Modelling and predicting exchange rate volatility via power ARCH models: the role of long-memory.

C. CONRAD[a], F. JIANG[b] and M. KARANASOS[b]

[a] *University of Mannheim, Mannheim, Germany*

[b] *University of York, Heslington, York, UK*

First draft: December 2003

Abstract

Tse (1996, 1998) proposes a model which combines the fractionally integrated GARCH (FIGARCH) model of Baillie, Bollerslev and Mikkelsen (1996) with the asymmetric power ARCH (APARCH) model of Ding, Granger and Engle (1993). This paper analyzes the applicability of this fractionally integrated APARCH (FIAPARCH) model to exchange rate returns for ten countries. We find this GARCH model to be generally applicable once power, leverage and long-memory effects are taken into consideration. In addition, we find that both the optimal fractional differencing parameter and power transformation are remarkably similar across countries. Out-of-sample evidence for the superior forecasting ability of the FIAPARCH specification for exchange rate volatility is provided in terms of forecast error statistics and tests for equal forecast accuracy of nested models.

Keywords: Fractional integration, Asymmetric Power ARCH, Exchange rate volatility, Volatility forecast evaluation.

JEL Classification: C13, C22, C52.

We would like to thank M. Karanassou for her helpful comments and suggestions.

Address for correspondence: M. Karanasos, Department of Economics and Related Studies, University of York, York, Y010 5DD; email: mk16@york.ac.uk, tel: +44 (0)1904 433799, fax: +44 (0)1904 433759.

1 Introduction

A common finding in much of the empirical finance literature is that although the returns on speculative assets contain little serial correlation, the absolute returns and their power transformations are highly correlated (see, for example, Dacorogna et al. 1993, Granger and Ding, 1995a, 1995b and Breidt et al. 1998). In particular, Ding et al. (1993) investigate the autocorrelation structure of $|r_t|^\delta$, where r_t are the daily S&P 500 stock market returns, and δ is a positive number. They found that $|r_t|$ has significant positive autocorrelations for long lags. Ding and Granger (1996) analyse the autocorrelation structure of the Deutschmark-U.S. dollar exchange rate and find the strongest autocorrelation pattern for $\delta = \frac{1}{4}$. Motivated by this empirical result they propose a new general class of ARCH models, which they call the Asymmetric Power ARCH (APARCH) model. In addition, they show that the APARCH model comprises seven other models in the literature[1]. Bollerslev and Mikkelsen (1996) provide strong evidence that the conditional variance for the S&P 500 composite index is best modeled as a mean-reverting fractionally integrated process. Baillie et al. (1996) apply the fractionally integrated GARCH model to the Deutschmark-U.S. dollar exchange rate. They illustrate how this model captures the long-run dynamics of the series better than either the stable GARCH or the integrated GARCH (IGARCH) model. Tse (1998) analyzed the yen-dollar exchange rate using a fractionally integrated APARCH (FIAPARCH) type of model. This model is constructed by extending the APARCH model of Ding et al. (1993) to a process that is fractionally integrated, as defined by Baillie et al. (1996)[2].

The FIAPARCH model increases the flexibility of the conditional variance specification by allowing (a) an asymmetric response of volatility to positive and negative "shocks", (b) the data to determine the power of returns for which the predictable structure in the volatility pattern is the strongest, and (c) long-range volatility dependence. These three features in the volatility processes of exchange rate returns have major implications for many paradigms in modern financial economics. Optimal portfolio decisions, the pricing of long-term options and optimal portfolio allocations must take into account all of these three findings. Value-at-Risk is computed by utilizing volatility forecasts. Giot and Laurent (2003) have shown that APARCH volatility forecasts outperform those obtained from the RiskMetrics model which is equivalent to an integrated GARCH model with pre-specified autoregressive parameter value. Thus, the FIAPARCH model may lead to further improvement, if FIAPARCH forecasts are more accurate than

[1]These models are: the ARCH(p) model (see Engle, 1982), the GARCH(p,q) model (see Bollerslev, 1986), the Taylor/Schwert GARCH in standard deviation model (see Taylor, 1986, and Schwert, 1900), the GJR model (see Glosten et al., 1993), the TARCH model (see Zakoian, 1994), the NARCH model (see Higgins and Bera, 1992) and the log-ARCH model (see Geweke, 1986, and Pantula, 1986).

[2]The Fractionally integrated GARCH (FIGARCH) model is closely related to the long- memory GARCH model introduced by Robinson (1991) and the component GARCH model introduced by Ding and Granger (1996).

those obtained from the stable APARCH.

Another important advantage of having a long-memory asymmetric power ARCH model is that it nests the two major classes of ARCH models, namely the APARCH and FIGARCH models, as special cases of the general model. This provides an encompassing framework for these two broad classes of models and facilitates comparison between them. The main contribution of this paper is to enhance our understanding of whether and to what extent this type of model improves upon its simpler counterparts.

The evidence provided by Tse (1996, 1998) suggests that the fractionally integrated APARCH model is applicable to the yen-dollar exchange rate. An interesting research issue is to explore how generally applicable the FIAPARCH model is to a wide range of currencies. In this paper we attempt to address this issue in two steps. First, by estimating a FIAPARCH model for ten series of exchange rate returns we provide in-sample evidence. These countries are Canada, France, Germany, Hong Kong, Japan, Singapore, the United Kingdom and the United States. As the FIAPARCH specification adopted in this paper nests the FIGARCH and APARCH formulations, the relative ranking of each of these models can be considered using the standard Wald testing procedure. Furthermore, standard information criteria, such as the Hannan-Quinn and the Shibata ones, can be used to provide a ranking of the models. Second, we provide out-of-sample evidence, i.e. the ability of the FIAPARCH model to forecast exchange rate volatility is assessed by a variety of forecast error statistics. Moreover, a direct comparison of the forecast accuracy of nested models will be carried out by encompassing tests.

The remainder of the paper is structured as follows. In section 2 we detail the FIAPARCH model and discuss how various GARCH models are nested within it. Section 3 discusses the data and presents the empirical results. Quasi-maximum Likelihood (QML) parameter estimates for the FIAPARCH model are presented, as are the results of the Wald testing procedure. The robustness of these results is assessed using four alternative information criteria. To test for the apparent similarity of the power and fractional differencing terms across countries pairwise Wald tests are performed. Section 4 provides out-of-sample evidence. The different model specifications are evaluated in terms of their forecast ability. For each country and each model fourteen forecast error measures are calculated and evaluated against each other. Moreover, we test for equal forecast accuracy of the nested models by utilizing two encompassing test statistics. Section 5 concludes the analysis.

2 FIAPARCH Model

One of the most common models in finance and economics to describe a time series r_t of exchange rate returns is the white noise process

$$r_t = c + \varepsilon_t, \quad t \in \mathbb{N}, \tag{2.1}$$

with

$$\varepsilon_t = e_t \sqrt{h_t},$$

where $\{e_t\}$ are independent, identically distributed random variables with $E(e_t) = E(e_t^2 - 1) = 0$. h_t is positive with probability one and is a measurable function of Σ_{t-1}, which in turn is the sigma-algebra generated by $\{r_{t-1}, r_{t-2}, \ldots\}$. That is h_t denotes the conditional variance of the returns $\{r_t\}$ $(r_t|\Sigma_{t-1}) \sim IID(c + \rho r_{t-1}, h_t)$.

Tse (1998) examined the conditional heteroskedasticity of the yen-dollar exchange rate using a long-memory volatility specification. The paper by Tse extends the APARCH model to a fractionally integrated APARCH process to represent long-memory in volatility. Accordingly, we consider the FIAPARCH $(1, d, 1)$ model represented by the following conditional variance equation:

$$(1 - \phi L)(1 - L)^d f(\varepsilon_t) = \omega + (1 - \beta L)\xi_t, \tag{2.2}$$

with

$$\xi_t \equiv f(\varepsilon_t) - h_t^{\frac{\delta}{2}},$$

and

$$f(\varepsilon_t) \equiv [|\varepsilon_t| - \gamma\varepsilon_t]^\delta,$$

where γ $(-1 < \gamma < 1)$ is the leverage parameter, δ is the parameter for the power term, $|\phi| < 1$, $\omega > 0$, and $0 \leq d \leq 1$[3].

On rearranging the terms, (2.2) can be written as follows:

$$h_t^{\frac{\delta}{2}} = (1 - \beta)^{-1}\omega + \lambda(L) f(\varepsilon_t),$$

[3]The fractional differencing operator, $(1 - L)^d$ is most conveniently expressed in terms of the hypergeometric function

$$(1 - L)^d \equiv F(-d, 1; 1; L) = \sum_{j=0}^{\infty} \frac{\Gamma(j - d)}{\Gamma(-d)\Gamma(j + 1)} L^j = \sum_{j=0}^{\infty} \binom{d}{j} (-1)^j L^j,$$

where

$$F(a, b; c; z) \equiv \sum_{j=0}^{\infty} \frac{(a)_j (b)_j}{(c)_j} \frac{z^j}{j!}$$

is the Gaussian hypergeometric series, $(b)_j$ is the shifted factorial defined as $(b)_j \equiv \prod_{i=0}^{j-1}(b + i)$ (with $(b)_0 \equiv 1$), and $\Gamma(\cdot)$ is the gamma function.

where $\lambda(L)$ is defined as[4]

$$\lambda(L) \equiv \sum_{i=1}^{\infty} \lambda_i L^i \equiv \left[1 - (1-\beta L)^{-1}(1-\phi L)(1-L)^d\right]$$

When $d = 0$, the process in (2.2) reduces to the APARCH(1,1) model. Two major classes of ARCH models are nested in the APARCH specification. Specifically, a Taylor/ Schwert type of model[5] is specified when $\delta = 1$, and a Bollerslev type of model is specified when $\delta = 2$. There seems to be no obvious reason why one should assume that the conditional standard deviation is a linear function of lagged absolute returns or the conditional variance a linear function of lagged squared returns. As Brooks et al. (2000) point out "The common use of a squared term in this role ($\delta = 2$) is most likely to be a reflection of the normality assumption traditionally invoked regarding financial data. However, if we accept that (high frequency) data are very likely to have a non-normal error distribution, then the superiority of a squared term is lost and other power transformations may be more appropriate. Indeed, for non-normal data, by squaring the returns one effectively imposes a structure on the data which may potentially furnish sub-optimal modelling and forecasting performance relative to other power term".

Since its introduction by Ding et al. (1993), the APARCH model has been frequently applied. It is also worth noting that Fornari and Mele (1997) showed the usefulness of the APARCH scheme in approximating models developed in continuous time as systems of stochastic differential equations. This feature of GARCH schemes has usually been overshadowed by their well-known role as simple econometric tools providing reliable estimates of unobserved conditional variances (Fornari and Mele, 2001).

When $\gamma = 0$ and $\delta = 2$ the process in (2.2) reduces to the FIGARCH$(1, d, 1)$ model[6]. The FIGARCH specification includes Bollerslev's (1986) GARCH (when $d = 0$) and the Integrated GARCH (IGARCH) model (when $d = 1$) as special cases. Baillie et al. (1996) mention that a striking empirical regularity that emerges from numerous studies of high-frequency, say daily, asset pricing data with ARCH-type models,

[4]The coefficients λ_i in the lag polynomial $\lambda(L)$ can be calculated using the following recursions: $\lambda_1 = \phi - \beta + d$, and $\lambda_i = \beta\lambda_{i-1} + [(i-1-d)i^{-1} - \phi]\zeta_{i-1}$ for $i = 2, \ldots, \infty$, where $\zeta_i = \zeta_{i-1}(i-1-d)i^{-1}$, with $\zeta_1 = d$ (see Bollerslev and Mikkelsen, 1996). From these recursions, it follows that the inequality constraints:

$$\begin{aligned} \beta - d &\leq \phi \leq (2-d)(0.333), \\ d[\phi - (1-d)(0.5)] &\leq \beta(\phi - \beta + d) \end{aligned}$$

are sufficient to ensure that the parameters in the infinite ARCH representation are all nonnegative (see Bollerslev and Mikkelsen, 1996).

[5]Taylor (1986) and Schwert (1990) have suggested that the conditional standard deviation obeys a GARCH specification.

[6]When $d \in (0, 1)$, the cumulative impulse response coefficients for the optimal linear forecast of the future conditional variance are eventually dominated by a hyperbolic rate of decay rather than the exponential rate that is characteristic of covariance-stationary GARCH processes. Furthermore, the FIGARCH model has a strictly stationary and ergodic solution, which is not, however, square integrable. Whether this result holds for the FIAPARCH model is an open question (Tse, 1998).

concerns the apparent widespread finding of integrated GARCH behavior. This property has been found in stock returns, exchange rates, commodity prices and interest rates (see Bollerslev et al. 1992). Yet unlike I(1) processes for the mean, there is less theoretical motivation for truly integrated behavior in the conditional variance (see Baillie et al. 1996 and the references therein)[7].

Finally, as noted by Baillie et al. (1996) for the variance, being confined to only considering the extreme cases of stable GARCH and IGARCH processes can be very misleading when long-memory (but eventually mean-reverting) processes are generating the observed data. They showed that data generated from a process exhibiting long-memory FIGARCH volatility may be easily mistaken for IGARCH behavior.

3 Data and basic estimation results

Daily exchange rate data for ten currencies against the US dollar were sourced from the Datastream database for the period 1st January 1990 to 18th November 2003, giving a total of 3,621 observations. The ten currencies are: the British pound (BRP), the German mark (GEM), the French franc (FRF), the Italian lira (ITL), the Swedish krona (SWK), the Spanish peseta (SPP), the Japanese yen (JAY), the Singapore dollar (SID), the Canadian dollar (CAD) and the Australian dollar (AUD). For each national currency, the continuously compounded return was estimated as $r_t = 100 \cdot [\log(p_t) - \log(p_{t-1})]$ where p_t is the price on day t. The FIAPARCH models are estimated for period 1st January 1990 to June 2002, i.e. we have 3256 in sample observations, while the period from July 2002 to November 2003 is used for out-of-sample forecasting providing 365 daily observations.

We proceed with the estimation of the FI(A)PARCH$(1, d, 1)$ model in equations (2.1) and (2.2) in order to take into account the serial correlation and the GARCH effects observed in our time series data, and to capture the possible long-memory in volatility. We estimate the FI(A)PARCH models using the quasi-maximum likelihood estimation (QMLE) method as implemented by Laurent and Peters (2002) in Ox. To obtain robust inference about the estimated models, we compute the robust standard errors as suggested by Bollerslev and Wooldridge (1992). Table 1 reports the results for the period 1st January 1990 to June

[7] In particular, the occurrence of a shock to the IGARCH volatility process will persist for an infinite prediction horizon. This extreme behavior of the IGARCH process may reduce its attractiveness for asset pricing purposes, where the IGARCH assumption could make the pricing functions for long-term contracts very sensitive to the initial conditions. This seems contrary to the perceived behavior of agents who typically do not frequently and radically change their portfolio compositions. In addition, the IGARCH model is not compatible with the persistence observed after large shocks such as the Crash of October 1987. A further reason to doubt the empirical reasonableness of IGARCH models relates to the issue of temporal aggregation. A data generating process of IGARCH at high frequencies would also imply a properly defined weak IGARCH model at low frequencies of observation. However, this theoretical result seems at odds with reported empirical findings for most asset categories (abstracted from Baillie et al. 1996).

2002. The $\hat{\phi}$ parameter is significant for all countries but Spain, while $\hat{\beta}$ is significant for all countries. The estimates for leverage term $(\widehat{\gamma})$ are statistically significant for the British pound, the Swedish krona, the Japanese yen, the Singapore dollar and the Canadian dollar, confirming the hypothesis that there is negative correlation between returns and volatility for these countries. For the other countries we estimated an insignificant leverage term and therefore reestimated the model excluding the γ parameter. In all countries the estimates for the power term $(\widehat{\delta})$ and the fractional differencing parameter $(\widehat{d})$ are highly statistically significant. In all cases, the ARCH parameters satisfy the set of conditions sufficient to guarantee the nonnegativity of the conditional variance (see footnote 3). According to the values of the Ljung-Box tests for serial correlation in the standardized and squared standardized residuals there is no statistically significant evidence of misspecification.

Table 1: FI(A)PARCH$(1, d, 1)$ models (QML Estimation).

	BRP	GEM	FRF	ITL	SWK	SPP	JAY	SID	CAD	AUD
$\hat{\omega}$	0.02	0.02	0.02	0.01	0.04	0.03	0.02	0.13	0.15	0.03
	(0.18)	(1.86)	(2.01)	(1.00)	(1.76)	(1.16)	(1.05)	(1.33)	(1.42)	(1.72)
$\hat{\phi}$	0.41	0.40	0.46	0.32	0.34	0.22	0.73	0.61	0.33	0.43
	(4.72)	(6.36)	(6.26)	(5.20)	(2.87)	(1.13)	(4.92)	(4.57)	(3.44)	(4.81)
$\hat{\gamma}$	-0.17	-	-	-	-0.22	-	0.48	-0.14	-0.12	-
	(-1.87)	-	-	-	(-2.10)	-	(2.47)	(-2.05)	(-1.54)	-
$\hat{\delta}$	1.92	1.78	1.81	1.97	1.73	2.08	1.76	1.89	1.91	1.94
	(10.00)	(5.46)	(7.97)	(10.74)	(5.90)	(8.70)	(5.85)	(12.00)	(7.10)	(7.96)
$\hat{\beta}$	0.64	0.732	0.72	0.71	0.53	0.46	0.80	0.80	0.78	0.63
	(7.78)	(8.05)	(8.94)	(8.25)	(3.91)	(1.91)	(7.88)	(7.60)	(6.92)	(5.41)
$\hat{d}$	0.30	0.39	0.35	0.45	0.30	0.27	0.19	0.45	0.51	0.27
	(3.63)	(3.02)	(3.63)	(4.06)	(5.05)	(3.54)	(2.45)	(3.90)	(2.49)	(3.50)
Q_{20}	14.36	14.65	16.08	27.26	15.02	21.16	26.83	32.60	18.46	22.21
	[0.81]	[0.80]	[0.71]	[0.13]	[0.78]	[0.39]	[0.14]	[0.04]	[0.56]	[0.33]
Q_{20}^2	12.13	24.57	13.69	28.67	19.32	15.81	17.78	7.09	12.58	13.15
	[0.84]	[0.14]	[0.75]	[0.05]	[0.37]	[0.61]	[0.47]	[0.99]	[0.82]	[0.78]

Notes: For each of the ten currencies , table 1 reports QML parameter estimates for the FI(A)PARCH$(1, d, 1)$ model: $r_t = c + \varepsilon_t$, with $\varepsilon_t = e_t\sqrt{h_t}$, $e_t \overset{i.d}{\sim} t(0,1)$, and $(1-\phi L)(1-L)^d f(\varepsilon_t) = \omega + (1-\beta L)\xi_t$, where $f(\varepsilon_t) \equiv [|\varepsilon_t| - \gamma\varepsilon_t]^\delta$, and $\xi_t \equiv f(\varepsilon_t) - h_t^{\frac{\delta}{2}}$. The numbers in parentheses are t-statistics. Robust standard errors are reported in $\{\cdot\}$. Q_{20} and Q_{20}^2 are the 20th order Ljung-Box tests for serial correlation in the standardized and squared standardized residuals respectively.

The numbers in brackets are p-values.

3.1 Tests of fractional differencing and power term parameters.

A large number of studies have documented the persistence of volatility in stock returns; see, e.g., Ding et al. (1993), Ding and Granger (1996), Engle and Lee (2000). Using daily data many of these studies have concluded that the volatility process is very persistent and appears to be well approximated by an IGARCH process. Similarly, Baillie et al. (1996) estimate a GARCH model for the Deutschmark-U.S. dollar exchange rate and obtain a persistence parameter which is not significantly different from one ($\hat{\alpha} + \hat{\beta} = 0.98$ (0.01)).

Accordingly, the persistence parameter we estimated in the (A)PARCH(1,1) specification (not reported here) is very close to unity, suggestive of "integrated" (A)PARCH ("I(A)PARCH") behavior. However, from the FI(A)PARCH$(1, d, 1)$ model estimates (reported in table 1), it appears that the long-run dynamics are better modelled by the fractional differencing parameter. To test for the persistence of the conditional heteroskedasticity models, we examine the Wald (W) statistic for the linear constraints $d = 1$ ("I(A)PARCH" model) and $d = 0$ (stable (A)PARCH model). As seen in table 2 the W statistics clearly reject both the (A)PARCH and "I(A)PARCH" null hypotheses against the FI(A)PARCH model. Thus, purely from the perspective of searching for a model that best describes the volatility in the exchange rate series, the FI(A)PARCH model appears to be the most satisfactory representation.

Following the work of Ding et al. (1993), Hentschel (1995), Brooks et al. (2000) and Tse (1998) among others, the Wald statistic can be used for model selection. Alternatively, the Akaike, Schwarz, Hannan-Quinn or Shibata information criteria (AIC, SIC, HQIC, SHIC respectively) can be applied to rank the various GARCH models[8]. These model selection criteria check the robustness of the Wald testing results discussed above[9]. According to the AIC, SIC, HQIC and SHIC, the optimal GARCH type model (i.e., FIAPARCH, APARCH or "IAPARCH") for all currencies was the FIAPARCH one[10].

Moreover, recall that the two common values of the power term imposed throughout much of the GARCH literature are the values of two (Bollerslev's model) and unity (the Taylor/Schwert model). The invalid imposition of a particular value for the power term may lead to sub-optimal modeling and forecasting performance (Brooks et al. 2000). Accordingly, we test whether the estimated power terms are significantly different from unity or two using Wald tests. As reported in table 2, all the estimated power coefficients are significantly different from unity (see column three). In sharp contrast, each of the power terms are not significantly different from two (see column four). Hence, on the basis of these results, in all the cases support is found for the Bollerslev FI(A)ARCH model. In other words, according to the Wald testing results, allowing the power term to take on values other than two does not significantly enhance the models.

The evidence obtained from the Wald test statistic is reinforced by the model ranking provided by the AIC, SIC, HQIC and SHIC model selection criteria. In all cases the criteria favor the Bollerslev FI(A)ARCH model over both Taylor/Schwert's and power fractionally integrated (asymmetric) GARCH models. That is, the optimal GARCH type model (i.e. the specification that produced the lowest AIC, SIC, HQIC or SHIC) according to each criterion was the Bollerslev FI(A)ARCH one for all currencies[11].

[8] As a general rule, the AIC, SIC, HQIC or SHIC approaches suggest selecting the model which produces the lowest AIC, SIC, HQIC or SHIC values.

[9] The use of the AIC, SIC, HQIC and SHIC techniques for comparing models has the advantage of being relatively less onerous compared to the Wald testing procedures, which only allow formal pairwise testing of nested models (Brooks et al. 2000).

[10] We do not report the AIC, SIC, HQIC or SHIC values for space considerations.

[11] We do not report the AIC, SIC, HQIC or SHIC values for space considerations.

Table 2: Tests of fractional differencing and power term parameters.

	H_0:(A)PARCH (d=0)	H_0: "I(A)PARCH" (d=1)	H_0: FI(A)ARCH (δ=1)	H_0: FI(A)ARCH (δ=2)
BRP	13.16 [0.00]	70.31 [0.00]	22.86 [0.00]	0.20 [0.66]
GEM	9.00 [0.00]	22.43 [0.00]	5.70 [0.02]	0.47 [0.50]
FRF	13.10 [0.00]	45.12 [0.00]	12.73 [0.00]	0.69 [0.41]
ITL	16.44 [0.00]	24.04 [0.00]	27.97 [0.00]	0.25 [0.62]
SWK	25.51 [0.00]	143.83 [0.00]	6.19 [0.02]	0.85 [0.36]
SPP	12.50 [0.00]	95.14 [0.00]	16.50 [0.00]	0.14 [0.91]
JAY	5.98 [0.02]	114.37 [0.00]	6.39 [0.01]	0.64 [0.43]
SID	15.19 [0.00]	22.71 [0.00]	31.66 [0.00]	0.52 [0.47]
CAD	6.18 [0.02]	5.83 [0.02]	11.44 [0.00]	0.11 [0.74]
AUD	12.22 [0.00]	86.54 [0.00]	14.79 [0.00]	0.71 [0.79]

Notes: For each of the ten currencies, columns 1 and 2 report the value of the following Wald test: W=$[\hat{d}_u - d_r)^2/(\text{SE}_d)^2$, where $\hat{d}_u$ denotes the estimated value of the long-memory parameter for the unrestricted FI[A)PARCH$(1, d, 1)$ model, SE_d is the corresponding standard error and d_r is either 0 or 1. Columns 3 and 4 report the corresponding Wald statistics for the power term parameter. The numbers in brackets are p values.

Furthermore, we test the apparent similarity of the optimal fractional differencing parameters estimated for each of the ten currencies using a pairwise Wald test:

$$\text{W} = \frac{(\hat{d}_1 - \hat{d}_2)^2}{(\text{SE}_1)^2 + (\text{SE}_2)^2},$$

where $\hat{d}_i$, $i = 1, 2$ is the fractional differencing parameter from the FI(A)PARCH model estimated for currency i and SE_i is the standard error associated with the FI(A)PARCH model estimated for currency i. The above Wald statistic tests whether the fractional differencing parameters of the two exchange rates are equal ($\hat{d}_1 = \hat{d}_2$), and is distributed as $\chi^2_{(1)}$. The following table presents the results of this pairwise testing procedure. The p values in the table provide support for the null hypothesis that the estimated fractional parameters are not significantly different from one another. For example, currencies which generated very similar fractional parameters, such as the BRP ($\hat{d} = 0.30$) and the SWK ($\hat{d} = 0.30$) or the ITL ($\hat{d} = 0.45$) and the CAD ($\hat{d} = 0.51$) were, as expected, not significantly different (W= 0.01 and W= 0.05 respectively). Furthermore, the null hypothesis could not be rejected even in the case of quite dissimilar estimated fractional differencing parameters, such as the JAY ($\hat{d} = 0.19$) and the CAD ($\hat{d} = 0.51$); the value of the Wald test (W= 2.17) is clearly insignificant at 5% level.

Table 3: Tests of fractional differencing parameters in FIAPARCH models.

	BRP	GEM	FRF	ITL	SWK	SPP	JAY	SID	CAD	AUD
BRP	-	0.32 [0.57]	0.15 [0.70]	1.17 [0.28]	0.01 [0.91]	0.10 [0.75]	1.05 [0.31]	1.081 [0.30]	0.87 [0.35]	0.06 [0.80]
GEM		-	0.06 [0.81]	0.14 [0.71]	0.43 [0.51]	0.68 [0.41]	1.84 [0.18]	0.12 [0.73]	0.24 [0.63]	0.59 [0.44]
FRF			-	0.48 [0.49]	0.23 [0.63]	0.48 [0.49]	1.79 [0.18]	0.43 [0.51]	0.48 [0.49]	0.39 [0.53]
ITL				-	1.53 [0.22]	1.92 [0.17]	3.89 [0.05]	0.01 [0.91]	0.05 [0.82]	1.74 [0.19]
SWK					-	0.10 [0.75]	1.31 [0.25]	1.41 [0.24]	0.99 [0.32]	0.06 [0.81]
SPP						-	0.56 [0.46]	1.78 [0.18]	1.23 [0.27]	0.00 [0.95]
JAY							-	3.64 [0.06]	2.17 [0.14]	0.64 [0.43]
SID								-	0.06 [0.81]	1.61 [0.21]
CAD									-	1.15 [0.28]
AUD										-

Notes: Table 3 presents a Wald test of the null hypothesis that the estimated fractional differencing parameters are not significantly different from one another.

The numbers in brackets are p-values.

Finally, we test the apparent similarity of the optimal power terms estimated for each of the ten currencies using a pairwise Wald test:

$$\mathrm{W} = \frac{(\hat{\delta}_1 - \hat{\delta}_2)^2}{(\mathrm{SE}_1)^2 + (\mathrm{SE}_2)^2},$$

where $\hat{\delta}_i$, $i = 1, 2$, is the power term from the FI(A)PARCH model estimated for currency i, and SE_i is the standard error associated with the FI(A)PARCH model estimated for currency i. The above Wald statistic tests whether the power terms of the two countries are equal ($\hat{\delta}_1 = \hat{\delta}_2$), and is distributed as $\chi^2_{(1)}$. The following table presents the results of this pairwise testing procedure. The p values in the table provide support for the null hypothesis that the estimated power parameters are not significantly different from one another. For example, currencies which generated very similar power terms, such as the BRP ($\hat{\delta} = 1.92$) and the CAD ($\hat{\delta} = 1.91$) or GEM ($\hat{\delta} = 1.78$) and JAY ($\hat{\delta} = 1.76$), were, as expected, not significantly different (W= 0.01 and W= 0.01 respectively). Furthermore, the null hypothesis could not be rejected even in the case of quite dissimilar estimated power terms, such as the SWK ($\hat{\delta} = 1.73$) and the SPP ($\hat{\delta} = 2.08$); the value of the Wald test (W= 0.83) is clearly insignificant at any conventional size of the test.

Table 4: Tests of power term parameters in FIAPARCH models.

	BRP	GEM	FRF	ITL	SWK	SPP	JAY	SID	CAD	AUD
BRP	-	0.13 [0.72]	0.12 [0.73]	0.44 [0.83]	0.28 [0.60]	0.27 [0.61]	0.19 [0.66]	0.01 [0.91]	0.01 [0.91]	0.00 [0.95]
GEM		-	0.01 [0.91]	0.27 [0.61]	0.01 [0.91]	0.54 [0.46]	0.01 [0.91]	0.09 [0.76]	0.10 [0.75]	0.15 [0.70]
FRF			-	0.30 [0.59]	0.05 [0.83]	0.64 [0.43]	0.02 [0.89]	0.07 [0.79]	0.08 [0.78]	0.14 [0.71]
ITL				-	0.49 [0.49]	0.12 [0.73]	0.36 [0.55]	0.12 [0.73]	0.03 [0.86]	0.01 [0.91]
SWK					-	0.83 [0.36]	0.01 [0.91]	0.22 [0.64]	0.21 [0.65]	0.29 [0.59]
SPP						-	0.67 [0.41]	0.43 [0.51]	0.21 [0.65]	0.17 [0.68]
JAY							-	0.14 [0.71]	0.14 [0.71]	0.21 [0.65]
SID								-	0.01 [0.91]	0.03 [0.87]
CAD									-	0.01 [0.91]
AUD										-

Notes: Table 4 presents a Wald test of the null hypothesis that the estimated power terms are not significantly different from one another.
The numbers in brackets are p-values.

4 Forecasting

4.1 Forecasting methodology and evaluation criteria

Our full sample consists of 3,621 trading days and each model is estimated over the first 3,256 observations of the full sample, i.e. over the period January 1990 to June 2002. As a result the out-of-sample period is from July 2002 to November 2003 providing 365 daily observations. The parameter estimates obtained with the data from the in-sample period are inserted in the relevant forecasting formulas and volatility forecasts $\widehat{h}_{t+1}$ calculated given the information available at time $t = T, \ldots, T + 364$, i.e. 365 one-step ahead forecasts are calculated.

The true underlying volatility process is unobservable, therefore in order to evaluate the predictive ability of various volatility models we need to have a valid proxy. Andersen and Bollerslev (1998) show that although the daily squared return is an unbiased estimator of the true volatility, it is also an extremely noisy estimator. They propose the construction of a volatility measure based on the cumulative squared returns obtained from high-frequency intraday data which they called realised volatility[12]. This

[12]See Andersen et al. (2001, 2003) and Barndorff-Nielsen and Shephard (2002) for a formal discussion of the properties of realised volatility.

integrated ex-post volatility measure allows for the construction of more accurate volatility forecast evaluation criteria. However, Awartani and Corradi (2003) argue that in the case of discrete time data generating processes, squared returns and realized volatility are both unbiased estimators of the true underlying volatility[13]. Therefore, when interested in the comparison of different models, in principle there is not much to choose between the use of squared returns and realized volatility as proxy of the true unobservable volatility process. In other words, the use of either squared returns or realized volatility, would lead to the choice of the 'right' model, as it does not alter the correct comparison of models, at least in terms of quadratic loss function.

We examine the ability of the various asymmetric GARCH models to forecast exchange rate return volatility. A complete description of the evaluation requires a specification of a loss function. As Andersen et al. (1999) pointed out, it is generally impossible to specify a forecast evaluation criterion that is universally acceptable (see also, e.g., Diebold et al., 1998). This problem is special acute in the context of nonlinear volatility forecasting. Accordingly, there is a wide range of evaluation criteria used in the literature. Following Andersen and et al. (1999) we shall not use any of the complex economically motivated evaluation criteria but instead we will report summary statistics based directly on the deviation between forecasts and realizations. A number of out-of-sample forecast performance measures will be used to evaluate and compare the various models.

First, the quality of the individual forecasts is assessed by regressing the squared returns (r^2) on the corresponding forecast:

$$r_t^2 = a + b\widehat{h}_t + e_t,$$

If the volatility forecasts are unbiased, then $a = 0$ and $b = 1$. The accuracy of the forecast can be assessed by the R^2 of the regression[14]. As the squared returns are a noisy measure of true volatility low values for R^2 should be expected. Andersen and Bollerslev (1998) have shown how the R^2 increases with the frequency from which realized volatility is constructed.

Although this is one of the most commonly employed criterion in the existing literature, it is not necessarily the best criterion adopt when evaluating nonlinear volatility forecasts (Andersen, et al., 1999).

[13]If the conditional mean is zero, then squared returns provide an unbiased estimator of the true underlying volatility process. But if the conditional mean in not zero, then one should use the squared residuals form the regression of r_t on a constant and on ρr_{t-1}. Of course, if we mispecify the conditional mean such squared residuals are no longer an unbiased estimator of the conditional variance (Awartani and Corradi, 2003).

[14]Andersen and Bollerslev (1998) showed that the maximum obtainable R^2 from the regression of squared returns on the volatility forecast is indeed very small. For example, under the null hypothesis that the returns are generated by a GARCH(1,1) model with zero conditional mean and conditional Gaussian errors the R^2 will be bounded from above by 1/3. Engle and Patton (2001) also point out that the heteroscedasticity of returns, implies even more heteroscedasticity in the squared returns so parameters are estimated inefficiently and the usual standard errors are misleading.

Consequently, we also report results from a number of alternative more robust evaluation criteria. First, we employ two error statistics which were used by Andersen et al. (1999). The first is the mean square error (MSE). The usual measure based on the MSE maybe unreliable in the presence of heteroscedasticity, so we also report the more robust mean absolute error (MAE). To better accommodate the heteroscedasticity in the forecast errors, we also calculate the corresponding hetroscedasticity adjusted statistics. These are the heteroskedasticity adjusted root mean square error (HRMSE) and the heteroskedasticity adjusted mean absolute error (HMAE), which were used by Andersen et al. (1999), Martens (2002) and Hol and Koopman (2002) (see table 5a below). Second, we employ four loss functions which are also based directly on the deviation between forecasts and realizations. These are the $HRMSE_1$ and $HMAE_1$ criteria. These are interesting because they are more robust to outliers, than the HRMSE and HMAE criteria (see Martens, 2001). The other two are the QLIKE and R^2LOG criteria which are discussed in Bollerslev et al. (1994). The former loss function corresponds to the loss implied by a Gaussian likelihood while the latter exaggerates the interest in predicting when residuals are close to zero. All these four statistics were employed by Hansen and Lunde (2001).

Table 5a: Evaluation criteria.

MSE:	$\frac{1}{k}\sum_{t=T+1}^{T+k}(\widehat{h}_t - r_t^2)^2$	MAE:	$\frac{1}{k}\sum_{t=T+1}^{T+k}\left\lvert\widehat{h}_t - r_t^2\right\rvert$
HRMSE:	$\sqrt{\frac{1}{k}\sum_{t=T+1}^{T+k}(1-\frac{r_t^2}{\widehat{h}_t})^2}$	HMAE:	$\frac{1}{k}\sum_{t=T+1}^{T+k}\left\lvert 1-\frac{r_t^2}{\widehat{h}_t}\right\rvert$
$HRMSE_1$:	$\sqrt{\frac{1}{k}\sum_{t=T+1}^{T+k}(1-\frac{rt}{\sqrt{\widehat{h}_t}})^2}$	$HMAE_1$:	$\frac{1}{k}\sum_{t=T+1}^{T+k}\left\lvert 1-\frac{r_t}{\sqrt{\widehat{h}_t}}\right\rvert$
QLIKE:	$\frac{1}{k}\sum_{t=T+1}^{T+k}[\ln(\widehat{h}_t)+\frac{r_t^2}{\widehat{h}_t}]$	R^2LOG:	$\frac{1}{k}\sum_{t=T+1}^{T+k}[\ln(\frac{r_t^2}{\widehat{h}_t})]^2$

Notes: k is the number of steps ahead, T is the sample size, $\widehat{h}_t$ is the forecasted variance and r_t^2 are the squared returns. MSE and MAE are the mean square and mean absolute errors. HRMSE and HMAE are the heteroscedasticity adjusted root MSE (RMSE) and MAE statistics.
QLIKE and R^2LOG are discussed in Bollerslev et al. (1994).

We also employ three error statistics which are used by Peters (2001). These are the median squared error (MedSE), the adjusted mean absolute percentage error (AMAPE), and the Theil inequality coefficient (TIC) (see table 5b). The Theil inequality coefficient is a scale invariant measure that always lies between zero and one, where zero indicates a perfect fit.

All the previous error statistics assume that the underling loss function is symmetric. Brailsford and Faff (1996) point out that many investors will not attribute equal importance to both over- and under-predictions of volatility of similar magnitude. For example, an over-prediction of stock price volatility will lead to an upward biased estimate of the corresponding call option price since the relationship between the volatility of underlying stock prices and call option prices is positive. This over-estimate of the price is more likely to be of greater concern to a buyer than a seller (see Brailsford and Faff, 1996). Therefore, we report a summary statistic which penalizes under-predictions more heavily and is called the mean mixed error (MME(U)). Similarly, the statistic MME(O) weights over-predictions more heavily. These two statistics were proposed by Brailsford and Faff (1996).

Table 5b: Evaluation Criteria.

AMAPE:	$\frac{1}{k+1}\sum_{t=T}^{T+k}\left\|\frac{\widehat{h}_t - r_t^2}{\widehat{h}_t + r_t^2}\right\|$
TIC:	$\sqrt{\frac{1}{k+1}\sum\limits_{t=T}^{T+k}\left(\widehat{h}_t - r_t\right)^2}\left[\sqrt{\frac{1}{k+1}\sum\limits_{t=T}^{T+k}\widehat{h}_t^2} + \sqrt{\frac{1}{k+1}\sum\limits_{t=N}^{N+k} r_t^2}\right]^{-1}$
MedSE:	Median of $\left\{(\widehat{h}_{T+1} - r_{T+1}^2), \ldots, (\widehat{h}_{T+1} - r_{T+K}^2)\right\}$
MME(U):	$\frac{1}{k}\left[\sum\limits_{j=1}^{O}\left\|\widehat{h}_j - r_j^2\right\| + \sum\limits_{j=1}^{U}\left(\left\|\widehat{h}_j - r_j^2\right\|\right)^{2^l}\right]$
MME(O):	$\frac{1}{k}\left[\sum\limits_{j=1}^{U}\left\|\widehat{h}_j - r_j^2\right\| + \sum\limits_{j=1}^{O}\left(\left\|\widehat{h}_j - r_j^2\right\|\right)^{2^l}\right]$

Notes: AMAPE is the adjusted mean absolute percentage error and TIC is the Theil inequality coefficient. MedSE is the median square error. MME(U) and MME(O) are the two mean mixed errors where O (U) is the number of over (under)-predictions. If the absolute value of the forecast error is greater (less) than unity , then $l = 1(l = -1)$, and MME(U)/MME(U) takes the square (root) of the errors in order to place a heavier weighting on the under-predictions/over-predictions.

4.2 Forecasting results

On the basis of several model selection techniques the fractionally integrated APARCH specification was the superior fitting model (see section 3.1). While such model fitting investigations provide useful insights into volatility, the models are usually selected on the basis of full sample information. For practical forecasting purposes, the predictive ability of these models needs to be examined out-of-sample. The aim of this section is to examine the relative ability of the various long-memory and power models to forecast daily exchange rate volatility. Table 6 gives a relative indication of overall forecasting performance. Calculated values are provided for fourteen different forecasting performance measures across ten exchange rates, so in total we have one hundred and forty cases.

An examination of table 6 reveals that the general FI(A)PARCH model is clearly superior. That is, there is strong evidence that both the restrictive stable and integrated (A)PARCH models are inferior to the fractionally integrated (A)PARCH specification. The more restrictive models are outperformed by the FI(A)PARCH specification for sixty four out of the one hundred and forty cases, whereas for only thirty six cases the best performing model is the (A)PARCH one. In particular, for all the currencies except the Canadian dollar the QLIKE loss function favors the long-memory (A)PARCH formulation. Moreover, the FI(A)PARCH specification has the best MSE statistic for eight out of the ten currencies. In addition, for seven currencies the best ranked model, as assessed by the MedSE, R^2LOG and the MME(O) statistics, is also the fractional integrated one. Finally, for six out of the ten exchange rates the MAE and AMAPE loss functions also favor the FI(A)PARCH formulation. In sharp contrast, for the majority of the currencies (six to seven out of the ten) the integrated (A)PARCH model has the best heteroscedasticity adjusted error statistics whereas the worst performing model, as assessed by these statistics, is the long-memory specification. However, in many cases these statistics do not allow for a clear distinction between the ranking models which is evidenced by the marginal difference in relative accuracy which separates the three models.

Notice also that the FI(A)PARCH model does extremely well in the Japanese yen since eleven out of the fourteen error statistics identify the fractionally integrated specification as superior. In other words, for this currency the results appear to be robust to the choice of the loss function. The FI(A)PARCH model does quite well in the Singapore dollar and the Spanish peseta. For these two exchange rates nine out of the fourteen error statistics indicate that the long-memory specification provides the most accurate forecasts. Similarly, for the British pound, the Swedish krona and the Australian dollar across the majority of the statistics (eight out of the fourteen) the best performing model is the fractionally integrated one. For the Italian lira while the FI(A)PARCH specification ranks second, for twelve out of the fourteen evaluation criteria, it is not substantially worse than the best ranked model.

Brooks et al. (2000) point out that invalid imposition of a particular value for the power term may lead to sub-optimal forecasting performance. Accordingly, we examine the relative ability of the three fractionally integrated models to forecast daily volatility. The Taylor/Schwert FI(A)ARCH model has the best heteroscedasticity adjusted and TIC error statistics for almost all the currencies. These five statistics rank the general FI(A)PARCH specification second for eight out of the ten currencies. For the majority of the currencies the worst performing model, as assessed by these five statistics, is the Bollerslev FI(A)ARCH model. In sharp contrast, the MAE error statistic in table 6 indicate the superiority of the Bollerslev FI(A)ARCH formulation for nine out of the ten currencies. Similarly, the AMAPE, R^2LOG and MME(O) statistics rank this specification first for eight out of the ten currencies. Further, the MedSE and MMe(U) loss functions identify the Bollerslev FI(A)ARCH model as superior for seven out of the ten currencies. According to these five statistics across the majority of the ten currencies the power FI(A)ARCH formulation is ranked second whereas the Taylor/Schwert FI(A)ARCH model provides the worst forecasts. Finally, for half of the currencies the Bollerslev FI(A)ARCH model is ranked first by the coefficient of determination (R^2), RMSE and QLIKE evaluation criteria. These last two statistics rank the general FI(A)PARCH specification first for four out of the ten currencies.

The FI(A)ARCH model does quite well in the Spanish peseta. For this exchange rate seven out of the fourteen statistics indicate that the fractionally integrated specification provides the most accurate forecasts. In sharp contrast, in the Canadian dollar the power formulation is ranked third by the majority of the evaluation criteria. However, for the other ten currencies the power FI(A)ARCH specification is the second best forecasting model according to the majority of the statistics.

In summary the ranking of the three fractionally integrated models varies depending upon the choice of the error statistic. In other words, our rankings for the three long-memory formulations are not robust between the error statistics. This contrast in the rankings of these specification illustrates that the forecasts are highly sensitive to the assessment criteria. Hence, as Brailsford and Faff (1996) point out, caution should be exercised in the interpretation of the obtained rankings.

Table 6: Error statistics from forecasting daily exchange rate volatility.

	BRITISH POUND					GERMAN MARK				
	P	IP	FIP	BFI	TSFI	P	IP	FIP	BFI	TSFI
R^2	0.00	**$5/10^4$**	0.00	$1/10^4$	0.00	$1/10^4$	$3/10^4$	**$4/10^4$**	$7/10^4$	0.00
MSE	0.19	0.20	**0.18**	0.18	0.20	0.37	0.38	**0.37**	0.37	0.40
MAE	0.30	0.33	**0.25**	0.25	0.34	**0.41**	0.45	0.42	0.41	0.50
HRMSE	1.48	**1.41**	1.90	1.96	1.26	1.55	**1.33**	1.49	1.59	1.17
HRMSE$_1$	0.67	**0.67**	0.71	0.72	0.66	0.66	**0.64**	0.65	0.66	0.62
HMAE	0.97	**0.95**	1.12	1.14	0.90	1.02	**0.92**	1.00	1.04	0.87
HMAE$_1$	**0.57**	0.57	0.58	0.58	0.57	0.55	**0.54**	0.55	0.56	0.54
AMAPE	0.57	0.59	**0.55**	0.54	0.59	**0.54**	0.55	0.54	0.54	0.57
TIC	0.53	**0.52**	0.59	0.59	0.51	0.53	**0.50**	0.53	0.54	0.48
QLIKE	-0.32	-0.28	**-0.34**	-0.34	-0.27	0.13	0.14	**0.13**	0.13	0.18
R2LOG	6.47	6.82	**5.65**	5.58	7.13	**5.57**	6.04	5.65	5.48	6.52
MedSE	0.06	0.08	**0.03**	0.03	0.10	**0.11**	0.17	0.12	0.11	0.22
MME(U)	0.39	0.41	**0.36**	0.36	0.41	**0.58**	0.60	0.59	0.58	0.64
MME(O)	0.48	0.51	**0.41**	0.41	0.53	**0.58**	0.61	0.58	0.57	0.65
	FRENCH FRANC					ITALIAN LIRA				
	P	IP	FIP	BFI	TSFI	P	IP	FIP	BFI	TSFI
R^2	$5/10^4$	$2/10^4$	**$6/10^4$**	$9/10^4$	$5/10^4$	$6/10^4$	**$1/10^3$**	0.00	0.00	$1/10^4$
MSE	0.37	0.38	**0.37**	0.37	0.39	**0.37**	0.38	0.37	0.37	0.39
MAE	**0.42**	0.45	0.42	0.41	0.47	**0.41**	0.45	0.42	0.42	0.48
HRMSE	1.52	**1.37**	1.48	1.55	1.24	1.57	**1.39**	1.52	1.53	1.23
HRMSE$_1$	0.65	**0.64**	0.65	0.66	0.63	0.66	0.64	**0.66**	0.66	0.63
HMAE	1.01	**0.93**	1.00	1.03	0.90	1.03	**0.95**	1.02	1.02	0.90
HMAE$_1$	0.55	**0.54**	0.55	0.55	0.54	0.56	**0.55**	0.55	0.56	0.54
AMAPE	**0.54**	0.55	0.54	0.54	0.56	**0.54**	0.55	0.54	0.54	0.56
TIC	0.53	**0.50**	0.53	0.54	0.49	0.53	**0.50**	0.53	0.53	0.49
QLIKE	0.13	0.14	**0.13**	0.12	0.15	0.13	0.15	**0.13**	0.13	0.16
R2LOG	**5.31**	5.68	5.34	5.22	5.90	**4.53**	4.95	4.57	4.56	5.23
MedSE	**0.12**	0.16	0.12	0.11	0.19	**0.11**	0.15	0.12	0.12	0.20
MME(U)	**0.59**	0.60	0.59	0.58	0.62	**0.58**	0.60	0.59	0.58	0.62
MME(O)	**0.58**	0.61	0.58	0.57	0.63	**0.57**	0.61	0.58	0.58	0.64

	SWEDISH KRONA					SPANISH PESETA				
	P	IP	FIP	BFI	TSFI	P	IP	FIP	BF	TSFI
R^2	0.00	**1/10³**	0.00	0.00	0.00	$3/10^4$	$2/10^4$	<u>**2/10³**</u>	$6/10^4$	$1/10^3$
MSE	0.48	0.47	<u>**0.46**</u>	0.46	0.47	0.37	0.39	**0.37**	<u>0.37</u>	0.38
MAE	0.52	0.51	**0.46**	<u>0.46</u>	0.52	0.42	0.47	<u>**0.40**</u>	0.41	0.46
HRMSE	**1.25**	1.30	1.53	1.60	<u>1.19</u>	1.52	**1.29**	1.61	1.57	<u>1.26</u>
$HRMSE_1$	**0.63**	0.63	0.66	0.66	<u>0.62</u>	0.65	**0.63**	0.67	0.66	<u>0.63</u>
HMAE	**0.91**	0.93	1.04	1.07	<u>0.89</u>	1.01	**0.90**	1.05	1.04	<u>0.90</u>
$HMAE_1$	**0.54**	0.54	0.56	0.56	<u>0.54</u>	0.55	**0.54**	0.56	0.56	<u>0.54</u>
AMAPE	0.56	0.56	**0.54**	<u>0.54</u>	0.56	0.54	0.56	<u>**0.54**</u>	0.54	0.56
TIC	**0.48**	0.49	0.53	0.54	<u>0.49</u>	0.53	**0.49**	0.54	0.54	<u>0.49</u>
QLIKE	0.30	0.29	<u>**0.27**</u>	0.27	0.28	0.13	0.16	**0.12**	<u>0.12</u>	0.15
R2LOG	6.25	6.11	**5.56**	<u>5.46</u>	6.22	5.63	6.28	<u>**5.44**</u>	5.50	6.18
MedSE	0.23	0.21	**0.14**	<u>0.12</u>	0.22	0.12	0.18	<u>**0.10**</u>	0.10	0.18
MME(U)	0.71	0.70	**0.68**	<u>0.68</u>	0.70	0.59	0.62	<u>**0.58**</u>	0.58	0.61
MME(O)	0.67	0.66	**0.62**	<u>0.61</u>	0.67	0.58	0.63	<u>**0.56**</u>	0.57	0.62

	JAPANESE YEN					SINGAPORE DOLLAR				
	P	IP	FIP	BFI	TSFI	P	IP	FIP	BFI	TSFI
R^2	**0.01**	0.01	$4/10^3$	$4/10^3$	<u>0.01</u>	0.01	**0.01**	<u>0.01</u>	$4/10^3$	0.01
MSE	0.41	0.50	**0.32**	<u>0.30</u>	0.38	333.06	226.31	**198.06**	<u>191.26</u>	218.58
MAE	0.56	0.63	**0.46**	<u>0.43</u>	0.52	15.80	11.77	**10.25**	<u>9.78</u>	11.85
HRMSE	0.96	**0.95**	1.07	1.13	<u>0.98</u>	**1.08**	1.28	1.36	1.39	<u>1.11</u>
$HRMSE_1$	0.66	0.67	<u>**0.66**</u>	0.66	0.66	0.68	0.68	**0.68**	0.67	<u>0.66</u>
HMAE	0.83	**0.82**	0.87	0.89	<u>0.83</u>	**0.88**	0.93	0.96	0.97	<u>0.88</u>
$HMAE_1$	0.59	0.60	**0.59**	<u>0.58</u>	0.59	0.62	0.60	**0.59**	<u>0.58</u>	0.59
AMAPE	0.65	0.66	**0.63**	<u>0.62</u>	0.64	0.67	0.63	**0.61**	<u>0.60</u>	0.63
TIC	0.50	0.51	**0.50**	0.50	<u>0.50</u>	0.52	**0.51**	0.53	0.53	<u>0.50</u>
QLIKE	0.08	0.15	**-0.05**	<u>-0.08</u>	0.37	3.37	3.22	**3.14**	<u>3.11</u>	3.20
R2LOG	9.71	10.29	**8.64**	<u>8.29</u>	9.39	9.57	8.18	**7.53**	<u>7.29</u>	8.30
MedSE	0.35	0.45	**0.21**	<u>0.18</u>	0.31	251.64	108.00	**81.42**	<u>73.80</u>	132.56
MME(U)	0.64	0.70	**0.56**	<u>0.53</u>	0.61	**113.99**	133.80	142.38	145.41	<u>125.97</u>
MME(O)	0.71	0.75	**0.63**	<u>0.61</u>	0.69	234.87	104.29	**65.95**	<u>55.64</u>	104.47

	CANADIAN DOLLAR					AUSTRALIAN DOLLAR				
	P	IP	FIP	BFI	TSFI	P	IP	FIP	BFI	TSFI
R^2	0.01	**0.01**	0.01	<u>0.01</u>	0.01	$1/10^4$	$\mathbf{1/10^3}$	$2/10^4$	$3/10^4$	0.00
MSE	**103.88**	1795.88	1843.81	1817.80	<u>1784.90</u>	0.37	0.37	**0.37**	<u>0.37</u>	0.38
MAE	**10.18**	22.28	22.33	<u>22.30</u>	22.31	0.42	0.45	**0.42**	<u>0.42</u>	0.48
HRMSE	**0.98**	3.80	4.39	4.03	<u>3.70</u>	1.45	**1.34**	1.45	1.48	<u>1.17</u>
$HRMSE_1$	**0.89**	0.97	1.05	1.00	<u>0.96</u>	0.65	**0.64**	0.65	0.65	<u>0.62</u>
HMAE	**0.98**	2.02	2.29	2.13	<u>1.97</u>	0.99	**0.93**	0.99	1.00	<u>0.88</u>
$HMAE_1$	0.88	**0.72**	0.77	0.74	<u>0.71</u>	0.55	**0.54**	0.55	0.55	<u>0.54</u>
AMAPE	0.95	**0.53**	0.54	0.53	<u>0.53</u>	0.54	0.55	**0.54**	<u>0.54</u>	0.56
TIC	0.93	**0.72**	0.75	0.74	<u>0.72</u>	0.52	**0.50**	0.52	0.52	<u>0.49</u>
QLIKE	**2.37**	4.71	4.90	4.78	<u>4.68</u>	0.12	0.14	**0.12**	<u>0.12</u>	0.15
R2LOG	29.30	4.62	<u>**4.56**</u>	4.60	4.66	5.54	5.84	**5.49**	<u>5.45</u>	6.15
MedSE	106.26	91.68	<u>**78.14**</u>	86.63	95.13	0.12	0.16	**0.12**	<u>0.12</u>	0.20
MME(U)	**10.18**	1765.14	1821.59	1791.06	<u>1751.10</u>	0.58	0.60	**0.58**	<u>0.58</u>	0.61
MME(O)	103.88	53.03	<u>**44.57**</u>	49.06	56.13	0.59	0.61	**0.58**	<u>0.58</u>	0.64

Notes: For the ten currencies, calculated values are provided for ten different error statistics across five ARCH-type models used to forecast daily volatility. FIP, IP and P denote the fractionally integrated, integrated and stable (A)PARCH models respectively. BFI and TSFI denote the Bollerslev and Taylor/Schwert fractionally integrated (A)ARCH models respectively. R^2 is the coefficient of determination of the regression: $r_t^2 = a + b\widehat{h}_t + e_t$. MSE and MAE are the mean square and mean absolute errors.
HRMSE and HMAE are the heteroscedasticity adjusted RMSE and MAE errors. AMAPE is the adjusted mean absolute percentage error and TIC is the Theil inequality coefficient. The QLIKE and R^2LOG criteria are discussed in Bollerslev et al. (1994). MME are the mean mixed error statistics.
Bold (underline) numbers indicate the power (fractionally integrated) model with the best performance.

4.3 Encompassing Tests

In this section we utilize two encompassing tests proposed by Ericsson (1992) and Harvey et al. (1998). Before moving to the two tests some notation is needed. First, we denote the 1-step ahead forecast errors as $\xi_{U,t} \equiv r_t^2 - \widehat{h}_{U,t}$ and $\xi_{R,t} \equiv r_t^2 - \widehat{h}_{R,t}$ $(t = T+1, \ldots, T+k)$, for the unrestricted (FIAPARCH) and restricted models, respectively. Forecasts of the squared returns are generated using the fixed forecasting scheme (described in West and McCracken, 1998, p. 819). Next, let $\psi_t \equiv \xi_{R,t}^2 - \xi_{R,t}\xi_{U,t}$ and $\overline{\psi} \equiv k^{-1}\sum_{j=T+1}^{T+k}\psi_j$. The first encompassing test, proposed by Harvey et al. (1998) and denoted by ENC-T, is formed as

$$\text{ENC-T} = \frac{\sqrt{k-1\psi}}{\sqrt{\frac{1}{k}\sum_{j=1}^{k}(\psi_{T+j}-\overline{\psi})^2}}. \tag{4.1}$$

The second test statistic proposed by Ericsson (1992) is a regression-based variant of the ENC-T test. This forecast encompassing test, denoted by ENC-REG, can be expressed as

$$\text{ENC-REG} = \frac{\sqrt{k-1\psi}}{\sqrt{\frac{1}{k}\sum_{j=1}^{k}(\xi_{R,T+j}-\xi_{U,T+j})^2(\frac{1}{k}\sum_{j=1}^{k}\xi_{R,T+j}^2)-\overline{\psi}^2}}. \tag{4.2}$$

Clark and McCracken (2001) (see footnote 2 in their paper), show that for the fixed scheme the ENC-T and ENC-REG tests are asymptotically standard normal. The null hypothesis is that of equal predictive accuracy, whereas the alternative is that the unrestricted model provides a more accurate prediction than the restricted.

Next, we move to the pairwise comparison of nested models using the two aforementioned test statistics. In particular, the FI(A)PARCH formulation is compared against the (A)PARCH, I(A)PARCH, and the two restricted FI(A)ARCH specifications with $\delta = 1, 2$.

For nine out of the ten currencies the ENC-T test rejects the null that the integrated (A)PARCH forecasts encompass the FI(A)PARCH forecasts. The evidence obtained from the ENC-T test statistic is reinforced by the ENC-REG test. That is, for eight out of the ten currencies the ENC-REG test rejects the null of equal accuracy. Moreover, both the ENC-T and ENE-REG tests reject the null that the stable (A)PARCH forecast encompasses the FI(A)PARCH forecast for four out of the ten exchange rates.

Clearly, assuming integrated (A)PARCH behavior results in volatility forecasts which are less accurate compared to those forecasts obtained from the FI(A)PARCH model.

The Taylor/Schwert FI(A)ARCH is clearly rejected in favor for the FI(A)PARCH model by the ENC-T (in nine cases) and by the ENC-REG test (in eight cases). In sharp contrast, the two encompassing statistics indicate the superiority of the power model over the Bollerslev one only for the British pound and the Swedish krona. This result however is in accordance with the results we obtained from the tests on the δ parameter in section 3.1. While the $\delta = 1$ hypothesis was clearly rejected for all the countries, the $\delta = 2$ hypothesis could not be rejected for any country. Naturally, the forecasts of the FI(A)PARCH model should be superior to those obtained from the Taylor/Schwert FI(A)ARCH model, but should not be able to improve upon those obtained from the Bollerslev FI(A)ARCH model. Interestingly, for the Swedish krona where the Bollerslev FI(A)ARCH forecasts are dominated by the more general FI(A)PARCH forecasts, we estimated a δ parameter ($\hat{\delta} = 1.73$) for which the distance to $\delta = 2$ was the largest.

Hence, on the basis of these results, in the majority of the cases the FI(A)PARCH formulation outperforms the integrated (A)PARCH model and the restricted Taylor/Schwert FI(A)ARCH specifications, but - as expected - does not improve upon the Bollerslev FI(A)ARCH model.

Table 7. ENC-T and ENC-REG test statistics.

H_0 :	APARCH		IAPARCH		FIAARCH $(\delta=2)$		FIAARCH $(\delta=1)$	
	ENC-T	ENC-R	ENC-T	ENC-R	ENC-T	ENC-R	ENC-T	ENC-R
BRP	2.94***	3.04***	4.89***	5.38***	1.23*	1.30*	6.05***	6.00***
GEM	0.30	0.29	2.85***	2.83***	0.53	0.51	5.71***	5.46***
FRF	0.77	0.88	2.89***	2.88***	0.20	0.19	3.95***	3.77***
ITL	0.04	0.03	2.82***	2.88***	0.01	0.01	4.38***	4.19***
SWK	3.73***	3.77***	2.87***	3.00***	1.26*	1.28*	3.01***	2.99***
SPP	0.71	0.72	4.62***	4.53***	-0.22	-0.21	3.57***	3.42***
JAY	14.11***	11.26***	16.40***	12.95***	-6.39	-5.76	12.30***	10.20***
SID	15.79***	-10629.66	9.85***	-306.80	-4.88	-12.59	7.01***	-180.71
CAD	-7.26	-131.43	-5.20	-747.58	-6.09	-166.63	-5.20	-747.58
AUD	1.12	1.28*	2.85***	2.89***	-0.15	-0.15	4.38***	4.21***

Notes: The test statistics ENC-T and ENC-REG are defined in (4.1) and (4.2) respectively. The stable and integrated (A)PARCH models and the Taylor/Schwert and Bollerslev FI(A)ARCH models are compared against the FI(A)PARCH model. The null hypothesis is that of equal predictive accuracy, whereas the alternative hypothesis is that the FI(A)PARCH model provides a more accurate prediction than the restricted models. *,**,***indicate significance at the10%, 5% and 1% level respectively.

4.4 The empirical evidence

In this paper we proceeded in two steps. In a first step strong evidence has been put forward suggesting that the conditional volatility for ten US bilateral exchange rates is best modelled as a mean-reverting fractionally integrated APARCH process. On the basis of Wald tests the FI(A)PARCH model provides statistically significant improvement over its integrated counterpart. One can also reject the more restrictive stable (A)PARCH model, and consequently all the existing models (see Ding et al. 1993) nested by this specification in favor of the fractionally integrated parameterization. Hence, the in-sample analysis has shown that the FI(A)PARCH formulation is preferred to both the stable (A)PARCH and the integrated (A)PARCH models. In other words, the fractionally integrated model appeared to have superior ability to differentiate between stable models and their integrated alternatives.

Moreover, both the (asymmetric) Taylor/Schwert and Bollerslev GARCH formulations are nested within the (A)PARCH model. According to our analysis for all ten currencies the Taylor/Schwert specification was rejected in favor of the FI(A)PARCH formulation. In sharp contrast, for all ten currencies tested the Wald statistics indicated a preference for the (asymmetric) Bollerslev GARCH formulation over the FI(A)PARCH model. That is, allowing the power term to take on values other than two did not significantly enhance the model. In other words there is a lack of evidence to suggest the need of power effects even in the presence of long-range volatility dependence.

Finally, comparing the pairwise testing results of the Wald procedures to the relative model rankings provided by the four alternative criteria we observe that the findings were generally robust. That is, where the Wald testing results provided unanimous support for the FI(A)PARCH specification over either the stable (A)PARCH or integrated (A)PARCH formulations, the model selection criteria concurred without exception. Thus, the inclusion of a a fractional unit root in the conditional variance equation appear to augment the model in a worthwhile fashion. The same conclusion holds for the (asymmetric) Bollerslev GARCH formulation which was supported by all the information criteria.

Baillie et al. (1996) analyzed the Deutschmark-U.S. dollar exchange rate by different GARCH specifications. While in the simple GARCH the persistence coefficient was quite close to one and suggestive of integrated behavior, the FIGARCH specification seemed to capture the long-run dynamics of the series quite well. The estimated fractional differencing parameter was $\hat{d} = 0.652$ (0.160). Beine et al. (2002) argue that this value may be a local maximum[15] and estimate a FIGARCH model with fractional differencing parameter of $\hat{d} = 0.43$ (4.613) (including closing days effects) which is quite similar to the value we estimated for the Deutschmark-U.S. dollar exchange rate. The estimates they obtain for the BRP-, Yen- and FRF-U.S. dollar exchange rates are again close to the ones we obtained. Thus, our results imply that the fractional differencing parameter is much closer to stable than to integrated behavior.

Tse (1998) applied the FIAPARCH model to the Yen-U.S. dollar exchange rate. While he found significant power and fractional differencing parameters, he estimated an insignificant leverage parameter. In our analyzes for five out of the ten countries we find significant leverage parameters (including Japan). Moreover, while he can not reject both hypothesis of $\delta = 1, 2$ we clearly reject the $\delta = 1$ hypothesis. Tse (1998) also compares the stable APARCH and the FIAPARCH models by examining the coefficients of their infinite series representations. He finds the differences between models to be small. This again shows that the FIAPARCH specification is much closer to the stable APARCH specification than to the integrated one which would produce constant coefficients.

In a second step we provided out-of-sample evidence for the FI(A)PARCH model. Its forecasting accuracy was compared to the one of the stable and integrated specifications as well as to the one of the Taylor/Schwert and Bollerslev GARCH formulations. Clearly, the forecast evaluation criteria favored the more general FI(A)PARCH compared to the stable and integrated specifications. The ranking of the three long-memory specifications heavily depended on the choice of the error statistic. The encompassing tests again clearly rejected the integrated specification, while the FI(A)PARCH and the stable

[15] See Beine et al. (2002), p. 697, fig. 2.

(A)PARCH nearly performed equivalently. This result is in line with the finding that the estimated fractional differencing parameters are much closer to stable than to integrated behavior. In accordance with the Wald tests the encompassing tests reject the Taylor/Schwert model, but can not reject the Bollerslev GARCH formulation.

5 Conclusion

The purpose of the current paper was to consider the applicability of the fractionally integrated (asymmetric) PARCH model to ten US bilateral exchange rates. It was found that the FI(A)PARCH specification captures the temporal pattern of volatility for observable returns better than previous parameterizations. Our analyzes has shown that this improves forecasts for volatility and thus is useful for financial decisions which utilize such forecasts.

We provided an interesting comparison to the stable and integrated (A)PARCH models. The results reject both the hypotheses of a stable and an integrated model. This is consistent with the conditional volatility profiles in Gallant et al. (1993), which suggests that shocks to the variance are very slowly damped, but do die out.

All ten currencies show strong evidence in favor of the Bollerslev FI(A)ARCH model when long-memory persistence in the conditional volatility have been taken into account, as both the power and Taylor/Schwert (asymmetric) FIGARCH specifications were rejected in favor of the FI(A)ARCH formulation with power term equal to two.

Moreover, we find that both the optimal fractional differencing parameter and power transformation are remarkably similar across countries.

The in-sample estimation and testing results were reinforced by the out-of-sample forecast evaluation. The forecast error measures supported the FI(A)PARCH model as the one with the highest forecast accuracy. The Bollerslev FI(A)ARCH model can be seen as equal accurate compared with the more general FI(A)PARCH specification in terms of encompassing tests, while the FI(A)PARCH dramatically outperforms the integrated specification.

References

[1] Andersen, T. G., Bollerslev, T., 1988. Answering the skeptics: yes, standard volatility models do provide accurate forecasts. *International Economic Review* 39, No. 4, 885-905.

[2] Andersen, T. G., Bollerslev, T., Lange, S., 1999. Forecasting financial market volatility: sample frequency vis-a-vis forecast horizon. *Journal of Empirical Finance* 6, 457-477.

[3] Andersen, T. G., Bollerslev, T., Diebold, F. X., Ebens, H., 2001. The distribution of realized stock return volatility. *Journal of Financial Economics* 61, 43-76.

[4] Andersen, T. G., Bollerslev, T., Diebold, F. X., Labys, P., 2003. Modelling and forecasting realized volatility. *Econometrica* 71, 597-625.

[5] Awartani, B. M. A., Corradi, V., 2003. Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries. Unpublished manuscript.

[6] Baillie, R. T., Bollerslev, T., Mikkelsen, H. O., 1996. Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 74, 3–30.

[7] Barndorff-Nielsen, O. E., Shepard, N., 2002. Econometric analysis of realized volatility its use in estimating stochastic volatility models. *Journal of Royal Statistical Society* B 64, 253-280.

[8] Beine, M., Laurent, S., Lecourts, C., 2002. Accounting for conditional leptokurtosis and closing days effects in FIGARCH models of daily exchange rates. *Applied Financial Economics* 12, 589-600.

[9] Bollerslev, T., 1986. Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics* 31, 307–327.

[10] Bollerslev, T., Mikkelsen, H. O., 1996. Modeling and pricing long memory in stock market volatility. *Journal of Econometrics* 73, 151–184.

[11] Bollerslev, T., Wooldridge, J. M., 1992. Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews* 11, 143-172.

[12] Bollerslev, T., Chou, R. Y., Kroner, K. F., 1992. ARCH modeling in finance: a review of the theory and empirical evidence. *Journal of Econometrics* 52, 5-59.

[13] Bollerslev, T., Engle R. F., Nelson, D., 1994. ARCH models. In Engle, R. F., McFadden, D. L., eds, *Handbook of Econometrics*, Vol. IV, Elsevier Science B. V., p. 2961-3038.

[14] Brailsford, T. J., Faff, R. W., 1996. An evaluation of volatility forecasting techniques. *Journal of Banking & Finance* 20, 419-438.

[15] Breidt, F. J., Crato, N., de Lima, P., 1998. On the detection and estimation of long memory in stochastic volatility. *Journal of Econometrics* 83, 325–348.

[16] Brooks, R. D., Faff, R. W., McKenzie, M. D., 2000. A multi-country study of power ARCH models and national stock market returns. *Journal of International Money and Finance* 19, 377-397.

[17] Clark, T. E., McCracken, M. W., 2001. Tests of equal forecast accuracy and encompassing for nested models. *Journal of Econometrics* 105, 85-110.

[18] Dacorogna, M. M., Muller, U. A., Nagler, R. J., Olsen, R. B., Pictet, O. V, 1993. A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance* 12, 413–438.

[19] Diebold, F. X., Gunther, T. A., Tay, A. S., 1998. Evaluating density forecasts with applications to financial risk management. *International Economic Review* 39, No. 4, 863-883.

[20] Ding, Z., Granger, C. W. J., 1996. Modeling volatility persistence of speculative returns: a new approach. *Journal of Econometrics* 73, 185–215.

[21] Ding, Z., Granger, C. W. J., Engle, R. F., 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance* 1, 83–106.

[22] Engle, R. F., 1982. Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica* 50, 987–1008.

[23] Engle, R. F., Lee, G. G. J., 2000. A long-run and short-run component model of stock return volatility, in: Engle, R. F., White, H., Cointegration, causality, and forecasting. Festschrift in honor of Clive W. J. Granger, Oxford University Press, Oxford, 475-497.

[24] Engle, R. F., Patton, A. J., 2001. What good is a volatility model ? *Quantitative Finance* 1, 237-245.

[25] Ericsson, N. R. 1992. Parameter constancy, mean square forecast errors, and measuring forecast performance: an exposition, extensions, and illustration. *Journal of Policy Modeling* 14, 465-495.

[26] Fornari, F., Mele, A., 1997. Weak convergence and distributional assumptions for a general class of nonlinear ARCH models. *Econometric Reviews* 16, 205-229.

[27] Fornari, F., Mele, A., 2001. Recovering the probability density function of asset prices using GARCH as diffusion approximations. *Journal of Empirical Finance* 8, 83-110.

[28] Gallant, A. R., Rossi, P. E., Tauchen, G., 1993. Nonlinear dynamic structures. *Econometrica* 61, 871-907.

[29] Giot, P., Laurent, S., 2003. Value at Risk for long and short trading positions. *Journal of Applied Econometrics* 18.6, 541-663.

[30] Glosten, L. R., Jaganathan and Runkle, D., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance* 48, 1779–1801.

[31] Granger, C. W. J., Ding, Z., 1995a. Some properties of absolute returns: an alternative measure of risk. *Annales d'Economie de Statistique* 40, 67-95.

[32] Granger, C. W. J., Ding, Z., 1995b. Stylized facts on the temporal and distributional properties of daily data from speculative returns. Department of Economics, University of California San Diego, unpublished manuscript.

[33] Geweke, J., 1986. Modeling the persistence of conditional variances: a comment. *Econometric Review* 5, 57–61.

[34] Hansen, P. R., Lunde, A., 2002. A forecast comparison of volatility models: does anything beat a GARCH(1,1). Brown University, Department of Economics, Working Paper 02-04.

[35] Harvey, D. I., Leybourne, S. J., Newbold, P. 1998. Tests for forecast encompassing. *Journal of Business and Economic Statistics* 16, 254-259.

[36] Hentschel, L., 1995. All in the family. Nesting symmetric and asymmetric GARCH models. *Journal of Financial Economics* 39, 71-104.

[37] Higgins, M., Bera, A., 1992. A class of nonlinear ARCH models. *International Economic Review* 33, 137–158.

[38] Hol, E., Koopman, S. J., 2002. Stock index volatility forecasting with high frequency data. Unpublished manuscript.

[39] Laurent, S., Peters, J. P., 2002. G@RCH 2.2: an Ox package for estimating and forecasting various ARCH models. *Journal of Economic Surveys* 3, 447-485.

[40] Martens, M., 2002. Measuring and forecasting S&P-500 index-futures volatility using high frequency data. *Journal of Futures Markets* 22, 497-518.

[41] McCurdy, T. H., Michaud, P. K., 1996. Capturing long memory in the volatility of equity returns: a fractionally integrated asymmetric power ARCH model. Unpublished manuscript.

[42] Pantula, S., 1986. Modeling the persistence in conditional variances: a comment. *Econometric Review* 5, 71–74.

[43] Peters, J.-P., 2001. Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (skewed) student-t densities. Unpublished manuscript.

[44] Robinson, P. M., 1991. Testing for strong serial correlation and dynamic conditional heteroskedasticity in multiple regression. *Journal of Econometrics* 47, 67–84.

[45] Schwert, W., 1990. Stock volatility and the crash of '87. *Review of Financial Studies* 3, 77-102.

[46] Taylor, S., 1986. Modeling Financial Time Series. Wiley, New York.

[47] Tse, Y. K., 1996. The conditional heteroscedasticity of the yen-dollar exchange rate. Unpublished manuscript.

[48] Tse, Y. K., 1998. The conditional heteroscedasticity of the yen-dollar exchange rate. J*ournal of Applied Econometrics* 13, 49-55.

[49] West, K. D., McCracken, M. W., 1988. Regression-based tests of predictive ability. *International Economic Review* 39, No.4, 817-840.

[50] Zakoian, J. M., 1994. Threshold Heteroskedastic Model. *Journal of Economic Dynamics and Control* 18, 931 - 995.