A conditionally heteroskedastic model with time-varying coefficients for daily gas spot prices

Nazim Regnard and Jean-Michel Zakoïan
A conditionally heteroskedastic model with time-varying coefficients for daily gas spot prices

Nazim Regnard and Jean-Michel Zakoïan\textsuperscript{a,b,c}\textsuperscript{1}

\textsuperscript{a}CREST, Paris, France
\textsuperscript{b}FIME, Paris, France
\textsuperscript{c}EQUIPPE Université Lille 3, France

Abstract

A novel GARCH(1,1) model, with coefficients function of the realizations of an exogenous process, is considered for the volatility of daily gas prices. A distinctive feature of the model is that it produces non-stationary solutions. The probability properties, and the convergence and asymptotic normality of the Quasi-Maximum Likelihood Estimator (QMLE) have been derived by Regnard and Zakoian (2009). The prediction properties of the model are considered. We derive a strongly consistent estimator of the asymptotic variance of the QMLE. An application to daily gas spot prices from the Zeebrugge market is presented. Apart from conditional heteroskedasticity, an empirical finding is the existence of distinct volatility regimes depending on the temperature level.

Key words: GARCH, Nonstationary models, Periodic models, Quasi-maximum likelihood estimation, Time-varying coefficients.

1 Introduction

Following the deregulation of natural gas markets in Europe, natural gas supplying historically ran by long term contracts indexed on crude oil between producing countries and retailers was diversified trough new financial markets (National Balancing Point in UK, Zeebrugge market in Belgium), where it

\textsuperscript{1} Corresponding author. CREST, 15, Boulevard G. Péri, 92245 Malakoff Cedex, France. E-mail adress: zakoian@ensae.fr. Recherche réalisée avec le soutien de la chaire "Finance et développement durable".
can be freely sold at different time horizons. This restructuring has generated uncertainty, requiring the development of appropriate valuation and risk management strategies.

Such strategies require an appropriate modeling of the prices volatility. The standard GARCH models of Engle (1982) and Bollerslev (1986), which arguably constitute the most important class of models for financial data may be inadequate for energy prices. The reason is that energy prices are subject to pronounced daily seasonal patterns, which may not only concern the conditional mean but also the volatility. The periodic ARCH model introduced by Bollerslev and Ghysels (1996) is able to capture those seasonal behaviors in the conditional variance. However, in this model the different regimes appear in a purely periodic succession and it may be worth introducing more flexibility. A GARCH model with regression effects and scaled by seasonal factors has been recently proposed for electricity prices by Koopman, Ooms and Cornaro (2009).

The purpose of this article is to develop a new class of volatility models, introduced in a companion paper by Regnard and Zakoian (2009) (hereafter RZ), for characterizing the seasonal patterns induced by other variables such as temperature. In this model, the parameters associated with the volatility dynamics depend on an exogenous variable, similarly to papers dealing with the conditional mean by Azrak and Mélard (2006), Bibi and Francq (2003), Francq and Gautier (2004a, 2004b).

The article is organized as follows. Section 2 introduces the model and its main probability properties. It is shown how the model can be used for prediction purposes and the QML (Quasi-Maximum Likelihood) estimation is discussed. Section 3 proposes an application to gas prices. A preliminary treatment based on a vector error correction model, involving daily gas prices, brent prices and the temperature, is discussed. Finally, the proposed model is fitted with up to five volatility regimes depending on the temperature. The different specifications are tested, and compared via out-of-sample predictions. Section 4 concludes. A technical proof is given in the appendix.

2 A nonstationary GARCH(1,1) model

The model we consider in this paper is given by

$$
\epsilon_t = \sigma_t \eta_t, \quad \sigma_t^2 = \omega(s_t) + \alpha(s_t) \epsilon_{t-1}^2 + \beta(s_t) \sigma_{t-1}^2, \quad t \in \mathbb{Z}
$$

(1)

where $\{\eta_t\}$ is a sequence of independent and identically distributed (iid), centered variables with unit variance; $\{s_t\}$ is the realization of a process $\{S_t\}$ with
values in a finite set $E = \{e_1, \ldots, e_d\}$; the functions $\omega(\cdot), \alpha(\cdot), \beta(\cdot)$ are defined on $E$ with values in $\mathbb{R}^+$ with $\omega(\cdot) > 0$.

In our application, $s_t$ will correspond to a level of temperature, observed at time $t$. For each level of temperature, the volatility is that of a standard GARCH(1,1) model. Thus, if this level remains constant in some period, the volatility is governed by a standard GARCH. When another level of temperature is reached, the specification of the volatility changes. The existence of different regimes for the volatility, is a common feature between this model and the so-called Markov-switching GARCH models (see ...). However, the models interpretations are completely different. In Markov-switching models, the mechanism of regime change is governed by an non observable variable. In our model, it is governed by an observable process which is exogenous to the model. The dynamics of $\epsilon_t$ is conditional to $(S_t)$.

2.1 Probability properties

The probabilistic properties of this model have been established by Regnard and Zakoïan (2009). Assuming

**A0:** $(s_t)$ is a realization of a process $(S_t)$ which is stationary, ergodic, defined on the same probability space $(\Omega, \mathcal{A}, \mathbb{P})$ as $(\eta_t)$, and independent of $(\eta_t)$,

and letting

$$\pi_j = P(S_t = e_j), \quad j = 1, \ldots, d \quad \text{and} \quad a(x, y) = \alpha(x)y^2 + \beta(x),$$

RZ established that if

$$\gamma_0 := \sum_{j=1}^d \pi_j E\{\log a(e_j, \eta_0)\} < 0,$$

Model (1) admits a nonanticipative nonexplosive solution $(\epsilon_t)$. When $\gamma_0 > 0$, the process is explosive: for any initial value $\sigma_0^2$, we have

In the ARCH(1) case (no coefficients $\beta$), Condition (2) takes the more explicit form $\Pi_{j=1}^d \alpha^{\pi_j}(j) < e^{-E \log \sigma_0^2}$. It can also be noted that the stability of the GARCH(1,1) in each regime, that is

$$E\{\log \alpha(j)\eta_0^2 + \beta(j)\} < 0, \quad j = 1, \ldots, d$$

is sufficient (but not necessary) for the global stability. A necessary condition for (2) is given by $\Pi_{j=1}^d \beta^{\pi_j}(j) < 1$. 

3
Moreover
\[ \gamma_1 := \prod_{j=1}^{d} \{Ea(e_j, \eta_0)\}^\pi_j < 1 \quad \Rightarrow \quad E\epsilon_t^2 < \infty. \tag{3} \]

2.2 Predictions of the squares

For standard GARCH(1,1) models, the optimal prediction of \( \epsilon_t^2 \) in the \( L^2 \) sense, \( E(\epsilon_t^2 \mid \{\epsilon_{t-\ell}^2, \ell > 0\}) \), is obtained from the ARMA(1,1) representation for the squares. Similarly, for Model (1), letting \( u_t = \epsilon_t^2 - \sigma_t^2 = (\eta_t^2 - 1)\sigma_t^2 \) we have
\[ \epsilon_t^2 = \omega(s_t) + (\alpha + \beta)(s_t)\epsilon_{t-1}^2 + u_t - \beta(s_t)u_{t-1}. \]
Letting \( \delta_t = \epsilon_t^2 - \omega(s_t) - (\alpha + \beta)(s_t)\epsilon_{t-1}^2 \), we thus have,
\[ \epsilon_t^2 = \omega(s_t) + (\alpha + \beta)(s_t)\epsilon_{t-1}^2 - \sum_{k \geq 0} \beta(s_t) \ldots \beta(s_{t-k})\delta_{t-k-1} + u_t. \tag{4} \]

This representation is valid because (2) implies
\[ \sum_{j=1}^{d} \pi_j \log \beta(e_j) \leq \sum_{j=1}^{d} \pi_j E\{\log a(e_j, \eta_0)\} < 0, \]
from which the existence of the infinite sum in (4) is deduced, by the arguments used to establish the stationarity condition. Note that the expectation of \( u_t \) conditional on \( \epsilon_t^2 \) past values is zero. The optimal predictor \( \hat{\epsilon}_t^2 \) of \( \epsilon_t^2 \), in the \( L^2 \) sense, is then
\[ \hat{\epsilon}_t^2 = \omega(s_t) + (\alpha + \beta)(s_t)\epsilon_{t-1}^2 - \sum_{k \geq 0} \beta(s_t) \ldots \beta(s_{t-k})\delta_{t-k-1}. \]

Predictions at higher horizons can be derived similarly. Contrary to standard GARCH models, predictions formulas are time dependent through the coefficients \( s_t \).

2.3 QML Estimation

The consistency and asymptotic normality of the QMLE have been proven under mild conditions by Berkes, Horváth and Kokoszka (2003) and Francq and Zakoïan (2004) for standard GARCH and ARMA-GARCH models. NZ showed that these properties can be extended to the model of this paper under assumptions which we now detail.
Let $\theta$ denote the vector of parameters,
\[
\theta = (\omega(e_1), \ldots, \omega(e_d), \alpha(e_1), \ldots, \alpha(e_d), \beta(e_1), \ldots, \beta(e_d))',
\]
with true value $\theta_0$. The parameter is assumed to belong to a parameter space $\Theta \subset [0, +\infty]^d \times [0, \infty]^d$. The sequence $(s_t)$ is observed, and the orders $p, q$ and $d$ are known a priori. Let $(\epsilon_1, \ldots, \epsilon_n)$ be a realization of length $n$ of the nonanticipative solution $(e_t)$. Conditionally on initial values $\tilde{\epsilon}_0$ and $\tilde{\sigma}_0^2$ the gaussian quasi-likelihood is given by
\[
L_n(\theta) = L_n(\theta; \epsilon_1, \ldots, \epsilon_n) = \prod_{t=1}^{n} \frac{1}{\sqrt{2\pi \tilde{\sigma}_t^2}} \exp\left(-\frac{\epsilon_t^2}{2\tilde{\sigma}_t^2}\right),
\]
where the $\tilde{\sigma}_t^2$ are defined recursively, for $t \geq 2$, by
\[
\tilde{\sigma}_t^2 = \tilde{\sigma}_{t-1}^2(\theta) = \omega(s_{t-1}) + \alpha(s_t)\epsilon_{t-1}^2 + \beta(s_t)\tilde{\sigma}_{t-1}^2.
\]
with $\tilde{\sigma}_1^2 = \omega(s_1) + \alpha(s_1)\epsilon_0^2 + \beta(s_1)\tilde{\sigma}_0^2$. A QMLE (quasi-maximum likelihood estimator) of $\theta$ is defined as any measurable solution $\hat{\theta}_n$ of
\[
\hat{\theta}_n = \arg\max_{\theta \in \Theta} L_n(\theta) = \arg\min_{\theta \in \Theta} \hat{I}_n(\theta),
\]
where
\[
\hat{I}_n(\theta) = n^{-1} \sum_{t=1}^{n} \tilde{\ell}_t, \quad \text{and} \quad \tilde{\ell}_t = \tilde{\ell}_t(\theta) = \frac{\epsilon_t^2}{\tilde{\sigma}_t^2} + \log \tilde{\sigma}_t^2.
\]
Indexing the true parameter values by 0, we make the following assumptions.

**A1:** $\theta_0 \in \Theta$ and $\Theta$ is compact.

**A2:** $\sum_{j=1}^{d} \pi_j E\{\log a_0(e_j, \eta_0)\} < 0$ and $\prod_{j=1}^{d} \tilde{\beta}_j^{p_j} < 1$, where $\tilde{\beta}_j = \sup_{\theta \in \Theta} \beta(e_j)$.

**A3:** There exist $r, \rho \in (0, 1)$ and $C > 0$ such that
\[
\forall i > 0, \quad E\{a'_0(S_i, \eta_{t-1}) \ldots a'_0(S_{t-i}, \eta_{t-1-i})\} < C \rho^{i+1}.
\]

**A4:** $\eta_t^2$ has a nondegenerate distribution with $E\eta_t^2 = 1$.

**A5:** For all $i$, $\alpha_0(e_i) + \beta_0(e_i) \neq 0$ and there exist $\ell \in \{1, \ldots, d\}$ and $k > 0$ such that $\alpha_0(e_{\ell}) \mathbb{P}(S_{t-k} = e_{\ell}, S_t = e_i) > 0$.

**A6:** $\theta_0$ belongs to the interior of $\Theta$.

**A7:** $\kappa_\eta = E\eta_t^2 < \infty$.

Then, RZ showed that

1. under **A0-A5**, almost surely $\hat{\theta}_n \to \theta_0$, as $n \to \infty$,

2. under **A0-A7**, $\sqrt{n}(\hat{\theta}_n - \theta)$ is asymptotically $\mathcal{N}(0, (\kappa_\eta - 1)J^{-1})$ distributed, where
\[
J := E_{S, \eta} \left( \frac{1}{\sigma_{S,t}^2(\theta_0)} \frac{\partial \sigma_{S,t}^2(\theta_0)}{\partial \theta} \frac{\partial \sigma_{S,t}^2(\theta_0)}{\partial \theta'} \right)
\]
is a positive-definite matrix, and \((\sigma^2_{S,t}(\theta_0))\) is the process obtained by replacing \(s_t\) by \(S_t\) in the second equation of (1).

The following examples illustrate the variability of the asymptotic covariance matrix, for given ARCH(1) coefficients with two regimes, depending on the distributions of \((S_t)\) and \((\eta_t)\). Let the model

\[
\epsilon_t = \begin{cases} 
(1 + 0.1\epsilon^2_{t-1})^{1/2} \eta_t & \text{if } s_t = 1 \\
(3 + 0.1\epsilon^2_{t-1})^{1/2} \eta_t & \text{if } s_t = 2 
\end{cases}
\]  

and suppose that \((S_t)\) is a Markov chain with transition probabilities \(p(i, j)\).

Then, if

- \(p(1, 1) = p(2, 2) = 0.5; \eta_t \sim \mathcal{N}(0, 1)\):

\[
\text{Var}_{as}(\sqrt{n}(\hat{\theta}_n - \theta)) = \begin{pmatrix} 7.41 & 0 & -1.62 & 0 \\
0 & 56.78 & 0 & -8.96 \\
-1.62 & 0 & 1.30 & 0 \\
0 & -8.96 & 0 & 5.28 \end{pmatrix};
\]

- \(p(1, 1) = p(2, 2) = 0.95; \eta_t \sim \mathcal{N}(0, 1)\):

\[
\text{Var}_{as}(\sqrt{n}(\hat{\theta}_n - \theta)) = \begin{pmatrix} 3.83 & 0 & -1.33 & 0 \\
0 & 300.51 & 0 & -53.24 \\
-1.33 & 0 & 1.58 & 0 \\
0 & -53.24 & 0 & 32.39 \end{pmatrix};
\]

- \(p(1, 1) = p(2, 2) = 0.95, \eta_t\) is distributed as a mixing og Gaussian distributions (with \(\kappa_\eta \approx 9\)):

\[
\text{Var}_{as}(\sqrt{n}(\hat{\theta}_n - \theta)) = \begin{pmatrix} 11.39 & 0 & -1.92 & 0 \\
0 & 918.26 & 0 & -77.02 \\
-1.92 & 0 & 4.21 & 0 \\
0 & -77.02 & 0 & 87.99 \end{pmatrix};
\]

where these asymptotic covariance matrices have been obtained numerically, by simulation of Model (8). The presence of asymptotic covariances equal to zero for parameters of different regimes is due to the absence of coefficients \(\beta\) in the model.
2.4 Consistent estimation of the asymptotic variance of the QMLE

To build tests and confidence intervals for the parameters of Model (1), it is essential to have a consistent estimator of the asymptotic covariance matrix of the QMLE. In view of (7), this matrix depends on the distribution of \((S_t)\) which is unknown. However, the following result provides a consistent estimator which can be easily computed.

**Proposition 1** Under Assumptions A0-A7, a strongly consistent estimator of the matrix \(J\) is given by

\[
\hat{J}_n = \frac{1}{n} \sum_{t=1}^{n} \frac{1}{\hat{\sigma}_t^4(\hat{\theta}_n)} \frac{\partial \hat{\sigma}_t^2(\hat{\theta}_n)}{\partial \theta} \frac{\partial \hat{\sigma}_t^2(\hat{\theta}_n)}{\partial \theta'},
\]

and a strongly consistent estimator of \((\kappa_n - 1)J^{-1}\) is

\[
(\hat{\kappa}_n - 1)\hat{J}_n^{-1}, \quad \text{where} \quad \hat{\kappa}_n = \frac{1}{n} \sum_{t=1}^{n} \frac{\epsilon_t^4}{\hat{\sigma}_n^4(\hat{\theta}_n)}.
\]

**Proof.** See appendix.

3 Application to gas prices volatility

We now turn to an example with real data, namely the daily series of gas spot prices from the Zeebruge market. Before modeling the volatility we filter the series from the conditional mean. To capture the joint behavior of the series of gas, Brent prices and temperatures, we consider a VAR model.

We have a sample of daily prices and temperatures from April, 5, 2000 through March, 3, 2008. Figure 1 displays a plot of the gas and Brent series, \((G_t)\) and \((B_t)\) on the whole period. Let \(g_t = \log G_t\) and \(b_t = \log B_t\) denote the log prices and let \((T_t)\) denote the temperature series. Augmented Dickey-Fuller and KPSS (Kwiatowski,Phillips, Schmidt and Shin, 1992) unit-root tests not reported here suggest that the series \(g_t, b_t\) and \(T_t\) are integrated of order one.

To filter the gas price conditional mean from the influence of the Brent oil price and the temperature, we use a vector error correction model (VECM). There is a growing literature examining the cointegration relationships between different energy prices. Asche, Osmundsen, and Sandsmark (2006) discuss the cointegration between UK natural gas, Brent oil and electricity prices before and after the opening of the Interconnector in 1998. Bachmeier and Griffin (2006) found evidence of cointegration between crude oil, natural gas and coal
in the USA. Panagiotidis and Rutledge (2007) found evidence of a cointegration relationship between the UK wholesale gas prices and the Brent over the period 1996-2003, contradicting the assumption that gas prices and oil prices are decoupled since the liberalisation of gas markets in Europe.

3.1 A VECM for gas and Brent prices

We begin the analysis with an error correction approach. Recall that, in Johansen’s (1988, 1995) notation, a $p$-dimensional VECM takes the form

$$\Delta y_t = \sum_{i=1}^{k-1} \Gamma \Delta y_{t-i} + \Pi y_{t-1} + \mu + u_t$$

where $\Delta$ is the difference operator, $y_t$ is a $p \times 1$ vector of I(1) variables, $\mu$ is a drift parameter, $(u_t)$ is a white noise, $\Pi = \alpha \beta'$ is a $p \times p$ matrix where $\alpha$ and $\beta$ are $p \times r$ full-rank matrices, with $\beta$ containing the $r$ cointegrating vectors and $\alpha$ carrying the loadings in each of the $r$ vectors. A preliminary analysis suggest that oil prices have an impact on gas prices with a delay of 13 weeks. Let $y_t = (g_t, b_{t-\tau}, T_t)$ where $\tau = 91$ days. The Johansen test rejects the null of zero cointegrating vectors between the components of $y_t$. The existence of $r = 1$ cointegrating relation is not rejected and the estimated cointegration vector is, by renormalizing so that the first element be unity, $\hat{\beta} = (1, -1.0809, 0.0194)$.

The estimated VECM is as follows. For ease of presentation, unsignificant coefficients, at the 5% level, have been omitted.
\[ \Delta g_t = -0.077 (g_t - 1.080 b_{t-\tau} + 0.019 T_t + 4.46) + 0.056 \Delta g_{t-1} - 0.010 \Delta T_{t-4} \]

\[ -0.103 \Delta g_{t-5} - 0.091 \Delta g_{t-6} - 0.087 \Delta g_{t-8} - 0.003 \Delta T_{t-8} + \epsilon_t \]

\[ \Delta b_{t-\tau} = \zeta_t \]

\[ \Delta T_t = -0.218 \Delta T_{t-1} - 0.280 \Delta T_{t-2} - 0.225 \Delta T_{t-3} - 0.207 \Delta T_{t-4} \]

\[ -0.135 \Delta T_{t-5} - 0.107 \Delta T_{t-6} - 0.067 \Delta T_{t-8} + \xi_t \]

It is worth noting that for the brent prices, no significant linear influence of the past variables is detected. Results not reported here show that the process \((\epsilon_t, \zeta_t, \xi_t)\) pass the diagnostic tests for the absence of autocorrelation.

### 3.2 Modelling the volatility of gas prices

Figure 2 displays the series of residuals \(\epsilon_t\) for the gas prices. From Figure 3, displaying the empirical autocorrelation function (ACF) and partial autocorrelation function (PACF) of \((\epsilon_t)\), it is seen that this series has the characteristics of a white noise. Caution is needed, however, in the interpretation of the significance bands under conditional heteroskedasticity (see Francq and Zakoian, 2009). The same graphs, displayed in Figure 4 for the series \((\epsilon_t^2)\), show that a GARCH effect is present in the data.

The volatility models for the series \(\epsilon_t\) were estimated over the period April 2000 to December 2004, involving 1192 observations. To have a gauge, the following standard one-regime GARCH(1,1) model was fitted

\[ \sigma_t^2 = 0.0003 + 0.13 \epsilon_{t-1}^2 + 0.79 \sigma_{t-1}^2 \]  

where the standard errors appear in parenthesis. The GARCH coefficients are close to those generally obtained for financial series, with a strong persistence in volatility \((\alpha + \beta = 0.92)\).

Next, we turn to multi-regimes GARCH(1,1) models, where the regimes are determined by the temperature level. We start by a three-regimes model, where the three classes of temperatures correspond to approximately the same num-
Fig. 2. Series \((\epsilon_t)\) for the gas prices.

Fig. 3. Empirical ACF and PACF of the series \((\epsilon_t)\). The bands \(\pm 1.96/\sqrt{n}\) are displayed in dotted lines.

The number of observations. This leads to choose \(s_t = 1\) when \(T_t < 9\), \(s_t = 2\) when \(T_t \in [9, 14]\), and \(s_t = 3\) when \(T_t > 14\), with frequencies in the sample

\[
\pi_1 = 0.35, \quad \pi_2 = 0.32, \quad \pi_3 = 0.33.
\]  (10)
The fitted three-regimes GARCH(1,1) model is as follows.

\[
\begin{align*}
\sigma_t^2 &= \begin{cases} 
0.0003 + 0.13 \epsilon_{t-1}^2 + 0.80 \sigma_{t-1}^2 & \text{when } T_t < 9, \\
0.0011 + 0.37 \epsilon_{t-1}^2 + 0.36 \sigma_{t-1}^2 & \text{when } 9 \leq T_t \leq 14, \\
0.0004 + 0.14 \epsilon_{t-1}^2 + 0.76 \sigma_{t-1}^2 & \text{when } T_t > 14.
\end{cases}
\end{align*}
\]

(11)

All coefficients, except the intercept in the first regime, are significant at the 5% level. The most striking point is the difference between the volatility dynamics in the middle regime, compared to the volatilities of the two extreme regimes. The volatility of the second regime is less persistent \((\alpha(2) + \beta(2) = 0.73)\) with a more convex "news-impact curve". The impact of recent observations on the volatility is stronger than in the low- and high-temperature regimes. It can be noted that the three GARCH(1,1) models are second-order stationary, which entails the global stability with a finite time-dependent variance for \(\epsilon_t\). Note also that the marginal variances within each regimes \((\omega(j)/(1 - \alpha(j) - \beta(j)))\) are roughly the same (around 0.04).

The next model is based on a decomposition of the lower and upper regimes in (11). Letting \(s_t = 1\) when \(T_t < 6\), \(s_t = 2\) when \(T_t \in [6, 9]\), \(s_t = 3\) when \(T_t \in [9, 14]\), \(s_t = 4\) when \(T_t \in [14, 16]\), and \(s_t = 5\) when \(T_t > 16\), the regimes
frequencies are given by

$$\pi_1 = 0.16, \quad \pi_2 = 0.19, \quad \pi_3 = 0.32, \quad \pi_4 = 0.15, \quad \pi_5 = 0.18.$$  \hspace{1cm} (12)

Using the estimated parameters of Model (11) as initial values in the numerical optimization routine, we get the fitted model

$$\sigma_t^2 = \begin{cases} 
0.0008 + 0.15 \epsilon_{t-1}^2 + 0.80 \sigma_{t-1}^2 & \text{when } T_t < 6, \\
0.0010 + 0.00 \epsilon_{t-1}^2 + 0.80 \sigma_{t-1}^2 & \text{when } 6 \leq T_t \leq 9, \\
0.0015 + 0.46 \epsilon_{t-1}^2 + 0.21 \sigma_{t-1}^2 & \text{when } 9 < T_t \leq 14, \\
0.0007 + 0.32 \epsilon_{t-1}^2 + 0.62 \sigma_{t-1}^2 & \text{when } 14 < T_t \leq 16, \\
0.0003 + 0.04 \epsilon_{t-1}^2 + 0.81 \sigma_{t-1}^2 & \text{when } T_t > 16.
\end{cases}$$  \hspace{1cm} (13)

The effects already noticed for the middle regime (little persistence and strong convexity of the news impact curve) is more pronounced with this five-regimes model. A strong coefficient \(\alpha\) is also obtained in the fourth regime. On the contrary, the volatility in all other regimes mainly does not much depend on the last observation. Again, the model is globally stable in the second order sense.

The next model is aimed to detect the effect of extremely low or high temperatures. Letting \(s_t = 1\) when \(T_t < 3.2\), \(s_t = 2\) when \(T_t \in [3.2, 9]\), \(s_t = 3\) when \(T_t \in [9, 14]\), \(s_t = 4\) when \(T_t \in [14, 18.5]\), and \(s_t = 5\) when \(T_t > 18.5\), the regimes frequencies are given by

$$\pi_1 = 0.06, \quad \pi_2 = 0.29, \quad \pi_3 = 0.32, \quad \pi_4 = 0.28, \quad \pi_5 = 0.05.$$  \hspace{1cm} (14)
The fitted model is
\[
\sigma_t^2 = \begin{cases} 
0.0036 + 0.38 \epsilon_{t-1}^2 + 0.47 \sigma_{t-1}^2 & \text{when } T_t < 3.2, \\
0.0007 + 0.04 \epsilon_{t-1}^2 + 0.68 \sigma_{t-1}^2 & \text{when } 3.2 \leq T_t \leq 9, \\
0.0004 + 0.30 \epsilon_{t-1}^2 + 0.62 \sigma_{t-1}^2 & \text{when } 9 < T_t \leq 14, \\
0.0004 + 0.20 \epsilon_{t-1}^2 + 0.72 \sigma_{t-1}^2 & \text{when } 14 < T_t \leq 18.5, \\
0.0000 + 0.00 \epsilon_{t-1}^2 + 0.90 \sigma_{t-1}^2 & \text{when } T_t > 18.5.
\end{cases}
\] (15)

However, many coefficients are found insignificant at the 5% level. Finally, we estimated a model in which the extreme temperatures (low and high) are gathered in the same regime. Letting \( s_t = 1 \) when \( T_t < 3.2 \) or \( T_t > 18.5 \), \( s_t = 2 \) when \( T_t \in [3.2, 9] \), \( s_t = 3 \) when \( T_t \in [9, 14] \), and \( s_t = 4 \) when \( T_t \in [14, 18.5] \), the regimes frequencies deduced from (14) are
\[
\pi_1 = 0.11, \quad \pi_2 = 0.29, \quad \pi_3 = 0.32, \quad \pi_4 = 0.28
\] (16)

and the estimated model is
\[
\sigma_t^2 = \begin{cases} 
0.0026 + 0.34 \epsilon_{t-1}^2 + 0.41 \sigma_{t-1}^2 & \text{when } T_t < 3.2 \text{ or } T_t > 18.5, \\
0.0004 + 0.08 \epsilon_{t-1}^2 + 0.75 \sigma_{t-1}^2 & \text{when } 3.2 \leq T_t \leq 9, \\
0.0011 + 0.38 \epsilon_{t-1}^2 + 0.35 \sigma_{t-1}^2 & \text{when } 9 < T_t \leq 14, \\
0.0004 + 0.08 \epsilon_{t-1}^2 + 0.75 \sigma_{t-1}^2 & \text{when } 14 < T_t \leq 18.5.
\end{cases}
\] (17)

The likelihoods of the different models, displayed in Table 1 allow to compare the different fits. From likelihood ratio tests, at the 5% significance level,
- the standard GARCH(1,1) model is not rejected against the 3 regimes model;
- the GARCH(1,1) model is however rejected against any model with \( d > 3 \);
- the 3 regimes model is rejected against the 5 regimes Model (13).

Wald tests not reported here lead to the same conclusions. In the same table, the estimated kurtosis of the variable \( \eta_t = \epsilon_t / \sigma_t \) are reported. The biggest kurtosis reduction is obtained with the 5-regimes Model (13).
Table 1

Likelihoods of the estimated models and Kurtosis of the standardized returns

<table>
<thead>
<tr>
<th></th>
<th>GARCH</th>
<th>Model (11)</th>
<th>Model (13)</th>
<th>Model (15)</th>
<th>Model (17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\log L_n$</td>
<td>5173</td>
<td>5179</td>
<td>5206</td>
<td>5210</td>
<td>5187</td>
</tr>
<tr>
<td>$\hat{\kappa}_\eta$</td>
<td>6.00</td>
<td>5.76</td>
<td>5.43</td>
<td>5.68</td>
<td>5.63</td>
</tr>
</tbody>
</table>

Table 2

MSE ($\times 10^{-5}$) of predictions (last 500 observations)

<table>
<thead>
<tr>
<th></th>
<th>GARCH</th>
<th>Model (11)</th>
<th>Model (13)</th>
<th>Model (15)</th>
<th>Model (17)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$(d = 1)$</td>
<td>7.66</td>
<td>7.57</td>
<td>7.29</td>
<td>7.47</td>
<td>7.47</td>
</tr>
<tr>
<td>$(d = 3)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(d = 5)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(d = 5)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$(d = 4)$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2 reports Mean-Squared Errors (MSE) of prediction. We re-estimated the different models over the same sample minus the last 500 observations, which were used for the predictions. The estimated models over the sample were very close to those estimated on the whole sample. From the prediction point of view, the 5-regime Model (13) is again the preferred specification.

4 Conclusion

This paper reviewed a class GARCH models, which allow the volatility to depend on an observed exogenous process. This observability of the state variable makes the model much easier to use than the so-called Markov-switching models in which the regime change is governed by a latent Markov chain. The model can be estimated by QML and a consistent estimator of the asymptotic covariance matrix has been proposed. The methodology has been applied to daily gas prices using the temperature as the exogenous variable. We found evidence of five regimes, with very different dynamics for the volatility in the moderate-temperature regime. The model could be used for prediction purposes, using temperature scenarios. Many extensions, by including more lags in the volatility dynamics or by considering multivariate series, are left for future research. It is hoped that the article will broaden the use of time series models driven by exogenous variables.
A Technical details

Proof of Proposition 1. For all \( \theta \in \Theta \), let

\[
\tilde{J}_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \frac{1}{\sigma_t^4(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'}, \\
J_n(\theta) = \frac{1}{n} \sum_{t=1}^{n} \frac{1}{\sigma_t^4(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'}.
\]

Note that \( \hat{J}_n = \tilde{J}_n(\hat{\theta}_n) \). We have, letting \( \theta = (\theta_i)_{i=1}^{3d} \),

\[
\frac{1}{n} \sum_{t=1}^{n} \left\{ \frac{1}{\sigma_t^4(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \right\}_{\theta = \hat{\theta}_n}
= \frac{1}{n} \sum_{t=1}^{n} \left\{ \frac{1}{\sigma_t^4(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \right\}_{\theta = \theta_0} + \frac{1}{n} \sum_{t=1}^{n} \left\{ \frac{1}{\sigma_t^4(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \right\}_{\theta = \theta_{ij}^*} (\hat{\theta}_n - \theta_0).
\]

(A.1)

where \( \theta_{ij}^* \) is between \( \hat{\theta}_n \) and \( \theta_0 \). Denote by \( (\sigma_{\delta,t}^2(\theta)) \) the process recursively defined under A2 by \( \sigma_{\delta,t}^2(\theta) = \omega(S_t) + \alpha(S_t)\epsilon_{t-1}^2 + \beta(S_t)\sigma_{\delta,t-1}^2(\theta) \). We have, for almost all sequence \( (s_t) \),

\[
\limsup_{n \to \infty} \left\| \frac{1}{n} \sum_{t=1}^{n} \frac{\partial}{\partial \theta'} \left( \frac{1}{\sigma_t^4(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \right) \right\|
\leq \limsup_{n \to \infty} \sup_{\theta \in \mathcal{V}(\theta_0)} \left\| \frac{\partial}{\partial \theta'} \left( \frac{1}{\sigma_t^4(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \right) \right\|
= E_{\theta_0} \sup_{\theta \in \mathcal{V}(\theta_0)} \left\| \frac{\partial}{\partial \theta'} \left( \frac{1}{\sigma_{\delta,t}^4(\theta)} \frac{\partial \sigma_{\delta,t}^2(\theta)}{\partial \theta} \frac{\partial \sigma_{\delta,t}^2(\theta)}{\partial \theta'} \right) \right\| < \infty.
\]

where \( \| \cdot \| \) denotes any norm on \( \mathbb{R}^{3d} \). The equality follows from Lemma 5.2 in RZ and the fact that \( \sigma_t^2(\theta) \) and \( \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \) are measurable functions of \( (s_t, s_{t-1}, \ldots, \eta_t, \eta_{t-1}, \ldots) \). The last inequality is a consequence of iii), in the proof of Theorem 4.2 in RZ. Since \( \hat{\theta}_n - \theta_0 \to 0 \) a.s., the last term in (A.1) converges to zero in probability as \( n \) tends to infinity.

Using again Lemma 5.2 in RZ, we obtain the a.s. convergence to \( J \) of the first term in the right-hand side of (A.1). Thus we have shown that

\[
\frac{1}{n} \sum_{t=1}^{n} \left\{ \frac{1}{\sigma_t^4(\theta)} \frac{\partial \sigma_t^2(\theta)}{\partial \theta} \frac{\partial \sigma_t^2(\theta)}{\partial \theta'} \right\}_{\theta = \hat{\theta}_n} \to J, \quad \text{a.s.}
\]

Since, by FZ, Proof of Theorem 4.2,

\[
\sup_{\theta \in \mathcal{V}(\theta_0)} \left\| \frac{1}{n} \sum_{t=1}^{n} \left( \frac{\partial^2 \hat{\ell}_t(\theta)}{\partial \theta \partial \theta'} - \frac{\partial^2 \hat{\ell}_t(\theta)}{\partial \theta \partial \theta'} \right) \right\| \to 0 \quad \text{a.s.}
\]

15
where $\ell_t(\theta)$ is defined as $\tilde{\ell}_t(\theta)$ with $\tilde{\sigma}_t$ replaced by $\sigma_t$, we thus have

$$
\hat{J}_n = \frac{1}{n} \sum_{t=1}^{n} \left\{ \frac{1}{\tilde{\sigma}_t^4} \frac{\partial \tilde{\sigma}_t^2}{\partial \theta_i} \frac{\partial \tilde{\sigma}_t^2}{\partial \theta_j} \right\}_{\theta = \hat{\theta}_n} \rightarrow J, \quad \text{a.s.}
$$

By the same arguments we prove that

$$
\hat{\kappa}_\eta = \frac{1}{n} \sum_{t=1}^{n} \frac{\epsilon_t^4}{\tilde{\sigma}(\hat{\theta}_n)^4} \rightarrow E\eta_t^4
$$

and the proposition is proved.

References


Francq, C., Gautier, A., 2004b. Estimation of time-varying ARMA models