A proposal of a new specification for a conditionally heteroskedastic variance model: The Quadratic Moving-Average Conditional Heteroskedasticity and an application to the D. Mark-US Dollar Exchange Rate.

Daniel Ventosa Santaularia*

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 [i.e. Bollerslev's (1986) $GARCH$; Nelson's (1990) $EGARCH$]. This recent domain has achieved very successful developments. Nevertheless, several empirical studies seem to show that the performance of such models is not always appropriate [Boulier(1992)].

In this paper we propose a new specification: the Quadratic Moving Average Conditional heteroskedasticity model. Its statistical properties, such as the kurtosis and the symmetry, as well as two estimators (Method of Moments and Maximum Likelihood) are studied. Two statistical tests are presented, the first one tests for homoskedasticity and the second one, discriminates between $ARCH$ and $QMACH$ specification. A Monte Carlo study is presented in order to illustrate some of the theoretical results. An empirical study is undertaken for the $DM - US\$$ exchange rate.

Keywords: Conditionally heteroskedastic models, Quadratic Moving Average Conditionally heteroskedasticity model, Homoskedasticity tests, Volatility, Truncated Volterra developments.

JEL classification: C22, C12, C13.

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

*I would like to thank Dr. Russell Davidson for his comments and encouragement. I am also grateful to the GREQAM laboratory, and to the Universitat Autònoma de Barcelona for their support.

$ARCH$ type models appeared, some of them possessing statistical properties extremely appealing to financial econometrics. Among them, Bollerslev's (1986) $GARCH$ model, Engle Lilien and Robins(1987) $ARCH-M$ and Nelson's(1990) $EGARCH$ have succeeded in generalizing $ARCH$ models, incorporating the volatility of a variable in its value equation and taking into account asymmetric effects respectively. Other $ARCH$-class models, more recent, seem very promising, such as $QGARCH$ [Sentana(1995)], the $GJR-GARCH$ [Glosten, Jagannathan and Runkle(1993)], and the $LSTGARCH$ [Hagerud(1997) and Gonzalez-Rivera(1998)] but are still to recent to be sure.

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. The infamous $GARCH(1,1)$ model remains one of the best options for practitioners of financial econometrics. Yet, some studies, such as the one carried out by Boulier(1994), indicate that this class of models doesn't always perform well when dealing with forecasts.

When dealing with conditionally heteroskedastic models, the accent has always been put in Autoregressive specifications, neglecting the potential usefulness of 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 development. He also gave some of the statistical properties of such model as well a 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. $NLMA$ models are nowadays seen as being ineffectual for empirical purposes [e.g. Tong(1990), Guégan(1994) and Granger(1998)]. Among its defects, non-invertibility and difficulties in estimation are widely known.

By using the same source of inspiration (Volterra developments) but developing it within an $ARCH$ framework, we define a different specification; the Quadratic Moving Average Conditionally heteroskedastic model, $QMACH$. This specification can reproduce several of the typical characteristics of financial variables, such as leptokurticity, asymmetric effects of shocks, and, of course, heteroskedasticity. It is not necessary to impose conditions on the parameters to ensure the existence of all moments.

The $QMACH$ can be easily estimated. We present two different estimators; a Method of Moments estimator and a Maximum Likelihood estimator, the latter being the better one. We also propose two statistical tests. The first one is based on the classic Likelihood Ratio test and discriminates between homoskedasticity and $QMACH$-type heteroskedasticity. The second one is also a LR test with an artificially-nested null hypothesis. It distinguishes between $ARCH$-type heteroskedasticity and $QMACH$-type heteroskedasticity. Using Monte Carlo simulations, we present evidence that both the estimators and the specification tests perform well.

Finally, we present an empirical study of the $DM - US\$$ exchange rate. We adjust two models, the $QMACH(1)$ model, and also the $GARCH(1,1)$. Several statistics, aimed at evaluating intra-sample forecasts as well as the adjustment, are calculated in order to compare the models.

This paper is divided in five sections. The second introduces the $QMACH(1)$ model and some of its statistical properties. Section three deals with the estimation problem; two statistical tests are also proposed to identify heteroskedasticity. The fourth section presents the empirical study. Conclusions appear in section five.

2 The QMACH model

2.1 The QMACH specification

As it has been already said, the $QMACH$ model is inspired by the Volterra expansion which can be expressed as:

$$\begin{aligned} X_t = \mu + \sum_{i=-\infty}^{+\infty} \theta_i \epsilon_{t-i} + \sum_i \sum_j \theta_{ij} \epsilon_{t-i} \epsilon_{t-j} \\ + \sum_i \sum_j \sum_k \theta_{ijk} \epsilon_{t-i} \epsilon_{t-j} \epsilon_{t-k} + \cdots \end{aligned} \quad (1)$$

Such an expansion can be truncated in order to make it feasible. Robinson's (1977) model, $NLMA$ and, indeed, the $QMACH$ can be encompassed by the following equation, proposed by Guégan(1994):

$$X_t = \mu + \sum_{i=0}^{p} k_i \epsilon_{t-i} + \sum_{i=0}^{p} \sum_{j=0}^{q} k_{ij} \epsilon_{t-i} \epsilon_{t-j} + \cdots \quad (2)$$

These model have several properties, among which non-invertibility [Granger(1998)] and high non-linearity [Guégan(1994)] stand out. We shall concentrate on a particular version of (2) which is different from the one developed by Robinson(1977) but still possesses some very appealing characteristics; the $QMACH(1)$:

$$\begin{aligned} X_t &= V_t h_t^{\frac{1}{2}} \\ h_t &= (\delta_0 + \delta_1 V_{t-1})^2 \\ &\Leftrightarrow \delta_0 V_t + \delta_1 V_t V_{t-1} \\ V_t &\sim {}_{iid}\mathcal{N}(0,1) \end{aligned} \quad (3)$$

As can be inferred from (3), the $QMACH(1)$ is a particular case of (2). Yet, it is presented in the usual $ARCH$ style, so that its conditional heteroskedasticity stands out. Figure (1) shows a simulation of this process.

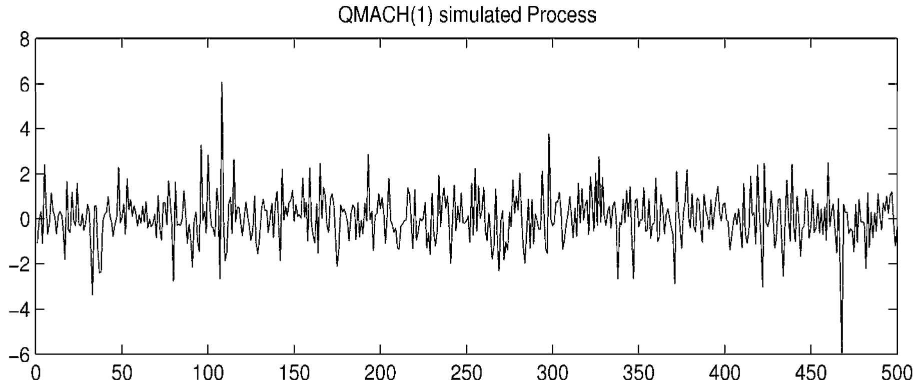

Figure 1: $h_t = 1.00 + 0.34V_{t-1}$ $\quad n = 500$

2.2 Distribution of the first-order QMACH process

The $QMACH(1)$ has the advantage of being a very simple specification. Most of its properties can be inferred easily. In order to make a brief comparison with the $ARCH(1)$, we present the first two -unconditional and conditional-moments of the process:

$$\begin{aligned} E(X_t) &= 0 \\ E(X_t X_{t-j}) &= \begin{cases} \delta_0^2 + \delta_1^2 & for\, j = 0 \\ 0 & otherwise \end{cases} \\ E_{t-1}(X_t) &= 0 \\ E_{t-1}(X_t^2) &= \delta_0^2 + 2\delta_0\delta_1 V_{t-1} + \delta_1^2 V_{t-1}^2 \end{aligned} \tag{4}$$

It can be seen that, contrary to most of the specifications of conditionally heteroskedastic models, there are not conditions for the existence of the second moment. The previous results let us conclude that the $QMACH(1)$ is weackly stationary. We have calculated the autocorrelation function of the process, which is: $g_x(z) = \delta_0^2 + \delta_1^2$. The invertibility problems appears now clearly, since there are four distinct $QMACH(1)$ process that satisfy this function having each one the following parameters: δ_0, δ_1; $\delta_0, -\delta_1$; $-\delta_0, \delta_1$ and ;$-\delta_0, -\delta_1$. Such problem could be seen as a major one when estimating the process, but, it will be seen in the next section that it is not.

We have calculated also the autocorrelation function of the squared process, which is:

$$\rho_i = \begin{cases} \frac{\delta_0^2(\delta_0^2 + \delta_1^2)}{\delta_0^4 + 8\delta_0^2\delta_1^2 + 4\delta_1^4} & for\, i = 1 \\ 0 & \forall\, i > 1 \end{cases} \tag{5}$$

The form of the autocorrelation function seems to differ greatly from the one proposed by the stylized facts in finance theory, where the autocorrelations decay slowly and not abruptly as in the $QMACH(1)$ process. This limitation can be tackled by different means. The first one could be the use of a $QMACH(q)$ where $q > 1$. This case is studied below. The second alternative is to generalize the process by including lags of h_t in the conditional variance specification. By doing this we could eventually insert an exponential decay of the autocorrelations. this idea is currently being developped succesfully in a not yet published working paper.

Contrary to $ARCH$ processses, the determination of the conditions that render the $QMACH$ process stationary does not need a recursive argument. The symmetry of the normal distribution makes all the odd moments zero. In the case of even moments, there are no conditions for their existence:

theorem 1 *For integer r, the 2rth moment of a first-order $QMACH$ process exists always and it's equal to:*

$$E(X_t^{2r}) = \prod_{j=1}^{2r}(2j-1) \cdot \left[\sum_{i=0}^{r} \binom{2r}{2i} \delta_0^{2r-2i} \delta_1^{2i} \prod_{k=1}^{i}(2k-1)\right] \tag{6}$$

proof.
The even moments of the $QMACH(1)$ process can be defined as follows:

$$\begin{aligned} E(X_t^{2r}) &= E(V_t^{2r}) \cdot E(h_t^r) \\ &= \prod_{j=1}^{2r}(2j-1) \cdot E\left[(\delta_0 + \delta_1 V_{t-1})^{2r}\right] \end{aligned}$$

By using the Newton's formulae, we can develop the most-right expectation,

$$\begin{aligned} E\left[(\delta_0 + \delta_1 V_{t-1})^{2r}\right] &= E\left[\binom{2r}{0}\delta_0^{2r} + \binom{2r}{1}\delta_0^{2r-1}\delta_1 V_{t-1} + \ldots \right. \\ &\quad \ldots + \binom{2r}{c}\delta_0^{2r-c}(\delta_1 V_{t-1})^c + \ldots \\ &\quad \left. + \binom{2r}{2r-1}\delta_0(\delta_1 V_{t-1})^{2r-1} + \ldots + \binom{2r}{2r}(\delta_1 V_{t-1})^{2r}\right] \\ &= E\left[\sum_{i=0}^{2r}\binom{2r}{i}\delta_0^{2r-i}(\delta_1 V_{t-1})^i\right] \end{aligned}$$

Recognising that the expectations of odd powers are zero, we simplify the expression by omitting such cases:

$$\begin{aligned} E\left[(\delta_0 + \delta_1 V_{t-1})^{2r}\right] &= \sum_{i=0}^{r}\binom{2r}{2i}\delta_0^{2r-2i}\delta_1 E(V_{t-1})^{2i}) \\ &= \sum_{i=0}^{r}\binom{2r}{2i}\delta_0^{2r-2i}\delta_1^{2i}\prod_{k=1}^{i}(2k-1) \end{aligned}$$

Which is the expression given in the statement of **Theorem 1**.

Q.E.D.

The unconditional variance of the $QMACH(1)$ model can be easily generalized for the $QMACH(q)$:

$$\begin{aligned} X_t &= V_t h_t^{\frac{1}{2}} \\ h_t &= \left(\delta_0 + \sum_{i=1}^{q} \delta_i V_{t-i}\right)^2 \\ V_t &\sim {}_{iid}\mathcal{N}(0,1) \end{aligned} \tag{7}$$

The unconditional variance is:

$$E(X_t^2) = \sum_{i=0}^{q} \delta_i^2 \tag{8}$$

In particular, we have also calculated the autocorrelation function of a squared $QMACH(2)$ $[X_t = V_t\,(\delta_0 + \delta_1 V_{t-1} + \delta_2 V_{t-2})]$ process:

$$\rho_i = \begin{cases} \gamma_1 & for\, i = 1 \\ \gamma_2 & for\, i = 2 \\ 0 & \forall\, i > 2 \end{cases} \tag{9}$$

Where,

$$\begin{aligned} \gamma_1 &= \frac{2\delta_1^4 + 2\delta_0^2\delta_1^2 + 4\delta_1^2\delta_2^2 + 4\delta_0^2\delta_1\delta_2}{3\left[\delta_0^4 + 3(\delta_1^4 + \delta_2^4) + 6(\delta_0^2\delta_1^2 + \delta_0^2\delta_2^2 + \delta_1^2\delta_1^2) - (\delta_0^2 + \delta_1^2)^2\right]} \\ \gamma_2 &= \frac{2\delta_2^4 + 2\delta_0^2\delta_2^2 + 2\delta_1^2\delta_2^2}{3\left[\delta_0^4 + 3(\delta_1^4 + \delta_2^4) + 6(\delta_0^2\delta_1^2 + \delta_0^2\delta_2^2 + \delta_1^2\delta_1^2) - (\delta_0^2 + \delta_1^2)^2\right]} \end{aligned}$$

It can be seen that the order of the $QMACH(q)$ process, q, can be inferred by means of its squared autocorrelation function in exactly the same way a Moving-Average model is identified with its autocorrelation function. A theorem, as well as a proof of the latter statement can be constructed as follows:

theorem 2 *Let X_t be a $QMACH(q)$ process such as the one indicated in (7). For any integer $k > q$, the kth autocovariance of the squarred procces is equal to zero:*

$$E\left[\left(X_t^2 - E(X_t^2)\right)\left(X_{t-k}^2 - E(X_{t-k}^2)\right)\right] = 0 \quad \forall\, k > q \tag{10}$$

proof.
Knowing that:

$$E\left[\left(X_t^2 - E(X_t^2)\right)\left(X_{t-k}^2 - E(X_{t-k}^2)\right)\right] = E\left(X_t^2 X_{t-k}^2\right) - \left[E\left(X_t^2\right)\right]^2 \quad (11)$$

The first element in the right term of equation (11) can be developed, when $k > q$, as follows:

$$\begin{aligned}
E\left(X_t^2 X_{t-k}^2\right) &= E\left(V_t^2 V_{t-k}^2\right) \cdot E\left(h_t h_{t-k}\right) \\
&= E\left(h_t h_{t-k}\right) \\
&= E\left[\left(\delta_0 + \sum_{i=1}^{q} \delta_i V_{t-i}\right)^2 \cdot \left(\delta_0 + \sum_{i=1}^{q} \delta_i V_{t-k-i}\right)^2\right] \\
&= E\Bigg[\delta_0^4 + 2\delta_0^3 \left(\sum_{i=1}^{q} \delta_i V_{t-k-i}\right) + \delta_0^2 \left(\sum_{i=1}^{q} \delta_i V_{t-k-i}\right)^2 + \\
&\quad 2\delta_0^3 \left(\sum_{i=1}^{q} \delta_i V_{t-i}\right) + 4\delta_0^2 \left(\sum_{i=1}^{q} \delta_i V_{t-k-i}\right)\left(\sum_{i=1}^{q} \delta_i V_{t-i}\right) + \\
&\quad 2\delta_0 \left(\sum_{i=1}^{q} \delta_i V_{t-k-i}\right)^2 \left(\sum_{i=1}^{q} \delta_i V_{t-k-i}\right) + \delta_0^2 \left(\sum_{i=1}^{q} \delta_i V_{t-i}\right)^2 + \\
&\quad 2\delta_0 \left(\sum_{i=1}^{q} \delta_i V_{t-k-i}\right)\left(\sum_{i=1}^{q} \delta_i V_{t-k-i}\right)^2 + \\
&\quad \left(\sum_{i=1}^{q} \delta_i V_{t-k-i}\right)^2 \left(\sum_{i=1}^{q} \delta_i V_{t-i}\right)^2\Bigg]
\end{aligned}$$

This expression can be simplified because of the independence between V_{t-i} and V_{t-k-i}.

$$E\left(X_t^2 X_{t-k}^2\right) = \left(\sum_{i=1}^{q} \delta_i^2\right) \cdot \left(\sum_{i=1}^{q} \delta_i^2\right) \quad (12)$$

The second element in the right term of equation (11) can also be developed, by using the equation (8):

$$\left[E\left(X_t^2\right)\right]^2 = \left(\sum_{i=1}^{q} \delta_i^2\right) \cdot \left(\sum_{i=1}^{q} \delta_i^2\right)$$

Since both terms are identical, equation(11) is equal to zero.

Q.E.D.

2.3 Relevant statistical properties for financial applications

Several characteristics shared by many financial variables have been identified. In this section we deal with two of the most prominent ones; leptokurticity and asymmetry. We show that the $QMACH(1)$ can encompass such characteristics. To study the first one, we calculate the theoretical value of the kurtosis coefficient (K) of a $QMACH(1)$ distribution and we compare it to the one yielded by a normal distribution. For asymmetry, defined as the difference in effects of positive and negative shocks on conditional volatility,we use the News Impact Curve (NIC), introduced by Pagan and Schwert(1990). The NIC measures how new information is incorporated into volatility, e.g. the relationship between V_t and h_{t+1} [see Franses and van Dijk(2000)].

The fourth Central moment of the $QMACH(1)$ is:

$$E\left(X_t^4\right) = 3 \cdot \left[\delta_0^4 + 6\delta_0^2\delta_1^2 + 3\delta_1^2\right] \quad (13)$$

Which yields a kurtosis of,

$$\mathcal{K} = \frac{3 \cdot \left[\delta_0^4 + 6\delta_0^2\delta_1^2 + 3\delta_1^2\right]}{\left(\delta_0^2 + \delta_1^2\right)^2} \quad (14)$$

$\mathcal{K}$ is always greater than 3 and so the $QMACH(1)$ distribution is leptokurtik.
proof.

$$\begin{aligned} \frac{3\cdot\left[\delta_0^4+6\delta_0^2\delta_1^2+3\delta_1^2\right]}{\left(\delta_0^2+\delta_1^2\right)^2} &> 3 \\ \Leftrightarrow \quad 4\delta_0^2\delta_1^2 + 2\delta_1^4 &> 0 \end{aligned}$$

The latter expression is always true unless δ_1^2, is equal to zero.

The $QMACH(1)$ model possesses a particular NIC curve. Contrary to that of an $ARCH$-type model, the $QMACH(1)$'s NIC is not centered at zero. It's this characteristic that produces an asymmetry effect. $GQARCH$ models, proposed by Sentana(1995) has the same property. Restating expression (3), we obtain:

$$h_t = \delta_0^2 + \left(\frac{2\delta_0\delta_1}{V_{t-1}} + \delta_1^2\right) V_{t-1}^2 \quad (15)$$

this expression shows that the impact of V_{t-1}^2 on h_t is equal to $\left(\frac{2\delta_0\delta_1}{V_{t-1}} + \delta_1^2\right)$. If $\delta_1\delta_0 < 0$, the effect of negative shocks on h_t will be larger than the effect of a positive shock of the same size. Additionally, the effect depends on the sign of the shock. By taking the partial derivative of h_t on V_{t-1} and setting it equal to zero,

$$
\begin{aligned}
\frac{\partial h_t}{\partial V_{t-1}} &= 2\delta_1 \left(\delta_0 + \delta_1 V_{t-1}\right) \\
&= 0 \qquad (16) \\
\Rightarrow \quad V_{t-1}^{*} &= -\frac{\delta_0}{\delta_1}
\end{aligned}
$$

Which is the point around which the effect of shocks on the conditional variance is symmetric. We can now build the NIC curve for the $QMACH(1)$ model:

$$NIC\left(X_t\right) = \delta_0^2 + 2\delta_0\delta_1 V_{t-1} + \delta_1^2 V_{t-1}^2 \qquad (17)$$

Figure(2) shows a tipical example of a $QMACH(1)$ NIC curve.
In the last ten years, a number of $ARCH$-class models have been proposed to

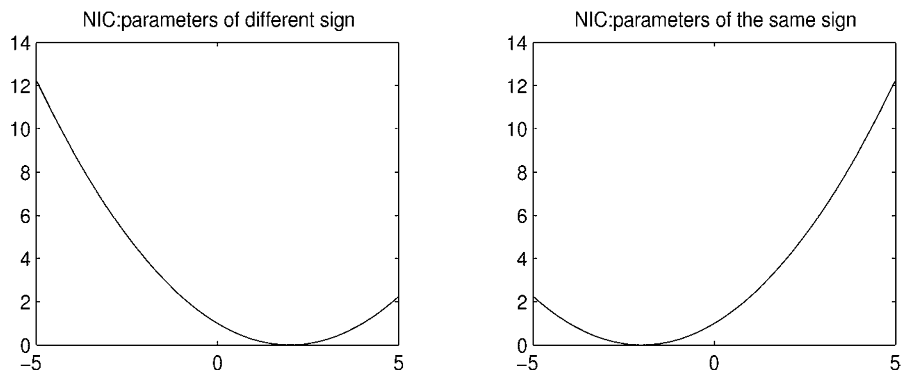

Figure 2: The center of symmetry in the QMACH(1) NIC curve is $-\frac{\delta_0}{\delta_1}$, and depends of the sign of the parameters.

deal with asymmetric effects on variance. Among them, Engle's $AGARCH$ [see Engle an Ng(1993)], $GJRGARCH$, proposed by Glosten, Jagannathan and Runkle(1993), the $LSTGARCH$ presented by Hagerud(1997) and Gonzalez-Rivera(1998) and the $QGARCH$ [see Sentana(1995)]stand out. Several variants of the $LSTGARCH$ have appeared, but the latter seems to be the most successful specification. It has to be said that, the $AGARCH$ and the $QGARCH$ have roughly the same structure of asymmetry than the $QMACH$; the $LSTGARCH$, on the other hand, possesses a more sophisticated asymmetric mechanism, consisting basically in a threshold model that allows the parameters shift depending on the sign of the shock. Although it can't be denied this strategy offers a more versatile asymmetric structure than $QMACH$ does, we should have in mind the fact that the latter could easily incorporate the same kind of asymmetry by transforming it the same way, $LSTGARCH$ transforms $GARCH$. Several advantages could be found by doing so, among which, the absence of stationarity and positiveness conditions are important. In the figure (3), we present the NIC curve of the $GARCH$ as well as the one yielded by an $LSTGARCH$.

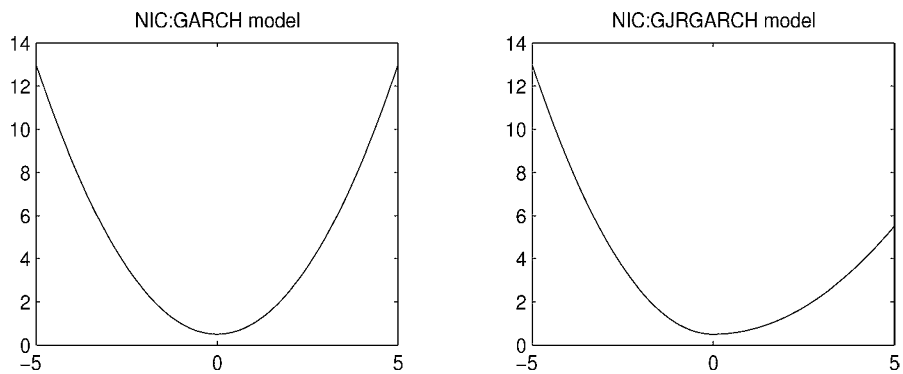

Figure 3: NIC curves for the GARCH model and for the GJRGARCH model: It can be noted the lack of asymmetry in the first one.

3 Estimation of the first-order $QMACH$ and statistical inference

3.1 Estimation of the $QMACH(1)$

Once the main statistical properties established, next step is estimation. The $QMACH(1)$ estimation is simple despite the fact of being a highly non-linear model. Its non-invertibility property doesn't interfere significantly and can be practically neglected. In order to show this, we present first a very simple estimator, based in the Method of Moments (MOM). With this method, we can estimate consistent estimators of the two parameters but the estimates are inefficient relative to the Maximum Likelihood (ML) estimates, our second estimating technique.

For the MOM, we present the two empirical moments used to compute the estimated parameters. This two-equation system is rather too complicated to be solved analytically, so we let an ordinary gradient algorithm solve it for us.

The theoretical moments used to match the empirical ones are the second and the fourth central moment, appearing in equations (4) and (13). The MOM procedure allows for consistent parameters estimates. The GMM technique should be able to improve these results, but we rather prefer the ML estimator. In the case of the MOM estimator, it has to be said that, although numerical precision of the estimated parameters is good, their sign can be wrong (the problem of invertibility). If all the parameters are of the opposite sign there is no problem since the volatility equation is squared and provides exactly the same result. This is why, when studying the precision and accuracy of parameter estimates by means of a Monte Carlo experiment, we take the absolute value, in order to not bias the results [1].

The ML technique works in the same way as with $ARCH$ models. Letting Ψ_t be the information set available at time t, we can use conditional densities:

$$X_t|\psi \quad \sim \quad \mathcal{N}(0, h_t)$$

This property can be used to define the loglikelihood function:

$$\begin{aligned} l &= \frac{1}{T}\sum_{t=1}^{T} l_t \\ l_t &= -\frac{1}{2}\log h_t - \frac{1}{2}\frac{X_t^2}{h_t} \qquad (18) \\ &= -\frac{1}{2}\log h_t - \frac{1}{2}V_t^2 \end{aligned}$$

The ML estimators can be obtained by maximizing this function. We use a gradient algorithm to do so. Initial values for the procedure can be the MOM estimates. The underlying theory of the (pseudo) ML estimator allows us to infer the asymptotic variance of the parameters. Under some regularity conditions [Gallant(1987), Gourieroux-Monfort(1989), Gourieroux-Monfort-Trognon(1984) and White(1981)], this estimator is consistent, and asymptotically Normal. Its asymptotic Covariance matrix is:

$$\begin{aligned} V_a s\left[\sqrt{T}\left(\hat{\theta} - \theta\right)\right] &= \mathcal{I}_{op}^{-1} \\ \mathcal{I}_{op} &= \underset{0}{E}\left[\frac{\partial \log l_t}{\partial \theta}\frac{\partial \log l_t}{\partial \theta'}\right] \end{aligned}$$

Where $\theta = \begin{bmatrix} \delta_0 \\ \delta_1 \end{bmatrix}$ and $\underset{0}{E}$ is the expectance under the true distribution.
In the case of the $QMACH(1)$, the first derivative of the log likelihood of the tth observation is:

$$\frac{\partial \log l_t}{\partial \theta} = \frac{1}{h_t^{1/2}}\begin{bmatrix} 1 \\ V_{t-1} \end{bmatrix}$$

And the outer-product information matrix estimate is:

$$\frac{\partial \log l_t}{\partial \theta}\frac{\partial \log l_t}{\partial \theta'} = \frac{1}{h_t}\begin{bmatrix} 1 & V_{t-1} \\ V_{t-1} & V_{t-1}^2 \end{bmatrix}$$

[1] Another solution, that offered excellent results can be sumarized as follows: once the the MOM estimation done, create three combinations of the obtained parameters, by changing the signs of either one or both parameters. Evaluate its likelihood and choose the one maximising it.

The information matrix, which is the expectation of the *outer-product* matrix, can be replaced by its empirical counterpart:

$$\widehat{\mathcal{I}_{op}} = \frac{1}{T}\sum_{t=1}^{T}\frac{1}{h_t}\begin{bmatrix} 1 & V_{t-1} \\ V_{t-1} & V_{t-1}^2 \end{bmatrix} \tag{19}$$

Expression (19) allows us to perform inference on the estimated parameters, by t statistics, for instance.

We compare the two different estimation methods of the parameters. The comparison is performed through a Monte Carlo study with the real parameters being $\delta_0 = 0.8$ and $\delta_0 = 0.34$; there are $10,000$ replications. The number of observations is $T = 200$. Table (1) gives the empirical mean and the standard deviation for each of these estimators. For both, the finite sample bias is small but it has to be said that maximum likelihood performs much better. efficiency gain, on the other side is rather small.

Estimator	δ_0 mean	δ_0 st. Deviation	δ_1 mean	δ_1 st. Deviation
MOM	0.7730	0.1592	0.3209	0.1605
ML	0.8060	0.1705	0.3428	0.1505

Table 1: Monte Carlo Simulation of estimates perfomance

3.2 Statistical inference: specification tests

The $QMACH(1)$ specification's statistical properties and estimation performance can be considered as evidence of the viability of the model. Nevertheless, the huge availability of alternative specifications for the volatility makes it necessary to propose tools to discriminate among different models, in particular, among existing specifications and the $QMACH(1)$ specification. For this purpose, we propose two tests. The first one, constructed as a likelihood rate, allows one to discriminate between a white noise process and a conditionally heteroskedastic process of type $QMACH(1)$. The second test is a non-nested hypothesis test and allows one to distinguish between a $QMACH(1)$ DGP and an $ARCH(1)$ DGP.

3.2.1 The white noise test

The test is constructed as a likelihood ratio (LR). It can be developed as follows. According to the null, in expression (3) the parameter $\delta_1 = 0$, and thus, we have:

$$f(X_t/\theta_0) = \frac{1}{\sqrt{2\pi}\delta_0}\exp\left[\frac{-X_t^2}{2\delta_0^2}\right]$$

$$L_1 = \frac{1}{(2\pi)^{\frac{n}{2}}\delta_0^n}\exp\left[\sum_{t=1}^{T}\frac{-X_t^2}{2\delta_0^2}\right]$$

On the other side, under the alternative hypothesis, $\delta_1 \neq 0$, we have:

$$L_2 = \frac{1}{(2\pi)^{\frac{n}{2}} \prod_{t=1}^{T} h_t^{\frac{1}{2}}} \exp\left[\sum_{t=1}^{T} \frac{-X_t^2}{2h_t}\right]$$

We have only to calculate the likelihood ratio, take the logarithm and multiply by minus two, in order to obtain the statistic test:

$$\lambda_1 = 2T\log(\delta_0) - \sum_{t=1}^{T}\log(h_t) - \sum_{t=1}^{T} X_t\left(\frac{\delta_0^2 - h_t}{\delta_0^2 h_t}\right) \quad (20)$$

$$\lambda_1 \sim \chi^2_{1d.f.} \quad under\,\mathcal{H}_0$$

A Monte Carlo study has been carried out in order to measure the power-level relationship. The latter tell us if the test is not biased (against this particular alternative hypothesis) and, in a certain manner, how good is it. As shown in figure (4), where 10,000 replications of samples including 300 observations have been simulated,the curve "sticks" close to the axes, which tells us that the test is not biased.

Figure(5) shows the same relation power-level, but estimated with bigger

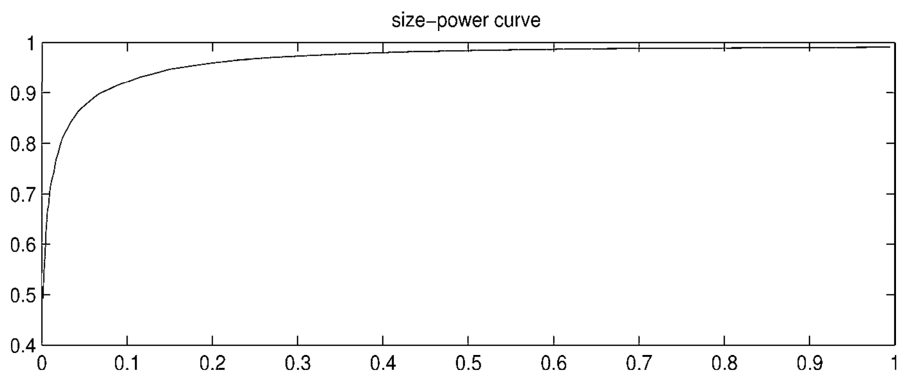

Figure 4: Monte Carlo Simulation: 10,000 replications of $T = 300$ Power is measured in the y-axis and level in the x-axis

samples containing each one 600 simulated observations. The improvement of the power-lever relationship. The test has a 98% power for almost any level.

3.2.2 The QMACH(1)-ARCH(1) test

The aim of the test, as has already been said is to distinguish between a $QMACH(1)$ process and an $ARCH(1)$ process. Following Pollak and Wales (1991), the test is based on the LR test statistic theory. Comparing loglikelihoods values of competing models is an attractive idea. The first step is to

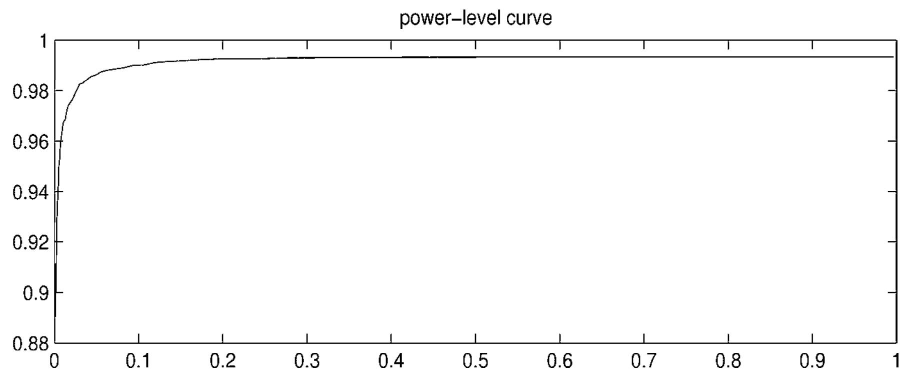

Figure 5: Monte Carlo Simulation: 10,000 replications of $T = 600$ Power is measured in the y-axis and level in the x-axis

construct a general model embedding both specifications, which is done as follows: We specify both the $QMACH(1)$ and the $ARCH(1)$ models as:

- $QMACH(1)$ specification:

$$\begin{aligned} X_t &= V_t\left(\delta_0 + \delta_1 V_{t-1}\right) \\ &= V_t \cdot H_{1t} \end{aligned}$$

- $ARCH(1)$ specification:

$$\begin{aligned} X_t &= V_t\left(\alpha_0 + \alpha_1 X_{t-1}^2\right)^{\frac{1}{2}} \\ &= V_t \cdot H_{2t} \end{aligned}$$

Then, we construct this artificial general model:

$$\begin{aligned} X_t &= V_t H_{1t}^{\phi} \cdot H_{2t}^{1-\phi} \\ &= \begin{cases} V_t H_{1t} & if \quad \phi = 1 \\ V_t H_{2t} & if \quad \phi = 0 \end{cases} \end{aligned} \tag{21}$$

As can be seen, the general model reduces to a $QMACH(1)$ if $\phi = 1$ and to an $ARCH(1)$ if $\phi = 0$. Since we know that the embedding model (21) must fit at least as well as whichever of $ARCH(1)$ and $QMACH(1)$ fits best, the unrestricted maximum of the loglikelihood function must be at least as great as the greater loglikelihood between the specifications. Thus an LR test statistic of the hypothesis yielding the lower loglikelihood (null hypothesis) against the embedding model (alternative hypothesis) must be not less than:

$$\begin{aligned} \lambda_2 &= -2\,(l_l - l_h) \\ &\sim \chi^2_{1d.f.}\; under\, \mathcal{H}_0 \\ &\quad \Rightarrow l_h : greatest\, loglikelihood \\ &\quad \Rightarrow l_l : lowest\, loglikelihood \end{aligned}$$

This feature of the LR test is very convenient. The LR test allows one to put a lower bound on the test statistic without estimating the unrestricted model. This test is consistent with the classical statistical approach. Pollak and Wales (1991) define it in terms of a fictive experiment in wich the two competing hypotheses are nested in a composite. It must be said however that, if one of the specifications performs well, we are allowed to reject the other one, but nothing can said about the former. It might well be rejected too if we tested it against the embedding model. The null hypothesis of the test must be established *a posteriori*, after the estimations of both models. Of course, there is always the possibility of finding the opposite result of what it should be. In order to investigate such possibility, we performed two Monte Carlo experiments. In the first one, we simulated both DPG's, the $QMACH(1)$ and the $ARCH(1)$ (number of replications:10, 000, containing individually 300 observations) and we calculated the test statistic in both cases. Table (2) shows the percentage of "wrong decisions[2]" made by comparing the test statistic to a $\chi^2_{1d.f.}$ (the asymptotic distribution of the test statistic) at a 5% level: In table (3), the very

True DGP	Percentage of correct decisions
QMACH(1)	97.24%
ARCH(1)	81.44%

Table 2: Monte Carlo Simulation of test power. $n = 300$

same experiment but made with samples containing 500 observations, is carried on. The improvement of the test statistic is notorious in both cases:

True DGP	Percentage of correct decisions
QMACH(1)	99.17%
ARCH(1)	95.37%

Table 3: Monte Carlo Simulation of test power. $n = 500$

[2]This means that if, based on the test statistic, we reject the true DGP, the test fails. When the estimated model, made under the asumption of the right DGP, yields a loglikelihood inferior to the one yielded by the other model, we asume the test statistic makes an Type-1 error.

4 Empirical application: the D.Mark-U.S. Dollar exchange rate

The theoretical results presented so far let us think of the $QMACH(1)$ model as a rather attractive one. Nevertheless, its empirical application should be seen as the most important proof of the viability of the specification. The critical aspect is, of course, the comparison of its performance with the most popular available models. When dealing with conditionally heteroskedastic models, the $GARCH(1,1)$ specification is a benchmark.
Empirical exchange rate model are known to fail to adjust adequately during long periods. This temporal inconsistency of the models in the explanation of floating rates may be provoked by a number of structural changes, which suggests dividing the sample into subperiods. For the $DM-US\$$ rate Goldberg and Frydman(2001) found evidence of seven structural breaks during the period $1973-1998$. Most important, they also find evidence that the models used to fit the $DM-US\$$ rate have residuals which are non-normal and heteroskedastic. In this section, we adjust the $QMACH(1)$ model and the $GARCH(1,1)$ model to the $DM-US\$$ rate [3]. The test statistics proposed in section four, as well as the Augmented Dickey-Fuller unit root test (ADF) are performed in order to elucidate the properties of the variable and its DGP. Finally, several statistics aiming to evaluate intra-sample forecasts are presented. Table (4) presents the results of the ADF applied to the $DM-US\$$ rate. As can be seen, there is no evidence (at a 5% level) to reject the presence of a unit root. Since both mod-

Test Equation	no lag. change	1 lag. change	Crit. Values
No constant	-1.6738	-1.4642	-1.9639
Constant	-2.3935	-2.4951	-2.8628
Const. plus trend	-1.3655	-2.0890	-3.4200

Table 4: ADF test

els, $QMACH(1)$ and $GARCH(1,1)$ require stationarity, we transformed the variable using the relation $X_t = 100 \cdot [\log(DM-US\$_t) - \log(DM-US\$_{t-1})]$. The evolution of this variable during the studied period is shown in figure (6).

When we perform the homoskedasticity test proposed in section??, we have no evidence, at a 5% level, to reject the null hypothesis $\left(\lambda_1 = 2.1242; \chi^2_{1.d.f.} = 3.841\right)$. Such inference does not back up Goldberg and Frydman(2001) evidence. This seemingly contradictory result should be attenuated because of methodological differences; the test used here may be committing a type II error. There is more evidence of the existence of the heteroskedasticity problem in this variable than

[3] D. Mark/US$ (1971:01-2001:08): Frecuency: Mensual.Average of daily figures, Noon buying rates in New York City for cable transfers payable in foreign currencies, starting January 1999 derived using the official fixed conversion rates.
Source: Federal Reserve Bank of St. Louis and Federal Reserve Board of Governors

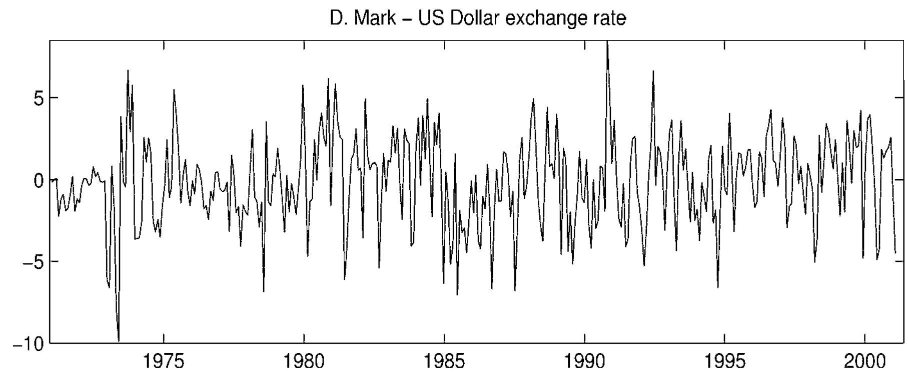

Figure 6: $X_t = 100 \cdot [\log(DM - US\$_t) - \log(DM - US\$_{t-1})]$

the contrary and we thus decide, despite this result, to continue our empirical study.

The second test proposed may be seen as more illuminating. Since the test can be easily adjusted so it compares the $QMACH$ specification with the $GARCH(1,1)$ model rather than the $ARCH(1)$ model, we decided to do it so. As a matter of fact, once the loglikelihoods of both the $QMACH(1)$ and the $GARCH(1,1)$ models have been calculated, we can test the null according to which the underlying DGP is a $GARCH(1,1)$ against the alternative of the embedding model (21). There is enough evidence to reject the null, so we prefer the alternative model $\left(\lambda_2 = 17,02; \chi^2_{1d.f.} = 3.841\right)$.

Finally, we present the estimations of the $QMACH(1)$ and the $GARCH(1,1)$ models. Eventually, a $QMACH$ specification with more lags ($QMACH(q)$, $\forall$ $q > 1$) should yield better results. In order to verify this assumption, we present the estimated parameters of a $QMACH(2)$ specification. Two statistics are included in order to evaluate the intra-sample conditional variance forecasts performance: the Mean Square Error (MSE) and the Mean Absolute Error (MAE). Additionally, we show the loglikelihood so we can compare the adjustment of the models to the data. As can be seen in Table 5, all models have similar performance, which is extremely encouraging, since the $QMACH(1)$ specification is more parsimonious and easier to estimate than the $GARCH(1,1)$. On the adjustment side, the $GARCH(1,1)$ is outperformed by the $QMACH(1)$, and by the $QMACH(2)$ wich offers an attractive alternative. The first two parameters of the $QMACH(2)$ are similar to the ones yielded by the $QMACH(1)$. This could be seen as a sign of robustness of $QMACH$-type specifications. The overall performance of $QMACH$ specifications compares well with the benchmark model, $GARCH(1,1)$. Yet, a warning should be stated, since the $GARCH(1,1)$ autocovariance function is more appropriate when dealing with this kind of variables. Of course, it must remain clear that this is a specific case; a more ambitious empirical study would be more conclusive.

Model	Const.	2d par.	3d par.	MSE	MAE	likel.
QMACH(1)	2.7217	-0.1229		133.97	7.51	-551.13
QMACH(2)	2.7210	-0.1468	0.1244	133.41	7.46	-550.35
GARCH(1,1)	1.4936	-0.4083	0.4456	149.38	8.03	-559.64

Table 5: Comparation of models

5 Conclusions

This paper has presented a new model, deeply inspired by the Non-Linear Moving Average models, but with the approach tipically used when dealing with conditionally heteroskedastic models. The $QMACH$ does not belong to the $ARCH$ class model. It should indeed be seen as a new instrument to deal with heteroskedasticity. This paper provides some of the most important tools for that purpose. On the one hand, the $QMACH(1)$ model can be consistently estimated either by MOM or ML and, on the other hand, the theoretical correlation function of the squared process, as well as the two statistical tests, should facilitate identification, and provide statistical evidence of either the presence or the absence of $QMACH$ type variables.

Yet, despite the empirical application presented in this paper, the $QMACH$ specification still needs to demonstrate its usefulness in real world. Nevertheless, the $DM - US\$$ application lets us know that the $QMACH$ specification was able to fit as well as the benchmark model, the $GARCH(1,1)$, but has in parallel several advantages over the latter, among which stand out the lack of stationarity conditions, the absence of sign restrictions over the parameters, the parsimony of the model and the asymmetric treatment of shocks.

This new specification will have to compete with the many variants belonging to the $ARCH$ class. Such competitors vary in complexity and robustness; the $QMACH$ model, despite its simplicity, still offers extremely interesting characteristics. A generalization of this model, similar to the one made by Bollerslev(1986) for the $ARCH$ model could possibly enhance its robustness, and thus increase its empirical interest, which has been virtually neglected in the case of Non-Linear Moving Average models.