Financial Engineering and the Japanese Markets **4:** 1–35, 1997.

Volatility Clustering, Asymmetry and Hysteresis in Stock Returns: International Evidence

MICHEL CROUHY and MICHAEL ROCKINGER[1] *
[1]*HEC School of Management, 78351 Jouy-en-Josas, France*
e-mail: rockinger@gwsmtp.hec.fr

Abstract. Encompassing a very broad family of ARCH-GARCH models, we show that the ATGARCH $(1, 1)$ model, where volatility rises more in response to bad news than to good news, and where news are considered bad only below a certain level, is a remarkably robust representation of worldwide stock market returns. The residual structure is then captured by extending ATGARCH $(1, 1)$ to an *hysteresis* model, HGARCH, where we model structured memory effects from past innovations. Obviously, this feature relates to the psychology of the markets and the way traders process information. For the French stock market we show that votalitity is affected differently, depending on the recent past being characterized by returns all above or below a certain level. In the same way a longer term trend may also influence volatility. It is found that bad news are discounted very quickly in volatility, this effect being reinforced when it comes after a negative trend in the stock index. On the opposite, good news have a very small impact on volatility except when they are clustered over a few days, which in this case reduces volatility.

Key words: GARCH, Hysteresis, market psychology

1. Introduction

The ARCH methodology developed by Engle (1982) and generalized by Bollerslev (1986) has been very successful to model time varying volatility. A huge literature has emerged from those models and several surveys are available (see Bollerslev, Chou, and Kroner (1993), Bera and Higgins (1993), as well as Bollerslev, Engle, and Nelson (1994). A casual analysis of this literature shows that there exists a multitude of competing models, that they are mostly applied to the U.S. market, and that none of them attempts to capture possible residual structure beyond ARCH features of the first order.

In this paper we start by nesting popular models within a common framework, then by using a likelihood ratio test strategy we show how a best possible model can be found. There exists other research which attempts to compare ARCH models. Pagan and Schwert (1990) and Engle and Ng (1993) compare models by evaluating their ability to forecast volatility. Whereas Pagan and Schwert consider

* The first author is at the Canadian Imperial Bank of Commerce and the second author is at the HEC School of Management. This research was done while the first author was at the HEC School of Management. This research was partially supported by the HEC Foundation and FNEGE (Fondation Nationale pour l'Enseignement de la Gestion des Entreprises).

the mean square errors of volatility forecasts as the criterion for evaluation, Engle and Ng adopt the closeness of forecasted volatility to a benchmark provided by a non-parametric model as a criterion. Encompassing of models with non stochastic volatility has already been performed by Higgins and Bera (1992), Ding *et al.* (1993), and Hentschel (1991, 1995). Higgins and Bera propose a non-linear generalization of ARCH models, such that ARCH, GARCH and LOG ARCH models are special cases. The model of Ding *et al.* allows for an asymmetric impact of news, whereby positive past returns do not have the same impact as negative returns. They also consider modelling a power of volatility rather than just its square or its log. The model of Hentschel, which allows for asymmetries and a level effect comes closest to our research. However, our basis model is somewhat more flexible and allows easier interpretation.

An application to 21 countries in the FT-Goldman Sachs data base at the daily frequency shows the usefulness of a model allowing for asymmetries and a level effect.[1]

We establish that for all these countries the way market participants react to extreme price movements depends on the direction and magnitude of events. For positive events, it seems that traders and investors process information in an efficient manner. On the contrary, stock markets tend to overreact to negative events. Asymmetry is sometimes referred to as the 'leverage effect', following Black (1976), but this phenomenon is also well documented in the 'stock market overreacton' literature (see, e.g., De Bondt and Thaler (1985)). Interestingly, we show that this overreaction occurs, at a world-wide scale, when the magnitude of the bad news goes beyond some level.

An additional contribution of this research is the quest for a more complicated structure in residuals in addition to the first order ARCH effect. We call such a structure *hysteresis*, and the associated econometric model the HGARCH. The idea behind hysteresis is to perform a 'technical analysis' for second moments of returns. It characterizes the fact that a positive or a negative shock will not have the same impact on volatility, whether it comes after several consecutive days where all the innovations were above or below a certain level, or whether the longer term trend of past innovations was above or below a certain level. This is a sort of 'size' or 'threshold' effect in a dynamic context, and it relates to the way traders process information. The impact of a shock on volatility will depend on the cumulative size of past innovations. If it goes beyond a threshold level, then volatility reacts

[1] From a casual analysis of trading activities, one might suspect that this phenomenon has become more pronounced in the recent years since banks started to sell huge nominal amounts of complex derivative instruments on the OTC markets. As a result, most of the banks carry on their books the same market exposure. Either the stock market is flat and there is no need for the banks to readjust their hedging position. Or, on the contrary, the market moves unexpectedly in one direction and, simultaneously, all the banks need to buy the stock market (futures contracts) when the market goes up, and they have to sell when the market goes down. These trading patterns may trigger liquidity problems and amplify market movements, and are most likely to generate sudden bursts in volatility (cf. Group of Thirty report (1993)).

more strongly. Since psychological features may play an important role we test the HGARCH model for France, which is the country we know best,[2] and where, according to French officials, the market is characterized by the highest ratio of notional principal in OTC equity derivatives to stock market turnover. French banks have been very active in manufacturing guaranteed capital and performance funds, and other structured products where the CAC40 is the benchmark index. Therefore, we suspect that if we can detect this hysteresis feature, the French stock market must be the ideal candidate.

In Section 2 we study statistical properties of the CAC40 and show that returns are fat-tailed, not normally distributed, and that volatility is heteroskedastic. In Section 3 we present a generic ARCH model of the asymmetric type, show how popular models are encompassed, and consider a statistical procedure to select the most appropriate model. We then extend this model to accommodate for hysteresis by proposing a Hysteresis-GARCH (HGARCH) model where we differentiate between short term and longer term cumulative innovations. In Section 4 we estimate all the models using maximum likelihood techniques. Focusing on the French stock market, we show that all the nested asymmetric models still have fat-tailed residuals. As a consequence, the non linearities apparent in the data are not fully captured by the existing models. Our new specification permits to capture some of them. In Section 5 we discuss the estimation results for the 21 countries in the FT-Goldman Sachs database. Finally, we summarize the main contributions of this paper and make suggestions for further research.

2. Data Description for the French Stock Market

2.1. The data

Our primary data consist of daily closing levels for the new French stock index, CAC40, for the period from July 1987 to November 1995 for a total of 2077 observations.[3] Figure 1 plots daily compound returns of the CAC40 for the full sample period.

We can decompose the series of daily rates of return $r_t, t = 1, \ldots, T$ as

$$r_t = E(r_t|\mathcal{I}_{t-1}) + y_t,$$

[2] Actually, traders suggested to us the possible existence of such a psychological phenomenon.

[3] The CAC40 is a value weighted index composed of France's 40 most liquid blue chips traded on the Paris stock exchange. This index is adjusted for splits and stock dividends but not for cash dividends. This index constitutes the underlying asset for futures and options contracts traded on the MATIF (Marché à Terme International de France) and the MONEP (Marché des Options Négociables de Paris), respectively, as well as for structured derivative products sold on the OTC market. The older SBF240 is a broader index composed of 240 stocks. It is apparently more representative of the French stock market, although most of the stocks which compose this index are very thinly traded. As a consequence the SBF240 suffers from severe serial correlation due to non syncronous trading of its components. Our source for the CAC40 series is the SBF (Société des Bourses Françaises).

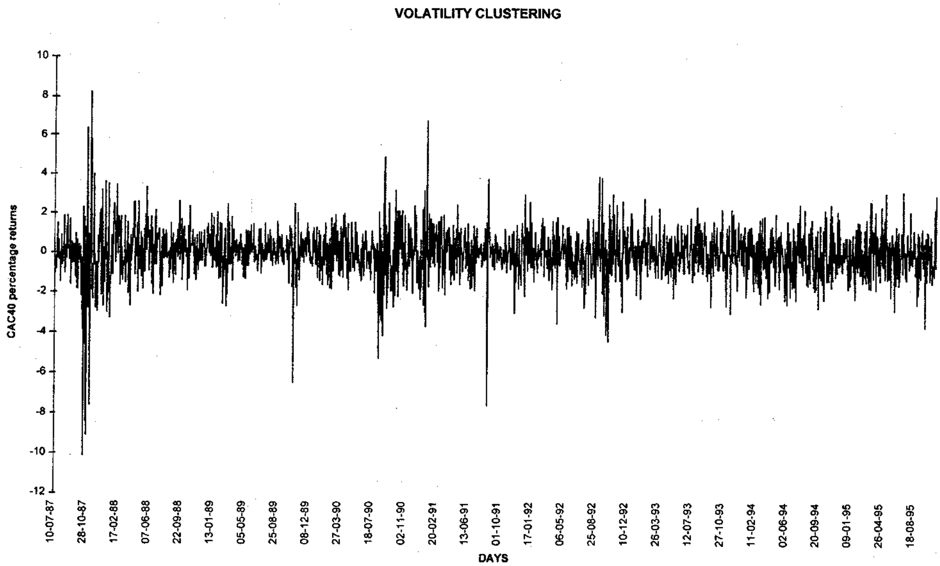

Figure 1. Volatility clustering. Daily continuously compounded returns on the CAC40 over the time period July 1987 to November 1995.

where $E(r_t|\mathcal{I}_{t-1})$ represents the conditional mean given the information set $\mathcal{I}_{t-1}$ available at time $t-1$ and where y_t is a non-predictable component (meaning that $E(y_t|\mathcal{I}_{t-1}) = 0$).

In this section we will characterize the empirical properties of the two components for the CAC40. By drawing on the existing literature, we first examine which variables may forecast returns.

French (1980), Rogalsky (1984), Gibbons and Hess (1981), and Harris (1986) document negative mean returns for U.S. stocks on Mondays, while Fama (1965) and Godfrey, Granger and Morgenstern (1964) document higher return variances for U.S. stocks on Mondays.

As Table I shows, the same 'Monday effect' and in addition a 'Thursday effect' affect rates of return of the French CAC40. In order to capture a potential day of the week effect we have included dummies in the conditional mean equation. Even though, day of the week effects are globally significant, later estimation results do not change for series filtered for a Monday or any other day of the week effect. In this light we decided not to filter the returns data for any day of the week effect.

Other specific calendar effects characterize the French stock market, namely a settlement day and an end of the month effect. The Paris Stock Exchange is organized as a forward market with a monthly settlement day. Payment for transactions concluded during the period between the previous settlement and the current one are due at the end of the month.[4] In addition, at the end of each month

[4] See Crouhy and Galai (1992).

Table I. Search for a deterministic component in the CAC40 data

Coefficient	Day of week	End of month	Settlement
Constant	—	0.0135	0.0417
		(0.0349)	(0.0594)
Monday	−0.2290	—	—
	(0.0227)*		
Tuesday	0.0144	—	—
	(0.0607)		
Wednesday	0.0519	—	—
	(0.0604)		
Thursday	0.1252	—	—
	(0.0608)*		
Friday	0.0662	—	—
	(0.0612)		
End of month	—	0.0963	—
		(0.1583)	
Settlement day	—	—	0.1465
			(0.2688)

Results from the OLS regression of r_t on a set of day-of-the-week dummies, and end-of-the-month dummy, and a settlement-day dummy.

In this Table we report heteroskedasticity robust standard erros. In all Tables,* (**) denotes parameter estimates statistically different from 0 at the 5% (10%) level.

option contracts on the CAC40 expire which may be the source of trading patterns.[5] To take into account these effects we introduced corresponding dummies in the conditional mean equation, as shown in Table I, the estimates were small and non significant.

Thus, faced with the risk of just introducing spurious noise by filtering the data for calendar effects, we decided instead to continue our research with the raw returns series.

Bollerslev (1986), French, Schwert, and Stambaugh (1987), Baillie and DeGennaro (1990), and Hamao, Masulis, and Ng (1990) adjust the conditional mean return for a first moving average, MA(1). Alternatively, provided that only the first autocorrelation coefficients are mildly significant, a low order autoregressive adjustment can be adopted, e.g. AR(1). [See, Lo and MacKinlay (1988), Akgiray (1989), Engle and Ng (1993), and Nelson (1991)].[6]

Figure 2 and Table II show that the pattern of autocorrelations indicates little serial correlation. For the sample, however, the traditional Ljung–Box statistic

[5] See de Jong *et al*. (1989).

[6] As a practical matter there is little difference between the AR(1) and the MA(1) adjustment for returns when the AR or MA coefficients are small as it is the case for the CAC40.

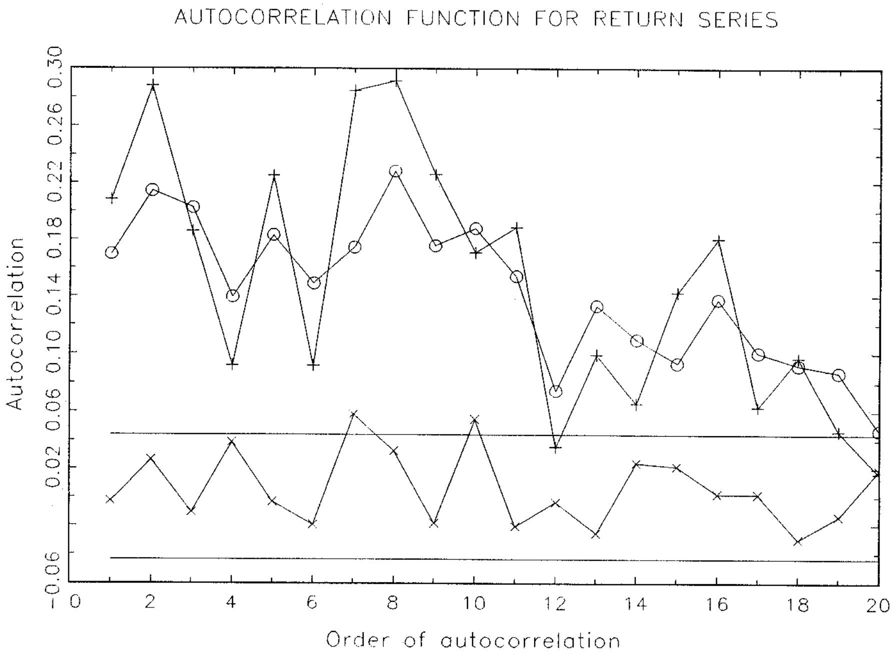

Figure 2. Autocorrelation function for return series. Daily continuously compounded returns on the CAC40 over the time period July 1987 to November 1995. The band is centered at zero with width $\pm 2/\sqrt{T}$. $\times$: Raw CAC40 returns (r_t); $\circ$: r_t^2; $+$: $|r_t|$.

rejects the null hypothesis that the correlation coefficients are jointly zero. This effect can be attributed to heteroskedasticity. Indeed, when we perform a test for joint autocorrelation with the correction for heteroskedasticity suggested by White (1980), we accept that returns are not autocorrelated at the 5 percent confidence level.[7]

Having established that our returns data does not need to be filtered for a conditional mean, we show in Table II that the data is left skewed, leptokurtic (and thus non-normal), as well as heteroskedastic. As Table II further shows, there is substantially more correlation between squared and absolute returns, than there is between the raw CAC40 returns themselves. This means that large absolute (or squared) returns are more likely to be followed by a large absolute return than by a small absolute return: big shocks are indeed clustered together as confirmed in Figure 1. More generally, the distribution of the next absolute return can depend on several past absolute returns. No linear stochastic process can provide a satisfactory explanation for this structure of autocorrelations. In this paper we explore various models of conditional changes in the variance of returns.

[7] See Table II for the definitions.

Table II. Summary statistics for the French CAC40 index over the period from July 1987 to November 1995

Statistic	Raw CAC40 r_t	Squared CAC40 r_t^2	Absolute CAC40 $\|r_t\|$
Number of observations (T)	2077	2077	2077
Mean	0.0092	1.5575	0.8908
Standard deviation	1.2483	4.8783	0.8742
Median	0.0112	0.4507	0.6714
Minimum	−10.1376	0.0000	0.0000
Maximum	8.2254	102.7704	10.1376
Skewness (S)	−0.5906	12.0544	3.2901
	(−10.9887)*	(224.2783)*	(61.2138)*
Kurtosis (K)	7.8241	186.9931	21.1937
	(72.7861)*	(1739.5547)*	(197.16)*
J–B	5418.57*	3.10^{7*}	4.10^{5*}
$\rho(1)$	−0.0029	0.2085*	0.1696*
$\rho(2)$	0.0262	0.2882*	0.2145*
$\rho(3)$	−0.0108	0.1858*	0.2027*
$\rho(4)$	0.0385	0.0918*	0.1395*
$\rho(5)$	−0.0035	0.2253*	0.1831*
$\rho(6)$	−0.0197	0.0915*	0.1488*
$\rho(7)$	0.0578*	0.2847*	0.1744*
$\rho(8)$	0.0323	0.2917*	0.2284*
$\rho(9)$	−0.0186	0.2285*	0.1756*
$\rho(10)$	0.0545*	0.1710*	0.1880*
$\rho(11)$	−0.0209	0.1886*	0.1540*
$\rho(12)$	−0.0039	0.0352*	0.0747*
$\rho(13)$	−0.0259	0.0993*	0.1335*
$\rho(14)$	0.0236	0.0657*	0.1095*
$\rho(15)$	−0.0212	0.1428*	0.0932*
$\rho(16)$	−0.0020	0.1806*	0.1378*
$\rho(17)$	−0.0016	0.0631*	0.1005*
$\rho(18)$	−0.0303	0.0970*	0.0918*
$\rho(19)$	−0.0139	0.0463*	0.0870*
$\rho(20)$	0.0186	0.0164	0.0470*
$2/\sqrt{T}$	0.0439	0.0439	0.0439
B–P	220.11*	87.72*	198.10*
L–B(10)	21.75*	992.74*	710.15*
L–B(20)	29.34*	1245.30*	952.77*
L–B–W(10)	6.65	21.85*	38.65*
L–B–W(20)	10.86	39.35*	49.43*

Note: Given a time series of rates of return: $\{r_t\}, t = 1, \ldots, T$, with $r_t = 100\ln(S_t/S_{t-1})$ where S_t is the closing level of the index on day t, the skewness, S, and excess kurtosis, K, of the sample are defined as follows: (Note continues on next page)

Note Table II (continued)

$$S = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})^3/\sigma^3 \quad \text{and} \quad K = \frac{1}{T-1}\sum_{t=1}^{T}(r_t - \bar{r})^4/\sigma^4 - 3,$$

where $\bar{r}$ and σ respectively denote the sample mean and standard deviation. If the sample distribution is normal, S and K are asymptotically normally distributed: $S \sim \mathcal{N}(0, 6/T)$ and $K \sim \mathcal{N}(0, 24/T)$. Under the normality assumption, the skewness and excess kurtosis coefficient should be zero. In addition, if we denote by S' and K' the standardized skewness and kurtosis, i.e. $S' = S/\sqrt{6/T}$ and $K' = K/\sqrt{24/T}$, then the Jarque–Bera (J–B) statistic $S'^2 + K'^2$ follows a chi-square distribution with 2 degrees of freedom. The t statistics of the skewness and kurtosis are in parenthesis.

The Breusch–Pagan test (B–P), obtained by regressing a squared variable on two of its squared lags is distributed as a χ^2 with two degrees of freedom.

The asymptotic distribution of the autocorrelation coefficient at lag h, $\rho(h)$, is normal with mean 0 and standard deviation $1\sqrt{T}$ (see Anderson (1942)).

The joint test for autocorrelation among H lags, given by the Ljung–Box statistic is defined by

$$\text{L–B(H)} = T(T+2)\sum_{h=1}^{H}\frac{\rho(h)^2}{T-h},$$

where $\rho(h) = \Sigma_{t=1}^{T-h} r_t r_{t-h}/\Sigma_{t=1}^{T} r_t^2$. The Ljung–Box statistic corrected for heteroskedasticity as suggested by White (see Bera and Higgins (1993) p. 358) is

$$\text{L–B–W(H)}$$
$$= [\gamma(1), \ldots, \gamma(H)]' \begin{bmatrix} \sum_t r_{t-1}^2 r_t^2 & \cdots & \sum_t r_{t-1} r_{t-H} r_t^2 \\ \vdots & \vdots & \vdots \\ \sum_t r_{t-1} r_{t-H} r_t^2 & \cdots & \sum_t r_{t-H}^2 r_t^2 \end{bmatrix}^{-1} \begin{bmatrix} \gamma(1) \\ \vdots \\ \gamma(H) \end{bmatrix} * T,$$

where $\gamma(h) = \frac{1}{T}\sum_{t=1}^{T-h} r_t r_{t-h}$. Under the null hypothesis of no autocorrelation of the r_t both statistics are asymptotically distributed as a chi-square with H degrees of freedom.
The critical values for the χ^2 with 1, 2, 10 and 20 degrees of freedom are respectively 3.84, 5.99, 18.31 and 31.41 at the 5 percent confidence level.

3. Methodology

3.1. The Encompassing Model

From the data analysis presented in the previous section, we conclude that variance shows predictability. Many models in the literature try to capture this predictability. As a matter of fact there are so many models that several surveys exist (see Bollerslev, Chou and Kroner (1993), Bera and Higgins (1993), Bollerslev, Engle and Nelson (1994)) and soon we can expect surveys of surveys.

There exist several possibilities to compare these models. Rather than to compare forecasting abilities of models as in Pagan and Schwert (1990), or as in Engle and Ng (1993), we think that models should be nested and that formal tests should be used to select the best one.

The following encompassing model allows us not only to compare a restrictive class of models, but also to estimate models which so far haven't gotten any attention in the literature.

(i) *conditional return:*

$$r_t = f(\sigma_t) + y_t, \quad \text{with } y_t = \sigma_t \varepsilon_t, \tag{1}$$

$$\varepsilon_t \text{ are i.i.d.} \quad \text{and} \quad \varepsilon_t \sim \mathcal{N}(0,1). \tag{2}$$

f denotes the conditional expected return at time t. Volatility may play a role in the mean equation due to our discrete time approximation of a continuous time process.[8]

(ii) *conditional volatility:*

Current volatility is linearly conditional on lagged innovations (or returns) and conditional volatility:

$$\begin{aligned}\frac{\sigma_t^{\beta_1}-1}{\beta_1} &\\ &= \alpha_0 + \sum_{i=1}^{p} (\alpha_{1i}^{+} \mathbb{I}_{\{y_{t-i}/\sigma_{t-i}^{\beta_2} > \gamma\}} + \alpha_{1i}^{-} \mathbb{I}_{\{y_{t-1}/\sigma_{t-i}^{\beta_2} \leqslant \gamma\}}) \left| \frac{y_{t-1}}{\sigma_{t-i}^{\beta_2}} - \gamma \right|^{\beta_3} \\ &\quad + \sum_{j=1}^{q} \alpha_{2j} \frac{\sigma_{t-j}^{\beta_1}-1}{\beta_1},\end{aligned} \tag{3}$$

where $\alpha_0, \alpha_{1i}^{+}, \alpha_{1i}^{-}, \alpha_{2j}, \beta_1, \beta_2, \beta_3$ and γ are parameters. This model will be extended in Section 3.4 to allow for current volatility to depend on past innovations with a more complex structure which, from our point of view, better corresponds to the way traders process information. Positivity and stationarity conditions for the parameters are considered in an Appendix.

Expression (3) can also be written in the more compact form

$$\frac{\sigma_t^{\beta_1}-1}{\beta_1} = \alpha_0 + \sum_{i=1}^{p} c_i(y_{t-i}, \sigma_{t-i}) + \sum_{j=1}^{q} \alpha_{2j} \frac{\sigma_{t-j}^{\beta_1}-1}{\beta_1},$$

[8] If in continuous time the stock index S_t, follows a geometric Brownian motion $dS_t = \mu S_t dt + \sigma S_t dw_t$, where dw_t is a Wiener process, then $d\ln(S_t) = (\mu - 0.5\sigma^2)dt + \sigma dw_t$. In a discrete time approximation the left-hand side becomes $\ln(S_t/S_{t-1})$ which is exactly the way continuous compound returns are computed

Table III. Models encompassed by the general specification

β_1	β_2	β_3	Model
0	0	1	$\ln(\sigma_t) = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}\leqslant\gamma\}})\lvert y_{t-1} - \gamma\rvert + \alpha_2 \ln(\sigma_{t-1})$
0	0	2	$\ln(\sigma_t) = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}\leqslant\gamma\}})(y_{t-1} - \gamma)^2 + \alpha_2 \ln(\sigma_{t-1})$
0	1	1	$\ln(\sigma_t) = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{\varepsilon_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{\varepsilon_{t-1}\leqslant\gamma\}})\lvert\varepsilon_{t-1} - \gamma\rvert + \alpha_2 \ln(\sigma_{t-1})$
0	1	2	$\ln(\sigma_t) = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{\varepsilon_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{\varepsilon_{t-1}\leqslant\gamma\}})(\varepsilon_{t-1} - \gamma)^2 + \alpha_2 \ln(\sigma_{t-1})$
1	0	1	$\sigma_t = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}\leqslant\gamma\}})\lvert y_{t-1} - \gamma\rvert + \alpha_2\sigma_{t-1}$
1	0	2	$\sigma_t = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}\leqslant\gamma\}})(y_{t-1} - \gamma)^2 + \alpha_2\sigma_{t-1}$
1	1	1	$\sigma_t = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{\varepsilon_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{\varepsilon_{t-1}\leqslant\gamma\}})\lvert\varepsilon_{t-1} - \gamma\rvert + \alpha_2\sigma_{t-1}$
1	1	2	$\sigma_t = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{\varepsilon_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{\varepsilon_{t-1}\leqslant\gamma\}})(\varepsilon_{t-1} - \gamma)^2 + \alpha_2\sigma_{t-1}$
2	0	1	$\sigma_t^2 = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}\leqslant\gamma\}})\lvert y_{t-1} - \gamma\rvert + \alpha_2\sigma_{t-1}^2$
2	0	2	$\sigma_t^2 = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}\leqslant\gamma\}})(y_{t-1} - \gamma)^2 + \alpha_2\sigma_{t-1}^2$
2	1	1	$\sigma_t^2 = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{\varepsilon_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{\varepsilon_{t-1}\leqslant\gamma\}})\lvert\varepsilon_{t-1} - \gamma\rvert + \alpha_2\sigma_{t-1}^2$
2	1	2	$\sigma_t^2 = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{\varepsilon_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{\varepsilon_{t-1}\leqslant\gamma\}})(\varepsilon_{t-1} - \gamma)^2 + \alpha_2\sigma_{t-1}^2$

For the model $(\beta_1, \beta_2, \beta_3) = (1, 0, 1)$, if we set $\gamma = 0$ we obtain the TGARCH(1, 1) model.
The EGARCH(1, 1) model can be recovered from $\beta_1, \beta_2, \beta_3) = (0, 1, 1)$ and by setting $\gamma = 0$.
For $(\beta_1, \beta_2, \beta_3) = (2, 0, 2)$ if we impose $\alpha_1^+ = \alpha_1^-$ we get an AGARCH(1, 1). Further restricting $\gamma = 0$ yields a GARCH(1, 1).
For $(\beta_1, \beta_2, \beta_3) = (2, 1, 2)$ and $\alpha_1^+ = \alpha_1^-$ we get an AVGARCH(1, 1).
The model with $(\beta_1, \beta_2, \beta_3) = (1, 0, 1)$ could be called an ATGARCH(1, 1).

where $c_i(y_{t-i}, \sigma_{t-i})$ will be referred to as the *core*.

Table III summarizes the various specifications which this model encompasses. This Table and a formal proposition in the Appendix shows that this family includes popular models as well as several others which have not received attention in the literature. It should be noticed that this nesting procedure helps to obtain a better taxonomy of existing models.

This model contains the linear ARCH model and its generalized version, the GARCH model, introduced by Engle (1982) and Bollerslev (1986) respectively, i.e.

$$\text{ARCH}(p):\quad \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2, \tag{4}$$

and

$$\text{GARCH}(p, q):\quad \sigma_t^2 = \omega + \sum_{i=1}^{p} \alpha_i y_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2, \tag{5}$$

where the system's coefficients ω, the α_i's and the β_j's are non-negative. For covariance stationary of the volatility process the coefficients of the lagged errors and lagged conditional variances must sum to less than one. Since GARCH models involve past volatility and returns, it follows, as noted by Mandelbrot (1963):

...large changes tend to be followed by large changes – of either sign – and small changes by small changes... In other words volatility comes in waves, i.e. calm periods follow turbulent ones.

However, with GARCH models, positive and negative innovations of the same magnitude produce the same amount of volatility. This feature contradicts empirical evidence where stock returns appear to be negatively correlated with changes in returns volatility, i.e. volatility tends to rise in response to *bad news* (returns lower than expected) and to fall in response to *good news* (returns higher than expected). As a consequence GARCH models will underpredict (overpredict) volatility following bad (good) news. This suggests that a model in which σ_t^2 responds asymmetrically to positive and negative innovations should be preferable to standard GARCH models.

The EGARCH (Exponential-GARCH) model developed by Nelson (1994) explicitly accommodates for the asymmetric relationship between innovations and volatility changes. The conditional variance in the EGARCH(1, 1) is modeled as

$$\text{EGARCH}(1,1):$$

$$\ln(\sigma_t^2) = \omega + a\frac{y_{t-1}}{\sigma_{t-1}} + b\left[\frac{|y_{t-1}|}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}}\right] + \beta\ln(\sigma_{t-1}^2). \tag{6}$$

By modeling the log of the conditional volatility, it is not necessary to restrict parameter values to avoid negative conditional variances as in the ARCH and GARCH models. It involves the normalized shocks $\varepsilon_t = y_t/\sigma_t$. The contribution of a positive innovation on the log of the conditional variance is: $(b+a)y_{t-1}/\sigma_{t-1}$ while for a negative innovation it is $(b-a)|y_{t-1}|/\sigma_{t-1}$. A negative a would confort the hypothesis that negative shocks generate more volatility than positive shocks. The use of logs increases the impact of large shocks on the next period conditional variance.

The existence of a strong leverage effect, i.e. bad news affecting more volatility than good news is thus well supported by the empirical evidence, since the estimated coefficient $\hat{a}$ is statistically negative for several stock markets: the U.S. (see Nelson (1991)), Japan (see Engle and Ng (1993)), and the U.K. (see Poon and Taylor (1992)), and in all cases a negative shock has an impact on volatility almost twice as big as a positive one.

Engle (1990) and Sentana (1991) propose the Asymmetric GARCH or AGARCH model which is a GARCH model where the minimum of the News Impact Curve (NIC), first defined by Engle and Ng (1993) occurs at a nonzero level, λ

$$\text{AGARCH}(1,\ 1):\quad \sigma_t^2 = \omega + \alpha(y_{t-1}-\lambda)^2 + \beta\sigma_{t-1}^2. \tag{7}$$

The leverage effect hypothesis requires that λ be positive so that the minimum of the parabola occurs for a positive level λ of the innovation. In that case we should observe a negative correlation between innovations and next period volatility.[9]

A closely related model is the AVGARCH where

$$\text{AVGARCH}(1,1):\quad \sigma_t^2 = \omega + \alpha(\varepsilon_{t-1} - \lambda)^2 + \beta\sigma_{t-1}^2 \tag{8}$$

such that the minimum of the NIC occurs at $y_{t-1} = \lambda\sigma_{t-1}$. The model differs from AGARCH in that not past returns, but rather past normalized shocks matter as in the case of EGARCH.

Glosten, Jagannathan and Runkle (1989) with their GJR model, and Zakoïan (1991) with his Threshold GARCH model or TGARCH, allow both sides of the NIC to have different slopes

$$\text{GJR}(1,1):\quad \sigma_t^2 = \omega + \alpha y_{t-1}^2 + \gamma y_{t-1}^2 \mathbb{I}_{\{y_{t-1}<0\}} + \beta\sigma_{t-1}^2, \tag{9}$$

where $\mathbb{I}_{\{\text{condition}\}}$ is an indicator function which is equal to one when the condition is satisfied, and 0 otherwise.

$$\text{TGARCH}(p,q):\quad \sigma_t = \omega + \sum_{i=1}^{p}(\alpha_i^+ y_{t-i}^+ - \alpha_i^- y_{t-i}^-) + \sum_{j=1}^{q}\beta_j\sigma_{t-j}, \tag{10}$$

where $y_{t-i}^+ = \max(y_{t-i}, 0)$ and $y_{t-i}^- = \min(y_{t-i}, 0)$.

This encompassing model contains also models which have not obtained attention in the literature. For instance, the model obtained by setting in model (3) $\beta_1 = 1, \beta_2 = 0, \beta_3 = 1, p = q = 1$ could be called an ATGARCH$(1,1)$ since there is both an asymmetry and a threshold effect:[10]

$$\sigma_t = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}\leqslant\gamma\}}|y_{t-1} - \gamma| + \alpha_2\sigma_{t-1}.$$

If instead we select $\beta_1 = 1, \beta_2 = 1, \beta_3 = 2$ then we get what might be called an AVT-GARCH$(1,1)$.

3.2. Comparison with existing models

Hentschel (1991, 1995) proposes the model

$$\frac{\sigma_t^\lambda - 1}{\lambda} = \omega_0 + \alpha\sigma_{t-1}^\lambda(|\varepsilon_t - b| - c(\varepsilon_t - b))^\nu + \beta\frac{\sigma_{t-1}^\lambda - 1}{\lambda}.$$

Our model is different from his since:

---

[9] See also Campbell and Hentschel (1992).

[10] In the following, we will often consider restrictions of model (3) and regroup all constant terms in α_0 without changing notation. Also, for the case $p = q = 1$ we will suppress the second index of the α's.

(i) by allowing a different power for the Box–Cox transform of the volatility and for the core, here $\alpha\sigma_{t-1}^{\lambda}(\cdot)^{\nu}$, we are able to generate more models,

(ii) we believe that the interpretation of past innovations is easier within our framework where we directly model asymmetries.

In order to interpret our model it is hence forth not necessary to trace the NIC. Our work also differs from his, in that we investigate the possibility that higher order lags may explain volatility.

Ding, Granger and Engle (1993) propose the Asymmetric Power ARCH (A-PARCH). Where

$$\sigma_t^{\delta} = \alpha_0 + \sum_{i=1}^{p} \alpha_i (|y_{t-i}| - \gamma y_{t-i})^{\delta} + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^{\delta}.$$

Again, there are differences with our model since

(i) we allow not only past returns to have an impact on current volatility, but we allow both for past innovations y_{t-i} and past standardized innovations $y_{t-i}/\sigma_{t-i}^{\beta_2}$ to matter,

(ii) we allow for a threshold parameter (γ) which we show is a crucial parameter in characterizing asymmetry,

(iii) their model does not truly impose a Box–Cox power transform since their model does not allow for $\delta = 0$ corresponding to the log case,

(iv) their model does not allow different powers for σ_t and the core $(|y_{t-i}| - \gamma y_{t-i})$.

In the model of Higgins and Bera (1992), there is no different impact on volatility of positive and negative past returns, therefore, their model omits asymmetries, which we show to play an important role.

Our model aims at providing easily interpretable estimates and, therefore, it does not include a non-parametric part as in Gallant and Tauchen (1989), Gallant, Hsieh and Tauchen (1991), or Gourieroux and Monfort (1992).[11] The reasons are

(i) it is hard to interpret the non-parametric estimates,

(ii) the estimates are not efficient, and

(iii) the estimation procedure is quite slow.[12]

[11] See also Bollerslev *et al*. (1992, pp. 12–14).

[12] We experimented with several non-parametric models where we approximated the core by a series of Hermite polynomials. This allowed us to confirm that the shape of the NIC obtained out of the parametric model is similar to the one obtained out of the encompassing model. An alternative approach to ARCH-type models is given by Harvey and Shephard (1993) or Ruiz (1994) where volatility is stochastic.

3.3. Statistical selection of a best model

Maximum-likelihood estimation allows selection of the best performing model within a nested family. Nested hypothesis can be tested using the generalized likelihood ratio

$$\Lambda = \frac{\sup_{\theta \in \Theta_0} \mathcal{L}(\theta; x)}{\sup_{\theta \in \Theta} \mathcal{L}(\theta; x)} \tag{11}$$

of the maximized-log likelihood functions $\mathcal{L}$ under the null and the enlarged parameter space (Θ_0 and Θ, respectively). Under the null Θ_0, the statistic $-2 \ln \Lambda$ follows a chi-square distribution with degrees of freedom equal to the number of restrictions on the parameters in the constrained model.[13] This is an asymptotic result which holds under mild assumptions as long as the hypothesis Θ_0 is nested within Θ.

A parameter θ can also be tested individually for being equal to some parameter θ_0. If we have an estimate $\hat{\theta}$ then the asymptotic distribution of $\sqrt{T}(\hat{\theta} - \theta_0)$ is $\mathcal{N}(0, J^{-1} I J^{-1})$ where J is the Hessian matrix and I the matrix of cross products of the likelihood (see also Gourieroux, Monfort and Trognon (1984)). This result holds in particular if the true distribution of the residuals is not normal.

For numerical purposes it may be difficult to directly estimate a model as rich as (3) because of possible over-parameterization. For this reason we recommend estimation of restricted versions of (3) followed by the general estimation where starting values correspond to the parameter estimates of the restricted model with the largest likelihood value.

3.4. Extension to models which accommodate for hysteresis

In the previous section we have examined how lagged returns and conditional voltality may influence today's volatility, and whether volatility responses are symmetrical or not. We now extend the model to allow for hysteresis.

Because of psychological reasons market makers may behave differently if they have, day after day, lost or won money. For this reason we now consider the possibility that richer structures may affect volatility. The first possibility we consider is that the conditional volatility may be affected by past innovations having been above or below a certain threshold level during a certain time period, such as 2 or 3 days. What could also matter is the overall magnitude of certain events: volatility might be more intense if for several days traders have lost beyond a certain threshold, moreover, this intensity could depend on the magnitude of cumulative losses during the days considered.

A second possibility is that a longer time trend, above or below a certain threshold, may affect the conditional volatility. If a trader has had a losing position

[13] See, for example, Graybill, Mood and Boes (1982, pp. 419–421 and 440–442).

for the last weeks he might be reacting more nervously to any news arrivals than if the trend has been a good one.

This phenomenon relates to the economic and psychological rationales already discussed in the introduction concerning market overreaction. In all previous models past returns or innovations may have an impact on volatility, but they are modelled as independent shocks, and no consideration is made concerning the influence some patterns of past returns might have on current and future volatility.

In order to model the impact on volatility of historical patterns we propose to extend the previous models by differentiating the very short term, up to a few days, and the longer term, up to a few weeks. For the short term effect we will only consider trading patterns where all the innovations are above or below a certain threshold. For the longer term we distinguish between trends above or below a certain level. We adopt the following specification for modelling hysteresis where

$$
\begin{aligned}
& z_t = y_t/\sigma_t^{\beta_2} - \gamma \\
& y_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1)
\end{aligned}
\tag{12}
$$

$$
\begin{aligned}
& \frac{\sigma_t^{\beta_1} - 1}{\beta_1} \\
& \quad = \alpha_0 + \sum_{i=1}^{p} (\alpha_{1i}^{+} \mathbb{I}_{\{z_{t-i}>0\}} + \alpha_{1i}^{-} \mathbb{I}_{\{z_{t-i}\leqslant 0\}}) |z_{t-i}|^{\beta_3} + \sum_{j=1}^{q} \alpha_{2j} \frac{\sigma_{t-j}^{\beta_1} - 1}{\beta_1} \\
& \qquad + \alpha_3^{+} \mathbb{I}_{\{z_{t-2}>0,\ldots,z_{t-a-1}>0\}} + \alpha_3^{-} \mathbb{I}_{\{z_{t-2}\leqslant 0,\ldots,z_{t-a-1}\leqslant 0\}} \\
& \qquad + (\alpha_4^{+} \mathbb{I}_{\{z_{t-2}>0,\ldots,z_{t-a-1}>0\}} + \alpha_4^{-} \mathbb{I}_{\{z_{t-2}\leqslant 0,\ldots,z_{t-a-1}\leqslant 0\}}) \left| \sum_{k=2}^{a+1} z_{t-k} \right|^{\beta_3} \\
& \qquad + (\alpha_5^{+} \mathbb{I}_{\{\sum_{l=2}^{b+1} z_{t-l}>0\}} + \alpha_5^{-} \mathbb{I}_{\{\sum_{l=2}^{b+1} z_{t-l}\leqslant 0\}}) \left| \sum_{l=2}^{b+1} z_{t-l} \right|^{\beta_3}.
\end{aligned}
\tag{13}
$$

The parameters a and b are selected based on economic priors. The second and third term in Equation (13) associated with parameters α_3^+, α_3^- and α_4^+, α_4^- correspond to the short term effect, while the last term associated with α_5^+ and α_5^- allows for the longer term effect. In an Appendix we will consider positivity and stationarity conditions for this model.

4. Empirical Estimation

In this section we are going to present the results of the estimations. First, we will estimate a large set of models to get a good prior on what a satisfactory set of starting values could be for the general model (3). Once we have estimated the general model we will test restrictions of it. We will then investigate the robustness of this model. As a last step we will present the results of the Hysteresis GARCH model. [14]

4.1. Estimation results

Experiments with various estimation methods and different values for the parameters p and q, for the given sample size, indicate a serious problem of over-parameterization. It was never possible to achieve convergence of the general model (3) with values of p and q larger than one.

In Table IV we first present the results of various restricted versions of our general model. We notice that the largest likelihood is obtained for the ATGARCH model where $\beta_1 = 1, \beta_2 = 0$ and $\beta_3 = 1$ corresponding to a linear specification of volatility and where innovations rather than their standardized version are used. If the time required for an estimation can be considered a criterion for valuing a model, then, for this alternative measure too, this model turns out to be the best.

We use the estimates of this model as starting values for the general model, the results of which are reported in the last column of Table IV.[15]

4.1.1. *Linear versus quadratic or log specification*

The general model has a likelihood of -3177.70. Since all other models in this Table correspond to restrictions of $\beta_1, \beta_2, \beta_3$ we can test those restrictions which a chi-square with three degrees of freedom. A simple computation yields that models where the likelihood is smaller than – 3181.61 can be rejected against the general model.

Among all estimates we notice that the only model with a likelihood larger than this value is the model where $\beta_1 = 1, \beta_2 = 0, \beta_3 = 1$. From now on we will focus on this ATGARCH model.

4.1.2. *Is $\gamma = 0$?*

We notice that for the ATGARCH model γ is highly significant which suggests that there is an important threshold effect at around -0.85 percent.

---

[14] All estimations were carried out using the Constrained Maximum Likelihood GAUSS subroutines. All estimations were performed on a Pentium computer running at 133 Mhz and require at most a few minutes. The Berndt, Hall, Hall and Hausman (1974) BHHH algorithm was used most of the time.

[15] Clearly, once we have estimates for the general model we use alternative starting values to convince ourselves that the estimates correspond to a global maximum.

Table IV. Pseudo maximum likelihood estimates for nested models

β_1	0	0	0	0	1	1	1	1	2	2	2	2	0.4985
													(0.1445)
β_2	0	0	1	1	0	0	1	1	0	0	1	1	0.0279
													(0.4946)
β_3	1	2	1	2	1	2	1	2	1	2	1	2	0.8117
													(0.2995)
α_0	−0.0115	−0.0533	−0.0754	−0.0358	0.0955	0.1225	0.0270	0.0501	0.0277	0.0558	0.0000	0.0000	0.1973
	(0.0139)	(0.0748)	(0.0220)*	(0.0259)	(0.0261)*	(0.0569)*	(0.0262)	(0.0241)*	(0.0137)*	(0.0291)	(0.0301)	(0.0105)	(0.2995)
α_1^+	0.0007	−0.0729	0.0737	0.0609	0.0128	0.1244	0.0283	0.1392	0.0172	0.0329	0.0152	0.1340	0.0084
	(0.0151)	(0.1587)	(0.0473)*	(0.0871)	(0.0146)	(0.0211)	(0.0376)	(0.0676)*	(0.0164)	(0.0181)	(0.0243)	(0.0397)*	(0.0154)
α_1^-	0.1200	0.0107	0.1292	0.0268	0.2226	0.0793	0.2241	0.1068	0.3249	0.1989	0.1497	0.1635	0.1701
	(0.0369)*	(0.0234)	(0.0291)*	(0.0303)	(0.0624)*	(0.0411)	(0.1015)*	(0.0417)*	(0.0875)*	(0.1283)	(0.0725)*	(0.0306)*	(0.0483)*
α_2	0.8839	0.9564	0.9354	0.9551	0.8759	0.8524	0.9294	0.9020	0.8726	0.8266	0.9189	0.8907	0.8892
	(0.0430)*	(0.0755)*	(0.0260)*	(0.0425)*	(0.0322)*	(0.0543)*	(0.0402)*	(0.0255)*	(0.0314)*	(0.0497)*	(0.0530)*	(0.0168)*	(0.0586)*
γ	−0.6697	3.0928	0.1738	1.4372	−0.8546	0.2963	−0.6680	0.3710	−0.9451	−0.2606	−1.1704	0.1825	−0.9118
	(0.0032)*	(5.7792)	(0.1318)	(1.7101)	(0.1750)*	(0.4074)	(0.2473)*	(0.2412)	(0.1403)*	(0.4785)	(0.3101)*	(0.2045)	(0.3149)*
$\mathcal{L}$	−3183.52	−3189.48	−3193.88	−3205.99	−3205.99	−3187.44	−3201.13	−3208.19	−3184.84	−3184.77	−3243.89	−3210.38	−3177.70
$J--B$	985.33*	1214.16*	1999.95*	1587.60*	1005.4*	1145.91*	1555.56*	1136.82*	1058.53*	1006.75*	1523.65*	963.67*	1023.48*
B–P(2)	2.76	6.27	2.71	2.52	2.69	2.91	2.28	2.64	2.33	2.96	1.45	2.36	2.73
L–B–W(10)	15.63	12.54	14.41	12.57	15.59	15.86	13.21	13.68	14.27	16.02	9.72	12.57	15.61

The unrestricted model estimated here is

$$r_t = \mu + y_t,$$

$$y_t = \sigma_t \varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0,1),$$

where

$$\frac{\sigma_t^{\beta_1}}{\beta_1} = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}/\sigma_{t-1}^{\beta_2} > \gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}/\sigma_{t-1}^{\beta_2} \leqslant \gamma\}}) \left| \frac{y_{t-1}}{\sigma_{t-1}^{\beta_2}} - \gamma \right|^{\beta_3} + \alpha_2 \frac{\sigma_{t-1}^{\beta_1}}{\beta_1}, \quad \text{if } \beta_1 \neq 0$$

and

$$\ln(\sigma_t) = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}/\sigma_{t-1}^{\beta_2} > \gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}/\sigma_{t-1}^{\beta_2} \leqslant \gamma\}}) \left| \frac{y_{t-1}}{\sigma_{t-1}^{\beta_2}} - \gamma \right|^{\beta_3} + \alpha_2 \ln(\sigma_{t-1}), \quad \text{if } \beta_1 = 0.$$

The last column contains the unrestricted model. Other columns contain restrictions of the general model determined by the values of $(\beta_1, \beta_2, \beta_3)$.

$\mathcal{L}$ is the value of the likelihood function. Statistics are as in Table II. All standard errors are corrected for possible non-normality of the likelihood function.

4.1.3. *Is α_1^+ equal to α_1^-?*

To what extent do past returns have an asymmetric impact on volatility? The likelihood principle requires again to reestimate the model by imposing the restriction $\alpha_1^+ = \alpha_1^-$. We then obtain

$$\begin{array}{rcccccc} \sigma_t & = & 0.0702 & + & 0.1362\ |z_{t-1}| & + & 0.8316\ \sigma_{t-1}, \\ & & (0.0313)^* & & (0.0255)^* & & (0.0357)^* \end{array}$$

$$\begin{array}{rcl} \text{with } \gamma & = & 0.3075, \quad \mathcal{L} = -3197.36, \\ & & (0.1314)^* \end{array}$$

where $z_t = y_t - \gamma$. Here we have one restriction which can be tested with a chi-square with one degre of freedom. The generalized likelihood ratio statistic is 37.2, high enough to allow rejection of symmetry at any level of significance.

4.1.4. *Conclusion for the nesting procedure*

From the statistical inference we reach the conclusion that ATGARCH(1, 1), i.e. model 101, is the best specification for the general model (3).

For this model the parameter α_2 is 0.8759 which shows the important impact of past volatility on current and future volatility.[16] High past volatility tends to be followed by high volatility and similarly for small volatility. The difference between α_1^+ and α_1^- reveals strong asymmetry. Since α_1^+ is insignificant, good news have very little importance on volatility changes. The large value of α_1^- indicates that bad news, on the contrary, tend to have a strong impact on volatility.

The level parameter γ is negative, $\gamma = -0.85$ percent and statistically significant, i.e. the minimum of the NIC is reached for a negative return. This reinforces the asymmetric impact of bad news. Only when returns are negative beyond the critical level of -0.85 percent they are considered as bad news and have a stronger impact on volatility. This result comforts the intuition we developed earlier in the introduction.

Those two results clearly document that for daily data on the CAC40, a significnat asymmetry and a shift in the news impact curve exist. As Figure 3 shows, the effect of asymmetry shifts the NIC towards negative values and the TGARCH effect rotates the NIC. This suggests that returns below the -0.85 percent have a strong impact on volatility. Each additional point of 'bad news' increases volatility by about 0.2 point. For returns above -0.85 percent and, of course, for positive returns, volatility increases positively but at the negligible rate of 0.01.

4.2. How good is ATGARCH(1, 1)?

We have shown in the previous section that ATGARCH(1, 1) is a quite satisfactory model since it removes heteroskedasticity from the data. However, residuals remain

[16] The half life of a shock to volatility is 5 days.

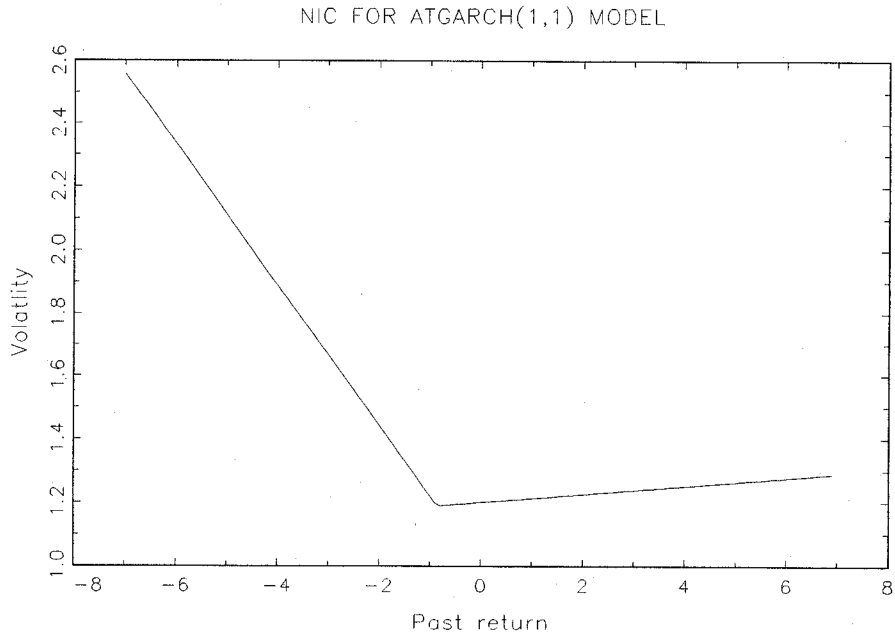

Figure 3. News Impact Curve for the ATGARCH(1, 1) model. Plot of how current volatility is affected by past returns *ceteris paribus* in the ATGARCH(1, 1) model.

fat tailed. This suggests that a more general model allowing either for volatility in the mean or for more lags might lead to an improvement.

4.2.1. *ATGARCH* $(1, 1)$ *in mean*

We consider for the conditional mean a constant, as well as a linear, and a quadratic function of volatility. The estimation results are shown in Table V.

Inspection of the estimates of θ_1 and θ_2 indicate that both coefficients are not significant. Also, the linear and quadratic mean both nest the constant volatility model. The generalized likelihood ratio for testing the linear model is $-2\ln(\Lambda) = 0.04$ and for testing the quadratic model is 0.18, thus, adding a linear or a quadratic volatility term to the mean, does not improve the model.

A possible reason why we don't capture any mean effect may be that the daily series are very noisy in comparison with the smoother estimated volatility series. Omission of volatility from the mean can, therefore, not explain the remaining fat tailedness of the residuals.

4.2.2. *Extensions to ATGARCH*(p, q)

Even though one could extend the ATGARCH(1, 1) to a higher order we limit

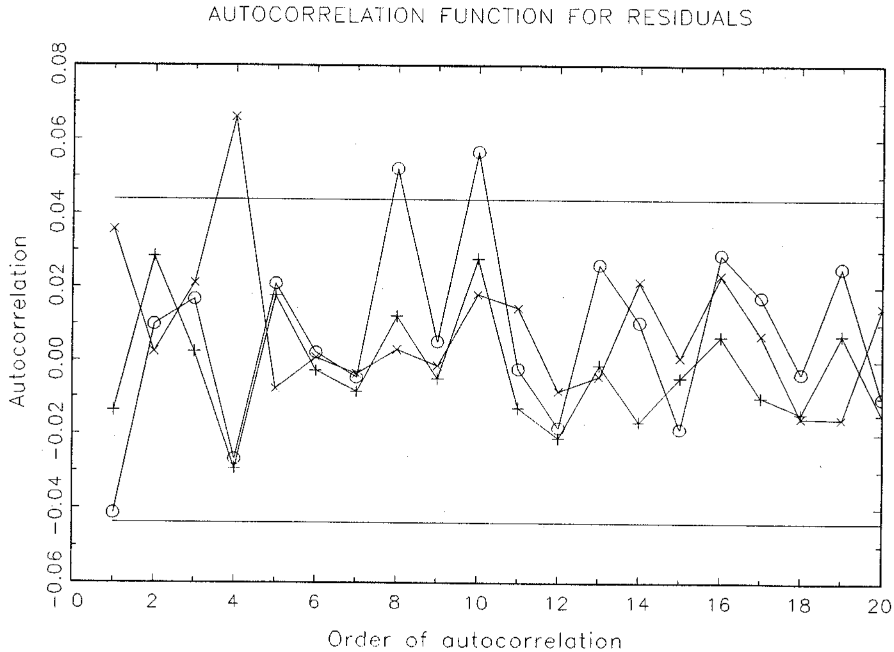

Figure 4. Autocorrelation function for the residual series of the ATGARCH(1, 1). The band is centered at zero with width $\pm 2/\sqrt{T}$. $\times$: residuals (ε_t); o: ε_t^2; $+$: $|\varepsilon_t|$.

our analysis to the second order $(p, q \leqslant 2)$, since this is enough to show that this extension does not yield any improvement.

Again, it is possible to conduct the estimations in nested form. An ATGARCH(2, 2) can be viewed as the general, less restricted model, where ATGARCH(2, 1) or ATGARCH(1, 2) represent a first level of restriction. A further restriction yields the ATGARCH(1, 1). Estimation results are presented in Table VI.[17]

First, computations of the generalized likelihood ratios to test the ATGARCH(2, 1) and ATGARCH(1, 2) against an ATGARCH(2, 2) are 0.42 and 5.48, respectively. Since those two statistics follow respectively a chi-square distribution with 1 and 2 degrees of freedom, we cannot reject the restrictions. Second, to test the further restriction of an ATGARCH(1, 1) against ATGARCH(2, 1) or ATGARCH(1, 2) we obtain as statistics 3.04 and 0.84. Since the restrictions involve a chi-square with one and two degrees of freedom, we cannot reject the restriction corresponding to ATGARCH(1, 1). Those results suggest that the strategy of simply adding more lags to the ATGARCH is not appropriate.

[17] We give consideration to positivity and stationarity constraints in an Appendix.

Table V. Extension of ATGARCH(1, 1) to allow for volatility in the mean

Param.	With no volatility in mean	ATGARCH(1, 1) With linear volatility in mean	With quadratic volatility in mean
θ_0	0.0175	0.0430	0.0336
	(0.0232)	(0.1372)	(0.0475)
θ_1	—	− 0.0243	—
		(0.1271)	
θ_2	—	—	−0.0142
			(0.0350)
α_0	0.0966	0.0960	0.0954
	(0.0265)*	(0.0265)*	(0.0265)*
α_1^+	0.0116	0.0115	0.0115
	(0.0142)	(0.0143)	(0.0143)
α_1^-	0.2252	0.2253	0.2273
	(0.0644)*	(0.0658)*	(0.0659)*
α_2	0.8764	0.8772	0.8776
	(0.0323)*	(0.0323)*	(0.0325)*
γ	−0.8863	−0.8832	−0.8866
	(0.1776)*	(0.1796)*	(0.1765)*
$\mathcal{L}$	−3178.67	−3178.65	−3178.58
J–B	986.38*	998.14*	999.47*
B–P(2)	2.62	2.51	2.42
L–B–W(10)	15.44	15.03	14.68

The estimated model is

$$r_t = \theta_0 + \theta_1\sigma_t + \theta_2\sigma_t^2 + \sigma_t\varepsilon_t$$

$$\sigma_t = \alpha_0 + (\alpha_1^+\mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^-\mathbb{I}_{\{y_{t-1}\leqslant\gamma\}})|y_{t-1} - \gamma| + \alpha_2\sigma_{t-1}.$$

4.3. Testing for short and longer term hysteresis

Before estimating the general model we performed an *hysteresis-bias test* in the spirit of Engle and Ng (1991) to check if hysteresis patterns can be detected in the residuals of the ATGARCH(1, 1). Let ε_t be the estimated residual, a positive-hysteresis-bias test can be defined as the t-ratio for the coefficient δ_1 in the OLS regression $|\varepsilon_t| = \delta_0 + \delta_1\mathbb{I}_{[\sum_{s=1}^{b}\varepsilon_{t-s}>0]}\frac{1}{b}\sum_{s=1}^{b}\varepsilon_{t-s} + \nu_t$ and a negative-hysteresis-bias test can be defined in the same way for a negative trend. If residuals were white noise, all tests should yield an insignificant δ_1. On the other hand, a significant δ_1 means that the past trend can help in explaining current volatility. As indicated in Table VII, the positive-hysteresis-bias tests yield significant negative statistics for small values of b, suggesting the possibility that historical patterns may have some explanatory power for the residuals. Clearly, this test is not very powerful and in a

Table VI. Extension of ATGARCH(1, 1) to ATGARCH(p, q) with $p, q \leqslant 2$

Param.	ATGARCH(2, 2)	ATGARCH(2, 1)	ATGARCH(1, 2)
α_0	0.1166	0.1111	0.0963
	(0.0404)*	(0.0343)*	(0.0284)*
α_{11}^+	0.0174	0.0151	0.0119
	(0.0337)	(0.0290)	(0.0139)
α_{11}^-	0.1259	0.1252	0.2251
	(0.0606)*	(0.0631)**	(0.0610)*
α_{12}^+	0.0000	0.0000	—
	(0.0422)	(0.0295)	—
α_{12}^-	0.1394	0.1276	—
	(0.1120)	(0.0894)	—
α_{21}	0.7804	0.8568	0.8764
	(0.3065)*	(0.0425)*	(0.3047)*
α_{22}	0.0671	—	0.0000
	(0.2632)	—	(0.2875)
γ	−0.8437	−0.8702	−0.8779
	(0.1720)*	(0.1743)*	(0.1789)*
$\mathcal{L}$	−3176.81	−3177.23	−3179.17
J–B	919.3*	891.69*	983.81*
B–P(2)	2.77	2.74	2.59
L–B–W(10)	15.39	15.43	15.57

The general ATGARCH(2, 2) model is

$$\begin{aligned}\sigma_t = \alpha_0 &+ (\alpha_{11}^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_{11}^- \mathbb{I}_{\{y_{t-1}\leqslant\gamma\}}|y_{t-1} - \gamma| \\ &+ (\alpha_{12}^+ \mathbb{I}_{\{y_{t-2}>\gamma\}} + \alpha_{12}^- \mathbb{I}_{\{y_{t-2}\leqslant\gamma\}})|y_{t-2} - \gamma| \\ &+ \alpha_{21}\sigma_{t-1} + \alpha_{22}\sigma_{t-2}.\end{aligned}$$

We further restrict it to ATGARCH(2, 1) and ATGARCH(1, 2)

complete likelihood estimation results may be different. For the moment we only consider those results encouraging.

The choice of a value for a follows from Table VIII where the number of series of consecutive positive and negative returns is displayed. We see that it is not reasonable to select a greater than 3 since when a is equal to 4 the percentage of zeros in the corresponding dummy would be almost 95 percent. Therefore, we will restrict a to values of 2 and 3. For convenience we call this pattern *isosigned returns* of order 2 and 3. We chose b to correspond to one, two or three trading weeks.

The general model (13) contains many parameters. We use a restriction of the β_1, β_2 and β_3 parameters corresponding to the ATGARCH. Table IX presents the results of the estimation.

Table VII. Hysteresis-bias-tests for the residuals of the ATGARCH(1, 1)

b	Hysteresis-bias-test	Estimate
5	positive	−2.147
5	negative	−1.275
10	positive	−2.138
10	negative	0.250

Let ε_t be the estimated residual and $\mathbb{I}_{\text{condition}}$ an indicator function taking the value 1 when condition is true and 0 otherwise. The positive-hysteresis-bias test is defined as the t-ratio for the coefficient δ_1 in the OLS regression $|\varepsilon_t| = \delta_0+\delta_1\mathbb{I}_{\left[\sum_{s=1}^{b}\varepsilon_{t-s}>0\right]}\frac{1}{b}\sum_{s=1}^{b}\varepsilon_{t-s}+\nu_t$ and the negative-hysteresis-bias test is defined in the same way for a negative trend.

Table VIII. Frequency of series of consecutive positive and negative returns

Length of monotonicity	Positive history	Negative history
1	1053	1024
2	551	523
3	275	250
4	133	113

First we observe that the significance and magnitude of the ATGARCH coefficients found in the previous estimation is preserved and stays unaffected by the extension of the model to a richer structure.

When we allow for an indicator variable corresponding to isosigned innovations $(b = 0)$ we notice that returns which are constantly larger than our -0.86 percent threshold decrease volatility, whereas returns constantly smaller increase volatility for very short patterns $(a = 2)$. Only the constantly larger returns decrease volatility when $a = 3$.

This result confirms that there remain nonlinearities in the residuals which can be captured by tendencies of markets. It also shows that positive news can decrease market's volatility. Negative patterns matter only for a shorter horizon. This means that when rsults are constantly bad then volatility adjusts after 2 days. A bad result for a third day does no longer matter.

Table IX. Maximum likelihood estimates of the Hysteresis GARCH

Param.	$a=2$	$a=3$	Param.	$a=2$	$a=3$	$b=5$	$b=10$	$b=15$	$a=2, b=10$
γ	−0.8695	−0.9213	γ	−1.0864	−0.9957	−0.7269	−0.7496	−0.7573	−0.7992
	(0.1302)*	(0.1306)*		(0.2339)*	(0.1776)*	(0.1096)*	(0.1322)*	(0.1337)*	(0.1498)*
α_0	0.1090	0.1048	α_0	1.1297	0.1219	0.1108	0.1169	0.1068	0.1143
	(0.0159)*	(0.0178)*		(0.0311)*	(0.0258)*	(0.0428)*	(0.0186)*	(0.0189)*	(0.0185)*
α_1^+	0.0300	0.0213	α_1^+	0.0060	0.0115	0.0182	0.0221	0.0201	0.0306
	(0.0147)*	(0.0126)		(0.0088)	(0.0103)	(0.0234)	(0.0126)	(0.0123)	(0.0153)*
α_1^-	0.1744	0.2237	α_1^-	0.2137	0.2255	0.1754	0.1977	0.2089	0.1681
	(0.0263)*	(0.0229)*		(0.0270)*	(0.0245)*	(0.0623)*	(0.0247)*	(0.0245)*	(0.0295)*
α_2	0.8689	0.8723	α_2	0.8756	0.8697	0.8587	0.8572	0.8673	0.8596
	(0.0176)*	(0.0173)*		(0.0161)*	(0.0182)*	(0.0198)*	(0.0206)*	(0.0198)*	(0.0201)*
α_3^+	−0.0130	−0.0056	α_4^+	−0.0276	−0.0243				−0.0220
	(0.1302)**	(0.0029)**		(0.0146)**	(0.0147)**				(0.0083)**
α_3^-	0.0791	−0.0091	α_4^-	0.0743	0.0715				0.0625
	(0.0283)*	(0.0559)		(0.0268)*	(0.1794)				(0.0271)*
			α_5^+			−0.007	−0.0014	−0.0009	0.001
						(0.0026)	(0.0012)	(0.0008)	(0.0013)
			α_5^-			0.0384	0.0227	0.0050	0.0267
						(0.0139)*	(0.0105)*	(0.0108)	(0.0142)**
$\mathcal{L}$	−3164.30	−3167.81	$\mathcal{L}$	−3168.91	−3167.76	−3166.42	−3166.52	−3168.02	−3164.28
J–B	710.77*	789.81*	J–B	750.59*	889.05*	763.41*	727.19*	789.62*	728.39*
B–P(2)	2.86	2.69	B–P(2)	2.77	2.72	2.98	2.88	2.91	2.92
L–B–W(10)	16.44	16.37	L–B–W(10)	17.12	16.67	16.52	16.39	16.67	16.3

Estimates correspond to the model

$$\sigma = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{y_{t-1}>\gamma\}} + \alpha_1^- \mathbb{I}_{\{y_{t-1}>\gamma\}})|y_{t-1} - \gamma| + \alpha_2\sigma_{t-1}$$
$$+ \alpha_3^+ \mathbb{I}_{\{y_{t-2}>\gamma,\dots,y_{t-a-1}>\gamma\}} + \alpha_3^- \mathbb{I}_{\{y_{t-2}\leqslant\gamma,\dots,y_{t-a-1}\leqslant\gamma\}}$$
$$+ (\alpha_4^+ \mathbb{I}_{\{y_{t-2}>\gamma,\dots,y_{t-a-1}>\gamma\}} + \alpha_4^- \mathbb{I}_{\{y_{t-2}\leqslant\gamma,\dots,y_{t-a-1}\leqslant\gamma\}})|y_{t-2} + \cdots + y_{t-a-1} - a\gamma|$$
$$+ (\alpha_5^+ \mathbb{I}_{\{y_{t-2}+\cdots+y_{t-b-1}>b\gamma\}} + \alpha_5^- \mathbb{I}_{\{y_{t-2}+\cdots+y_{t-b-1}\leqslant b\gamma\}})|y_{t-2} + \cdots + y_{t-b-1} - b\gamma|.$$

When the regression incorporates not only an indicator variable but also the sum of over or under performance, estimates remain similar (the estimates correspond to coefficients α_4^+, α_4^-). Only very short time horizons matter for negative returns.

All this is consistent with the view that bad news are rapidly incorporated in volatility revisions but good news tend to lower volatility over a longer time horizon provided that they are clustered.

Estimations where only the trend effect is allowed $(a = 0, b > 0)$ show that a negative trend in the index tends to increase volatility, while a positive trend has no impact. We notice that for horizons of 5 and 10 days, corresponding to one or two trading weeks, a positive trend does not affect volatility. A negative trend, however, tends to increase volatility. For an horizon of 3 trading weeks this phenomenon ceases to exist.

We chose to regroup the models where $a = 2, b = 10$ to verify that both features are orthogonal. Indeed, when both features are combined as in the last columns of Table IX, they appear with similar coefficients than when considered separately. Of course, we have a model with many parameters and the likelihood ratio test would allow us to accept the restricted models.

Thus the impact on volatility of a shock today will not depend only on its sign and magnitude, but also on the way past returns are distributed. Returns seem to adjust very quickly to negative events with a strong increase in volatility. This finding seems to corroborate the economic and psychological rationales to stock market overreaction. The short term pattern of returns has no further influence on volatility. However, we find that a negative event has a higher impact on volatility when it comes in the context of a negative two-week trend. Positive shocks have almost no effect on volatility, except if they come in a series where, in such instances, volatility decreases.

Negative events may well induce traders to quickly cut their losses, thereby creating market pressures that depress prices further down, which in turn trigger additional margin calls which force traders to sell more securities in order to generate the required liquidities. During this process prices may reach levels below what is justified by the information itself. In addition to this economic rational, there also are psychological reasons why markets overreact. Losses may indeed affect investment decisions to a greater extent than an equivalent amount of gains, because risk aversion increases or investors become irrational.[18]

5. International Evidence

Having found that the ATGARCH model is convenient to estimate, we applied this model to 21 stock market indices extracted from the FT-Goldman Sachs data base. We use data covering the period between January 1986 and November 1995.The French index in the FT-Goldman Sachs data base is different from the French official indices CAC40 and SBF240, it includes approximately 100 stocks.

[18] See, e.g., Kahneman and Tversky (1979) and Arrow (1982).

Table X contains sample statistics for the compound returns. The number of observations varies across countries due to non-trading days. For emerging markets, such as Hong-Kong, Japan, Malaysia, Singapore and Mexico we typically find a higher volatility than for other markets.

Many countries appear to have a high first order autocorrelation of returns. Also the distribution is highly nonnormal. All markets, except Australia, seem to have heteroskedastic stock returns. The Ljung–Box–White statistic for joint non-autocorrelation among the first ten autocorrelations is significant for many countries even though this measure has been corrected for heteroskedasticity. For this reason we decided to AR(1) filter all those markets where this statistic turned out to be significant.

In Table XI we present the results of the ATGARCH estimation. The value for α_2 in the range of 0.6 to 0.92 indicates that volatility tends to persist not only for France but for all countries in the database. Countries with lower α_2 are mostly located in the Asian–Australian hemisphere. Those countries economies emerged only recently and since their stock markets are rather small, one can expect that any change in news from those countries tends to imply rapid changes in volatility. Volatility is going to be dominated by current news rather than by the level of past volatility.

The coefficients of asymmetry α_1^+ and α_1^- also exhibit the same pattern as for France. Good news have less impact on changes in volatility than bad news. This phenomenon is exacerbated for smaller or younger stock markets. For instance Australia, Canada, Hong Kong, Malaysia and Singapore show a strong increase in volatility after bad news.

That negative news do not necessarily mean bad news is indicated by the uniformly negative γ ranging between -0.2 and -1.04.[20] At a world-wide level, only news below a certain level, lead to a strong increase in future volatility. There is no apparent correlation between a country's stock market age and depth on one side, and the magnitude of the critical level (γ) on the other.

After the ATGARCH(1, 1) effect has been accounted for, residuals don't show any trace of heteroskedasticity, but still reveal strong nonlinearities. However, the hysteresis effects, as they were modelled for the French market, are, most of the time, insignificant for the countries besides France. This shows that information is processed in the same way in most markets, except for some complex patterns. It is beyond the scope of this paper to relate the psychology of each market to the corresponding *ad hoc* hysteresis model.

6. Conclusion

This paper shows that heteroskedasticity in stock returns as already documented for the U.S. stock market, also prevails internationally. In an exhaustive study we nest the most popular models already proposed in the literature and test for the

[20] Italy and the United Kingdom are the exception with a statistically insignificant γ.

Table X. Sample statistics of compound returns

Country	Nobs.	Mean	Std	$\rho(1)$	$\rho(2)$	J–B	B–P(2)	L–B–W(10)
					Europe			
Belgium	2476	0.034	0.831	0.2080	−0.0050	5.36E + 4*	241.88*	57.0*
Denmark	2513	0.027	0.841	0.1950	0.0317	9.70E + 3*	77.86*	43.2*
France	2507	0.030	1.124	0.0744	0.0239	9.02E + 3*	242.84*	16.9
Germany	2510	0.013	1.197	0.0369	0.0460	2.07E + 4*	148.16*	9.62
Ireland	2511	0.046	1.231	0.1569	0.0112	4.65E + 4*	88.12*	18.9*
Italy	2517	0.018	1.385	0.1581	−0.0100	1.44E + 3*	184.56*	35.9*
Netherland	2543	0.030	0.974	−0.0450	−0.0310	6.04E + 4*	704.99*	6.78
Norway	2529	0.023	1.478	0.1012	0.0280	1.00E + 5*	302.63*	16.9
Spain	2493	0.043	1.157	0.2035	0.0307	4.99E + 3*	236.56*	35.7*
Sweden	2524	0.057	1.232	0.1617	0.0050	5.70E + 3*	318.80*	21.6*
Switzerland	2508	0.024	1.086	0.0544	0.0000	5.35E + 4*	211.86*	7.01
United Kingdom	2536	0.039	0.919	0.1023	0.0309	5.44E + 4*	729.36*	15.0
					America			
Canada	2527	0.019	0.740	0.1359	0.0566	2.52E + 5*	1140.4*	11.1
Mexico	2475	0.238	2.058	0.2296	−0.0200	5.45E + 3*	381.09*	41.5*
USA	2539	0.042	0.991	0.0548	−0.0690	1.27E + 6*	85.61*	8.28
					Asia			
Hong Kong	2505	0.072	1.724	0.0955	0.0046	1.97E + 6*	15.46*	8.77
Japan	2477	0.018	1.307	0.1099	−0.0940	2.74E + 4*	155.87*	27.7*
Malaysia	2531	0.071	1.408	0.2005	0.0339	2.79E + 4*	360.30*	21.4*
Singapore	2524	0.063	1.450	0.1394	0.0570	6.89E + 5*	657.52*	5.31
					Australia			
Australia	2543	0.035	1.143	0.1111	0.0740	3.37E + 6*	5.36	24.1*
New Zealand	2519	0.009	1.274	0.0715	0.0261	2.77E + 4*	74.68*	13.1

Elementary statistics for the compound returns of stock indices in local currency. The data has been provided by Goldman Sachs. The number of observations for each country is nobs. All other statistics are as in Table I.

Table XI. ATGARCH(1, 1) model for a set of countries

Country	γ	α_0	α_1^+	α_1^-	α_2	$\mathcal{L}$
Belgium	−0.5179	0.0614	0.0794	0.3769	0.8000	−2591.42
	(0.1275)*	(0.0351)	(0.0184)*	(0.1651)*	(0.0709)*	
Denmark	−0.5789	0.2095	0.0924	0.4112	0.6077	−2940.10
	(0.1183)*	(0.0929*)	(0.0346)*	(0.1447)*	(0.1425)*	
France	−0.6007	0.1041	0.0298	0.2351	0.8372	−3582.2
	(0.1732)*	(0.0309)*	(0.0140)*	(0.0548)*	(0.0360)*	
Germany	−0.5355	0.1089	0.0159	0.2243	0.8482	−3727.38
	(0.0004)*	(0.0577)	(0.0157)	(0.0759)*	(0.0552)*	
Ireland	−0.0406	0.1079	0.0507	0.3679	0.8160	−3704.94
	(0.1875)*	(0.0498)*	(0.0155)*	(0.1286)*	(0.0508)*	
Italy	0.2177	0.0133	0.0605	0.0890	0.9288	−4099.27
	(0.1845)	(0.0062)*	(0.0186)*	(0.0167)*	(0.0149)*	
Netherland	−0.2291	0.0409	0.0331	0.1704	0.8889	−3036.06
	(0.1156)**	(0.0169)*	(0.0155)*	(0.0538)*	(0.0307)*	
Norway	−0.7657	0.1313	0.0491	0.2980	0.8177	−4174.35
	(0.2025)*	(0.0794)	(0.0238)*	(0.1271)*	(0.0862)*	
Spain	−0.6518	0.0624	0.0530	0.2429	0.8615	−3549.05
	(0.1977)*	(0.0313)	(0.0157)*	(0.0731)*	(0.0371)*	
Sweden	−0.7440	0.0965	0.0178	0.2491	0.8623	−3753.54
	(0.1982)*	(0.0357)*	(0.0222)	(0.0598)*	(0.0441)*	
Switzerland	−0.5760	0.1377	0.0000	0.3458	0.8016	−3342.53
	(0.2191)*	(0.0536)*	(0.0317)	(0.1339)*	(0.0649)*	
United Kingdom	−0.0496	0.0633	0.0885	0.1762	0.8262	−3106.75
	(0.1060)	(0.0369)	(0.0332)*	(0.0702)*	(0.0758)*	
			America			
Canada	−0.5163	0.0775	0.0439	0.3043	0.7977	−2312.20
	(0.0020)*	(0.0383)*	(0.0149)*	(0.1174)*	(0.0818)*	
Mexico	−0.3989	0.1186	0.1228	0.2363	0.8081	−4834.30
	(0.1447)*	(0.0337)*	(0.0264)*	(0.0417)*	(0.0368)*	
USA	−0.2913	0.0336	0.0325	0.1755	0.9028	−3032.21
	(0.1081)*	(0.0301)	(0.0137)*	(0.1009)	(0.0583)*	
			Asia			
Hong Kong	−0.7390	0.1749	0.0486	0.4386	0.7712	−4249.45
	(0.1548)*	(0.0697)*	(0.0247)*	(0.1236)*	(0.0618)*	
Japan	−0.2025	0.0853	0.0845	0.3337	0.7938	−3730.84
	(0.1148)**	(0.0532)	(0.0242)*	(0.1431)*	(0.0791)*	
Malaysia	−0.9399	0.1955	0.0579	0.4655	0.7340	−4017.22
	(0.1804)*	(0.0747)*	(0.0178)*	(0.1430)*	(0.0638)*	
Singapore	−0.8107	0.3755	0.0931	0.6538	0.5243	−3925.77
	(0.1443)*	(0.1069)*	(0.0379)*	(0.2193)*	(0.1106)*	
			Australia			
Australia	−0.6045	0.2550	0.0813	0.5950	0.5828	−3381.04
	(0.2180)*	(0.2054)	(0.0354)*	(0.4562)	(0.2782)*	
New Zealand	−0.3622	0.1378	0.1084	0.2875	0.7581	−3947.36
	(0.1208)*	(0.0638)*	(0.0265)*	(0.1100)*	(0.0796)*	

statistically most relevant specification for the stock markets of 21 countries. On purely statistical grounds the model which allows for asymmetry and a shift of the NIC (the ATGARCH$(1,1)$ model) appears to be an amazingly robust description of volatility for all these countries. The asymmetry in the impact of positive and negative returns on volatility is strong. Negative returns below a certain level strongly increase volatility, while returns above this level, and of course positive returns, have a negligible effect on volatility. These empirical results comfort the economic and psychological rationales for market overreaction.

After the ATGARCH$(1,1)$ effect has been accounted for, residuals still reveal strong nonlinearities and patterns in more remote information still have predictive power. An attempt to capture the remaining structure in volatility within the ATGARCH(p,q) framework fails. Extensions including volatility in the conditional mean were not fruitful, either. However, a trend in the stock index and patterns of recent returns have some descriptive power, at least for the French market which we described, as the best candidate for exhibiting volatility clustering, asymmetry and hysteresis features. For France we propose an extension to the simple ATGARCH$(1,1)$ where we allow for hysteresis. A shock of either sign may affect volatility differently depending on the recent past being characterized by either all positive or all negative returns (the *isosigned returns* effect). In the same way, a longer term trend of either sign may also influence the impact on volatility of current innovations. This is some sort of size, or threshold effect in a dynamic context. The impact of a shock on volatility will depend on the cumulative size of past innovations. If it goes beyond a certain threshold level, then volatility will react more strongly.

The estimation results lead to the conclusion that bad news are discounted very quickly in volatility, while this effect is reinforced when it comes after a negative trend in the stock index. On the contrary, good news have a very small impact on volatility except when they are clustered over a few days, which in this case substantially reduces volatility.

This research can be extended by looking at the psychology of various markets and by attempting to capture this through specific *ad hoc* patterns.

Appendix

PROPOSITION. The encompassing model (3) contains as special cases, when $p = q = 1$

(i) GARCH$(1,1)$ for $\beta_1 = 2, \alpha_1^+ = \alpha_1^-, \beta_2 = 0, \gamma = 0, \beta_3 = 2$.

(ii) AGARCH$(1,1)$ for $\beta_1 = 2, \alpha_1^+ = \alpha_1^-, \beta_2 = 0, \beta_3 = 2$.

(iii) AVGARCH$(1,1)$ for $\beta_1 = 2, \alpha_1^+ = a_1^-, \beta_2 = 1, \beta_3 = 2$.

(iv) TGARCH$(1,1)$ for $\beta_1 = 1, \beta_2 = 0, \gamma = 0, \beta_3 = 1$.

(v) EGARCH$(1,1)$ for $\beta_1 = 0, \beta_2 = 1, \gamma = 0, \beta_3 = 1$.

Proof. First set in (3) $\alpha_{1i}^{+} = \alpha_{i1}^{-} = 0$ for $i > 1$, and $\alpha_{2j} = 0$ for $j > 1$.

(i) Setting $\beta_1 = 2, \alpha_1^+ = \alpha_1^- = \alpha_1, \beta_2 = 0, \gamma = 0, \beta_3 = 2$ in (3) we obtain

$$\tfrac{1}{2}\sigma_t^2 - \tfrac{1}{2} = \alpha_0 + \alpha_1 y_{t-1}^2 + \alpha_2(\tfrac{1}{2}\sigma_{t-1}^2 - \tfrac{1}{2})$$

and therefore $\sigma_t^2 = 2\alpha_0 + 1 - \alpha_2 + 2\alpha_1 y_{t-1}^2 + \alpha_2\sigma_{t-1}^2$ which is indeed the expression for a GARCH(1, 1) process as defined in (5) with an obvious change in notation.

(ii) Setting $\beta_1 = 2, \alpha_1^+ = \alpha_1^- = \alpha_1, \beta_2 = 0, \beta_3 = 2$ in formula (3) gives

$$\tfrac{1}{2}\sigma_t^2 - \tfrac{1}{2} = \alpha_0 + \alpha_1(y_{t-1} - \gamma)^2 + \alpha_2(\tfrac{1}{2}\sigma_{t-1}^2 - \tfrac{1}{2})$$

and therefore $\sigma_t^2 = 2\alpha_0 + 1 - \alpha_2 + 2\alpha_1(y_{t-1} - \gamma)^2 + \alpha_2\sigma_{t-1}^2$ which is the expression for an AGARCH(1, 1) process as defined in (7).

(iii) Setting $\beta_1 = 2, \alpha_1^+ = \alpha_1^- = \alpha_1, \beta_2 = 1, \beta_3 = 2$ in (3) yields

$$\tfrac{1}{2}\sigma_t^2 - \tfrac{1}{2} = \alpha_0 + \alpha_1\left(\frac{y_{t-1}}{\sigma_{t-1}} - \gamma\right)^2 + \alpha_2(\tfrac{1}{2}\sigma_{t-1}^2 - \tfrac{1}{2})$$

and therefore

$$\sigma_t^2 = 2\alpha_0 + 1 - \alpha_2 + 2\alpha_1\left(\frac{y_{t-1}}{\sigma_{t-1}} - \gamma\right)^2 + \alpha_2\sigma_{t-1}^2,$$

which is the expression for an AVGARCH(1, 1) process as defined in (8).

(iv) Setting $\beta_1 = 1, \beta_2 = 0, \gamma = 0, \beta_3 = 1$ in (3) gives

$$\sigma_t - 1 = \alpha_0 + (\alpha_1^+\mathbb{I}_{\{y_{t-1}>0\}} + \alpha_1^-\mathbb{I}_{\{y_{t-1}\leqslant 0\}}|y_{t-1}| + 2\alpha_2(\sigma_{t-1} - 1)$$

and therefore $\sigma_t = 1 + \alpha_0 - \alpha_2 + (\alpha_1^+\mathbb{I}_{\{y_{t-1}>0\}} + \alpha_1^-\mathbb{I}_{\{y_{t-1}\leqslant 0\}}|y_{t-1}| + \alpha_2\sigma_{t-1}$ which is the expression for a TGARCH(1, 1) process as defined in (10).

(v) Notice that $\lim_{\beta_1\to 0}(\sigma^{\beta_1} - 1)/\beta_1 = \ln(\sigma)$ and therefore setting the parameters $\beta_1 = 0, \beta_2 = 1, \gamma = 0, \beta_3 = 1$ we obtain

$$\ln(\sigma_t) = \alpha_0 + (\alpha_1^+\mathbb{I}_{\{y_{t-1}>0\}} + \alpha_1^-\mathbb{I}_{\{y_{t-1}\leqslant 0\}})\left|\frac{y_{t-1}}{\sigma_{t-1}}\right| + \alpha_2\ln(\sigma_{t-1}). \tag{14}$$

Since an EGARCH is given by

$$\begin{aligned}\ln(\sigma_t) &= \tilde{\alpha}_0 + a\frac{y_{t-1}}{\sigma_{t-1}} + b\left[\frac{|y_{t-1}|}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}}\right] + \beta\ln(\sigma_{t-1}) \\ &= \tilde{\alpha}_0 - b\sqrt{\frac{2}{\pi}} + (a+b)\frac{y_{t-1}}{\sigma_{t-1}}\mathbb{I}_{\{y_{t-1}>0\}} \\ &\quad +(a-b)\frac{y_{t-1}}{\sigma_{t-1}}\mathbb{I}_{\{y_{t-1}<0\}} + \beta\ln(\sigma_{t-1}),\end{aligned} \tag{15}$$

by identifying terms in (14) and (15) one also obtains that an EGARCH process an be recovered from the general model. □

Appendix

Here we present restrictions on the parameters of empirically relevant models so that the conditional variance remains well defined. Empirically we notice that a model where $p = q = 1$ in model (3) is most relevant, for this reason we will now focus on this restriction.

For the case where $\beta_1 \neq 0$ we can rewrite our general model, with an obvious change of notation, as

$$\sigma_{t+1}^{\beta_1} = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{z_t>0\}} + \alpha_1^- \mathbb{I}_{\{z_t\leqslant 0\}})|z_t|^{\beta_3} + \alpha_2 \sigma_t^{\beta_1}, \tag{16}$$

where we have $z_t \equiv y_t \sigma_t^{\beta_2} - \gamma$.

When $\beta_1 = 0$ our model becomes

$$\ln(\sigma_{t+1}) = \alpha_0 + (\alpha_1^+ \mathbb{I}_{\{z_t>0\}} + \alpha_1^- \mathbb{I}_{\{z_t\leqslant 0\}}|z_t|^{\beta_3} + \alpha_2 \ln(\sigma_t). \tag{17}$$

Positivity constraints

We always assume $\beta_1, \beta_2, \beta_3$ greater than 0. Then, if $\alpha_0, \alpha_1^+, \alpha_1^-$ and α_2 are all positive for (16) we are assured that our model is well defined.

For model (17), since the ln function has the entire real line as range there is no need to impose positivity restrictions on the parameters.

At some point we estimate the ATGARCH model with 2 lags

$$\begin{aligned}\sigma_{t+2} \;=\; & \alpha_0 + (\alpha_{11}^+ \mathbb{I}_{\{z_{t+1}>0\}} + \alpha_{11}^- \mathbb{I}_{\{z_{t+1}\leqslant 0\}})|z_{t+1}|^{\beta_3} \\ & + (\alpha_{12}^+ \mathbb{I}_{\{z_t>0\}} + \alpha_{12}^- \mathbb{I}_{\{z_t\leqslant 0\}})|z_t|^{\beta_3} + \alpha_{21}\sigma_{t+1} + \alpha_{22}\sigma_t.\end{aligned}$$

With an obvious analogue when $\beta_1 = 0$.

We could follow Nelson and Cao (1992) who give less restrictive assumptions than Bollerslev (1986) for positivity in a model with lags and search for general conditions guaranteeing positivity. Empirically, we found that all lagged models could be rejected using a likelihood test with or without the strict positivity constraints of Bollerslev. For this reason, we decided not to investigate the general case, but to report in Table VI only the result for the model where $\alpha_0, \alpha_{1i}^+, \alpha_{1i}^-$ and α_{2j} are all positive for $i = 1, 2$ and $j = 1, 2$.

For the Hysteresis Garch model of Equation (13), positivity constraints would require $\alpha_3^+, \alpha_3^-, \alpha_4^+, \alpha_4^-$ and α_5^+, α_5^- to be positive. We decided, however, to estimate our model without restrictions on the parameters to see, if at least in sample, we can find additional elements explaining residuals.

Stationarity conditions for the general model

Define $w_t = (\alpha_1^+ \mathbb{I}_{\{z_t>0\}} + \alpha_1^- \mathbb{I}_{\{z_t>0\}})|z_t|^{\beta_3}$ so that (16) can be rewritten as $\sigma_{t+1}^{\beta_1} = \alpha_0 + w_t + \alpha_2 \sigma_t^{\beta_1}$. Forward iteration yields

$$\sigma_{t+s}^{\beta_1} = \alpha_0 \left(\frac{1-\alpha_2^s}{1-\alpha_2} \right) + w_{t+s-1} + \alpha_2 w_{t+s-2} + \cdots + \alpha_2^{s-1} w_t + \alpha_2^s \sigma_t^{\beta_1}.$$

We will have stationarity if σ_{t+s} remains finite as s goes to infinity. Since further realizations of w_t are unknown we can replace them by their expectations. Consider now any future w_t. We have

$$\mathrm{E}[w_t] = \alpha_1^+ \mathrm{E}[z_t^{\beta_3}|z_t > 0]\Pr[z_t > 0] + \alpha_1^- \mathrm{E}[z_t^{\beta_3}|z_t \leqslant 0]\Pr[z_t \leqslant 0]. \tag{18}$$

Since $z_t = \sigma_t^{1-\beta_2}\varepsilon_t - \gamma$, we obtain under the assumption of normality that

$$\Pr[z_t > 0] = \Pr\left[\varepsilon_t > \frac{\gamma}{\sigma_t^{1-\beta_2}}\right] = 1 - \Phi\left(\frac{\gamma}{\sigma_t^{1-\beta_2}}\right), \tag{19}$$

and

$$\Pr[z_t \leqslant 0] = \Phi\left(\frac{\gamma}{\sigma_t^{1-\beta_2}}\right), \tag{20}$$

where $\Phi(\,\cdot\,)$ represents the cumulative distribution function of a normal distribution.

We notice that

$$\mathrm{E}[z_t^{\beta_3}|z_t > 0] = \int_{u=\gamma/\sigma_t^{1-\beta_2}}^{+\infty} (\sigma_t^{1-\beta_2}u - \gamma)^{\beta_3} \frac{1}{\sqrt{2\pi}} \exp\{-\tfrac{1}{2}u^2\}\,\mathrm{d}u, \tag{21}$$

and

$$\mathrm{E}[z_t^{\beta_3}|z_t \leqslant 0] = \int_{-\infty}^{u=\gamma/\sigma_t^{1-\beta_2}} (\gamma - \sigma_t^{1-\beta_2}u)^{\beta_3} \frac{1}{\sqrt{2\pi}} \exp\{-\tfrac{1}{2}u^2\}\,\mathrm{d}u. \tag{22}$$

To obtain stationarity one should have parameters so that the fixed point σ defined by

$$\begin{aligned} \sigma^{\beta_1} \; = \; & \frac{\alpha_0}{1-\alpha_2} + \frac{1}{1-\alpha_2}\left(\alpha_1^+ \mathrm{E}[z^{\beta_3}|z>0]\left(1-\phi\left(\frac{\gamma}{\sigma^{1-\beta_2}}\right)\right)\right. \\ & \left. + \alpha_1^- \mathrm{E}[z^{\beta_3}|z \leqslant 0]\phi\left(\frac{\gamma}{\sigma^{1-\beta_2}}\right)\right) \end{aligned} \tag{23}$$

remains finite.

For the particular case where $\beta_1 = 2, \beta_2 = 0, \beta_3 = 2, \alpha_1^+ = \alpha_1^- = \alpha_1, \gamma = 0$ corresponding to a GARCH(1, 1) model we obtain the condition

$$\sigma^2 = \frac{\alpha_0}{1-\alpha_2} + \frac{\alpha_1}{1-\alpha_2}\sigma^2 \Rightarrow \sigma^2 = \frac{\alpha_0}{1-\alpha_1-\alpha_2}.$$

The last expression is well defined if $\alpha_1 + \alpha_2 < 1$ the well known condition for stationarity of the GARCH(1, 1).

Having shown that our condition is reasonable, we also notice that the constraint (23) could be imposed numerically since (19), (20), (21) and (22) can all be easily computed using quadrature methods.

Clearly, imposing those constraints is numerically extremely demanding. For practical purposes, one can convince oneself that the final estimates correspond to a stationary solution.

The case where $\beta_1 = 0$ can be treated in similar manner but by using (17).

For the Hysteresis GARCH model a very similar analysis can be performed. The main difference resides in the conditioning of expectations since then we have expressions such as

$$\begin{aligned}
&\alpha_3^+ \mathrm{E}[\mathbb{I}_{\{z_{t-1}>0,\ldots,z_{t-a-1}>0\}} | z_{t-1} > 0, \ldots, z_{t-a-1} > 0] \\
&\quad \cdot \Pr[z_{t-1} > 0, \ldots, z_{t-a-1} > 0] \\
&+\alpha_3^- \mathrm{E}[\mathbb{I}_{\{z_{t-1}\leqslant 0,\ldots,z_{t-a-1}\leqslant 0\}} | z_{t-1} \leqslant 0, \ldots, z_{t-a-1} \leqslant 0] \\
&\quad \cdot \Pr[z_{t-1} \leqslant 0, \ldots, z_{t-a-1} \leqslant 0].
\end{aligned}$$

Assuming normality and independence of the z_t such probabilities could again be easily evaluated through quadrature.

Acknowledgements

The data used in this work was extracted from Datastream which is kindly sponsored by the BNP. We would like to acknowledge helpful comments and suggestions from Christian Gouriéroux, Franz Palm, Esther Ruiz and an anonymous referee.

References

Akgiray, Vedat (1989), ‘Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts’, *Journal of Business*, **62**, 55–80.

Anderson, Ronald L. (1942), ‘Distribution of the Serial Correlation Coefficient’, *Annals of Mathematical Statistics*, **13**, 1–13.

Arrow, Kenneth (1982), ‘Risk Perception in Psychology and Economics’, Economic Inquiry, January, 1–8.

Baillie, Richard T. and Ramon P. DeGennaro (1990), ‘Stock Returns and Volatility’, *Journal of Financial and Quantitative Analysis*, **25**(2), 203–214.

Bera, Anil K. and Matthew L. Higgins (1993), ‘ARCH Models: Properties, Estimation and Testing’, *Journal of Economic Surveys*, **7**(4), 305–366.

Berndt, Ernst, Bronwyn Hall, Robert Hall and Jerry Hausman (1974), ‘Estimation and Inference in Nonlinear Structural Models’, *Annals of Economic and Social Measurement*, **3**, 653–665.
Black, Fischer (1976), ‘Studies of the Stock Market Volatility Changes, Proceedings of the American Statistical Association’, *Business and Economic Statistics Section*, 177–181.
Bollerslev, Tim (1986), ‘Generalized Autoregressive Conditional Heteroskedasticity’, *Journal of Econometrics*, **31**, 307–327.
Bollerslev, Tim and Jeffrey Woolridge (1991), ‘Quasi Maximum Likelihood Estimation and Inference in Dynamic Models with Time Varying Covariances’, *Econometric Review*, **11**, 143–172.
Bollerslev, Tim, Chou Ray Y. and Kenneth F. Kroner (1992), ‘ARCH Modeling in Finance: A Review of the Theory and Empirical Evidence’, *Journal of Econometrics*, **52**, 5–59.
Bollerslev, Tim, Robert, F. Engle and Daniel B. Nelson (1994), ARCH Models, *The Handbook of Econometrics*, Vol. 4, North Holland.
Breusch, T.S. and Pagan (1979), ‘A Simple Test for the Heteroskedasticity and Random Coefficient Variation’, *Econometrica*, **47**, 1287–1297.
Campbell, John Y. and Ludger Hentschel (1992), ‘No News is Good News: An Asymmetric Model of Changing Volatility in Stock Returns’, *Journal of Financial Economics*, **31**(3), 281–318.
Crouhy, Michel G. and Dan Galai (1992), ‘The settlement Day Effect in the Paris Bourse’, *Journal of Financial Services Research*, 417–435.
De Bondt, Werner F.M. and Richard Thaler (1985), ‘Does the Stock Market Overreact?’ *Journal of Finance*, **40**(3), 793–808.
Diebold, Francis X. (1988), *Empirical Modeling of Exchange Rate Dynamics*, Springer-Verlag.
Ding, Zhuanxin, Clive W.J. Granger and Robert F. Engle (1993), ‘A Long Memory Property of Stock Market Returns and a New Model’, *Journal of Empirical Finance*, **1**(1), 83–106.
Engle, Robert F. (1982), ‘Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of the United Kingdom Inflation’, *Econometrica*, **50**, 987–1008.
Engle, Robert F. (1990), ‘Discussion: Stock Market Volatility and the crash of ’87’, *Review of Financial Studies*, **3**, 103–106.
Engle, Robert F. and Victor K. Ng (1993), ‘Measuring and Testing the Impact of News on Volatility’, *Journal of Finance*, **48**(5), 1749–78.
Fama, Eugene F. (1965), ‘The Behavior of Stock Market Prices’, *Journal of Business*, **38**, 34–105.
French, Kenneth R. (1980), ‘Stock Returns and the Weekend Effect’, *Journal of Financial Economics*, **8**, 55–70.
French, Kenneth R., G. William Schwert and Robert F. Stambaugh (1987), ‘Expected Stock Returns and Volatility’, *Journal of Financial Economics*, **19** 3–29.
Gallant, A. Ronald and George E. Tauchen (1989), ‘Semi-nonparametric Estimation of Conditionally Constrained Heterogenous Processes: Asset Pricing Applications’, *Econometrica*, **57**, 1091–1120.
Gallant, A. Ronald, David A. Hsieh and George E. Tauchen (1991), ‘On Fitting a Recalcitrant Series: the Pound/Dollar Exchange Rate’, 1974–83, in: W.A. Barnett, J. Powell and G.E. Tauchen, eds., *Nonparametric and Semiparametric Methods in Econometrics and Statistics* Cambridge University Press, Cambridge, U.K.
Gibbons, M. and P. Hess (1981), ‘Day of the Week Effects and Asset Returns’, *Journal of Business*, **54**, 579–96.
Glosten, Lawrence R., Ravi Jagannathan and David Runkle (1993), ‘On the Relationship between the Expected Value and the Volatility of the Nominal Excess Return on Stocks’, *Journal of Finance*, **48**(5), 1779–1801.
Godfrey, M., Cliff Granger and Oscar Morgenstern (1964), ‘The Random Walk Hypothesis of Stock Market Behavior’, *Kyklos*, **17**, 1–30.
Gourieroux, Christian, Alain Monfort and Alain Trognon (1984), ‘Pseudo Maximum Likelihood Methods: Theory’, *Econometrica*, **52**(3), May 1984, 681–700.
Gourieroux, Christian and Alain Monfort (1992), ‘Qualitative Threshold ARCH Models’, *Journal of Econometrics*, **52**, 159–199.
Graybill, A.M., Mood, F.A. and Boes, D.C. (1982), *Introduction to the Theory of Statistics*, McGraw Hill.
Group of Thirty (1993), *Derivatives: Practices and Principles*, Washington, DC.

Hamao, Yasushi, Ronald W. Masulis and Victor Ng (1990), 'Correlations in Price Changes and Volatility Across International Stock Markets', *Review of Financial Studies*, **3**(2), 281–307.

Harris, Lawrence (1986), 'A Transaction Data Study of Weekly and Intradaily Patterns in Stock Returns', *Journal of Financial and Quantitative Analysis*, **24**, 29–45.

Harvey, Andrew C. and Neil Shephard (1993), The Econometrics of Stochastic Volatility', working paper, London School of Economics.

Hentschel, Ludger (1991), The Absolute Value GARCH Model and the Volatility of the U.S. Stock Returns, working paper. Princeton University, Department of Economics.

Hentschel, Ludger (1995), 'All in the Family: Nesting Symmetric and Asymmetric GARCH Models', *Journal of Financial Economics*, **39**, 71–104.

Higgins, Matthew L. and Anil K. Bera (1992), 'A Class of Nonlinear Arch Models', *International Economic Review*, **33**(1), 137–58.

de Jong, F., A. Kemna and T. Kloek (1989), 'The Impact of Option Expiration on Stock Returns on the Dutch Market', working paper, Econometric Institute, Erasmus University.

Kahneman, D. and A. Tversky (1979), 'Prospect Theory: An Analysis of Decision Under Risk', *Econometrica*, **47**, 263–291.

Lo, Andy and A.C. MacKinlay (1988), 'Stock Market Prices do not follow Random Walks: Evidence from a Simple Specification Test', *Review of Financial Studies*, **1**, 41–66.

Mandelbrot, Benoit (1963), 'The Variation of Certain Speculative Prices', *Journal of Business*, **36**, 394–419.

Nelson, Daniel B. (1991), 'Conditional Heteroskedasticity in Asset Returns: A New Approach', *Econometrica*, **59**(2), 347–370.

Nelson, Daniel B. and C.Q. Cao (1992), Inequality Constraints in the Univariate GARCH Model, *Journal of Business and Economic Statistics*, **10**, 229–235.

Pagan, Adrian R. and G. William Schwert (1990), 'Alternative Models for Conditional Stock Volatility', *Journal of Econometrics*, **45**, 267–290.

Poon, S.H. and Stephen J. Taylor (1991), 'Stock Returns and Volatility: An Empirical Study of the U.K. Market', forthcoming in the *Journal of Banking and Finance*.

Rogalski, R.J. (1984), 'New Findings Regarding Day of the Week Returns over Trading and Non-trading Periods', *Journal of Finance*, **39**, 1603–1614.

Ruiz, Esther (1993), Stochastic Volatility versus Autoregressive Conditional Heteroskedasticity, mimeo Universidad Carlos III de Madrid.

Sentana, Enrique (1991), Quadratic ARCH Models: a Potential Reinterpretation of ARCH Models as Second-order Taylor Approximations, working paper, London School of Economics.

Taylor, Stephen J. (1986), *Modelling Financial Time Series*, Wiley, New York.

White, Halbert (1980), 'A Heteroskedasticity Consistent Covariance Matrix Estimation and a Direct Test for Heteroskedasticity', *Econometrica*, **48**(4), 817–837.

White, Halbert (1982), 'Maximum Likelihood Estimation of Misspecified Models', *Econometrica*, **50**, 1–25.

Zakoïan, Jean-Michel (1991), Threshold Heteroskedastic Models, working paper, INSEE.