Non Linear Moving-Average Conditional Heteroskedasticity

Ventosa-Santaulària, Daniel and Mendoza V., Alfonso

Escuela de Economia, Universidad de Guanajuato., Deparatamento de Economia, Universidad de las Americas, Puebla

2005

Non Linear Moving-Average Conditional Heteroskedasticity

Daniel Ventosa-Santaulària*
Alfonso Mendoza V.†

Abstract

Ever since the appearance of the $ARCH$ model (Engle 1982a), an impressive array of variance specifications belonging to the same class of models has emerged. Despite numerous successful developments, several empirical studies seem to show that their performance is not always satisfactory see Boulier (1994).
In this paper a new alternative to model conditional heteroskedastic variance is proposed: the Non-Linear Moving Average Conditional Heteroskedasticity: (NLMACH). While NLMACH properties are similar to those of the ARCH-class specifications this new proposal represents a convenient alternative to modeling conditional volatility through a non-linear moving average process. The $NLMACH$ performance is investigated using a Monte Carlo experiment and modeling exchange rate returns. It is found that NLMACH outperforms GARCHs forecasts whereas the application to exchange rates provides mixed evidence.

Keywords: Conditionally heteroskedastic models, $NLMACH(q)$, Volatility, Fat tails.
JEL classification: C22, C13, C12.

1 Introduction

The $ARCH$ class models, introduced by Engle (1982a), quickly became an important domain in the econometric literature because of their potential usefulness in financial applications. During the last twenty years, a vast quantity of $ARCH$ type models appeared, some of them possessing statistical properties extremely appealing to financial econometrics. Among them, the $GARCH$ model (Bollerslev 1986) has proved to be a very useful tool in the modeling of a wide array of financial variables. Other extensions such as the $ARCH-M$ (Engle,

*Corresponding Author: Escuela de Economia, Universidad de Guanajuato. Address: UCEA-Campus Marfil Fracc. I, El Establo, Guanajuato Gto 36250 Mexico. e-mail: daniel@ventosa-santaularia.com

†Departamento de Economia, Universidad de las Americas, Puebla. Address: Santa Catarina Martir, Cholula, Puebla 72820 Mexico. e-mail: alfonso.mendoza@udlap.mx

Lilien, and Robins 1987) and the $EGARCH$ (Nelson 1991) have succeeded in generalizing $ARCH$ models by incorporating the volatility of a variable in its mean equation and taking into account asymmetric effects respectively.

The evolution of the $ARCH$ models seems to follow a pattern. Each new specification tries to incorporate more characteristics typical of financial series such as leptokurticity, asymmetry and different kinds of non-linearity. Such progress is made at a cost of increasing complexity. The latter eventually makes some of the specifications to appear as having little robustness in empirical studies. This is perhaps why the popular $GARCH(1,1)$ model remains one of the best options for practitioners of financial econometrics.

When dealing with conditionally heteroskedastic models, the accent has always been put in Autoregressive specifications, neglecting the potential usefulness of Non-Linear Moving Average type specifications (although some models, such as $GARCH$ can be reinterpreted as very particular Moving-Average specifications). In that sense, Robinson (1977) proposed a Non-Linear Moving-Average model ($NLMA$) inspired by a truncated version of a Volterra expansion. He also gave the statistical properties of such model as well as several properties of a maximum likelihood estimator. Sadly, he did not present an empirical application of the $NLMA$ and did not consider it a practical model for financial variables. Indeed, $NLMA$ models are nowadays seen as being ineffectual for empirical purposes (Tong 1990, Guégan 1994, Granger 1998).

Despite these criticisms, we believe $NLMA$ can play a role similar to the one played by MA in linear modeling, although the process must be redefined in order to avoid the main difficulties of Robinson's (1977) proposal, i.e. non-invertibility and difficult estimation due non-linearity. We define a different specification, the Nonlinear Moving Average Conditionally Heteroskedastic model, $NLMACH$. Basically, we replace the explanatory variable $X_{t-1}^2$of the conditional variance in an $ARCH$ model with a non-observed white noise and obtained a model with simple theoretical properties and, most importantly, easy to estimate. Such specification can reproduce several of the typical characteristics of financial variables, such as: (1) high frequency of large variations; (2) tendency of large variations (in absolute terms) to cluster, and very interestingly, (3) leptokurticity. There are important advantages of this model when compared to the ARCH-class ones. Stationarity conditions are, for example, less stringent. The $NLMACH$ is estimated using simulation techniques and a set of currencies. Its properties are then compared to $ARCH$ and $GARCH$. Also, using Monte Carlo simulations, we present evidence that the estimators perform well.

This paper is divided in four sections. The second introduces the NLMACH model and the third deals with the estimation and identification problem. Conclusions appear in section four.

2 New proposal: the NLMACH

Engle's (1982a) $ARCH$ model brought about an impressive array of variance specifications belonging to the same class. Despite ARCH's successful developments, it can be argued that the $NLMACH$may be more relevant for the study of some particular phenomena. Some variables may be heteroskedastic, and yet being poorly adjusted by $ARCH$ models. $NLMACH$ may be a suitable alternative in such cases.

This section proposes a new conditional heteroskedastic variance model: the Quadratic Moving-Average Conditional Heteroskedasticity ($NLMACH$). Its properties are roughly the same as those of ARCH-class specifications but our model has in addition several important advantages. It is simple, easy to estimate, captures the high kurtosis observed in financial returns and impose fewer and less stringent existence conditions (stationarity). Indeed, it represents an alternative to the $ARCH - class$ when dealing with heteroskedasticity. As it will be explained later, $NLMACH$ heteroskedasticity is fundamentally different to $ARCH$ one.

2.1 The NLMACH model

Although the $NLMACH$ model is a non-linear MA, it cannot be encompassed in Robinson's (1977) $NLMA$ specifications. The latter has several unappealing properties, among them non-invertibility (Granger and Andersen 1978, Granger 1998) stands out. We propose a different model still possessing some very attractive characteristics; the $NLMACH(1)$:

$$\begin{aligned} X_t &= V_t h_t^{1/2} \\ h_t &= \delta_0 + \delta_1 V_{t-1}^2 \end{aligned} \tag{1}$$

Where, $V_t \sim_{iid} \mathcal{N}(0,1)$ and $\delta_0, \delta_1 > 0$.

As can be inferred from (1), the $NLMACH(1)$ is deeply inspired from an $ARCH(1)$. Yet, in our case, the explanatory variable of the conditional variance is not X_t^2 but rather V_t^2. Parameters must satisfy a condition in order to ensure positiveness ($\delta_i > 0$ for $i = 1,2$) of the conditional variance. Normality - and unit variance- of the white noise can also be seen as a condition of the model[1]. Its interesting to notice that the $NLMACH(q)$ yields a naturally fat-tailed distribution, conveying automatically a must wanted characteristic among financial econometricians.

2.1.1 Distribution of the first-order NLMACH process

The $NLMACH(1)$ has the advantage of being a very simple specification. Most of its properties can be inferred straightforward. In order to make a brief com-

[1] It may be interesting to modify such condition (using a t distribution instead, for example), so the model can broaden its scope. This will be This will be address in the empirical section of this paper.

parison with the $ARCH(1)$, we present the first two - unconditional and conditional - moments of the process:

$$
\begin{aligned}
E(X_t) &= 0 \\
E(X_t X_{t-j}) &= \begin{cases} \delta_0 + \delta_1 & for\, j = 0 \\ 0 & otherwise \end{cases} \\
E_{t-1}(X_t) &= 0 \\
E_{t-1}(X_t^2) &= \delta_0 + \delta_1 V_{t-1}^2
\end{aligned} \tag{2}
$$

where $\delta_0, \delta_1 > 0$.

Expression (2) shows that the $NLMACH(1)$ is weakly stationary. Figure (1) shows a simulation of a first order $NLMACH$.

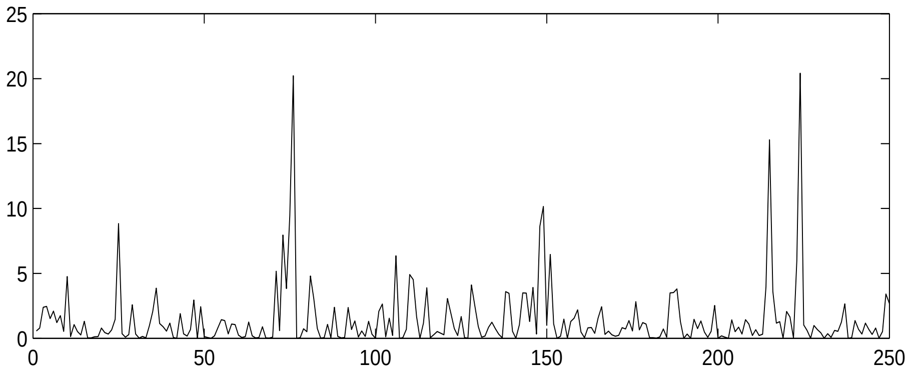

Figure 1: $NLMACH(1)$ Simulation: $h_t = 1 + 0.7V_{t-1}^2$

It can be seen that, contrary to most of the specifications of conditionally heteroskedastic models, there are fewer conditions for the existence of the second moment[2].

2.1.2 Stationarity of the NLMACH

Covariance stationarity of the $NLMACH$ specification was fairly easy to prove. In this section we demonstrate that, under the already mentioned hypothesis (normality of the white noise, and positiveness of the parameters), all the moments of a $NLMACH(q)$ exist.

theorem 1 *Let X_t be a $NLMACH(q)$ process satisfying the following equations:*

[2]Of course, we must not forget the hypothesis made on V_t. The latter must be a gaussian iid zero-mean white noise with unit variance. Also, there are positiveness constraints on the parameters.

$$
\begin{aligned}
X_t &= V_t h_t^{\frac{1}{2}} \\
h_t &= \delta_0 + \sum_{i=1}^{q} \delta_i V_{t-i}^2
\end{aligned} \tag{3}
$$

With $V_t \sim_{iid} \mathcal{N}(0,1)$ and $\delta_i > 0$ $\forall i = 1, 2, \cdots, q$.

Then, all the moments of X_t, $E\left(X_t^r\right)$ $\forall r = 1, 2, \cdots$ exist.

proof of theorem 1.

Odd moments can be easily calculated because of the properties of the gaussian white noise V_t. Indeed, all odd moments are equal to zero. We thus concentrate in even moments. The general formula of even moments is:

$$
\begin{aligned}
E(X_t^{2r}) &= E(V_t^{2r}) \cdot E(h_t^r) \\
&= \prod_{j=1}^{r} (2j-1) \cdot E\left[\left(\delta_0 + \sum_{i=1}^{q} \delta_i V_{t-i}^2\right)^r\right]
\end{aligned}
$$

It can be seen that the first term, $\prod_{j=1}^{r}(2j-1)$, has no conditions of existence. We have to develop the second term to look for "possible" conditions.

$$
\begin{aligned}
E(h_t^r) &= E\left[\left(\delta_0 + \sum_{i=1}^{q} \delta_i V_{t-i}^2\right)^r\right] \\
&= E\left[\sum_{j=0}^{r} \binom{r}{j} \delta_0^{r-j} \cdot \left(\sum_{i=1}^{q} \delta_i V_{t-i}^2\right)^j\right] \\
&= \sum_{j=0}^{r} \binom{r}{j} \delta_0^{r-j} \cdot E\left(\sum_{i=1}^{q} \delta_i V_{t-i}^2\right)^j
\end{aligned}
$$

We realize that we have to obtain the value of the second term, that is, $E\left(\sum_{i=1}^{q} \delta_i V_{t-i}^2\right)^j$. We can develop the latter by means of Newton's Formulae, as follows:

$$
\begin{aligned}
E\left(\sum_{i=1}^{q} \delta_i V_{t-i}^2\right)^j &= E\left[\sum_{z=0}^{j} \binom{j}{z} \left(\delta_1 V_{t-1}^2\right)^{j-z} \cdot \left(\sum_{i=2}^{q} \delta_i V_{t-i}^2\right)^z\right] \\
&= \sum_{z=0}^{j} \binom{j}{z} \delta_1^{j-z} E\left(V_{t-1}^{2(j-z)}\right) \cdot E\left(\sum_{i=2}^{q} \delta_i V_{t-i}^2\right)^z \\
&= \sum_{z=0}^{j} \binom{j}{z} \delta_1^{j-z} \prod_{k=1}^{j-z} (2k-1) \cdot E\left(\sum_{i=2}^{q} \delta_i V_{t-i}^2\right)^z
\end{aligned}
$$

We notice, once again that we should only worry about a single element, in this case $E\left(\sum_{i=2}^{q}\delta_i V_{t-i}^2\right)^k$. The sum has now fewer elements (it goes from $i = 2$ to q). This sum can indeed go over the same process (basically another application of Newton's Formulae) in order to reduce the number of elements. Eventually, we'll arrive to a sum with only one element:

$$E\left(\delta_q V_{t-q}^2\right)^s = \delta_q^s \cdot \prod_{l=1}^{s} 2l-1$$

So, we have "eliminated" all the expectation operators of the expression. There are thus, no conditions (except the normality of the white noise and the positiveness constraint) of existence for the unconditional moments of a $NLMACH(q)$.

Q.E.D.

We have also calculated the degree of Kurtosis, which is superior to 3, if $\delta_i > 0$ for at least one i, $i = 1, \cdots, q$ and if $\delta_i \geq 0 \, \forall \, i = 1, \cdots, q$:

$$\begin{aligned}\mathcal{K} &= \frac{(X_t)^4}{\sigma^4} \\ &= \frac{3\left[\left(\delta_0 + \sum_{i=1}^{q}\delta_i\right)^2 + 2\sum_{i=1}^{q}\delta_i^2\right]}{\left(\delta_0 + \sum_{i=1}^{q}\delta_i\right)^2} \\ &> 3\end{aligned} \tag{4}$$

proof.
By rearranging the terms of expression (4), we get:

$$\sum_{i=1}^{q}\delta_i^2 > 0$$

Which is true if, and only if $\delta_i \neq 0$ for at least one i, $i = 1, \cdots, q$.

Q.E.D.

2.1.3 Invertibility of the NLMACH

Invertibility has always been a problem when dealing with moving average processes, whether they are linear or not. As pointed out earlier, a $NLMACH(1)$ satisfying the normality hypothesis $V_t \sim_{iid} \mathcal{N}(0,1)$ yields the autocovariance structure stated in equation (2). The latter allows us to obtain the autocovariance function of the process, which is similar to the one yielded by a white noise:

$$g_x(z) \quad = \quad \delta_0 + \delta_1 \tag{5}$$

Thus, the autocovariance function is a constant. The invertibility of the specification may appear now clearly. On typical $NLMA$, it happens that different sets of parameters, yield the same autocovariance function (so the parameters are not identifiable). For the $NLMACH$ this does not occurs thanks to the positiveness constraint imposed on the parameters, $\delta_0, \delta_1 > 0$. It must be remembered that such condition appears naturally if we want the conditional variance to be always positive. Such condition not only ensures the positiveness of the conditional variance, but it also solves the identification problem of the parameters. We are thus able to reconstruct the unobserved white noise which can be seen as a proof of invertibility (Granger and Terasvirta 1993).
For the linear MA(q) process, conditions ensuring invertibility are well known. Our particular model, when manipulated algebraically, can exhibit analogous conditions. From the conditional variance expression stated in (1), we can get:

$$\begin{aligned} h_t &= \delta_0 + \sum_{i=1}^{q} \delta_i \left(V_{t-i}^2 - 1 \right) + \sum_{i=1}^{q} \delta_i \\ &= \varsigma + \sum_{i=1}^{q} \delta_i W_{t-i} \end{aligned} \tag{6}$$

where $\varsigma = \delta_0 + \sum_{i=1}^{q} \delta_i$ is a constant and $W_t = V_{t-i}^2 - 1$ is a non gaussian noise such that:

$$\begin{aligned} E(W_t) &= 0 \\ E(W_t W_{t-j}) &= \begin{cases} 2 & for\, j = 0 \\ 0 & otherwise \end{cases} \end{aligned} \tag{7}$$

We realize that h_t can be understood as a non gaussian $MA(q)$ and thus, the usual invertibility conditions apply, that is, the process is invertible if the roots of the polynomial $\left(1 + \delta_1 z + \delta_2 z^2 + \ldots + \delta_q z^q\right) = 0$ lie outside the unit circle.

2.1.4 Defining the value of q in a $NLMACH(q)$

In the next section, we present a estimation technique dealing in particular with a $NLMACH(q)$. Of course, once this model is to be used with real data, there is an additional requirement; the identification of the parameter q. The order of the $NLMACH(q)$ process can be inferred by means of its sample squares autocorrelation function. This is true because of the structural properties of the model we develop here. So identification of q must be done through the $SACF$ of the squares of the process. Undoubtedly, other tools allowing such inference can be found, but in this work we concentrate our efforts in the $SACF$. First of all, the theoretical shape of the autocorrelation function is to be developed:

Let X_t be a $NLMACH(q)$ specified in expression (3). Then, the autocorrelation function of the squares of X_t is:

$$\rho\left(X_t^2, X_{t-j}^2\right) = \begin{cases} \gamma_i & for\ j < q \\ \frac{\delta_q\left(\delta_0 + \sum_{i=1}^q \delta_i\right)}{\left(\delta_0 + \sum_{i=1}^q \delta_i\right)^2 + 3\sum_{i=1}^q \delta_i^2} & for\ j = q \\ 0 & \forall\, j > q \end{cases} \quad (8)$$

where,

$$\gamma i = \frac{\delta_j \sum_{i=0}^q \delta_i + \sum_{i=j+1}^q \delta_i \delta_{i-j}}{\delta_0^2 + 2\delta_0 \sum_{i=1}^q \delta_i + \left(\sum_{i=1}^q \delta_i\right)^2 + 3\sum_{i=1}^q \delta_i^2}$$

We now should be able to identify empirically the value of q by means of the sample autocorrelation function of the process's squares. In order to illustrate this, we simulated a $NLMACH(4)$ and plotted both, the sample and the theoretical autocorrelation function.

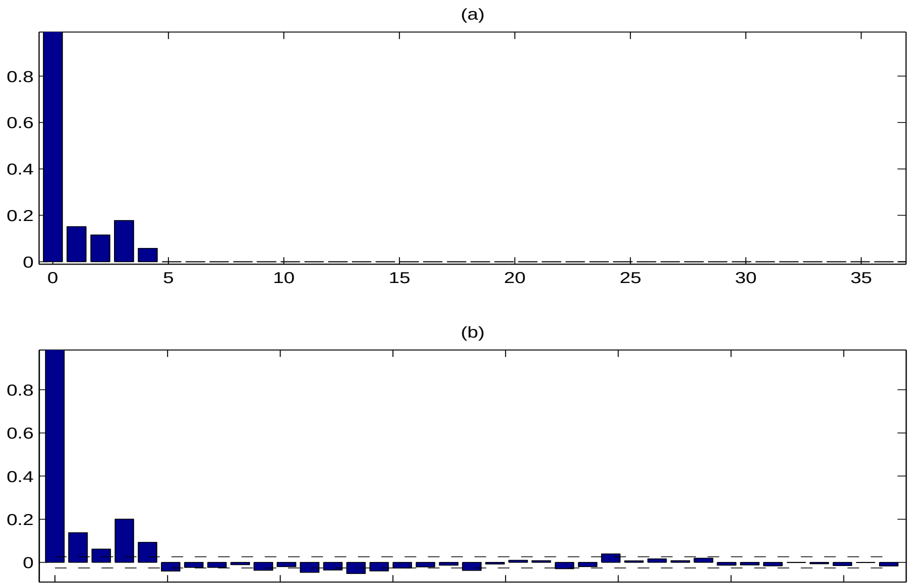

Figure 2: $NLMACH(4)$: (a) Theoretical ACF and (b) SACF

The autocorrelation function may yield a shape that approximates fairly well the one proposed by the stylized facts in finance theory. Yet, to achieve this we are forced to use a $NLMACH(q)$ with q greater than unity. An alternative to this is to generalize the process by including lags of h_t in the conditional variance specification. Although this seems to be an attractive option, it will not be done here.

3 Estimation of the $NLMACH(1)$

Once the main statistical properties have been established, the next step is estimation. The $NLMACH(1)$ estimation is simple despite the fact of being a highly non-linear model. In order to show the performance of the estimating technique, we present a Maximum Likelihood (ML) estimate. It works in the same way as with $ARCH$ models. The ML estimation of the $NLMACH(q)$ is straightforward. We take advantage of the fact that the conditional distribution is $\mathcal{N}(0, h_t^{1/2})$, that is, $X_t/\Psi_t \sim \mathcal{N}(0, h_t^{1/2})$, where Ψ_t is the past information set [3]. Under the usual regularity conditions, we are thus able to compute the corresponding Likelihood and maximize it using a gradient algorithm.
We performed a Monte Carlo Experiment to illustrate the ML estimator. Table (1) exhibits the estimation results for a variety of parameters(both parameters adopt the following values: 0.25, 0.50 and 0.75). 1,000 replications where made for each case. Table (1) shows the averages of such estimations as well as the standar deviations [4].

Parameters		Sample size					
		T=200		T=500		T=700	
δ_0	δ_1	$\widehat{\delta_0}$	$\widehat{\delta_1}$	$\widehat{\delta_0}$	$\widehat{\delta_1}$	$\widehat{\delta_0}$	$\widehat{\delta_1}$
0.25	0.25	0.250	0.248	0.251	0.249	0.251	0.249
		(0.04)	(0.07)	(0.04)	(0.05)	(0.02)	(0.04)
	0.50	0.249	0.494	0.251	0.498	0.250	0.506
		(0.04)	(0.12)	(0.03)	(0.08)	(0.02)	(0.07)
	0.75	0.251	0.742	0.253	0.748	0.250	0.749
		(0.04)	(0.18)	(0.03)	(0.11)	(0.02)	(0.09)
0.50	0.25	0.501	0.249	0.500	0.246	0.500	0.249
		(0.08)	(0.10)	(0.04)	(0.06)	(0.04)	(0.05)
	0.50	0.502	0.499	0.498	0.501	0.500	0.497
		(0.08)	(0.15)	(0.05)	(0.09)	(0.04)	(0.08)
	0.75	0.501	0.743	0.499	0.754	0.502	0.747
		(0.08)	(0.20)	(0.05)	(0.13)	(0.04)	(0.10)
0.75	0.25	0.756	0.243	0.754	0.247	0.749	0.251
		(0.11)	(0.12)	(0.07)	(0.07)	(0.06)	(0.06)
	0.50	0.753	0.504	0.751	0.492	0.749	0.502
		(0.11)	(0.18)	(0.07)	(0.11)	(0.06)	(0.09)
	0.75	0.759	0.752	0.752	0.748	0.747	0.745
		(0.12)	(0.22)	(0.07)	(0.14)	(0.06)	(0.12)

Table 1: Monte Carlo Simulation of estimates for a $NLMACH(1)$; N=200, 500 and 700

[3] $\Psi_t = \{X_{t-1}, X_{t-2}, \cdots, X_0, V_{t-1}, V_{t-2}, \cdots, V_0\}$
[4] Standard deviations are given in parentheses.

The Monte Carlo experiment reveals that, using a standard quasi-newton algorithm (Matlabs default) a convenient estimation can be performed, although its efficence could be improved. It is curious to notice that the standard deviation increases with the value of the parameter.

3.1 Forecasting capability of the Models

It must be said that our proposal ($NLMACH$) would not be particularly interesting if it was unable to offer good forecasting capabilities of the volatility of a variable. In order to study its performance in this domain, we simulate two DGPs; an NLMACH(1) and a GARCH(1,1). Over each simulated series we performed the estimation of both the NLMACH(1) and the GARCH(1,1) using only a fraction of the sample and constructed an out-of-sample forecast (one period ahead). Then we add an observation and rebuild the forecast until we use T-1 observations. Using these forecasts and knowing the real values we

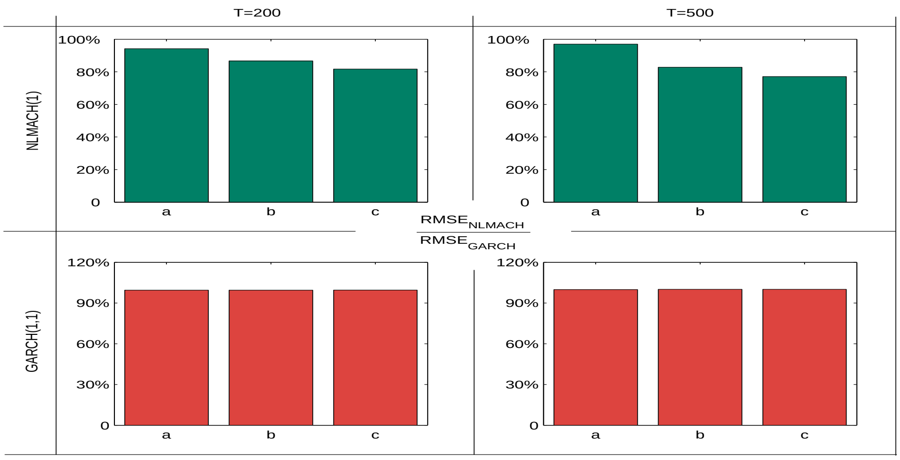

Figure 3: Monte Carlo Experiment: $NLMACH(1)$ and $GARCH(1,1)$ out-of-sample forecasts: NLMACH, cases a, b and c: $\delta_1 = 0.30$, GARCH, cases a, b and c: $\alpha = 0.30$ and $\beta = 0.30$
a) NLMACH: $\delta_1 = 0.15$; GARCH(1,1): $\gamma = 0.10$
b) NLMACH: $\delta_1 = 0.30$; GARCH(1,1): $\gamma = 0.30$
b) NLMACH: $\delta_1 = 0.45$; GARCH(1,1): $\gamma = 0.50$

compute the Root Mean Square Error for each specification and then compute the ratio: $\frac{RMSE_{NLMACH(1)}}{RMSE_{GARCH(1,1)}}$. We repeat the latter experiment 1000 times and show the results (averages) in figure (3)[5].

[5] We need to be cautious about this result. Bollerslev, Chou, and Kroner (1992) have warned that "[$\cdots$] out of sample forecasting is marred with difficulties and simply extrapolating the future vitality of the field based on past observations does not necessarily result in

The figure exhibits interesting results. when the real DGP is an $NLMACH(1,1)$, (first row of figure)the NLMACH(1) specifications forecasts outperforms the GARCH(1,1)s forecasts but the inverse is not completely true (second row of figure). If the DGP is a GARCH, there are several cases where the NLMACH(1), even if it is the wrong specification, yields better forecasts.

3.2 Application to Exchange Rates

In order to examine the $NLMACH$ performance using real market data, in this section we estimate the $NLMACH(1)$ model and compare it with the $ARCH(1)$ and $GARCH(1,1)$ processes. In addition we estimate the $ARCH(1)$ model assuming a Student-t distribution in order to capture the fat tails frequently observed in financial returns. Eight major currencies are employed for this exercise[6]. Daily exchange rate returns from January 2, 1991 to December 29, 1995 are calculated by taking the first log difference corresponding to a total of 1,303 observations for each currency. In particular the exchange rates under examination are the Australian Dollar (AUS), British Pound (GBR), Canadian Dollar (CAN), Dutch Guilder (NLG), French Franc (FRA), German Dmark (DEU), Japanese Yen (JPY) and Swiss Franc (CHF). Descriptive statistics are shown in Table (2) below.

Currency	Mean	Median	StdDev.	Min.	Max.	Skew.	Kurt.
Australian D.	0.0013	-0.0117	0.1997	-0.6954	0.8527	0.4171	1.6811
British P.	0.0075	-0.0079	0.2906	-1.2548	1.4271	0.3502	2.5506
Canadian D.	0.0055	0.0033	0.1213	-0.7040	0.5911	0.0320	2.7627
Dutch G.	-0.0016	-0.0083	0.3179	-1.2581	1.3060	0.0917	1.6173
French F	-0.0012	-0.0019	0.3004	-1.1734	1.1519	0.0700	1.6761
German M.	-0.0013	-0.0115	0.3190	-1.2578	1.3476	0.1233	1.6865
Japanese Y.	-0.0088	-0.0087	0.2828	-1.4727	1.4014	-0.2492	3.3761
Swiss F.	-0.0030	0.0000	0.3470	-1.6933	1.3517	-0.0223	1.5817

Table 2: Statistical characteristics of the exchange rate time series

Tables 3,4,5 and 6 present the estimation results for each currency for several model specifications. Different orders of the process were investigated. The NLMACH(1) was chosen according to the Akaike Information Criterion (AIC) and the Schwartz Bayesian Criterion (SBC). The optimization algorithm employed in the estimations was the BFGS and all the programs are written in RATS. Using robust standard errors it is observed that apart from the intercept all estimated parameters are highly significant–see δ_1 and δ_2 in each panel[7].

optimal predictions $[\cdots]$". Having said this, we should keep in mind the many limitations of time series forecasting performance.

[6] The data has been extensively examined by Franses and van Dijk (2000) for this subsample and from December 1979. The data is available in the authors' website.

[7] Note that δ_1 is associated to either the Nonlinear or ARCH effect respectively, whereas δ_2 is associated to GARCH effects.

	Dutch Guilder				Swiss Franc			
	NLMACH	ARCH	GARCH	ARCH-t	NLMACH	ARCH	GARCH	ARCH-t
				Estimated Coefficients				
C	-0.0019	-0.0019	-0.0046	-0.0056	-0.0011	-0.0010	-0.0070	-0.0036
	$(0.0079)^a$	(0.0074)	(0.0080)	(0.0093)	(0.0093)	(0.0098)	(0.0074)	(0.0082)
δ_0	0.0794*	0.0794*	0.0032*	0.0502*	0.0992*	0.0986*	0.0039*	0.0669*
	(0.0030)	(0.0055)	(0.0006)	(0.0027)	(0.0037)	(0.0034)	(0.0007)	(0.0025)
δ_1	0.0212*	0.2138*	0.0715*	0.1620*	0.0204*	0.1781*	0.0575*	0.1309*
	(0.0035)	(0.0466)	(0.0097)	(0.0289)	(0.0035)	(0.0271)	(0.0083)	(0.0281)
δ_2	–	–	0.8967*	–	–	–	0.9098*	–
	–	–	(0.0022)	–	–	–	(0.0015)	–
V^b	–	–	–	5.4836*	–	–	–	6.1327*
	–	–	–	(0.4327)	–	–	–	(0.7747)
				Decision Criteria				
$L(\theta)^c$	867.34	866.99	894.99	804.94	746.01	746.21	763.93	934.56
AIC^d	8753.71	8753.18	8796.28	8659.16	8558.86	8559.20	8591.54	8852.14
SBC^e	8769.20	8768.67	8816.94	8679.82	8574.35	8574.69	8612.20	8872.79

Table 3: Model Adjustment for the Dutch guilder and the Swiss Franc. *,** Significant at the 1% and 10% level respectively.[a] Robust Standard errors in parenthesis.[b] Shape parameter. [c]Optimizaed likelihood value.[d] AIC =Akaike information Criterion and [e] SBC =Schwartz Bayes Criterion

	French Franc				German Mark			
	NLMACH	ARCH	GARCH	ARCH-t	NLMACH	ARCH	GARCH	ARCH-t
				Estimated Coefficients				
C	-0.0043	-0.0039	-0.0057	-0.0062	-0.0016	-0.0016	-0.0039	-0.0058
	$(0.0081)^a$	(0.0079)	(0.0070)	(0.0070)	(0.0083)	(0.0081)	(0.0078)	(0.0079)
δ_0	0.0745*	0.0745*	0.0024*	0.0449*	0.0796*	0.0796*	0.0033*	0.0507*
	(0.0025)	(0.0051)	(0.0017)	(0.0038)	(0.0028)	(0.0029)	(0.0018)	(0.0017)
δ_1	0.0153*	0.1749*	0.0596*	0.1405*	0.0217*	0.2180*	0.0711*	0.01618*
	(0.0032)	(0.0459)	(0.0189)	(0.0269)	(0.0035)	(0.0319)	(0.0200)	(0.0253)
δ_2	–	–	0.9135*	–	–	–	0.8969*	–
	–	–	(0.0348)	–	–	–	(0.0337)	–
V^b	–	–	–	5.1436*	–	–	–	5.5298*
	–	–	–	(0.6332)	–	–	–	(0.4059)
				Decision Criteria				
$L(\theta)^c$	931.73	931.12	962.74	735.18	863.72	863.23	888.54	808.74
AIC^d	8846.30	8845.45	8890.63	8541.95	8748.29	8747.56	8786.93	8665.25
SBC^e	8861.79	8860.94	8891.29	8562.61	8763.79	8763.05	8807.59	8685.91

Table 4: Model Adjustment for the French Franc and the German Mark. *,** Significant at the 1% and 10% level respectively.[a] Robust Standard errors in parenthesis.[b] Shape parameter. [c]Optimizaed likelihood value.[d] AIC =Akaike information Criterion and [e] SBC =Schwartz Bayes Criterion

	Japanese Yen				Canadian Dollar			
	NLMACH	ARCH	GARCH	ARCH-t	NLMACH	ARCH	GARCH	ARCH-t
				Estimated Coefficients				
C	-0.0065	-0.0062	-0.0098	-0.0046	-0.0059	0.0059	0.0025	0.0029
	$(0.0073)^a$	(0.0076)	(0.0080)	(0.0065)	(0.0032)	(0.0031)	(0.0032)	(0.0029)
δ_0	0.0702*	0.0693*	0.0020*	0.0347*	0.0123*	0.0119*	0.001*	0.0067*
	(0.0022)	(0.0018)	(0.0014)	(0.0020)	(0.0004)	(0.0004)	(0.0001)	(0.0004)
δ_1	0.0089*	0.1286*	0.0484*	0.0751*	0.0024*	0.1992*	0.0517*	0.1178*
	(0.0021)	(0.0212)	(0.0164)	(0.0219)	(0.0004)	(0.0252)	(0.0152)	(0.0281)
δ_2	–	–	0.9251*	–	–	–	0.9404*	–
	–	–	(0.0279)	–	–	–	(0.0185)	–
V^b	–	–	–	3.8234*	–	–	–	4.3438*
	–	–	–	(0.3451)	–	–	–	(0.4418)
				Decision Criteria				
$L(\theta)^c$	1005.55	1006.56	1045.43	616.82	2100.63	2102.25	2139.88	451.57
AIC^d	8944.88	8946.19	8997.17	8319.99	8897.44	9898.43	9923.38	7912.92
SBC^e	8960.38	8961.68	9017.83	8335.65	9912.94	9913.93	9944.04	7933.58

Table 5: Model Adjustment for the Japanese Yen and the Canadian Dollar. *,** Significant at the 1% and 10% level respectively.[a] Robust Standard errors in parenthesis.[b] Shape parameter. [c]Optimizaed likelihood value.[d] AIC =Akaike information Criterion and [e] SBC =Schwartz Bayes Criterion

	British Pound				Australian Dollar			
	NLMACH	ARCH	GARCH	ARCH-t	NLMACH	ARCH	GARCH	ARCH-t
				Estimated Coefficients				
C	0.0026	0.0020	-0.0015	-0.0041	-0.0003	-0.0001	-0.0011	-0.0061
	$(0.0078)^a$	(0.0076)	(0.0069)	(0.0066)	(0.0050)	(0.0056)	(0.0055)	(0.0050)
δ_0	0.0701*	0.0672*	0.0008*	0.0329*	0.0367*	0.0367*	0.0025*	0.0206*
	(0.0022)	(0.0019)	(0.0005)	(0.0020)	(0.0012)	(0.0011)	(0.0023)	(0.0011)
δ_1	0.0144*	0.2138*	0.0507*	0.1490*	0.0032*	0.0813*	0.0595*	0.0723*
	(0.0025)	(0.0294)	(0.0143)	(0.0321)	(0.0014)	(0.0228)	(0.0355)	(0.0239)
δ_2	–	–	0.9403*	–	–	–	0.8794*	–
	–	–	(0.0187)	–	–	–	(0.0903)	–
V^b	–	–	–	3.8423*	–	–	–	4.4480*
	–	–	–	(0.3606)	–	–	–	(0.4853)
				Decision Criteria				
$L(\theta)^c$	975.20	979.79	1052.46	655.30	1441.20	1441.02	1457.70	219.87
AIC^d	8905.26	8911.33	9005.84	8393.00	9410.28	9410.13	9427.01	6981.21
SBC^e	8920.76	8926.83	9026.50	8413.89	9425.78	9425.62	9447.67	7001.87

Table 6: Model Adjustment for the British Pound and the Australian Dollar. *,** Significant at the 1% and 10% level respectively.[a] Robust Standard errors in parenthesis.[b] Shape parameter. [c]Optimizaed likelihood value.[d] AIC =Akaike information Criterion and [e] SBC =Schwartz Bayes Criterion

If we compare the NLMACH(1) against ARCH(1) or GARCH(1,1), the AIC and SBC criteria provide mixed evidence. For instance according to these criteria $NLMACH(1)$ is preferred to the $ARCH(1)$ for the Swiss Franc, the Japanese Yen, the Canadian Dollar and the British Pound. When the NLMACH(1) is compared to GARCH(1,1) it is observed that in all cases the NLMACH(1) is preferred to GARCH(1,1). Surprisingly, according to these criteria, the ARCH(1) model is also preferred to GARCH(1,1). However, we need to be cautious about using these criteria to discriminate between models. The use of these statistics might not be entirely appropriate since the two types of processes have a distinct nonlinear nature. Moreover, the statistical properties of AIC and SBC have not been investigated for the class of nonlinear models here proposed. Using the Optimized Likelihood value as the selection criterion, the GARCH(1,1) is the model that fits the data best. This however is not necessarily bad news for the NLMACH(1) since it only indicates that GARCH(1,1) captures well a specific type of conditional heteroskedasticity. One last case has been investigated: the $ARCH(1)$ with a t-distribution in order to capture the fat tails and non-normality of the data[8]. It turns out that, as indicated by the AIC and SBC, this model is strongly preferred to all other specifications including the $NLMACH(1)$ with the only exception of the Swiss Franc. As we have already shown, our $NLMACH(1)$ model reproduces the fat tails quite naturally without the need of imposing a different conditional distribution to replicate this behavior. However, as suggested by these results, imposing a conditional distribution different than a normal might capture other properties of the data. For instance, it might be that the source of non-normality is due to the existence of outliers; this feature is not obviously taken into account by the $NLMACH(1)$ model.

[8]Notice that, as required, the degrees of freedom parameter is significant and greater than four in all currencies except the Yen and the British Pound.

4 Conclusions

This paper has presented a new model, deeply inspired by the Non-Linear Moving Average models, but with the approach typically used when dealing with conditionally heteroskedastic models. A very simple specification modification solves the typical problems of this class. $NLMACH$ has simple statistical properties and is easy to estimate. It should indeed be seen as a new instrument to deal with heteroskedasticity. Several tools presented here aim to fulfill this purpose. On one hand, $NLMACH$ can be easily estimated by ML. This estimation technique proved to be efficient and reliable. On the other hand, the theoretical results, such as the autocorrelation function form of the squared process should facilitate identification, and provide statistical evidence of either the presence or the absence of $NLMACH$ behavior. for some particular cases (specified in the DGPs and the sample size of the Monte Carlo experiments) $NLMACH(1)$' forecasting capabilities outperform [9] the ones yielded by $GARCH(1,1)$ even when the true DGP is a $GARCH(1,1)$.

This new specification will have to compete with the many variants belonging to the $ARCH$ class. Such competitors vary in complexity and robustness. $NLMACH$ is the replication of fat tails; the estimation results indicate however that this process is preferred to ARCH models using a student-t as conditional distribution only in one case–the Swiss Franc. The $NLMACH$ model, despite its simplicity, still offers extremely interesting characteristics. All in all, the relative evidence in favor of NLMACH varies in complexity and robustness and all we hope is to increase empirical interest for Non-Linear Moving Average models, which have been virtually neglected along the past decades.

[9] comparison made using the $RMSE$ criterion

References

BOLLERSLEV, T. (1986): "Generalized Autoregressive Conditional Heteroskedasticity," *Journal of Econometrics*, 31, 309–28.

BOLLERSLEV, T., R. CHOU, AND K. KRONER (1992): "ARCH modeling in Finance: a Review of Theory and Empirical Evidence," *Journal of Econometrics*, 52, 5–59.

BOULIER, J. (1994): *Modélisation ARCH: Théories Statistiques et Applications dans le Domaine de la Finance*. Ellipses.

ENGLE, R. (1982a): "Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the U.K. Inflation," *Econometrica*, 50, 987–1008.

ENGLE, R., D. LILIEN, AND R. ROBINS (1987): "Estimating Time Varying Risk Premia in the Term Structure: the *ARCH-M* Model," *Econometrica*, 55, 391–407.

FRANSES, P., AND D. VAN DIJK (2000): *Non-Linear Time Series Models in Empirical Finance*. Cambridge.

GRANGER, C. W. (1998): "Overview of Nonlinear Time Series Specifications in Economics," UCSD, first draft.

GRANGER, C. W., AND A. ANDERSEN (1978): *Applied Time Series Analysis*. Academic Press, pp.25-38 Edited by David F. Findley.

GRANGER, C. W., AND T. TERASVIRTA (1993): *Modelling Nonlinear Economic Relationships*. Oxford University Press.

GUÉGAN, D. (1994): *Séries Chronologiques Non Linéaires à Temps Discret*. Economica.

NELSON, D. (1991): "Conditional Heteroskedasticity in Asset Returns: A New Approach," *Econometrica*, 59, 347–70.

ROBINSON, A. (1977): "The Estimation of a Non Linear Moving Average Model," *Stochastic Processes and their Applications*, 1, 81–90.

TONG, H. (1990): *Non-Linear Time Series: A Dynamical Approach*. Oxford Science Publications.