Fractionally Integrated APARCH Modeling of Stock Market Volatility: A multi-country study

C. CONRAD[a], M. KARANASOS[b*] and N. ZENG[b]

[a] *University of Heidelberg, Germany*

[b] *Brunel University, West London, UK*

First draft: December 2003

This draft: May 2008

Abstract

Tse (1998) propose a model which combines the fractionally integrated GARCH formulation of Baillie, Bollerslev and Mikkelsen (1996) with the asymmetric power ARCH specification of Ding, Granger and Engle (1993). This paper analyzes the applicability of a multivariate constant conditional correlation version of the model to national stock market returns for eight countries. We find this multivariate specification 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 multivariate FIAPARCH framework is provided in terms of forecast error statistics and tests for equal forecast accuracy of nested models.

Keywords: Asymmetric Power ARCH, Fractional integration, Stock returns, Volatility forecast evaluation.

JEL Classification:C13, C22, C52.

We would like to thank two anonymous referees and M. Karanassou for their helpful comments and suggestions.

*Address for correspondence: Economics and Finance, Brunel University, West London, UB3 3PH, UK; email: menelaos.karanasos@brunel.ac.uk, tel: +44 (0)1895265284, fax: +44 (0)1895269770.

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 is 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. Motivated by this empirical result they propose a new general class of ARCH models, which they call the Asymmetric Power ARCH (APARCH). In addition, they show that this formulation comprises seven other specifications in the literature.[1] Brooks et al. (2000) analyze the applicability of the PARCH models to national stock market returns for ten countries plus a world index. 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. McCurdy and Michaud (1996) analyze the CRSP value-weighted index using a fractionally integrated APARCH (FIAPARCH) type of model. McCurdy and Michaud (1996) and Tse (1996, 1998) extend the asymmetric power formulation of the variance to incorporate fractional integration, as defined by Baillie et al. (1996) (see also Robinson, 1991).[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 asset 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. Giot and Laurent (2003) have shown that APARCH volatility forecasts outperform those obtained from the RiskMetrics model which is equivalent to an integrated ARCH with pre-specified autoregressive parameter values. The fractionally integrated process may lead to further improvement, if its forecasts are more accurate than those obtained from the stable specification.

Another important advantage of having a FIAPARCH model is that it nests the formulation without power effects and the stable one as special cases. This provides an encompassing framework for these two broad classes of specifications and facilitates comparison between them. The main contribution of this

[1] These models are: the ARCH (Engle, 1982), the GARCH (Bollerslev, 1986), the Taylor/Schwert GARCH in standard deviation (Taylor, 1986, and Schwert, 1990), the GJR GARCH (Glosten et al., 1993), the TARCH (Zakoian, 1994), the NARCH (Higgins and Bera, 1992) and the log-ARCH (Geweke, 1986, and Pantula, 1986).

[2] The FIGARCH model of Baillie et al. (1996) is closely related to the long-memory GARCH process (see, Karanasos et al. 2003, and the references therein, and Conrad and Karanasos, 2006).

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 FIAPARCH model is applicable to the yen-dollar exchange rate. An interesting research issue is to explore how generally applicable this formulation is to a wide range of financial data. In this paper we attempt to address this issue by estimating multivariate versions of this framework for eight series of national stock market index returns. These countries are Canada, France, Germany, Hong Kong, Japan, Singapore, the United Kingdom and the United States. As the general multivariate specification adopted in this paper nests the various univariate formulations, the relative ranking of each of these models can be considered using the Wald testing procedures. In addition, standard information criteria can be used to provide a ranking of the specifications. Furthermore, the ability of the FIAPARCH formulation to forecast (out-of-sample) stock volatility is assessed by a variety of forecast error statistics. In order to verify whether the differences between the forecast error statistics from the different models is statistically significant we employ the Diebold and Mariano (1995) test.

The remainder of the paper is structured as follows. In section 2 we detail the FIAPARCH model and discuss how various ARCH specifications are nested within it. Section 3 discusses the data and presents the empirical results. Maximum likelihood parameter estimates for the various specifications are presented, as are the results of the Wald testing procedures. 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 evaluates the different specifications in terms of their out-of-sample forecast ability. For each country and each formulation four forecast error measures are calculated and evaluated against each other. Moreover, we test for equal forecast accuracy of the competing models by utilizing the Diebold and Mariano (1995) test statistic. Section 5 discusses our results and Section 6 concludes the analysis.

2 FIAPARCH Model

2.1 Univariate Process

One of the most common models in finance and economics to describe a time series r_t of stock returns is the AR(1) process

$$(1-\zeta L)r_t = c + \varepsilon_t, \quad t \in \mathbb{N}, \tag{2.1}$$

with

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

where $c \in (0, \infty)$, $|\zeta| < 1$ and $\{e_t\}$ are independently, identically distributed ($i.i.d$) student-t random variables with $\mathsf{E}(e_t) = \mathsf{E}(e_t^2 - 1) = 0$. 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 $\{r_t\}$ $(r_t|\Sigma_{t-1}) \sim i.i.d\ (c + \zeta r_{t-1}, h_t)$.

Tse (1998) examines the conditional heteroskedasticity of the yen-dollar exchange rate by employing the FIAPARCH$(1, d, 1)$ model. Accordingly, we utilize the following process

$$(1 - \beta L)(h_t^{\delta/2} - \omega) = [(1 - \beta L) - (1 - \phi L)(1 - L)^d](1 + \gamma s_t)|\varepsilon_t|^\delta, \tag{2.2}$$

where $\omega \in (0, \infty)$, $|\phi| < 1$, $0 \leq d \leq 1$,[3] $s_t = 1$ if $\varepsilon_t < 0$ and 0 otherwise, γ is the leverage coefficient, and δ is the parameter for the power term that takes (finite) positive values.

When $d = 0$, the process in equation (2.2) reduces to the APARCH(1,1) one which nests two major classes of ARCH models. Specifically, a Taylor/ Schwert type of formulation is specified when $\delta = 1$, and a Bollerslev type 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 modeling and forecasting performance relative to other power term".

Since its introduction by Ding et al. (1993), the APARCH formulation has been frequently applied. It is worth noting that Fornari and Mele (1997) show the usefulness of this scheme in approximating models developed in continuous time as systems of stochastic differential equations. This feature has usually been overshadowed by its well-known role as simple econometric tool providing reliable estimates of unobserved conditional variances (Fornari and Mele, 2001). Hentschel (1995) defines a parametric family of asymmetric models that nests the APARCH one.[4]

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

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

where

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

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

[4] For applications of the APARCH model in economics see Campos and Karanasos (2008), Campos et al. (2008a, 2008b)

When $\gamma = 0$ and $\delta = 2$ the process in equation (2.2) reduces to the FIGARCH$(1, d, 1)$ specification which includes Bollerslev's (1986) model (when $d = 0$) and the integrated specification (when $d = 1$) as special cases.[5] 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, concerns the apparent widespread finding of integrated 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).[6]

Finally, as noted by Baillie et al. (1996) for the variance, being confined to only considering the extreme cases of stable and integrated specifications 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 volatility may be easily mistaken for integrated behavior.

2.2 Multivariate Formulation

In this section we discuss the multivariate time series model for the stock returns and discuss its merits and properties. Let us define the N-dimensional column vector of the returns $\mathbf{r}_t$ as $\mathbf{r}_t = [r_{it}]_{i=1,\ldots,N}$ and the corresponding residual vector $\boldsymbol{\varepsilon}_t$ as $\boldsymbol{\varepsilon}_t = [\varepsilon_{it}]_{i=1,\ldots,N}$. Regarding $\boldsymbol{\varepsilon}_t$ we assume that it is conditionally student-t distributed with mean vector $\mathbf{0}$, variance vector $\mathbf{h}_t = [h_{1t}]_{i=1,\ldots,N}$ and constant conditional correlations (ccc), $\rho_{ij} = h_{ij,t}/\sqrt{h_{it}h_{jt}}$, $|\rho_{ij}| \leq 1$, $i, j = 1, \ldots, N$.

Next, the structure of the AR (1) mean equation is given by

$$\mathbf{Z}(L)\mathbf{r}_t = \mathbf{c} + \boldsymbol{\varepsilon}_t, \tag{2.3}$$

where $\mathbf{Z}(L) = \mathbf{I}_N \boldsymbol{\zeta}(L)$ with $\mathbf{I}_N$ being the $N \times N$ identity matrix and $\boldsymbol{\zeta}(L) = [1 - \zeta_i L]_{i=1,\ldots,N}$, $|\zeta_i| < 1$, and $\mathbf{c} = [c_i]_{i=1,\ldots,N}$ $[c_i \in (0, \infty)]$.

and Karanasos and Schurer (2008).

[5] An excellent survey of major econometric work on long-memory processes and their applications in economics and finance is given by Baillie (1996). For applications of the FIA(P)ARCH model to exchange rates see, among others, Karanasos et al., 2006 and Conrad and Lamla, 2007).

[6] 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).

Further, to establish terminology and notation, the multivariate FIAPARCH (M-FIAPARCH) process or order $(1, d, 1)$ is defined by

$$\mathbf{B}(L)(\mathbf{h}_t^{\wedge\frac{\delta}{2}} - \boldsymbol{\omega}) = [\mathbf{B}(L) - \boldsymbol{\Delta}(L)\boldsymbol{\Phi}(L)][\mathbf{I}_N + \boldsymbol{\Gamma}\mathbf{s}_t]|\boldsymbol{\varepsilon}_t|^{\wedge\delta}, \tag{2.4}$$

where $^\wedge$ denotes elementwise exponentiation and $|\boldsymbol{\varepsilon}_t|$ is the vector $\boldsymbol{\varepsilon}_t$ with elements stripped of negative values. Moreover, $\mathbf{B}(L) = \mathbf{I}_N\boldsymbol{\beta}(L)$ with $\boldsymbol{\beta}(L) = [1 - \beta_i L]_{i=1,\ldots,N}$, and $\boldsymbol{\Phi}(L) = \mathbf{I}_N\boldsymbol{\phi}(L)$ with $\boldsymbol{\phi}(L) = [1 - \phi_i L]_{i=1,\ldots,N}$, $|\phi_i| < 1$. In addition, $\boldsymbol{\omega} = [\boldsymbol{\omega}_i]_{i=1,\ldots,N}$ $[\omega_i \in (0, \infty)]$ and $\boldsymbol{\Delta}(L) = \mathbf{I}_N\mathbf{d}(L)$ with $\mathbf{d}(L) = [(1 - L)^{d_i}]_{i=1,\ldots N}$ $(0 \leq d_i \leq 1)$. Finally, $\boldsymbol{\Gamma} = \boldsymbol{\gamma}\mathbf{I}_N$ with $\boldsymbol{\gamma} = [\gamma_i]_{i=1,\ldots,N}$, and $\mathbf{s}_t = [s_{it}]_{i=1,\ldots,N}$ where $s_{it} = 1$ if $\varepsilon_{it} < 0$ and 0 otherwise.[7]

Dark (2004) applies a bivariate error correction FIAPARCH model to examine the relationship between stock and future markets. Kim et al. (2005) use a bivariate FIAPARCH-in-mean process to model the volume-volatility relationship.

3 Empirical Analysis

3.1 Data

Daily stock price index data for eight countries were sourced from the Datastream database for the period 1st January 1988 to 22nd April 2004, giving a total of 4,255 observations. We will use the period 1st January 1988 to 16th July 2003 for the estimation, while we produce 200 out-of-sample forecasts for the period 17th July 2003 to 22nd April 2004. The eight countries and their respective price indices are: UK: FTSE 100 (F), US: S&P 500 (SP), Germany: DAX 30 (D), France: CAC 40 (C), Japan: Nikkei 225 (N), Singapore: Straits Times (S), Hong Kong: Hang Seng (H) and Canada: TSE 300 (T). For each national index, the continuously compounded return was estimated as $r_t = 100[\log(p_t) - \log(p_{t-1})]$ where p_t is the price on day t.

3.2 Univariate Models

We proceed with the estimation of the AR(1)-FIAPARCH$(1, d, 1)$ model[8] in equations (2.1) and (2.2) in order to take into account the serial correlation[9] and the GARCH effects observed in our time series

[7] $\mathbf{Z}(L)$, $\mathbf{B}(L)$, $\boldsymbol{\Phi}(L)$ and $\boldsymbol{\Delta}(L)$ are $N \times N$ diagonal polynomial matrices with diagonal elements $1 - \zeta_i L$, $1 - \beta_i L$, $1 - \phi_i L$ and $(1 - L)^{d_i}$ respectively. Further, $\boldsymbol{\Gamma}$ is a $N \times N$ diagonal matrix with diagonal elements γ_i.

[8] The only exceptions are the Canadian and Singaporean indices, where an AR(1)-FIAPARCH$(0, d, 1)$ model is used. For these two indices the AR(1)-FIAPARCH$(1, d, 1)$ estimates for β were insignificant and the IC came out in favor for the $(0, d, 1)$ specification. In addition, for the Hang Seng index the criteria favor the $(1, d, 0)$ formulation.

[9] The 12th order Ljung-Box Q-statistics on the squared return series indicate high serial correlation in the second moment for all indices.

data, and to capture the possible long-memory in volatility. We estimate the various specifications using the maximum likelihood estimation (MLE) method as implemented by Davidson (2008) in Time Series Modelling (TSM). The existence of outliers, particularly in daily data, causes the distribution of returns to exhibit excess kurtosis.[10] To accommodate the presence of such leptokurtosis, we estimate the models using student-t distributed innovations. [11]

Table 1 reports the estimation results. If parameters turned out to be insignificant we reestimated the respective model without those parameters. In all countries the AR coefficient (ζ) is highly significant. The estimate for the $\phi(\beta)$ parameter is insignificant only in one(two) out of the eight cases. In three countries the estimates of the leverage term (γ) are statistically significant, confirming the hypothesis that there is negative correlation between returns and volatility. For all indices the estimates of the power term (δ) and the fractional differencing parameter (d) are highly significant. In all cases, the estimated degrees of freedom parameter (υ) is highly significant and leads to an estimate of the kurtosis which is different from three.[12]

In all cases, the ARCH parameters satisfy the set of necessary conditions sufficient to guarantee the non-negativity of the conditional variance (see Conrad and Haag, 2006). 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.

3.2.1 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. For the stable APARCH(1,1) model[13] the condition for the existence *of* the $\delta/2$ th moment of the conditional variance is $\mathsf{V} = \alpha \mathsf{E}(1+\gamma s)|e|^{\delta} + \beta < 1$ which depends on the density of e. For a student-t distributed innovation with υ degrees of freedom we have $\frac{\mathsf{V}-\beta}{a} = \frac{(1+\frac{\gamma}{2})}{\sqrt{\pi}}(\upsilon-2)^{\frac{\delta}{2}} \frac{\Gamma(\frac{\delta+1}{2})\Gamma(\frac{\upsilon-\delta}{2})}{\Gamma(\frac{\upsilon}{2})}$. Notice that if $\gamma = 0$ the expression for the $\frac{\mathsf{V}-\beta}{\alpha}$ is the one for the symmetric PARCH model (see Paolella, 1997 and Karanasos and Kim, 2006). In addition, if $\gamma = 0$, $\delta = 2$, $\mathsf{V} = \alpha + \beta < 1$ reduces to the usual stationarity condition of the GARCH(1,1) model.

Thus, estimating a V which is close to one is suggestive of integrated APARCH behaviour. Table 2

[10] For all indices the Jarque-Bera statistic rejects the normality hypothesis at the 1% level. The estimated kurtosis coefficient is significantly above three for all indices but FTSE 100 and Nikkei 225.

[11] Bai and Chen (2008) construct an asymptotically distribution-free test statistic for testing multivariate normal and t-distributions, which is applicable in vector autoregressive and GARCH processes.

[12] The kurtosis of a student-t distributed random variable with υ degrees of freedom is $3\frac{\upsilon-2}{\upsilon-4}$.

[13] Restricting d to be 0 in equation (2.2) leads to an APARCH(1,1) model with $\phi - \beta = \alpha$.

Table 1: Univariate AR-FI(A)PARCH models (ML Estimation)

	SP	T	C	D	F	H	N	S
c	0.03 (5.35)	0.02 (4.52)	0.02 (3.50)	0.02 (3.40)	0.02 (3.23)	0.02 (3.13)	0.01 (1.99)	0.01 (0.91)
ζ	-0.05^* (-3.28)	0.17 (10.94)	0.04 (2.34)	0.03^* (2.31)	0.04 (2.38)	0.06 (3.71)	-0.02 (-1.63)	0.15 (9.20)
ω	0.01 (1.59)	0.01 (2.54)	0.05 (3.63)	0.07 (3.65)	0.03 3.71	0.13 (5.12)	0.02 (2.50)	0.07 (4.47)
$\beta(MA)$	0.54 (5.81)	–	0.66 (6.94)	0.56 (5.65)	0.59 (5.32)	0.08 (2.01)	0.51 (4.83)	–
$\phi(AR)$	0.27 (4.11)	-0.11 (-2.95)	0.20 (3.84)	0.21 (4.82)	0.19 (3.77)	–	0.14 (2.03)	-0.07 (-2.23)
γ	–	–	–	0.46 (3.73)	–	0.69 (3.65)	–	0.76 (3.90)
δ	2.35 (23.50)	2.42 (17.28)	1.77 (12.64)	1.24 (11.46)	1.86 (14.31)	1.28 (12.80)	2.07 (18.81)	1.40 (12.73)
d	0.30 (6.00)	0.19 (6.33)	0.52 (4.33)	0.40 (4.34)	0.46 (4.60)	0.18 (4.50)	0.42 (6.00)	0.21 (5.25)
υ	5.60 (10.77)	5.38 (10.76)	8.53 (6.56)	6.83 (6.90)	10.70 (6.04)	4.56 (11.12)	5.80 (10.54)	4.86 (11.04)
Q_{12}	18.45 [0.10]	9.52 [0.66]	10.00 [0.61]	13.18 [0.36]	12.86 [0.38]	22.85 [0.03]	10.59 [0.56]	18.50 [0.10]
Q_{12}^2	5.12 [0.95]	19.47 [0.08]	11.74 [0.47]	8.13 [0.77]	18.00 [0.12]	33.24 [0.00]	20.90 [0.05]	2.20 [1.00]

Notes: For each of the eight indices, Table 1 reports ML parameter estimates for the AR(1)-FI(A)PARCH model. The numbers in parentheses are t-statistics. *For the S&P 500 and Dax 30 indices we estimate AR(3) and AR(4) models respectively. Q_{12} and Q_{12}^2 are the 12th order Ljung-Box tests for serial correlation in the standardized and squared standardized residuals respectively. The numbers in brackets are p-values.

presents the estimates for V from the AR-APARCH$(1,1)$ model with student-t distributed innovations. For all indices V is close to 1 indicating that $h_t^{\frac{\delta}{2}}$ may be integrated.[14]

Table 2: Estimates of V for AR-APARCH$(1,1)$ models

	SP	T	C	D	F	H	N	S
V	0.998	0.991	1.000	0.985	0.985	0.963	1.013	0.946

However, from the FI(A)PARCH estimates (reported in table 1), it appears that the long-run dynamics are better modeled by the fractional differencing parameter. To test for the persistence of the

[14] We do not report the estimated AR-APARCH$(1,1)$ coefficients for space considerations.

conditional heteroskedasticity models, we examine the Wald statistics for the linear constraints $d = 0$ (stable APARCH) and $d = 1$ (IAPARCH).[15] As seen in table 3 the W tests clearly reject both the stable and integrated null hypotheses against the FIAPARCH one.[16] Clearly, the results which emerged from table 2 where misleading, i.e. imposing the restriction $d = 0$ leads to parameter estimates which falsely suggest integrated behaviour. Thus, purely from the perspective of searching for a model that best describes the volatility in the stock return series, the fractionally integrated one appears to be the most satisfactory representation.[17]

Table 3: Tests for restrictions on fractional differencing and power term parameters

H_0:		$d=0$	$d=1$		$\delta=1$	$\delta=2$
	d	W	W	δ	W	W
S&P 500	0.30{0.05}	33[0.00]	173[0.00]	2.35{0.10}	178[0.00]	9[0.00]
TSE 300	0.19{0.03}	28[0.00]	522[0.00]	2.42{0.14}	102[0.00]	10[0.00]
CAC 40	0.52{0.12}	18 [0.00]	15[0.00]	1.77{0.14}	31[0.00]	3[0.09]
DAX 30	0.40{0.09}	18 [0.00]	39[0.00]	1.24{0.11}	15[0.00]	52[0.00]
FTSE 100	0.46 {0.10}	21 [0.00]	29[0.00]	1.86{0.13}	37[0.00]	1[0.30]
Hang Seng	0.18{0.04}	16 [0.00]	322[0.00]	1.28{0.10}	8[0.00]	72[0.00]
Nikkei 225	0.42{0.07}	35[0.00]	67[0.00]	2.07{0.11}	114[0.00]	0.50[0.54]
Straits Times	0.21{0.04}	32[0.00]	444[0.00]	1.40{0.11}	16[0.00]	36[0.00]

Notes: For each of the eight indices, Table 3 reports the value of the Wald (W) statistics for the unrestricted FI(A)PARCH and restricted ($d = 0, 1; \delta = 1, 2$) models respectively. The numbers in $[\cdot]$ are p values. The numbers in $\{\cdot\}$ are standard errors.

Following the work of Ding et al. (1993), Hentschel (1995), Tse (1998) and Brooks et al. (2000) among others, the Wald test 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 ARCH type of models.[18] These model selection criteria check the robustness of the Wald testing results

[15] Restricting d to be one leads to an IAPARCH(1,2) **model** (see equation (2.2)).

[16] Various tests for long-memory in volatility have been proposed in the literature (see, for details, Karanasos and Kartsaklas, 2008).

[17] It is worth mentioning the empirical results in Granger and Hyung (2004). They suggest that there is a possibility that, at least, part of the long-memory may be caused by the presence of neglected breaks in the series. We look forward to sorting this out in future work.

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

discussed above.[19] Specifically, according to the AIC, HQIC and SHIC, the optimal specification (i.e., FIAPARCH, APARCH or IAPARCH) for all indices was the FIAPARCH one.[20] The SIC results largely concur with the AIC, HQIC or SHIC results.[21]

Next, 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 specification). 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 3, all eight estimated power coefficients are significantly different from unity (see column six). Further, with the exception of the CAC 40, FTSE 100 and Nikkei 225 indices, each of the power terms are significantly different from two (see the last column of table 3). Hence, on the basis of these results, in the majority of cases support is found for the (asymmetric) power fractionally integrated model, which allows an optimal power transformation term to be estimated. The evidence obtained from the Wald tests is reinforced by the model ranking provided by the four model selection criteria.[22]

3.3 Multivariate Models

The analysis above suggests that the FIAPARCH formulation describes the conditional variances of the eight stock indices well. In this section within the framework of the multivariate ccc model we will analyze the dynamic adjustments of the variances for the various indices. Overall we estimate seven bivariate specifications. Three for the European countries: CAC 40-DAX 30 (C-D), CAC 40-FTSE 100 (C-F) and DAX 30-FTSE 100 (D-F); three for the Asian countries: Hang Seng-Nikkei 225 (H-N), Hang Seng-Straits Times (H-S) and Nikkei 225-Straits Times (N-S); one for the S&P 500 and TSE 300 indices (SP-T). Moreover, we estimate two trivariate models: one for the three European countries (C-D-F) and one for the three Asian countries (H-N-S).

3.3.1 Bivariate Processes

The best fitting bivariate specification is chosen according to likelihood ratio results and the minimum value of the information criteria (not reported). In the majority of the models the AR coefficients are significant at the 5% level or better. In almost all cases a $(1, d, 1)$ order is chosen for the FIAPARCH

[19] The use of the information criteria techniques for comparing models has the advantage of being relatively less onerous compared to Wald testing procedures, which only allow formal pairwise testing of nested models (Brooks et al., 2000).

[20] Caporin (2003) performs a monte carlo simulation study and verifies that information criteria clearly distinguish the presence of long memory.

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

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

formulation. Only for the H-S and N-S models we choose $(0, d, 1)$ order for the Straits Times index, and $(1, d, 0)$ order for the Hang Seng index. Note, that this in line with our findings for the univariate models where the β parameter was insignificant for Straits Times, while the ϕ parameter was insignificant for Hang Seng. In six out of the fourteen models the leverage term (γ) is significant. As in the univariate case it is significant in both indices for the H-S case and in the DAX 30 index for the D-F case. In addition, in the bivariate case it is also significant in the Tse 300 index for the SP-T model and in the Nikkei 225 for the N-S one. In almost all cases the power term (δ) and the fractional differencing parameter (d) are highly significant. In the D-F, H-S and N-S models the two countries generated very similar power terms: (1.28, 1.36), (1.42, 1.47) and (1.70, 1.62) respectively. In four out of the seven bivariate formulations the two countries generated very similar fractional parameters. These are the SP-T, the C-F, the H-N and the H-S models. The corresponding pair of values are: (0.22, 0.21), (0.24, 0.29), (0.36, 0.35) and (0.16, 0.13). Interestingly, in the majority of the cases the estimated power and fractional differencing parameters of the bivariate models take lower values than those of the corresponding univariate models. In all cases the estimated ccc (ρ) is highly significant. Most importantly, the conditional correlation is rather high among the American and European indices, while it is rather low among the Asian indices. Finally, the degrees of freedom (υ) parameters are highly significant and the ARCH parameters satisfy the set of necessary conditions sufficient to guarantee the non-negativity of the conditional variances (see Conrad and Haag, 2006). In the majority of the cases the hypothesis of uncorrelated standardized and squared standardized residuals is well supported (see the last two rows of table 4).

Next we examine the Wald statistics for the linear constraints $d = 0$ (stable APARCH) and $d = 1$ (IAPARCH). As seen in table 5 the W tests clearly reject both the stable and integrated null hypotheses against the FIAPARCH one. We also test whether the estimated power terms are significantly different from unity or two using Wald tests. The eight estimated power coefficients are significantly different from either unity or two (see the last two columns of table 5).

3.3.2 Trivariate Specifications

Table 6 reports the parameters of interest for the two trivariate FI(A)PARCH(1,1) models. In two out of the three Asian countries the leverage term (γ) is weakly significant. In all cases the power term (δ) and the fractional differencing parameter (d) are highly significant. Similarly, in all cases the estimated ccc (ρ) and degrees of freedom (υ) parameters are highly significant and the ARCH parameters satisfy the set of necessary conditions sufficient to guarantee the non-negativity of the conditional variances (see Conrad and Haag, 2006). In particular, the estimates of ρ confirm the results from the bivariate models, i.e. the conditional correlation between the European indices is considerably stronger than between the

Asian indices.

3.4 On the Similarity of the Fractional/Power Parameters

We test for the apparent similarity of the optimal fractional differencing and power term parameters for each of the eight country indices using pairwise Wald tests:

$$\mathrm{W}_d = \frac{(d_1 - d_2)^2}{\mathsf{Var}(d_1) + \mathsf{Var}(d_2) - 2\mathsf{Cov}(d_1, d_2)}, \quad \mathrm{W}_\delta = \frac{(\delta_1 - \delta_2)^2}{\mathsf{Var}(\delta_1) + \mathsf{Var}(\delta_2) - 2\mathsf{Cov}(\delta_1, \delta_2)},$$

where d_i (δ_i), $i = 1, 2$, is the fractional differencing (power term) parameter from the FIAPARCH model estimated for the national stock market index for country i, $\mathsf{Var}(d_i)$, $\mathsf{Var}(\delta_i)$ are the corresponding variances, and $\mathsf{Cov}(d_1, d_2)$, $\mathsf{Cov}(\delta_1, \delta_2)$ are the corresponding covariances. The above Wald statistics test whether the fractional differencing (power term) parameters of the two countries are equal $d_1 = d_2$ $(\delta_1 = \delta_2)$, and are distributed as $\chi^2_{(1)}$.

The following table presents the results of this pairwise testing procedure for the SP-T models. Several findings emerge from this table. The estimated long-memory parameters for the S&P 500(TSE 300) index are in the range $0.22(0.18) \leq d \leq 0.30(0.27)$ while the estimated power terms are in the range $1.86(1.47) \leq \delta \leq 2.35(2.24)$. In all cases the values of the two coefficients (d_i, δ_i) for the asymmetric models (see columns U^a, B^a) are lower than the corresponding values for the symmetric formulations (see columns U^s, B^s). The values of the Wald tests in the table support the null hypothesis that the two estimated fractional parameters and the two power term coefficients are not significantly different from another. For example, in the bivariate asymmetric model which generated very similar fractional parameters (0.22 and 0.23) the two coefficients were, as expected, not significantly different ($W = 0.04$). Furthermore, the null hypothesis could not be rejected even in the case of quite dissimilar estimated power terms, such as the bivariate asymmetric formulation (1.86 and 1.51): the value of the Wald test ($W = 2.59$) is clearly insignificant at the 5%.

The following table presents the results of the pairwise testing procedure for the three European countries. In the majority of the cases the values of the two coefficients (d_i, δ_i) for the bivariate (asymmetric) models are lower than the corresponding values for the univariate (symmetric) formulations. The values of the Wald tests in the table support the null hypothesis that the two estimated fractional parameters (d_1, d_2) and the two power term coefficients (δ_1, δ_2) are not significantly different from another. All specifications generated very similar long-memory coefficients between countries. The symmetric univariate models (see column U^s) for DAX 30 and FTSE 100 are those with the highest difference: $0.63 - 0.46 = 0.17$. Even for this case the value of the Wald test ($W = 1.44$) is clearly insignificant at any conventional size of the test. Moreover, in models which generated very similar power terms, such as the asymmetric D-F one (1.27, 1.39; see column B^a_{DF}) the two coefficients were, as expected, not significantly different ($W = 0.24$). The null hypothesis could not be rejected even in the case of quite dissimilar estimated power terms, such as the symmetric C-D formulation (1.55, 1.19; see column B^s_{CD}): the value of the Wald test (W =3.85) is insignificant at about the 5% level. The only exception to this general finding were the CD pairwise tests for the symmetric trivariate model whose CAC 40's estimated power term (1.84) and fractional parameter (0.11) were significantly different than those of the DAX 30 (1.22 and 0.25 respectively).

The following table presents the results of the pairwise testing procedure for the three Asian countries. In the majority of cases the values of the coefficients d_i for the bivariate/trivariate (asymmetric) models are lower than the corresponding values for the univariate (symmetric) formulations. Similarly, the estimated power terms for the bivariate/trivariate models are lower than the corresponding values for the univariate formulations. For the univariate/bivariate models the values of the Wald tests in the table support the null hypothesis that the two estimated fractional parameters (d_1, d_2) and the two power term coefficients (δ_1, δ_2) are not significantly different from another. All univariate/bivariate specifications generated very similar long-memory coefficients. The symmetric univariate models (see column U^s) for Nikkei 225 and Straits Times are those with the highest difference: $0.42 - 0.31 = 0.11$. Even for this case the value of the Wald test ($W = 1.43$) is clearly insignificant at any conventional size of the test. Moreover, in models which generated very similar power terms, such as the asymmetric H-S one (1.42, 1.47; see column B^a_{HS}) the two coefficients were, as expected, not significantly different ($W = 0.10$). The null hypothesis could not be rejected even in the case of quite dissimilar estimated power terms, such as the symmetric H-N formulation (1.50, 1.79; see column B^s_{HN}): the value of the Wald test ($W = 4.43$) is insignificant at about the 5% level. The only exception to this general finding was the pairwise test involving the symmetric univariate model for Hang Seng whose estimated power term (1.67) was significantly lower than that of the univariate symmetric formulation for Nikkei 225 (2.07). Finally, the trivariate model generated very similar long-memory and power terms coefficientsonly in three out of the six cases.

4 Forecasting Methodology

4.1 Evaluation Criteria

As Poon and Granger (2003) point out volatility forecasting is an important task in financial markets, and it has held the attention of academics and practitioners over the last two decades.[23] In this section we examine the ability of the various univariate/multivariate fractionally integrated and power asymmetric ARCH models to forecast stock return volatility.[24]

Our full sample consists of 4,255 trading days and each model is estimated over the first 4,055 observations of the full sample, i.e. over the period 1st January 1988 to 16th July 2003. As a result the out-of-sample period is from 17th July 2002 to 22nd April 2004 providing 200 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 $t = T(=4,055),\ldots,T+199(=4,254)$, i.e. 200 one-step ahead forecasts are calculated.

The true underlying volatility process is unobservable, therefore in order to evaluate the predictive ability of various volatility formulations we need to have a valid proxy. Awartani and Corradi (2005) point out that in comparing the relative predictive accuracy of various models, if the loss function is quadratic, the use of squared returns ensures that we actually obtain the correct ranking of models.

A complete description of the evaluation process requires a specification of a loss function. As Andersen et al. (1999) point 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 et al. (1999) we shall not use any of the complex economically motivated criteria but instead we will report summary statistics based directly on the deviation between forecasts and realizations. Four out-of-sample forecast performance measures will be used to evaluate and compare the various models.

First, we employ the mean square error (MSE) statistic. Although MSE is one of the most commonly employed criterion in the existing literature, it is not necessarily the best criterion adopted when evaluating nonlinear volatility forecasts (Andersen et al., 1999). Consequently, we also report results from two

[23] Several empirical studies examine the forecast performance of various GARCH models. The survey by Poon and Granger (2003) provides, among other things, an interesting and extensive synopsis of them.

[24] For the literature in the forecasting performance of univariate fractionally integrated and power ARCH models see, among others, Degiannakis (2004), Hansen and Lunde (2006), Trino-Manuel Ñíguez (2007). In addition, Angelidis and Degiannakis (2005) examine whether a simple GARCH specification or a complex FIAPARCH model generates the most accurate forecasts in three areas: option pricing, risk management and volatility forecasting.

alternative evaluation criteria. The first is the QLIKE which is discussed in Bollerslev et al. (1994). This statistic is also employed by Hansen and Lunde (2005). In addition, we employ an error statistic which is used by Peters (2001). This is the adjusted mean absolute percentage error (AMAPE) (see table 10 below).

On the basis of several model selection techniques the superior fitting specification was the FIAPARCH one (see section 3). While such model fitting investigations provide useful insights into volatility, the specifications 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 formulations to forecast daily stock return volatility. Table 11 gives a relative indication of overall forecasting performance. Calculated values are provided for four different forecasting performance measures across eight stock returns, so in total we have thirty two cases.

On the basis of several model selection techniques the superior fitting specification was the FIAPARCH one (see section 3). While such model fitting investigations provide useful insights into volatility, the specifications 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 formulations to forecast daily stock return volatility. For each index we calculated the three forecast error statistics for the specifications APARCH, IAPARCH, FIA(P)ARCH($\delta = 1$), FIA(P)ARCH($\delta = 2$) and FIAPARCH in the univariate, bivariate and (where possible) trivariate version. Hence, overall 15 values of each forecast error statistic are available for each index. Instead of presenting all the figures, we decided to present in Table 11 only the best and the worst specification for each index as identified by the forecast error statistic. In addition, we tested whether the values of the forecast error statistics from the best and the worst model are statistically significant using the Diebold and Mariano (2005) test. Table 11 contains the corresponding p-values.

The results can be summarized as follows. Only in three out of the 24 cases a univariate model was chosen best. In the other 21 case either bivariate are trivariate specifications were selected. Both, MSE and AMAPE uniformly favor multivariate specifications. For the two American Indices in five out of the six cases a bivariate model is selected as being best.

For the S&P 500 as well as the TSE 300 all the forecast error statistics rank a univariate specification as being worst. The integrated or fractionally integrated model is favored in four out of the six cases. The results for the European indices show the close connection between the volatilities of the three indices. In five out of the nine cases a trivariate specification is chosen, in three cases a bivariate specification. In particular, the FTSE 100 and the CAC 40 appear to be closely related, and to a somewhat lesser extend the DAX 30 and the FTSE 100. Interestingly, only in one of the nine cases a fractionally integrated specification is selected. Moreover, the restriction that $\delta = 2$ characterizes with one exception the worst performing specification. Also among the Asian indices the bivariate and trivariate specifications are ranked best. Here, in seven out of the nine cases a fractionally integrated specification is chosen.

In summary, the best specifications as ranked by the forecast error statistics are multivariate models. For the American and Asian indices the long memory property appears to be important for the forecast performance, while for the European indices short memory specifications are dominating.

5 Discussion

5.1 The Empirical Evidence

Brooks et al. (2000) analyzed the applicability of the stable APARCH model to national stock market returns for various industrialized countries. However, as in all cases the estimated values of the persistence coefficients were quite close to one, there was a need to examine closely the possibility of long-memory persistence in the conditional volatility.

In our paper, strong evidence has been put forward suggesting that the conditional volatility for eight national stock indices is best modeled as a FIAPARCH process. On the basis of Wald tests and information criteria the fractionally integrated model provides statistically significant improvement over its integrated counterpart. One can also reject the more restrictive stable process, and consequently all the existing specifications (see Ding et al. 1993) nested by it in favor of the fractionally integrated parameterization. Hence, our analysis has shown that the FIAPARCH formulation is preferred to both the stable and the integrated ones. In other words, the fractionally integrated process appeared to have superior ability to differentiate between stable specifications and their integrated alternatives.

The Bollerslev formulation is nested within the power specification. Brooks et al. (2000) applied the likelihood ratio test to this nested pair. The results of this test were mixed as far as supporting the presence of power effects is concerned. For the German and French indices there was strong evidence of power effects. For a further two countries (US and Japan) there was mild evidence and for Hong Kong there was only weak evidence in support of the power specification. In contrast, United Kingdom, Canada and Singapore show no evidence of power effects as the Bollerslev formulation could not be rejected in favor of the power one.

Moreover, the Taylor/Schwert specification is nested within the power model. For all countries tested, with the exceptions of Hong Kong and Singapore, the test statistics indicated a preference for the Taylor/Schwert formulation over the power specification. Accordingly, Brooks et al. (2000) concluded that allowing the power term to take on values other than unity, did not significantly enhance the model. In other words there was a lack of evidence to suggest the need of power effects in the absence of long-range volatility dependence, as the likelihood ratio tests produced insignificant calculated values indicating an inability to reject the Taylor/Schwert formulation over the power specification for eight of the national indices tested.

The results for the more general FIAPARCH model are in stark contrast. According to our analysis all eight countries show strong evidence (both the likelihood ratio and Wald tests produce significant calculated values) of power effects when long-memory persistence in the conditional volatility have been taken into account, as both the Bollerslev and Taylor/Schwert specifications were rejected in favor of the power formulation. Further, comparing the pairwise testing results of the log-likelihood procedures to the relative model rankings provided by the four alternative criteria we observed that the findings were generally robust. That is, where the log-likelihood results provided unanimous support for the FIAPARCH specification over either the Bollerslev or Taylor/Schwert (asymmetric) FIGARCH formulations, the model selection criteria concurred without exception. Thus, the inclusion of a power term and a fractional unit root in the conditional variance equation appear to augment the model in a worthwhile fashion.

Finally, we should also emphasize that the above results were robust to the dimension of the process. That is, the evidence obtained from the univariate models on the superiority of the FIAPARCH specification was reinforced by the multivariate processes. It is noteworthy that the results are not qualitatively altered by changes in the dimension of the model.

5.2 Possible Extensions

The main goal of this paper was to explore the issue of how generally applicable the ccc M-FIAPARCH formulation is to a wide range of national stock market returns. Possible extensions of this article can go in different directions. Kim et al. (2005) use a bivariate ccc FIAPARCH-in-mean process to model the volume-volatility relationship. In the context of our analysis, incorporating volumes either in the mean or in the variance specification or in both could be at work. We look forward to sorting this out in future work. He and Teräsvirta (1999) emphasize that if the standard Bollerslev type of model is augmented by the power term, the estimates of the other variance coefficients almost certainly change. More importantly, Karanasos and Schurer (2008) find that the relationship between the level of the process and its conditional variance, as captured by the in-mean parameter, is sensitive to changes in the values of the power term (see also Conrad and Karanasos, 2008b). Therefore, one promising avenue would be to adapt the multivariate model in a way that incorporates in-mean effects.

Moreover, Conrad and Karanasos (2008a) consider a formulation of the extended constant or time varying conditional correlation M-GARCH specification which allows for volatility feedback of either sign, i.e., positive or negative. We have not be able, in such a short space, to deal with the unrestricted extended (and/or time varying conditional correlation) version of the M-FIAPARCH model. We should also emphasize that the most commonly used measures of stock volatility apart from the conditional variance from an ARCH type of process is the realized volatility (see Andersen et al., 2003, and Conrad and Lamla, 2007) and the range-based intraday estimator (see Karanasos and Kartsaklas, 2008). In addition, Bai and Chen (2007) consider testing distributional assumptions in M-GARCH formulations based on empirical processes. To highlight the importance of using alternative measures of volatility and multivariate distributions in order to model the national stock market returns (and forecast their variances) we should have to go into greater detail than space in this paper permits.

Further, Baillie and Morana (2007) introduce a new long-memory volatility specification, denoted by Adaptive FIGARCH, which is designed to account for both long-memory and structural change in the conditional variance process. One could provide an enrichment of the M-FIAPARCH by allowing the intercepts of the two means and variances to follow a slowly varying function as in Baillie and Morana (2007). This is undoubtedly a challenging yet worthwhile task. Finally, Pesaran and Timmermann (2002) suggest an estimation strategy that takes into account breaks and provide gains in forecasting ability. Pesaran et al. (2006) provide a new approach to forecasting time series that are subject to discrete structural breaks. Their results suggest several avenues for further research.

6 Conclusion

The purpose of the current paper was to consider the applicability of the multivariate fractionally integrated asymmetric power ARCH model to the national stock market returns for eight countries. It was found that the M-FIAPARCH formulation captures the temporal pattern of volatility for observable returns better than previous parameterizations. It also 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 specifications. The results reject both the stable and integrated null hypotheses. This is consistent with the conditional volatility profiles in Gallant et al. (1993), which suggest that shocks to the variance are very slowly damped, but do die out. Moreover, all eight countries show strong evidence of power effects when asymmetries and/or long-memory persistence in the conditional volatility have been taken into account, as both the Bollerslev and Taylor/Schwert formulations were rejected in favor of the power specification. As convincingly argued by Brooks et al. (2000), for high frequency data which have a non-normal error distribution the presumption of an obvious superiority of a squared power term is lost. Other power transformations are more appropriate. Finally, the apparent similarity of the fractional differencing and power terms suggest that the M-FIAPARCH model has a quite general empirical validity across many different markets.

References

[1] Andersen, T. G., Bollerslev, Diebold, F. X., Labys, P. 2003. Modeling and forecasting relaized volatility. *Econometrica* 71(2), 579-625.

[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(5), 457-477.

[3] Angelidis, T., Degiannakis, S., 2005. Forecasting one-day-ahead VaR and intra-day realized volatility in the Athens stock exchange market. *Managerial Finance*.

[4] Awartani, B. M. A., Corradi, V., 2005. Predicting the volatility of the S&P-500 stock index via GARCH models: the role of asymmetries. *International Journal of Forecasting* 21, 167-183.

[5] Bai, J., Chen, Z., 2008. Testing multivariate distributions in GARCH models. *Journal of Econometrics* 143, 19-36.

[6] Baillie, R., 1996. Long memory processes and fractional integration in econometrics. *Journal of Econometrics* 73, 5–59.

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

[8] Baillie, R., Morana, C., 2007. Modeling long memory and structural breaks in conditional variances: an adaptive FIGARCH approach. ICER, Working Paper No. 11/2007.

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

[10] 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.

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

[12] 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.

[13] 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.

[14] Campos, N., Karanasos, M., 2008. Growth, volatility and political instability: non-linear time-series evidence for Argentina, 1890-2000. *Economics Letters*, forthcoming.

[15] Campos, N., Karanasos, M., Tan, B., 2008a. Two to tangle: growth effects of financial development and political instability in Argentina (1890-2000). Brunel University, Unpublished Paper.

[16] Campos, N., Karanasos, M., Tan, B., 2008b. From riches to rags and back? finance and growth in Argentina since 1890. Brunel University, Unpublished Paper.

[17] Caporin, B., 2003. Identification of long memory in GARCH models. *Statistical Methods and Applications* 12, 133-151.

[18] Conrad, C., Haag, B., 2006. Inequality constraints in the FIGARCH model. *Journal of Financial Econometrics* 3, 413-449.

[19] Conrad, C., Karanasos, M., 2006. The impulse response function of the long memory GARCH model. *Economics Letters* 90, 34-41.

[20] Conrad, C., Karanasos, M., 2008a. Negative volatility spillovers in the unrestricted ECCC-GARCH model. ETH Zurich, KOF Working Paper No. 189.

[21] Conrad, C., Karanasos, M., 2008b. Power ARCH in-mean model: a simulation study. In preparation.

[22] Conrad, C., Lamla, M. J., 2007. The High-frequency response of the EUR-US dollar exchange rate to ECB monetary policy Announcements. ETH Zurich, KOF Swiss Economic Institute, Working Paper No. 07-174.

[23] 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.

[24] Dark, J., 2004. Bivariate error correction FIGARCH and FIAPARCH models on the Australian All Ordinaries Index and its SPI futures. Monash University, Unpublished paper.

[25] Davidson, J., 2008. Time Series Modelling 4.26. http://www.timeseriesmodelling.com.

[26] Degiannakis, S., 2004. Volatility forecasting: evidence from a fractional integrated asymmetric power ARCH skewed-t model. *Applied Financial Economics* 14, 1333-1342.

[27] Diebold, F. X., and Mariano, R.S., 1995. Comparing Predictive Accuracy. *Journal of Business & Economic Statistics* 13(3), 253-263.

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

[29] 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.

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

[31] 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.

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

[33] 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.

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

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

[36] Giot, P., Laurent, S., 2003. Value-at-risk for long and short positions. *Journal of Applied Econometrics* 18, 641– 66

[37] 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.

[38] 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.

[39] 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.

[40] Granger, C., Hyung, N., 2004. Occasional structural breaks and long memory with an application to the S&P 500 absolute stock returns. *Journal of Empirical Finance* 11, 399-421.

[41] Hansen, P. R., Lunde, A., 2005. A forecast comparison of volatility models: does anything beat a GARCH(1,1). *Journal of Applied Econometrics,* 20, 873-889.

[42] Hansen, P. R., Lunde, A., 2006. Consistent ranking of volatility models. *Journal of Econometrics* 131, 97-121.

[43] He, C., Teräsvirta, T., 1999. Statistical properties of the asymmetric power ARCH model, in Engle, R. F., and H., White (eds), Cointegration, causality and forecasting. Festchrift in honour of Clive W. J. Granger. Oxford University Press, Oxford, 462-474.

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

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

[46] Karanasos, M., Kim, J., 2006. A re-examination of the asymmetric power ARCH model. *Journal of Empirical Finance* 13, 113-128.

[47] Karanasos, M., Kartsaklas, A., 2008. Dual long-memory, structural breaks and the link between turnover and the range-based volatility. Brunel University, Unpublished Paper.

[48] Karanasos, M., Psaradakis, Z., Sola, M., 2003. On the autocorrelations properties of long-memory GARCH processes. *Journal of Time Series Analysis* 25, 265-281.

[49] Karanasos, M., Schurer, S., 2008. Is the relationship between inflation and its uncertainty linear? *German Economic Review*, forthcoming.

[50] Kim, J., Kartsaklas, A., Karanasos, M., 2005. The volume-volatility relationship and the opening of the Korean stock market to foreign investors after the financial turmoil in 1997. *Asia-Pacific Financial Markets* 12, 245-271.

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

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

[53] Paolella, M. S., 1997. Using flexible GARCH models with asymmetric distributions. Institute of Statistics and Econometrics, Christian Albrechts University to Kiel, Working paper No. 111/97.

[54] Pesaran, M. H., Pettenuzo, D., Timmermann, A., 2006. Forecasting time series subject to multiple structural breaks. *Review of Economic Studies* 73, 1057-1084.

[55] Pesaran, M., Timermann, A., 2002. Market timing and return prediction under model instability. *Journal of Empirical Finance*, 9, 495-510.

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

[57] Poon, S. H., Granger, C. W. J., 2003. Forecasting volatility in financial markets: a review. *Journal of Economic Literature* XLI, 478-539.

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

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

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

[61] Trino-Manuel Ñíguez., 2007. Volatility and VaR forecasting in the Madrid Stock Exchange. *Spanish Economic Review*

[62] Tse, Y. K., 1996. The conditional heteroscedasticity of the yen-dollar exchange rate. Department of Economics and Statistics, National University of Singapore, unpublished manuscript.

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

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

Table 4: Bivariate AR-FI(A)PARCH models (ML Estimation)

	SP-T		C-D		C-F		D-F		H-N		H-S		N-S	
	SP	T	C	D	C	F	D	F	H	N	H	S	N	S
ζ_i	-0.05^*	0.17	-0.03	0.02^*	0.05	0.04	0.01^*	-0.03	0.05	-0.02	0.03	0.14	-0.03	0.15
	(-4.51)	(13.86)	(-2.63)	(1.53)	(3.88)	(2.70)	(0.37)	(-2.14)	(3.44)	(-1.07)	(1.79)	(9.16)	(-2.08)	(9.33)
β_i	0.46	0.33	0.50	0.62	0.35	0.45	0.55	0.42	0.57	0.46	0.08	–	0.43	–
	(4.78)	(2.27)	(3.94)	(9.00)	(1.55)	(1.48)	(5.51)	(1.93)	(3.70)	(3.82)	(2.85)		(3.41)	
ϕ_i	0.26	0.18	0.26	0.24	0.16	0.20	0.20	0.17	0.33	0.15	–	-0.02	0.14	-0.07
	(3.73)	(1.52)	(4.30)	(5.60)	(1.24)	(1.48)	(4.96)	(1.74)	(3.94)	(2.11)		(0.87)	(1.76)	(1.78)
γ_i	–	0.34	–	–	–	–	0.14	–	–	0.10	0.11	0.47	0.11	–
		(2.46)					(1.68)			(1.74)	(1.73)	(3.16)	(2.02)	
δ_i	1.85	1.59	1.55	1.23	1.76	1.55	1.28	1.36	1.49	1.69	1.42	1.47	1.70	1.62
	(8.81)	(8.37)	(9.12)	(9.84)	(7.65)	(5.54)	(11.64)	(8.00)	(17.71)	(13.75)	(12.07)	(12.01)	(13.75)	(15.71)
d_i	0.22	0.21	0.30	0.44	0.24	0.29	0.40	0.28	0.36	0.35	0.16	0.13	0.33	0.23
	(5.50)	(5.25)	(3.00)	(6.28)	(2.18)	(1.61)	(4.44)	(2.15)	(3.18)	(5.04)	(7.58)	(5.79)	(5.15)	(6.64)
ρ	0.65		0.65		0.67		0.54		0.33		0.43		0.26	
	(21.33)		(20.54)		(20.90)		(19.48)		(11.03)		(17.02)		(12.32)	
υ	13.69		16.69		18.96		18.13		12.62		11.31		12.42	
	(9.85)		(6.76)		(6.94)		(6.06)		(11.03)		(11.44)		(10.47)	
Q_{12}	18.08	10.74	34.92	10.17	10.33	15.80	12.48	36.27	22.30	9.78	39.18	20.69	12.93	16.36
	[0.11]	[0.55]	[0.00]	[0.60]	[0.59]	[0.20]	[0.41]	[0.00]	[0.03]	[0.63]	[0.00]	[0.06]	[0.37]	[0.17]
Q^2_{12}	2.77	2.81	20.28	5.17	24.51	40.18	3.31	17.44	35.79	54.52	108.29	5.03	58.83	1.50
	[0.99]	[0.99]	[0.06]	[0.95]	[0.02]	[0.00]	[0.99]	[0.13]	[0.00]	[0.00]	[0.00]	[0.96]	[0.00]	[1.00]

Notes: For each of the seven pairs of indices, Table 4 reports ML parameter estimates for the bivariate AR-FI(A)PARCH model. SP-T denotes the bivariate process for the S&P 500 and TSE 300 indices. C-D, C-F and D-F indicate the three bivariate models for the European indices. H-N, H-S and N-S stands for the three bivariate specifications for the Asian indices. * For the S&P 500 and DAX 30 indices we estimate AR models of order 3 and 4 respectively. The numbers in parentheses are t-statistics. Q_{12} and Q^2_{12} are the 12th order Ljung-Box tests for serial correlation in the standardized and squared standardized residuals respectively. The numbers in brackets are p-values.

Table 5: Tests for restrictions on fractional differencing and power term parameters

H_0:		d's= 0	d's= 1		δ's= 1	δ's= 2
	d's	W	W	δ's	W	W
SP-T	0.22{0.04}-0.21 {0.04}	37[0.00]	432[0.00]	1.85 {0.21}-1.59 {0.19}	141[0.00]	141[0.00]
C-D	0.30 {0.10}-0.44 {0.07}	39[0.00]	241[0.00]	1.55 {0.17}-1.23 {0.12}	97[0.00]	124[0.00]
C-F	0.24 {0.11}-0.29 {0.18}	5[0.10]	112[0.00]	1.76 {0.23}-1.55 {0.28}	30[0.00]	81[0.00]
D-F	0.40{0.09}-0.28 {0.13}	25[0.00]	279[0.00]	1.29 {0.11}-1.36 {0.17}	130[0.00]	155[0.00]
H-N	0.36 {0.11}-0.35 {0.07}	36[0.00]	65[0.00]	1.49 {0.08}-1.69 {0.12}	319[0.00]	318[0.00]
H-S	0.16 {0.02}-0.13 {0.02}	nc	255[0.00]	1.42 {0.12}-1.47 {0.12}	228[0.00]	247[0.00]
N-S	0.33 {0.06}-0.23 {0.03}	77[0.00]	158[0.00]	1.70 {0.12}-1.62 {0.10}	341[0.00]	284[0.00]

Notes: For each of the seven pair of indices, Table 5 reports the values of the Wald (W) statistics of the unrestricted bivariate FI(A)PARCH and restricted (d's$= 0, 1$; δ's$= 1, 2$) models respectively. SP-T denotes the bivariate model for the S&P 500 and TSE 300 indices. C-D, C-F and D-F indicate the three bivariate models for the European indices. H-N, H-S and N-S denote the three bivariate models for the Asian indices. The numbers in $[\cdot]$ are p values. The numbers in $\{\cdot\}$ are standard errors. nc indicates that there was no convergence.

Table 6: Trivariate AR-FI(A)PARCH$(1, d, 1)$ models (ML Estimation)

	C-D-F			H-N-S*		
	C	D	F	H	N	S
β_i	0.19 (1.40)	0.43 (4.61)	0.22 (1.15)	0.39 (1.92)	0.38 (2.50)	0.78 (18.85)
ϕ_i	0.11 (0.90)	0.22 (3.35)	0.09 (0.61)	0.28 (1.56)	0.15 (1.53)	0.81 (22.08)
γ_i	-	-	-	0.02 (1.46)	0.07 (1.60)	–
δ_i	1.83 (10.95)	1.25 (9.52)	1.56 (7.12)	1.47 (13.36)	1.53 (10.20)	1.88 (11.75)
d_i	0.11 (4.16)	0.25 (5.43)	0.15 (3.27)	0.18 (4.50)	0.26 (3.71)	0.08 (4.39)
	C-D	D-F	C-F	H-N	N-S	H-S
ρ	0.66 (21.07)	0.56 (19.86)	0.68 (21.70)	0.32 (14.84)	0.25 (12.19)	0.43 (16.92)
υ	9.60 (17.36)			8.42 (20.54)		

Notes: Table 6 reports ML parameter estimates for the two trivariate (white noise) FI(A)PARCH$(1, d, 1)$ models. C-D-F and H-N-S denote the models for the European and Asian countries respectively. *For the Nikkei 225 and Straits Times indices we estimate AR(1) models. The numbers in parentheses are t-statistics.

Table 7: Tests for similarity of fractional and power terms (America)

	d				δ			
	U^s	U^a	B^s	B^a	U^s	U^a	B^s	B^a
S	0.30	nc	0.24	0.22	2.35	nc	2.00	1.86
T	0.24	0.18	0.27	0.23	2.24	1.47	1.68	1.51
W	0.51	nc	0.25	0.04	0.20	nc	2.38	2.59

Notes: U(B) denotes univariate(bivariate) models. The superscript (a) stands for (a)symmetric specifications. S(T) denotes S&P 500(TSE 300) index. nc indicates that there was no convergence. The last row (W) reports the corresponding Wald statistics. The 5% and 1% critical values are 3.84 and 6.63 respectively.

Table 8: Tests for similarity of fractional and power terms (Europe)

	U^s	U^a	B^s_{CD}	B^a_{CD}	B^s_{CF}	B^a_{CF}	B^s_{DF}	B^a_{DF}	T^s	T^a
						d				
C	0.52	nc	0.30	0.19	0.24	0.26			0.11	0.06
D	0.63	0.40	0.45	0.29			0.48	0.36	0.25	0.13
F	0.46	0.33			0.29	0.33	0.41	0.23	0.15	0.07
W	$\underset{CD}{0.50};\underset{CF}{0.14};\underset{DF}{1.44}$	0.11	4.16	1.24	0.26	0.05	0.75	6.00	$\underset{CD}{14.77};\underset{CF}{1.34};\underset{DF}{4.40}$	$\underset{CD}{2.97};\underset{CF}{0.02};\underset{DF}{5.21}$
						δ				
C	1.77	nc	1.55	1.59	1.76	1.74			1.84	1.87
D	1.69	1.24	1.19	1.18			1.35	1.27	1.22	1.35
F	1.86	1.53			1.55	1.51	1.40	1.39	1.56	1.75
W	$\underset{CD}{0.18};\underset{CF}{0.24};\underset{DF}{0.86}$	2.15	3.85	6.57	1.078	1.55	0.10	0.24	$\underset{CD}{11.75};\underset{CF}{1.24};\underset{DF}{2.09}$	$\underset{CD}{5.33};\underset{CF}{0.18};\underset{DF}{2.93}$

Notes: U, B and T denote univariate, bivariate and trivariate models respectively. The subscripts C, D and F indicate the CAC 40, DAX 30 and FTSE 100 indices respectively. The superscript (a)s stands for (a)symmetric specification. nc indicates that there was no convergence. The W rows report the corresponding Wald statistics. The 5% and 1% critical values are 3.84 and 6.63 respectively.

Table 9: Tests for similarity of fractional and power terms (Asia)

	U^s	U^a	B^s_{HN}	B^a_{HN}	B^s_{HS}	B^a_{HS}	B^s_{NS}	B^a_{NS}	T^s	T^a
					d					
H	0.39	0.20	0.38	0.36	0.26	0.16			0.20	0.07
N	0.42	nc	0.36	0.35			0.34	0.32	0.26	0.26
S	0.31	0.20			0.20	0.13	0.24	0.22	0.07	0.07
W	$\underset{HN}{0.10}$;$\underset{HS}{0.58}$;$\underset{NS}{1.43}$	$\underset{HS}{2.00}$	0.04	0.02	0.85	1.61	1.46	1.62	$\underset{HN}{1.66}$;$\underset{HS}{5.75}$;$\underset{NS}{35.81}$	$\underset{HN}{12.20}$;$\underset{HS}{0.03}$;$\underset{NS}{29.00}$
					δ					
H	1.67	1.28	1.50	1.49	1.49	1.42			1.46	1.60
N	2.07	nc	1.79	1.70			1.80	1.66	1.62	1.50
S	1.91	1.41			1.68	1.47	1.68	1.58	2.01	1.87
W	$\underset{HN}{\mathbf{7.21}}$;$\underset{HS}{2.23}$;$\underset{NS}{0.83}$	$\underset{HS}{0.76}$	4.43	1.66	1.59	0.10	0.60	0.19	$\underset{HN}{2.84}$;$\underset{HS}{18.49}$;$\underset{NS}{27.59}$	$\underset{HN}{1.11}$;$\underset{HS}{5.31}$;$\underset{NS}{27.59}$

Notes: U, B and T denote univariate, bivariate and trivariate models respectively. The subscripts N, H and S indicate the Nikkei 225, Hang Seng and Straits Times indices respectively. The superscript (a)s stands for (a)symmetric specification. nc indicates that there was no convergence.The W rows report the corresponding Wald statistics. The 5% and 1% critical values are 3.84 and 6.63 respectively.

Table 10: **Forecasting** evaluation criteria

MSE:	$k^{-1}\sum_{t=T+1}^{T+k}(\widehat{h}_t - r_t^2)^2$
QLIKE:	$k^{-1}\sum_{t=T+1}^{T+k}[\ln(\widehat{h}_t) + r_t^2/\widehat{h}_t]$
AMAPE:	$(k+1)^{-1}\sum_{t=T}^{T+k}\left\|(\widehat{h}_t - r_t^2)/(\widehat{h}_t + r_t^2)\right\|$

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. .

Table 11: Best versus worst ranked models

	MSE	QLIKE	AMAPE
S&P 500	B-FIAP vs. U-FIAP	B-IAP vs. U-FIAP	B-AP vs. U-FIAP
	[0.00]	[0.03]	[0.02]
TSE 300	B-FIAP vs. U-IAP	U-FIP vs. U-IAP	B-AP vs. U-IAP
	[0.14]	[0.00]	[0.00]
CAC 40	T-P vs. B_F-FIAP($\delta = 2$)	T-IP vs. B_F-FIAP($\delta = 2$)	T-IP vs. B_F-FIAP($\delta = 2$)
	[0.00]	[0.15]	[0.00]
DAX 30	B_F-AP* vs. U-FIAP	U-FIA($\delta = 1$) vs. B_C-FIAP($\delta = 2$)	B_F-AP vs. B_F-FIAP($\delta = 2$)
	[0.00]	[0.08]	[0.17]
FTSE 100	T-P vs. B_C-FIAP($\delta = 2$)	T-P vs. B_C-FIAP($\delta = 2$)	B_D-AP vs. B_C-FIAP($\delta = 2$)
	[0.00]	[0.01]	[0.00]
Hang Seng	B_S-FIA vs. U-AP	B_N-AP vs. T-FIAP	T-FIA($\delta = 2$) vs. U-FIAP($\delta = 2$)
	[0.00]	[0.02]	[0.26]
Nikkei 225	B_S-FIA($\delta = 1$) vs. U-FIAP	U-FI($\delta = 1$) vs. T-AP	T-FIA($\delta = 2$) vs. U-AP
	[0.12]	[0.03]	[0.67]
Straits Times	B_H-FIAP vs. B_N-IAP	B_H-FIA($\delta = 2$) vs. U-AP	T-FIAP vs. U-AP
	[0.00]	[0.01]	[0.00]

Notes: U, B and T stand for univariate, bivariate and trivariate specifications respectively. (F)I, A and P indicate (fractionally) integrated, asymmetric and power models respectively. The subscript F refers to the bivariate model with the FTSE 100 index. **The numbers in brackets are the p-values from the Diebold and Mariano (1995) test.**