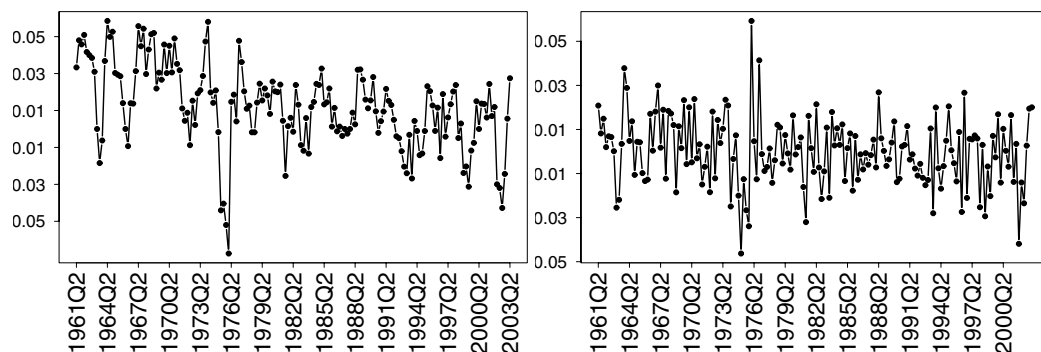
4. Analysis of real data

The purpose of this section is to explore properties of our method when applied to real time series, bearing in mind that series which show non-linear features are common in monthly sampled macroeconomic time series and daily financial time series.

4.1. Japanese industrial production

Output growth is one of the most analyzed macro-economic phenomena. While initially breaks in the mean of output growth were emphasized, more recently the focus has shifted to changes in output volatility (Blanchard and Simon, 2001; Stock and Watson, 2005). The occurrence of outliers in output has received considerably less attention, although Balke and Fomby (1994) and van Dijk et al. (1999) document the presence of aberrant observations in GDP and industrial production, respectively. In this section we analyze quarterly growth rates of industrial production in Japan over the period 1961Q2-2003Q2 which contains 168 observations (see left panel of Figure 2). The time series is taken from the OECD Main Economic Indicators, and is seasonally adjusted. The autocorrelation functions suggest the application of a AR(1) filter whose residuals are reported in the right panel of Figure 2. The presence of a short patch of outliers can be noticed over the period 1975Q2-1976Q2. Table 1 shows p -values of the standard and robust LM test on the AR(1) residuals from 1 to 5 ARCH components. Both versions of the test do not reject the hypothesis of homoscedasticity until lag 2, while from lag 3 to 5 they lead to opposite conclusions. In fact, the robust test does not still reject the hypothesis of homoscedasticity and the standard version rejects the same hypothesis being influenced by the presence of the patch of outliers previously pointed out.

Figure 2: Seasonal adjusted growth rate of Japanese (left panel) production and residuals after AR(1) filtering of the conditional mean (right panel)



4.2. Index of Italian Stock Exchange SPMIB

We describe the introduction of the weights for the series of Italian Stock index SPMIB, a synthetic index introduced in the Stock Exchange in June 2003. Observations P_t ,

lag	Standard LM test	Robust LM test
1	0.198976	0.889795
2	0.317027	0.936579
3	0.041661	0.897906
4	0.005227	0.828816
5	0.010462	0.814145

Table 1: *Japan Industrial production growth. Comparing standard LM test with the suggested robust version. The first column indicates the number of ARCH components considered, the second and the third column the p-values of the standard and robust LM test, respectively*

lag	Standard LM test	Robust LM test
1	0.938864	< 0.0001
2	0.960092	< 0.0001
3	0.593282	< 0.0001
4	0.74336	< 0.0001
5	0.820941	< 0.0001
6	0.788328	< 0.0001
7	0.804826	< 0.0001
8	0.653074	< 0.0001
9	0.718445	< 0.0001
10	0.106732	< 0.0001
11	0.106732	< 0.0001

Table 2: *Comparing standard LM test with the suggested robust version. The first column indicates the number of ARCH components considered, the second and the third column the p-values of the standard and robust LM test, respectively*

$t = 1, \dots, T$ have daily frequency for an overall number of days $T = 588$, and the reference price is the close price. We consider the transformed log-returns process $r_t = \log(P_t) - \log(P_{t-1})$, and the series r_t is plotted in Figure 3 (left panel), where also the square of the series, r_t^2 , is shown (right panel of Figure 3). According to the autocorrelation functions the series r_t seems to be a white noise, while analyzing the autocorrelation of the series r_t^2 it is unclear the dependence's structure of the series. The marginal distribution of the observed series seems to have leptokurtic behavior. These findings motivate the use of GARCH models, mainly suited at modeling the non-linear dependence of r_t shown in the right panel of Figure 3.

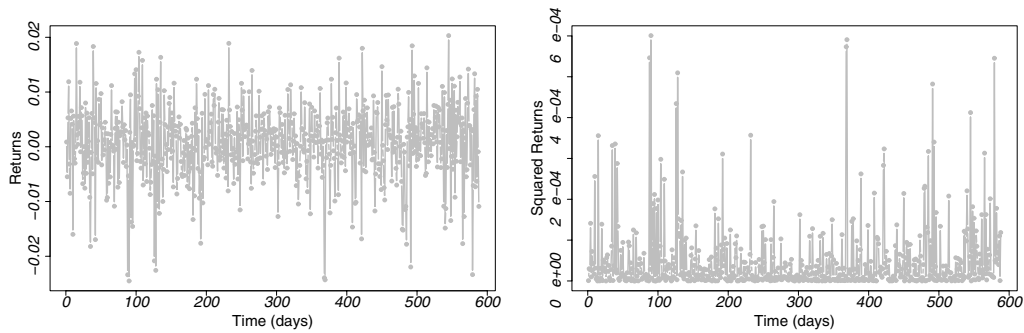
However, the LM test for ARCH effect suggests that there is no evidence for rejection of the null hypothesis of no ARCH effect at any lag. This striking result is mainly motivated by the poor fit of the auxiliary regression (1).

Applying the forward search on the series, we have observed the behaviour of the squares of studentized residuals, that we report in Figure 4. It turns out that many observations are influential and their effect is to reduce the overall goodness of fit for regression (1). Through the forward search we also have monitored the behavior of t -statistics for α_j coefficients of auxiliary regression (1) which lead to acceptance of null hypothesis of no ARCH effect in observed data. Figure 5 shows that t -statistics of α_j

are in the area of acceptance of the null hypothesis. The contribution of α_1 (the top solid line in Figure 5) is considerably reduced in the last steps of the forward search, where the joint inclusion of influential observations leads to masking effect.

The introduction of weights, as discussed in Section 2, leads to draw an opposite conclusion about the presence of non-linear features. By exploring Figure 4 there is evidence that some observations, that are the residuals of auxiliary regression (1), have greater influence. Their distance to the nominal F statistics is big, yielding to a weight of corresponding observations virtually equal to zero, due to transformation. Notice that almost all observations will be weighted given that the number of observations that are not labeled as influential during the forward search is very small. Despite differences of method, in this example the results found with the forward search weighting algorithm are broadly consistent to that of Van Dijk et al. (1999) (not reported), where the number of weighted observations is quite different.

Figure 3: *SPMIB index returns sampled at daily frequency and their squared values*



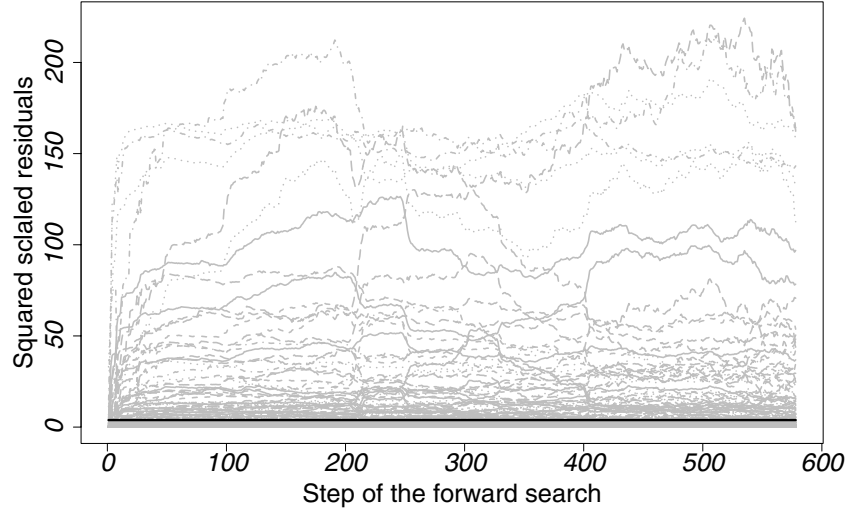
5. Discussion

The method described for weighting observations in time series is based on information of the forward search on the LM test of Engle for detection of ARCH components. Numerical results are quite good, at least in those circumstances where the test does not identify properly the presence of non-linear components.

From the simulation considered here it is clear that using the weighted regression is useful, at least in reducing the bias of the standard LM ARCH test. From real data analysis, the results given by the weighted LM test are overwhelming.

However, the aim of statistical tests is to provide conclusions concerning nonlinearity that are broader than what is usually studied. For instance, it seems reasonable to specify alternative hypotheses where the actual number of ARCH component is specified, rather than concluding that *there is at least one* ARCH component. As long as alternative hypotheses are specified appropriately, one possible extension to ARCH test of Engle is to study a likelihood ratio test, where one needs to evaluate the test under the null hypothesis H_0 and under the alternative, that may take many forms. To our understanding, a sensible way of specifying the alternative is to impose an increasing number of ARCH components. Thus, rather than specifying $H_1 : \exists_{i=1,\dots,q} \alpha_i \neq 0$, it would make sense to specify $H_1 : \alpha_1 > 0; \alpha_2 = \dots = \alpha_q = 0$ and study the likelihood under this last

Figure 4: Squared studentized residuals for the series SPMIB, under the assumption of no model for the conditional mean



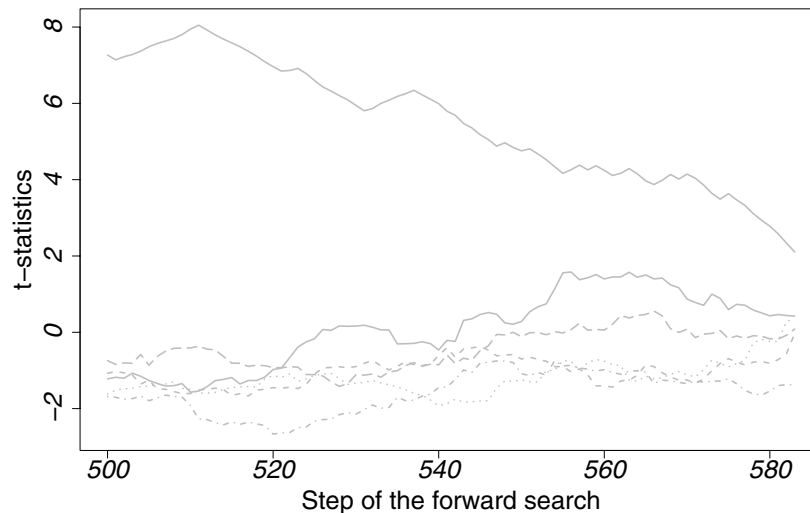
specification. If the likelihood ratio test is such that the null H_0 is accepted we stop, otherwise we extend to a broader $H_1 : \alpha_1 > 0, \alpha_2 > 0; \alpha_3 = \dots = \alpha_q = 0$, study the statistic under this last H_1 and compute the likelihood ratio test, and so on.

The method that we have suggested in this paper in order to weight observations needs further investigation. Monte Carlo simulations are required for assessing the accuracy, size and power of the weighted test; this issue has not been tackled at all in this paper. Moreover, there are possible extensions to the algorithm that we have proposed, and we want to focus onto two main directions. On the one hand the distribution of scaled residuals considered in the forward search is Student- t only toward the end of the search, therefore our method is actually approximated. In order to fully understand the distribution of scaled residuals at any step of the forward search it is required to find the distribution of quadratic forms for truncated samples. On the other hand, the transformation induced by $\exp(-x)I_{x>0}$ can be generalized to some parametric version, by allowing some flexibility, accounting differently the distance of the units to theoretical confidence bands.

References

- Atkinson, A. C. and Riani, M. (2000) *Robust Diagnostic Regression Analysis*. Springer Verlag, New York.
- Balke, N.S. and Fomby, T.B. (1994) Large shocks, small shocks, and economic fluctuations: Outliers in macroeconomic time series. *Journal of Applied Econometrics* 9, 181–200.
- Blanchard, O.J. and Simon, J. (2001) The long and large decline in US output volatility *Brooking Papers on Economic Activity*, 135–174.

Figure 5: t -statistics of LM ARCH test for the SPMIB index, where last steps of the forward search are monitored. Top line refers to the t -statistic for parameter α_1



- Bollerslev, T. (1986) Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics*, 31, 307–327.
- Engle, R. F. (1982) Autoregressive conditional heteroscedasticity with estimates of the variance of UK inflation. *Econometrica*, 50, 987–1007.
- Engle, R. F. and Ng, V. K (1993) Measuring and Testing the Impact of News on Volatility *The Journal of Finance*, 48, 1749–1778.
- Franses, P. H. and Van Dijk, D. (2000) *Non-Linear Time Series Models in Empirical Finance*. Cambridge University Press, Cambridge.
- Grossi, L. and Laurini, F. (2004) Analysis of economic time series: effects of extremal observations on testing heteroscedastic components. *Applied Stochastic Models in Business and Industry*, 20, 115–130.
- Grossi, L. and Riani, M. (2002) Robust Time Series Analysis Through the Forward search, Proceedings of Full Paper, COMPSTAT 2002, Physica-Verlag, Berlino, 521–526.
- Lee, J.H.H (1991) A Lagrange multiplier test for GARCH models. *Economics Letters*, 37, 265–271.
- Rousseeuw, P. J. (1984) Least median of squares regression. *Journal of the American Statistical Association*, 79, 871–880.
- Rousseeuw, P. J. and Leroy, A. M. (1987) *Robust Regression and Outlier Detection*. Wiley, New York.
- Stock, J.H. and Watson, M.W. (2005) Understanding changes in international business cycle dynamics. *Journal of the European Economic Association*, to appear.
- van Dijk, D., Franses, P. H. and Lucas, A. (1999) Testing for ARCH in the Presence of Additive Outliers. *Journal of Applied Econometrics*, 14, 539–662.