Kiel
Working Papers

Kiel Institute
for the World Economy

Forecasting Volatility under Fractality, Regime-Switching, Long Memory and Student-*t* Innovations

by Thomas Lux and Leonardo Morales-Arias

No. 1532 | July 2009

Web: www.ifw-kiel.de

Kiel Institute for the World Economy, Düsternbrooker Weg 120, 24105 Kiel, Germany

Kiel Working Paper 1532 | July2009

Forecasting Volatility under Fractality, Regime-Switching, Long Memory and Student-*t* Innovations

Thomas Lux and Leonardo Morales-Arias

Abstract:
We examine the performance of volatility models that incorporate features such as long (short) memory, regime-switching and multifractality along with two competing distributional assumptions of the error component, i.e. Normal vs Student-*t*. Our precise contribution is twofold. First, we introduce a new model to the family of Markov-Switching Multifractal models of asset returns (MSM), namely, the Markov-Switching Multifractal model of asset returns with Student-*t* innovations (MSM-*t*). Second, we perform a comprehensive panel forecasting analysis of the MSM models as well as other competing volatility models of the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) legacy. Our cross-sections consist of all-share equity indices, bond indices and real estate security indices at the country level. Furthermore, we investigate complementarities between models via combined forecasts. We find that: (i) Maximum Likelihood (ML) and Generalized Method of Moments (GMM) estimation are both suitable for MSM-*t* models, (ii) empirical panel forecasts of MSM-*t* models show an improvement over the alternative volatility models in terms of mean absolute forecast errors and that (iii) forecast combinations obtained from the different MSM and (FI)GARCH models considered appear to provide some improvement upon forecasts from single models.

Keywords: Multiplicative volatility models, long memory, Student-t innovations, international volatility forecasting

JEL classification: C20, G12

Thomas Lux
Kiel Institute for the World Economy
24100 Kiel, Germany
Phone: +49 431-8814 278
E-mail: thomas.lux@ifw-kiel.de
E-mail: lux@bwl.uni-kiel.de

Leonardo Morales-Arias
Department of Economics
University of Kiel
Olshausenstr. 40
24118 Kiel, Germany
Phone: +49 431-880 3171
E-mail: moralesarias@economics.uni-kiel.de

Forecasting Volatility under Fractality, Regime-Switching, Long Memory and Student-t Innovations

Thomas Lux * and Leonardo Morales-Arias**

Abstract

We examine the performance of volatility models that incorporate features such as long (short) memory, regime-switching and multifractality along with two competing distributional assumptions of the error component, i.e. Normal vs Student-t. Our precise contribution is twofold. First, we introduce a new model to the family of Markov-Switching Multifractal models of asset returns (MSM), namely, the Markov-Switching Multifractal model of asset returns with Student-t innovations (MSM-t). Second, we perform a comprehensive *panel* forecasting analysis of the MSM models as well as other competing volatility models of the Generalized Autoregressive Conditional Heteroskedasticity (GARCH) legacy. Our cross-sections consist of all-share equity indices, bond indices and real estate security indices at the country level. Furthermore, we investigate complementarities between models via combined forecasts. We find that: (i) Maximum Likelihood (ML) and Generalized Method of Moments (GMM) estimation are both suitable for MSM-t models, (ii) empirical *panel* forecasts of MSM-t models show an improvement over the alternative volatility models in terms of mean absolute forecast errors and that (iii) forecast combinations obtained from the different MSM and (FI)GARCH models considered appear to provide some improvement upon forecasts from single models.

JEL Classification: C20, G12

Keywords: Multiplicative volatility models, long memory, Student-t innovations, international volatility forecasting

*Institute of Economics. Chair of Monetary Economics and International Finance and Kiel Institute for the World Economy. Correspondence: Department of Economics, University of Kiel, Olshausenstrasse 40, 24118 Kiel, Germany. Email: lux@economics.uni-kiel.de.

**Corresponding author. Chair of Monetary Economics and International Finance and Kiel Institute for the World Economy. Correspondence: Department of Economics, University of Kiel, Olshausenstrasse 40, 24118 Kiel, Germany. Email: moralesarias@economics.uni-kiel.de. Tel: +49(0)431-880-3171. Fax: +49(0)431-880-4383. Financial support of the European Commission under STREP contract no. 516446 and the Deutsche Forschungsgemeinschaft is gratefully acknowledged. We would also like to thank participants of the 2nd International Workshop on Computational and Financial Econometrics, Neuchatel, Switzerland, June 2008 and of the Econophysics Colloquium, Kiel Institute for the World Economy, Kiel, Germany, August 2008 for helpful comments and suggestions. The usual disclaimer applies.

1 Introduction

Volatility forecasting is of crucial importance for financial practitioners and academics. Accurate forecasts of volatility allow analysts to build appropriate models for risk management such as portfolio allocation, Value-at-Risk, option and futures pricing, etc. For these reasons, scholars have devoted a great deal of attention to developing parametric as well as non-parametric models to forecast future volatility (cf. Andersen et al. (2005a) for a recent review on volatility modeling and Poon and Granger (2003) for a review on volatility forecasting).

In this paper, we are interested in the performance of a new type of volatility model, the so-called Markov-Switching Multifractal Model (MSM) vis-*à*-vis its more time honored competitors from the (Generalized) Autoregressive Conditional Heteroskedasticity (GARCH) family. The former model is a causal analog of the earlier non-causal Multifractal Model of Asset Returns (MMAR) due originally to Calvet et al. (1997).

In contrast to mainstream volatility models, the MSM model can accommodate, by its very construction, the feature of *multifractality* via its hierarchical, multiplicative structure with heterogeneous components. *Multifractality* refers to the variations in the *scaling behavior* of various moments or to different degrees of long-term dependence of various moments. Long-term dependence in various moments of (mainly financial) data have been reported in various studies by economists and physicists so that this feature now counts as a well established stylized fact (cf. Ding et al. (1993), Lux (1996), Mills (1997), Lobato and Savin (1998), Schmitt et al. (1999), Vassilicos et al. (2004)). Empirical research in finance also provides us with more direct evidence in favor of the hierarchical structure of multifractal cascade models (cf. Muller (1997)).

It seems plausible that the higher degree of flexibility of MSM models in capturing different degrees of temporal dependence of various moments could also facilitate volatility forecasting. Indeed, recent studies have shown that the MSM models can forecast future volatility more accurately than traditional long memory and regime-switching models of the (G)ARCH family such as Fractionally Integrated GARCH (FIGARCH) and Markov-Switching GARCH (MSGARCH) (cf. Calvet and Fisher (2004), Lux and Kaizoji (2007), Lux (2008)).

It is also worthwhile to emphasize the intermediate nature of MSM models between "true" long-memory and regime-switching. It has been pointed out that it is hard to distinguish empirically between both types of structures and that even single regime-switching models could easily give rise to apparent long memory (cf. Granger and Terasvirta (1999)). MSM models generate what has been called "long-memory over a finite interval" and in certain limits converges to a process with "true" long-term dependence.[1] The MSM model combines features of both types of generating mechanisms in a very parsimonious way. The flexible regime-switching nature of the MSM model might also allow to integrate seemingly unusual time periods such as the Japanese bubble of the 1980s in a very convenient manner without resorting to dummies or specifically designed regimes (cf. Lux and Kaizoji (2007)). Nevertheless, the finance literature has only scarcely exploited MSM models so far. Most efforts with respect to volatility modeling have been directed towards refinements of GARCH-type models, stochastic volatility models and more recently realized volatility models (cf. Andersen and Bollerslev (1998), Andersen et al. (2003), Andersen et al. (2005b), Abraham et al. (2007)).

Up until now, the scarce literature on MSM models of volatility has only considered the Gaussian distribution for return innovations. However, recent studies have shown that out-of-sample forecasts of volatility models with Student-t innovations might improve upon those resulting from volatility models with Gaussian innovations (cf. Rossi and Gallo (2006), Chuang et al. (2007), Wu and Shieh (2007)). In addition, there could also be an interaction between the modeling of *fat tails* and dependency in volatility: if more extreme realisations are covered by a fat-tailed distribution, the estimates of the parameters measuring serial dependence of higher moments might change which also alters the forecasting capabilities of an estimated model.

In this article we examine the performance of various volatility models from the MSM and GARCH families along with two competing distributional assumptions of the error component, i.e. Normal vs Student-t. Our precise contribution is twofold. First, we introduce a new model to the family of MSM models, the Markov-Switching Multifractal model of asset returns with Student-t innovations (MSM-t). This model is an extension of the MSM model with Normal innovations which can be estimated via Maximum Likelihood (ML) or Generalized Method of Moments (GMM) (cf. Calvet and Fisher (2004), Lux (2008)). Forecasting can be performed via

[1] In contrast to the combinatorial MMAR of Calvet et al. (1997), the MSM model has no asymptotic power-law behavior of its autocorrelation function. However, depending on the number of volatility components, a pre-asymptotic hyperbolic decay of the autocorrelation might be so pronounced as to be practically indistinguishable from "true" long memory (cf. Liu et al. (2007)).

Bayesian updating (ML) or best linear forecasts together with the generalized Levinson-Durbin algorithm (GMM). We investigate the in-sample and out-of-sample performance of the MSM-t model via Monte Carlo simulations.

Second, we perform a forecasting analysis of MSM vs (FI)GARCH models with Normal and Student-t innovations. By contrasting both sets of models we are able to empirically evaluate the performance of models which incorporate different characterizations of the latent volatility process: the MSM models which take account of multifractality, Markov-switching and (apparent) long memory against more traditional models of the GARCH legacy (GARCH, GARCH-t, FIGARCH and FIGARCH-t) which take account of short/long memory and autoregressive components. Furthermore, the wide variety of models considered here provides an interesting platform to study empirical out-of-sample complementarities between models via forecast combinations.

Given the recently witnessed international financial turmoil, it seems important to uncover the performance of various volatility models in international financial markets. Thus, the cross-sections chosen for our empirical analysis consist of all-share equity indices, bond indices and real estate security indices at the country level. We believe that the use of panel data is promising in two main aspects. First, in order not to generalize its usefulness, an interesting volatility model should perform adequately for a cross-section of markets and different asset classes. Second, testing volatility models for a cross-section of markets comes along with an augmentation of sample information and thus provides more power to statistical tests.

To preview some of our results, we confirm that ML and GMM estimation are both suitable for MSM-t models. We also find that using GMM plus linear forecasts leads to minor losses in efficiency compared to optimal Bayesian forecasts based on ML estimates. This justifies using the former approach in our empirical exercise which reduces computational costs significantly. Moreover, empirical *panel* forecasts of MSM-t models show an improvement over the alternative MSM models with Normal innovations in terms of mean absolute forecast errors while they seem to deteriorate for (FI)GARCH models with Student-t innovations in relation to their Gaussian counterparts. In terms of mean absolute errors, the MSM-t dominates all other models at long forecasting horizons for all asset classes. Lastly, forecast combinations obtained from the different MSM and (FI)GARCH models considered provide a clear improvement upon forecasts from single models.

The paper is organized as follows. The next section introduces the general framework of volatility modeling. Section three and four provide a short review of the MSM and (FI)GARCH volatility models. Section five presents the Monte Carlo experiments performed with respect to the MSM-t models. Section six addresses the results of our comprehensive *panel* empirical analysis of the different volatility models under inspection. The last section concludes. To save on space, technical details not discussed in the article can be provided upon request.

2 Theoretical framework of volatility

The following specification of financial returns is considered,

$$\Delta p_t = \upsilon_t + \sigma_t u_t, \tag{1}$$

where $\Delta p_t = \ln P_t - \ln P_{t-1}$, $\ln P_t$ is the log asset price and $\upsilon_t = E_{t-1}\Delta p_t$ is the conditional mean of the return series. A simple parametric model to describe the conditional mean is, for instance, a first order autoregressive model of the form $\upsilon_t = \mu + \rho\Delta p_{t-1}$. Different assumptions can be used for the distribution of u_t. For example, we may assume a Normal distribution, Student-t distribution, Logistic distribution, mixed diffusion, etc (cf. Chuang et al. (2007)). For the purpose of this article we consider two competing types of distributions for the innovations u_t, namely, a Normal distribution and a Student-t distribution. Defining $x_t = \Delta p_t - \upsilon_t$, the "centered" returns are modelled as,

$$x_t = \sigma_t u_t. \tag{2}$$

From the above general framework of volatility different parametric and non-parametric representations can be assumed for the latent volatility process σ_t. In what follows, we describe the new family of Markov-Switching Multifractal volatility models as well as the more time-honored

GARCH-type volatility models for the characterization of σ_t. Since the former models are a very recent addition to the family of volatility models, we devote most of the next sections to describing them and keep the explanation on the alternative volatility models short to save on space.

3 Markov-Switching Multifractal models

3.1 Volatility specifications

Instantaneous volatility σ_t in the MSM framework is determined by the product of k volatility components or multipliers $M_t^{(1)}, M_t^{(2)}, \ldots, M_t^{(k)}$ and a scale factor σ:

$$\sigma_t^2 = \sigma^2 \prod_{i=1}^{k} M_t^{(i)}. \tag{3}$$

Following the basic hierarchical principle of the multifractal approach, each volatility component will be renewed at time t with a probability γ_i depending on its rank within the hierarchy of multipliers and remains unchanged with probability $1 - \gamma_i$. Calvet and Fisher (2001) propose to formalize transition probabilities according to:

$$\gamma_i = 1 - (1 - \gamma_k)^{(b^{i-k})}, \tag{4}$$

which guarantees convergence of the discrete-time version of the MSM to a Poissonian continuous-time limit. In principle, γ_k and b are parameters to be estimated. Note that (4) or its restricted versions imply that different multipliers $M_t^{(i)}$ of the product (3) have different mean life times. However, previous applications have often used pre-specified parameters γ_k and b in equation (4) in order to restrict the number of parameters (cf. Lux (2008)). The MSM model is fully specified once we have determined the number k of volatility components and their distribution.

In the small body of available literature, the multipliers $M_t^{(i)}$ have been assumed to follow either a Binomial or a Lognormal distribution. Since one could normalize the distribution so that $E[M_t^{(i)}] = 1$, only one parameter has to be estimated for the distribution of volatility components. In this paper we explore the Binomial and Lognormal specifications for the distribution of multipliers. Following Calvet and Fisher (2004), the Binomial MSM (BMSM) is characterized by Binomial random draws taking the values m_0 and $2 - m_0$ $(1 \leq m_0 < 2)$ with equal probability (thus, guaranteeing an expectation of unity for all $M_t^{(i)}$). The model, then, is a Markov switching process with 2^k states. In the Lognormal MSM (LMSM) model, multipliers are determined by random draws from a Lognormal distribution with parameters λ and s, i.e.

$$M_t^{(i)} \sim LN(-\lambda, s^2). \tag{5}$$

Normalisation via $E[M_t^{(i)}] = 1$ leads to

$$\exp(-\lambda + 0.5s^2) = 1, \tag{6}$$

from which a restriction on the shape parameter can be inferred: $s = \sqrt{2\lambda}$. Hence, the distribution of volatility components is parameterized by a one-parameter family of Lognormals with the normalization restricting the choice of the shape parameter. It is noteworthy that the dynamic structure imposed by (3) and (4) provides for a rich set of different regimes with an extremely parsimonious parameterization. For increasing k there is, indeed, no limit to the number of regimes considered without any increase in the number of parameters to be estimated.

3.2 Estimation and forecasting

In a seminal study by Calvet and Fisher (2004), an ML estimation approach was proposed for the BMSM model. The log likelihood function in its most general form may be expressed as,

$$L(x_1, ..., x_T; \varphi) = \sum_{t=1}^{T} \ln g(x_t | x_1, ..., x_{t-1}), \tag{7}$$

where $g(x_t | x_1, ..., x_{t-1})$ is the likelihood function of the MSM model with various distributional assumptions. The parameter vector of the BMSM with Gaussian innovations is given by $\varphi = (m_0, \sigma)'$. On the other hand, the parameter vector of the BMSM with Student-t innovations is given by $\varphi = (m_0, \sigma, \nu)'$ where ν $(2 < \nu < \infty)$ is the distributional parameter accounting for the degrees of freedom in the density function of the Student-t distribution. When ν approaches infinity, we obtain a Normal distribution. Thus, the lower ν, the "fatter" the tail.

The greatest advantage of the ML procedure is that, as a by-product, it allows one to obtain optimal forecasts via Bayesian updating of the conditional probabilities $\Omega_t = \mathcal{P}(M_t = m^i | x_1, ..., x_t)$ for the unobserved volatility states $m^i, i = 1, ..., 2^k$. Although the ML algorithm was a huge step forward for the analysis of MSM models, it is restrictive in the sense that it works only for discrete distributions of the multipliers and is not applicable for, e.g., the alternative proposal of a Lognormal distribution. Due to the potentially large state space (we have to take into account transitions between 2^k distinct states), ML estimation also encounters bounds of computational feasibility for specifications with more than about $k = 10$ volatility components in the Binomial case.

To overcome the lack of practicability of ML estimation, Lux (2008) introduced a GMM estimator that is universally applicable to all possible specifications of MSM processes. In particular, it can be used in all those cases where ML is not applicable or computationally unfeasible. In the GMM framework for MSM models, the vector of BMSM parameters φ is obtained by minimizing the distance of empirical moments from their theoretical counterparts, i.e.

$$\widehat{\varphi}_T = \arg\min_{\varphi \in \Phi} f_T(\varphi)' A_T f_T(\varphi), \tag{8}$$

with Φ the parameter space, $f_T(\varphi)$ the vector of differences between sample moments and analytical moments, and A_T a positive definite and possibly random weighting matrix. Moreover, $\hat{\varphi}_T$ is consistent and asymptotically Normal if suitable "regularity conditions" are fulfilled (cf. Harris and Matyas (1999)). Within this GMM framework it becomes also possible to estimate the LMSM model. In the case of the LMSM model, the parameter vector $\vartheta = (\lambda, \sigma)'$ $(\vartheta = (\lambda, \sigma, \nu)')$ replaces φ in (8) when Normal (Student-t) innovations are assumed.

In order to account for the proximity to long memory characterizing MSM models, Lux (2008) proposed to use log differences of absolute returns together with the pertinent analytical moment conditions, i.e.

$$\xi_{t,T} = \ln |x_t| - \ln |x_{t-T}|. \tag{9}$$

The above variable only has nonzero autocovariances over a limited number of lags. To exploit the temporal scaling properties of the MSM model, covariances of various moments over different time horizons are chosen as moment conditions, i.e.

$$\mathrm{Mom}\,(T, q) = E\left[\xi^{q}_{t+T,T} \cdot \xi^{q}_{t,T}\right], \tag{10}$$

for $q = 1, 2$ and $T = 1, 5, 10, 20$ together with $E\left[x_t^2\right] = \sigma^2$ for identification of σ in the MSM model with Normal innovations. In the case of the MSM-t model, two sets of moment conditions are utilized in addition to (10), namely, one that considers $E\left[|x_t|\right]$ (GMM1) and the other one that considers $E\left[|x_t|\right]$, $E\left[x_t^2\right]$ and $E\left[|x_t^3|\right]$ (GMM2).

We follow most of the literature by using the inverse of the Newey-West estimator of the variance-covariance matrix as the weighting matrix for GMM1. We also adopt an iterative GMM scheme updating the weighting matrix until convergence of both the parameter estimates and the variance-covariance matrix of moment conditions is obtained. However, we note that including the third moment ($E\left[|x_t^3|\right]$) for data generated from a Student-t distribution would not guarantee convergence of the sequence of weighting matrices under our choice of the inverse of the Newey-West (or any other) estimate of the variance-covariance matrix. Therefore, estimates based on the usual choice of the weighting matrix would not be consistent. Thus we simply resort to using the identity matrix for GMM2 which guarantees consistency as all the regularity conditions required for GMM are met.

Since GMM does not provide us with information on conditional state probabilities, we cannot use Bayesian updating and have to supplement it with a different forecasting algorithm. To this end, we use best linear forecasts (cf. Brockwell and Davis (1991), c.5) together with the generalized Levinson-Durbin algorithm developed by Brockwell and Dahlhaus (2004). We first have to consider the zero-mean time series,

$$X_t = x_t^2 - E[x_t^2] = x_t^2 - \widehat{\sigma}^2, \tag{11}$$

where $\widehat{\sigma}$ is the estimate of the scale factor σ. Assuming that the data of interest follow a stationary process $\{X_t\}$ with mean zero, the best linear h-step forecasts are obtained as

$$\widehat{X}_{n+h} = \sum_{i=1}^{n} \phi_{ni}^{(h)} X_{n+1-i} = \boldsymbol{\phi}_n^{(h)} \mathbf{X}_n, \tag{12}$$

where the vectors of weights $\boldsymbol{\phi}_n^{(h)} = (\phi_{n1}^{(h)}, \phi_{n2}^{(h)}, ..., \phi_{nn}^{(h)})'$ can be obtained from the analytical auto-covariances of X_t at lags h and beyond. More precisely, $\boldsymbol{\phi}_n^{(h)}$ are any solution of $\boldsymbol{\Psi}_n \boldsymbol{\phi}_n^{(h)} = \boldsymbol{\kappa}_n^{(h)}$ where $\boldsymbol{\kappa}_n^{(h)} = (\kappa_{n1}^{(h)}, \kappa_{n2}^{(h)}, ..., \kappa_{nn}^{(h)})'$ denote the autocovariance of X_t and $\boldsymbol{\Psi}_n = [\kappa(i-j)]_{i,j=1,...,n}$ is the variance-covariance matrix.

4 Generalized Autoregressive Conditional Heteroskedasticity models

4.1 Volatility specifications

We shortly turn to the "competing" GARCH type volatility models to describe σ_t. The most common GARCH(1,1) model assumes that the volatility dynamics is governed by,

$$\sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta \sigma_{t-1}^2, \tag{13}$$

where the unconditional variance is given by $\sigma^2 = \omega(1-\alpha-\beta)^{-1}$ and the restrictions on the parameters are $\omega > 0, \alpha, \beta \geq 0$ and $\alpha + \beta < 1$. Various extensions to (13) have been considered in the financial econometrics literature. One of the major additions to the GARCH family are models that allow for long-memory in the specification of volatility dynamics. The FIGARCH model introduced by Baillie et al. (1996) expands the variance equation of the GARCH model by considering fractional differences. As in the case of (13) we restrict our attention to one lag in both the autoregressive term and in the moving average term. The FIGARCH(1,d,1) is given by,

$$\sigma_t^2 = \omega + \left[1 - \beta L - (1-\delta L)(1-L)^d\right] x_t^2 + \beta \sigma_{t-1}^2, \tag{14}$$

where L is a lag operator, d is the parameter of fractional differentiation and the restrictions on the parameters are $\beta - d \leq \delta \leq (2-d)3^{-1}$ and $d(\delta - 2^{-1}(1-d)) \leq \beta(d - \beta + \delta)$. The major advantage of model (14) is that the Binomial expansion of the fractional difference operator introduces an infinite number of past lags with hyperbolically decaying coefficients for $0 < d < 1$. For $d = 0$, the FIGARCH model reduces to the standard GARCH(1,1) model while for $d = 1$ the model reduces to an IGARCH(1,1) model. Note that in contrast to the MSM model, both GARCH and FIGARCH are *unifractal* models. While GARCH exhibits only short-term dependence (i.e. exponential decay of autocorrelations of moments) FIGARCH has homogeneous hyperbolic decay of the autocorrelation of its moments characterized uniquely by the parameter d.

4.2 Estimation and forecasting

The GARCH and FIGARCH models can be estimated via standard (Quasi) ML procedures as in (7). In the case of the GARCH(1,1) the parameter vector, say θ, replaces φ in (7), where $\theta = (\omega, \alpha, \beta)'$ ($\theta = (\omega, \alpha, \beta, \nu)'$) is the vector of parameters if Normal (Student-t) innovations are assumed. The h-step ahead forecast representation of the GARCH(1,1) is given by,

$$\hat{\sigma}^2_{t+h} = \hat{\sigma}^2 + (\hat{\alpha} + \hat{\beta})^{h-1} \left[\hat{\sigma}^2_{t+1} - \hat{\sigma}^2 \right], \tag{15}$$

where $\hat{\sigma}^2 = \hat{\omega}(1 - \hat{\alpha} - \hat{\beta})^{-1}$. In the case of the FIGARCH(1,d,1) the parameter vector, say ψ, replaces φ (7), where $\psi = (\omega, \alpha, \delta, d)'$ ($\psi = (\omega, \alpha, \delta, d, \nu)'$) is the vector of parameters if Normal (Student-t) innovations are assumed. Note that in practice, the infinite number of lags with hyperbolically decaying coefficients introduced by the Binomial expansion of the fractional difference operator $(1-L)^d$ must be truncated. We employ a lag truncation at 1000 steps as in Lux and Kaizoji (2007). The h-period ahead forecasts of the FIGARCH(1,d,1) model can be obtained most easily by recursive substitution, i.e.

$$\hat{\sigma}^2_{t+h} = \hat{\omega}(1 - \hat{\beta})^{-1} + \eta(L)\hat{\sigma}^2_{t+h-1}, \tag{16}$$

where $\eta(L) = 1-(1-\hat{\beta}L)^{-1}(1-\hat{\delta}L)(1-L)^{\hat{d}}$ can be calculated from the recursions $\eta_1 = \hat{\delta}-\hat{\beta}+\hat{d}$, $\eta_j = \hat{\beta}\eta_j + [(j-1-\hat{d})j^{-1} - \hat{\delta}]\pi_{j-1}$ where $\pi_j \equiv \pi_{j-1}(j-1-\hat{d})j^{-1}$ are the coefficients in the MacLaurin series expansion of the fractional differencing operator $(1-L)^d$.

5 Monte Carlo analysis

Monte Carlo studies with the MSM-t were performed along the lines of Calvet and Fisher (2004) and Lux (2008) in order to shed light on parameter estimation and out-of-sample forecasting via the MSM-t vis-*à*-vis the MSM model with Normal innovations. Monte Carlo experiments are reported in Tables 1 through 3.

insert Tables 1 and 2 around here

5.1 In-sample analysis

Table 1 shows the result of the Monte Carlo simulations of the BMSM-t via ML estimation and the two sets of moment conditions for GMM estimation (GMM1, GMM2) with a relatively small number of multipliers $k = 8$ for which ML is still feasible. The Binomial parameters are set to $m_0 = 1.3, 1.4, 1.5$ and the sample sizes are given by $T_1 = 2,500$, $T_2 = 5,000$ and $T_3 = 10,000$. As mentioned previously the admissible parameter range for m_0 is $m_0 \in [1,2]$ and the volatility process collapses to a constant if the latter parameter hits its lower boundary 1. The parameter corresponding to the Student-t distribution is set to $\nu = 5$ and $\nu = 6$. As in the case of Lux (2008), the main difference in our simulation set up to the one proposed in Calvet and Fisher (2004) is that we fix the parameters of the transition probabilities in (4) to $b = 2$ and $\gamma_k = 0.5$ which reduces the number of parameters for estimation to only three.

The simulation results show (as expected) that GMM estimates of m_0 are in general less efficient in comparison to ML estimates. The finite sample standard error (FSSE) and root mean squared error (RMSE) of the GMM estimates with $\nu = 5$ show that the estimated parameters for m_0 are more variable with lower T and smaller "true" values of m_0. As in the case of the MSM model with Normal innovations, biases and MSEs of the ML estimates for m_0 are found to be essentially independent of the true parameter values $m_0 = 1.3, 1.4, 1.5$. With respect to GMM estimates with the two different sets of moment conditions (GMM1, GMM2), both the bias and the MSEs decrease as we increase m_0 from 1.3 to 1.5. Interestingly, when the degrees of freedom are increased from $\nu = 5$ to $\nu = 6$ we find an overall increase in the bias and MSEs of m_0 via GMM1 while the bias and MSEs of m_0 via GMM2 decrease.

ML estimates of the distributional parameter ν show a relatively small bias although it seems to slightly increase for larger m_0 at $T = 2,500$. The variability of ν via ML is also found to be more pronounced than the variability of the Binomial parameter m_0. GMM estimates of ν have a larger bias and MSEs in comparison to ML estimates. As we move from $\nu = 5$ to $\nu = 6$, we find that the bias and MSEs of the parameter ν estimated via GMM1 and GMM2 become larger.

The quality of the estimates of the scale parameter σ at $\nu = 5$ is very similar under ML and GMM1 particularly when the sample size is increased. With respect to estimates of σ via GMM2, we find that the bias is somewhat larger in comparison to ML and GMM1. As in the case of the MSM model with Normal innovations, we find that the MSEs of σ increase for higher m_0 while they are more or less unchanged as we move from $\nu = 5$ to $\nu = 6$.

Table 2 displays the results of the Monte Carlo analysis of the MSM-t model with a setting that makes ML estimation computationally infeasible, that is, the BMSM-t and LMSM-t models with $k = 10$.[2] As in the previous experiments, the simulations are performed with $m_0 = 1.3, 1.4, 1.5$, $\nu = 5, 6$ and the same logic is applied to the LMSM model for which the location parameter of the continuous distribution is set to $\lambda = 0.05, 0.1, 0.15$. Note that the admissible space for λ is $\lambda \in [0, \infty)$. As in the case of the Binomial parameter m_0, when the Lognormal parameter hits its lower boundary at 0, the volatility process collapses to a constant. To save on space, the simulations are only presented with $T = 5,000$.

The results of the simulations indicate that the Binomial parameter m_0 estimated via GMM1 or GMM2 are practically invariant to higher number of components k, both in terms of bias and MSEs for the parameter values $m_0 = 1.3, 1.4, 1.5$ and $\nu = 5$. The bias and MSEs of m_0 usually increase in GMM1 as we increase the degrees of freedom from $\nu = 5$ to $\nu = 6$. As in the BMSM

[2]Simulation results for $k = 15, 20$ are qualitatively similar and can be provided upon request.

model with Normal innovations, the bias and the variability of σ increase with k as it becomes hard to discriminate between very long-lived volatility components and the constant scale factor (cf. Lux (2008)). The distributional parameter ν is found to be relatively invariant for GMM1 and GMM2 when $k = 10$ in relation to $k = 8$. Bias and MSEs of the distributional parameter ν increase for GMM1 and GMM2 as we move from $\nu = 5$ to $\nu = 6$. In the LMSM-t model, we find that biases and MSEs for λ at $\nu = 5$ are somewhat larger for GMM1 than GMM2. As for the BMSM-t, bias and MSEs of λ are relatively invariant for larger k buy they usually increase as we move from $\nu = 5$ to $\nu = 6$.

Summing up, we find that the Monte Carlo simulations for the in-sample performance of the Binomial and Lognormal MSM models with Student-t innovations "point" into the same direction as those of Lux (2008) for the MSM model with Gaussian innovations: while GMM is less efficient than ML, it comes with moderate biases and moderate standard errors. The efficiency of both GMM algorithms also appear quite insensitive with respect to the number of multipliers.

insert Table 3 around here

5.2 Out-of-sample analysis

Table 3 shows the forecasting results from optimal forecasts (ML) and best linear forecasts (GMM) of the BMSM-t model. The out-of-sample MC analysis is performed within the same framework as the in-sample analysis when comparing ML and GMM procedures. That is, we set $k = 8$ and evaluate the forecasts for the BMSM-t model with parameters $m_0 = 1.3, 1.4, 1.5$, $\sigma = 1$ and $\nu = 5, 6$. In our Monte Carlo experiments, we also imposed a lower boundary $\underline{\nu} = 4.05$ as a constraint in the GMM estimates as otherwise forecasting with the Levinson-Durbin algorithm would have been impossible.

In the forecasting simulations we set $T = 10,000$ and use $T = 5,000$ for in-sample estimation and $T = 5,000$ for out-of-sample forecasting in order compare them with the results of the Gaussian MSM models in Lux (2008). The forecasting performance of the models is evaluated with respect to their mean squared errors (MSE) and mean absolute errors (MAE) standardized relative to the in-sample variance which implies that values below 1 indicate improvement against a constant volatility model. Relative MSE and MAE are averages over 400 simulation runs.

The results basically show that, similarly as for the Gaussian MSM models, the loss in forecasting accuracy when employing GMM as opposed to ML is small particularly when compared against GMM2. Thus, the lower efficiency of GMM does not impede its forecasting capability in connection with the Levinson-Durbin algorithm. Both MSE and MAE measures show improvement when the parameter m_0 increases from 1.3 to 1.5 while they seem to deteriorate for longer horizons although only marginally. Interestingly, we find that GMM2 based forecasts even improve in terms of MAEs relative to ML based forecasts for $h \geq 5$ so that it appears entirely justified to resort to the computationally parsimonious GMM2 estimation and linear forecasts in our subsequent empirical part.[3]

6 Empirical analysis

insert Table 4 around here

In this section we turn to the results of our empirical application to compare the in-sample and out-of-sample performance of the different volatility models discussed previously. We follow a similar approach to the *panel* empirical analysis of volatility forecasting performed for the Tokyo Stock Exchange in Lux and Kaizoji (2007). However, here we concentrate on three new different cross-sections of asset markets, namely, all-share stock indices ($N = 25$), 10-year government bond market indices ($N = 11$), and real estate security indices ($N = 12$) at the cross-country level. The sample runs from 01/1990 to 01/2008 at the daily frequency which leads to 4697 observation from which 2,500 are used for in-sample estimation and the remaining observations for out-of-sample forecasting. The data is obtained from Datastream and the countries were chosen upon data availability for the sample period covered. Specific countries for each of the three asset markets are presented in Table 4. In the following discussions we refer to statistical significance at the 5% level throughout.

[3]Out-of-sample forecasts of the lognormal MSM-t models behave very similar, but they are not displayed here because of the lack of a ML benchmark.

insert Tables 5 and 6 around here

6.1 In-sample analysis

For our in-sample analysis we account for a constant and an AR(1) term in the conditional mean of the return data as in (2). Results of the Mean Group (MG) estimates of the parameters of the (FI)GARCH and MSM models explained in previous sections are reported in Table 5 and Table 6, respectively. Mean Group estimates are obtained by averaging individual market estimates. We also report minimum and maximum values of the estimates obtained to have an idea about the distribution of the parameters across the countries under inspection.

We find in the case of the GARCH model that there is on average a statistically significant effect of past volatility on current volatility ($\bar{\beta}$) and of past squared innovations on current volatility ($\bar{\alpha}$) in all three markets at the 5% significance level (Table 5). The results are qualitatively the same with respect to the estimates $\bar{\beta}$ and $\bar{\alpha}$ in the case of the GARCH-t. The distributional parameter ($\bar{\nu}$) is on average greater than 4 in all three markets and statistically significant.

Taking into account long memory and Student-t innovations via the FIGARCH specification we find that there is a statistically significant average effect of past volatility ($\bar{\beta}$) and past squared innovations ($\bar{\delta}$) on current volatility in all three markets. FIGARCH also provides evidence for the presence of long memory as given by the MG estimate of the differencing parameter $\bar{d}$ in the three cross-sections (Table 5). When we consider the FIGARCH-t we find the same qualitative results for the average impact of the parameters $\bar{\beta}$, $\bar{\delta}$ and $\bar{d}$ as in the FIGARCH and the same qualitative results of the distributional parameter $\bar{\nu}$ as with the GARCH-t model.

In-sample estimation of the BMSM and LMSM models with Normal and Student-t innovations is restricted to GMM since ML estimation with panel data requires a tremendous amount of time for $k > 8$. The estimation procedure for the MSM models consists in estimating the models for each country in each of the stock, bond and real estate markets for a cascade level of $k = 10$. The choice of the number of cascade levels is motivated by previous findings of very similar parameter estimates for all k above this benchmark (Liu et al. (2007), Lux (2008)). Note, however, that forecasting performance might nevertheless improve for $k > 10$ and proximity to temporal scaling of empirical data might be closer. Our choice of the specification $k = 10$ is, therefore, a relatively conservative one.[4]

For space considerations we only present the results of the MSM-t models estimated with the second set of moment conditions (GMM2) given that we found that this set of moment conditions produced more accurate forecasts in the Monte Carlo Simulations. We have also restricted the parameter ν by using 4.05 as a lower bound in order to employ best linear forecasts. Nevertheless, we found very few cases where $\nu < 4$ in both (FI)GARCH and MSM models.

With respect to the BMSM model, the mean Binomial parameter $\bar{m}_0$ is statistically different from the benchmark case $\bar{m}_0 = 1$ in all three markets (Table 6). In the LMSM model, we find that the mean Lognormal parameter $\bar{\lambda}$ has a value which is statistically different from zero in all three asset markets. In the case of the BMSM-t, the mean distributional parameter $\bar{\nu}$ obtains a value that is statistically significant in all the three markets. Considering the LMSM-t we obtain similar qualitative results as for the BMSM-t in terms of the average parameters $\bar{\sigma}$ and $\bar{\nu}$.

Summarizing the in-sample results at the aggregate level, we find that there is (on average) a statistically significant effect of past volatility and long-memory on current volatility as well as evidence of multifractality and *fat tails* in return innovations. It is also noteworthy, that in many cases, the mean multifractal parameters $\bar{m}_0$ and $\bar{\lambda}$ turn out to be different for the models with Student-t innovations from those with Normal innovations. Since higher m_0 and λ lead to more heterogeneity and, therefore, more extreme observations, we see a trade-off between parameters for the fat-tailed innovations and those governing temporal dependence of volatility. What differences these variations in multifractal parameters make for forecasting, is investigated below.

insert Tables 7 and 8 around here

[4] In our case, results for $k = 15$ and $k = 20$ are practically the same as with $k = 10$. Details are available upon request.

6.2 Out-of-sample analysis

In this section we turn to the discussion of the out-of-sample results. Forecasting horizons are set to 1, 5, 20, 50 and 100 days ahead. We have used only one set of in-sample parameter estimates and have not re-estimated the models via rolling window schemes because of the computational burden that one encounters with respect to ML estimation of the FIGARCH models. We have also experimented with different subsamples but we have found no qualitative difference with respect to the current in-sample and out-of-sample window split which is roughly about half for in-sample estimation and half for out-of-sample forecasting.

In order to compare the forecasts across models we use the principle of relative MSE and MAE as previously mentioned. That is, the MSE and MAE corresponding to a particular model are given in percentage of a naive predictor using historical volatility (i.e. the sample mean of squared returns of the in-sample period). We also report the number of statistically significant improvements of a particular model against a benchmark specification via the Diebold and Mariano (1995) test. The latter test allows us to test the null hypothesis that two competing models have statistically equal forecasting performance.[5]

6.2.1 Single models

Results of the average relative MSE and MAE and corresponding standard errors of the out-of-sample forecasts from the different models are reported in Table 7. We find that at short horizons, $h = 1$, GARCH and FIGARCH models obtain lower average MSE and MAE than the BMSM and LMSM models in all three markets. However, the variability of the MSE and MAE in the MSM models are in general lower than those of the (FI)GARCH models. At horizons over 20 days, GARCH models (with Normal or Student-t innovations) show a deterioration in MSE and MAE measures hinting at their inability to accurately forecast volatility at higher horizons. In contrast, MSE and MAE resulting from the FIGARCH models (with Normal or Student-t innovations) are in general lower and more stable across horizons.

With respect to the MSM models (with Normal or Student-t innovations) we find that they produce MSEs and MAEs which are lower than one in stock and real estate markets. We also

[5]Details about the computation of MSEs and MAEs with panel data and corresponding standard errors and the count test based on the Diebold and Mariano (1995) statistics can be provided upon request.

find that the forecasts from the MSM models are much more homogeneous and less variable across horizons than those of the (FI)GARCH models whose forecasts usually deteriorate as the horizon is increased beyond $h = 20$. Comparing forecasts of the BMSM versus LMSM we find that the models produce qualitatively similar forecasts in terms of MSEs and MAEs.

Diagnosing forecasts from the models with Normal vs. Student-t innovations, we find that neither GARCH nor FIGARCH models produce lower MSEs and MAEs on average over the three markets when Student-t innovations are employed. Results are different in MSM models for which we find Student-t innovations to improve forecasting precision over all three markets in particular at horizons $h \geq 20$. In fact, Table 8 shows that, in terms of MAEs, there is a larger number of statistically significant improvements against historical volatility with the BMSM-t and LMSM-t models in comparison to their Gaussian and (FI)GARCH counterparts. Interestingly, the LMSM-t outperforms all other models in all markets in terms of MAEs when $h \geq 50$ and seems to provide for a sizable gain in forecasting accuracy at long horizons.

Note that the MSM models also showed some sensitivity of parameter estimates on the distributional assumptions (Normal vs. Student-t). As it seems, the volatility models react quite differently to different distributions of innovations u_t: on the one hand, the transition to Student-t was not reflected in remarkable changes of estimated parameters for (FI)GARCH models and their forecasting performance, if anything, slightly deteriorates under fat-tailed innovations. On the other hand, the effect of distributional assumptions on MSM parameters was more pronounced and their forecasting performance appears to be superior under Student-t innovations throughout our samples. Taken together, we see different patterns of interaction of conditional and unconditional distributional properties. This indicates that alternative models may capture different facets of the dependency in second moments so that there would be a potential gain from combining forecasts (a topic explored below).

insert Tables 9 and 10 around here

6.2.2 Combined forecasts

A particular insight from the methodological literature on forecasting is that it is often preferable to combine alternative forecasts in a linear fashion and thereby obtain a new predictor (cf.

Granger (1989), Aiolfi and Timmermann (2006)). We analyze forecast complementarities of (FI)GARCH and MSM models by addressing the performance of combined forecasts. The forecast combinations are computed by assigning each single forecast a weight equal to a model's empirical frequency of minimizing the absolute or squared forecast error over realized past forecasts. To take account of structural variation we update the weighting scheme over the 20 most recent forecast errors so that despite linear combinations of forecasts, the influence of various components is allowed to change over time via flexible weights.[6]

Tables 9 and 10 report the results of the forecasting combination exercise. Our forecast combination strategy consists in considering whether forecast combinations of (FI)GARCH models, MSM models or both families of models lead to an improvement upon forecasts from single models. Our results put forward that they generally do. This is in line with the empirical result of Lux and Kaizoji (2007) that the rank correlations of forecasts obtained from certain volatility models are quite low, hinting at room for improvement upon forecasts from single models with forecast combinations.

We start by considering the results of the forecast combinations of the (FI)GARCH models (Tables 9 and 10). Three different combination strategies are presented denoted CO1, CO2 and CO3. The first combination strategy (CO1) is given by the (weighted) linear combination between FIGARCH and FIGARCH-t forecasts. The latter combination gives an idea how FIGARCH forecasts can be complemented by considering a fat tailed distribution. We find an improvement in terms of MSEs from CO1 over single forecasts of the FIGARCH and the FIGARCH-t models at horizons $h = 20$, $h = 50$ and $h = 100$ in the different asset markets under inspection. The same result applies when we consider the forecast combinations GARCH+FIGARCH+FIGARCH-t (CO2) and GARCH+GARCH-t+FIGARCH+FIGARCH-t (CO3) although only for stock and real estate markets at higher horizons. CO2 and CO3 hint at how forecast could be improved when considering short memory along with long memory and fat tails. In terms of MAEs, CO1 can improve upon forecasts of the single (FI)GARCH specifications at all horizons in all markets. The forecast combinations CO2 and CO3 can also improve upon forecasts of the single (FI)GARCH models in terms of MAEs at all horizons in

[6]Technical details on the algorithm for forecast combinations can be provided upon request.

stock markets and real estate markets.

The second set of forecasts combinations considered are those resulting from the MSM models. The forecast combinations are given by BMSM+LMSM-t (CO4), BMSM+BMSM-t+LMSM-t (CO5) and BMSM+BMSM-t+LMSM+LMSM-t (CO6) to analyze the complementarities that arise when one combines models with different assumptions regarding the distribution of the multifractal parameter as well as the tails of the innovations. The results indicate that there is an improvement upon forecasts of single models in all three markets particularly against those obtained from the MSM models with Normal innovations both in terms of MSEs and MAEs. The improvement obtained from forecasts combinations is immediately evident in the case of bond markets where the MSEs and MAEs become less than one. We also find that the combination of forecasts in the MSM models does not translate into more variable MSEs or MAEs, a feature that speaks in favor of optimally combining single models' ingredients.

The last set of forecasts combinations examined are those resulting from MSM models and FIGARCH models. The combinatorial strategies are given by FIGARCH+LMSM-t (CO7), BMSM-t+LMSM-t+FIGARCH (CO8), BMSM-t+LMSM-t+FIGARCH+FIGARCH-t (CO9). The latter forecast combinations allow us to analyze the complementarities of two families of volatility models which assume two distinct distributions of the innovations along with different characteristics for the latent volatility process: (FI)GARCH models which account for short/long memory and autoregressive components and MSM models which account for multifractality, regime-switching and apparent long memory. Interestingly, the improvement upon forecasts of single models from the MSM-FIGARCH strategy is somewhat more evident than in the previous strategies. In terms of MSEs, for instance, we generally find a statistically significant improvement over historical volatility more frequently in stock, bond and real estate markets when comparing CO7, CO8, CO9 against single models (Tables 8 and 10). We also find that the variability of the combinations of MSM and FIGARCH models does not change too much with respect to single models.

Summing up, we find that the forecast combinations between FIGARCH, MSM or both types of models lead to improvements in forecasting accuracy upon forecasts of single models. In particular, we find that the forecasting strategy FIGARCH-MSM seems to be the most successful one in relation to single models or the other combination strategies - a feature that could be exploited in real time for risk management strategies. The particular usefulness of this combination strategy appears plausible given the flexibility of the MSM model in capturing varying degrees of long-term dependence and the added flexibility of FIGARCH for short horizon dependencies via its AR and MA parameters which are not accounted for in MSM models.

7 Conclusion

In this paper we examined the in-sample and out-of-sample performance of volatility models that incorporate different features characterizing the latent volatility process (long vs. short memory, regime-switching and multifractality) as well as distributional regularities of returns (*fat tails*). More precisely, we consider two major sets of "competitors", namely, volatility specifications from the new MSM models and the popular (FI)GARCH models along with Normal or Student-t innovations. We introduce a new member to the family of MSM models that accounts for Student-t innovations. This new model allows us to study whether there is an improvement in forecasting accuracy vis-*à*-vis the existing MSM model with Normal innovations and the (FI)GARCH models with Normal or Student-t innovations. The MSM-t model can be estimated either via ML or GMM. The suitability of ML and GMM estimation for MSM models with Student-t innovations is analyzed via Monte Carlo simulations. We conduct a comprehensive empirical study using country data on all-share equity indices, 10-year government bond indices and real estate security indices. In addition, we explore whether we may improve forecasting accuracy by constructing forecast combinations of the various models under inspection.

In-sample Monte Carlo experiments of the MSM-t model behave similarly like the MSM model with Normal innovations indicating that ML and GMM estimation are both suitable for estimating the new Binomial (ML and GMM) and Lognormal (GMM) MSM-t models. The out-of-sample Monte Carlo analysis shows that best linear forecasts are qualitatively similar to optimal forecasts so that the computationally advantageous strategy of GMM estimation of parameters plus linear forecasts can be adapted without much loss of efficiency. The in-sample empirical analysis shows that there is strong evidence of long memory and multifractality in international equity markets, bond markets and real estate markets as well as evidence of *fat tails*. The out-of-sample empirical analysis puts forward that GARCH models are less precise in accurately forecasting volatility for horizons greater than 20 days. This problem is not encountered once long-memory is incorporated via the FIGARCH model which produces MSEs and MAEs that are generally less than one.

The recently introduced MSM models with Normal innovations produce forecasts that improve upon historical volatility, but are in some cases inferior to FIGARCH with Normal innovations. However, two additional observations shed more positive light on the capabilities of MSM models for forecasting volatility. First, adding *fat tails* typically improves forecasts from MSM models while the same change of specification has, if anything, a negative effect for the (FI)GARCH models. While MSM-t is somewhat inferior to FIGARCH under the MSE criterion, it is superior under the MAE criterion at long horizons across all markets. Second, our forecasting combination exercise showed particularly sizable gains from combining FIGARCH and MSM in various ways. Therefore, both models appear to capture somewhat different facets of the latent volatility and can be sensibly used in tandem to improve upon forecasts of single models.

References

Abraham, B., Balakrishna, N., and Sivakumar, R. (2007). Gamma stochastic volatility models. *Journal of Forecasting*, 25:153–171.

Aiolfi, M. and Timmermann, A. (2006). Persistence in forecasting performance and conditional combination strategies. *Journal of Econometrics*, 135:31–53.

Andersen, T. and Bollerslev, T. (1998). Answering the sceptics: Yes standard volatility models do provide accurate forecasts. *International Economic Review*, 39:885–905.

Andersen, T., Bollerslev, T., Christoffersen, P., and Diebold, F. (2005a). Volatility forecasting. *PIER Working Paper No. 05-011*, pages 1–100.

Andersen, T., Bollerslev, T., Diebold, F., and Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica*, 71:579–625.

Andersen, T., Bollerslev, T., and Meddahi, N. (2005b). Correcting the errors: volatility forecast evaluation using high frequency data and realized volatilities. *Econometrica*, 73:279–296.

Baillie, R., Bollerslev, T., and Mikkelsen, H. (1996). Fractionally Itegrated Generalized Autoregressive Conditional Heteroskedasticty. *Journal of Econometrics*, 74:3–30.

Brockwell, P. and Dahlhaus, R. (2004). Generalized Levinson−Durbin and Burg algorithms. *Journal of Econometrics*, 118:129–144.

Brockwell, P. and Davis, R. (1991). *Time Series: Theory and Methods*. Berlin: Springer.

Calvet, L. and Fisher, A. (2001). Forecasting multifractal volatility. *Journal of Econometrics*, 105:27–58.

Calvet, L. and Fisher, A. (2004). Regime-switching and the estimation of multifractal processes. *Journal of Financial Econometrics*, 2:44–83.

Calvet, L., Fisher, A., and Mandelbrot, B. (1997). A multifractal model of asset returns. *Cowles Foundation Discussion Papers: Yale University*, pages 1164–1166.

Chuang, I., Lu, J., and Lee, P. (2007). Forecasting volatility in the financial markets: a comparison of alternative distributional assumptions. *Applied Financial Economics*, 17:1051–1060.

Diebold, F. and Mariano, R. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics*, 13:253–265.

Ding, Z., Granger, C., and Engle, R. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1:83–106.

Granger, C. (1989). Combining forecasts–twenty years later. *Journal of Forecasting*, 8:167–173.

Granger, C. and Terasvirta, T. (1999). A simple nonlinear time series model with misleading linear properties. *Economic Letters*, 62:161–165.

Harris, D. and Matyas, L. (1999). *Introduction to the Generalized Method of Moments estimation*. ed.L.Matyas, Cambridge: University Press.

Liu, R., di Matteo, T., and Lux, T. (2007). True and apparent scaling: The proximities of the Markov-Switching Multifractal Model to long-range dependence. *Physica A*, 383:35–42.

Lobato, I. and Savin, N. (1998). Real and spurious long-memory properties of stock market data. *Journal of Business and Economics Statistics*, 16:261–283.

Lux, T. (1996). Long-term stochastic dependence in financial prices: Evidence from the German stock market. *Applied Economics Letters*, 3:701–706.

Lux, T. (2008). The Markov-Switching Multifractal Model of Asset returns: GMM estimation and linear forecasting of volatility. *Journal of Business and Economic Statistics*, 26:194–210.

Lux, T. and Kaizoji, T. (2007). Forecasting volatility and volume in the Tokyo stock market: Long memory, fractality and regime switching. *Journal of Economic Dynamics and Control*, 31:1808–1843.

Mills, T. (1997). Stylized facts of the temporal and distributional properties of daily FTSE returns. *Applied Financial Economics*, 7:599–604.

Muller, U. (1997). Volatilities of different time resolutions: Analyzing the dynamics of market components. *Journal of Empirical Finance*, 4:213–239.

Poon, S. and Granger, C. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature*, 26:478–539.

Rossi, A. and Gallo, G. (2006). Volatility estimation via hidden Markov models. *Journal of Empirical Finance*, 13:203–230.

Schmitt, F., Schertzer, D., and Lovejoy, S. (1999). Multifractal analysis of foreign exchange data. *Applied Stochastic Models and Data Analysis*, 15:29–53.

Vassilicos, J., Demos, A., and Tata, F. (2004). *No evidence of chaos but some evidence of multifractals in the foreign exchange and the stock market in Applications of Fractals and Chaos*. Berlin: Springer.

Wu, P. and Shieh, S. (2007). Value-at-Risk analysis for long-term interest rate futures: Fat tail and long memory in return innovations. *Journal of Empirical Finance*, 14:248–259.

		$\nu=5$									$\nu=6$								
		ML	GMM1	GMM2	ML	GMM1	GMM2	ML	GMM1	GMM2	ML	GMM1	GMM2	ML	GMM1	GMM2	ML	GMM1	GMM2
		$m_0=1.3$			$m_0=1.4$			$m_0=1.5$			$m_0=1.3$			$m_0=1.4$			$m_0=1.5$		
	$\bar{m}_0$	1.294	1.215	1.344	1.395	1.348	1.434	1.496	1.469	1.527	1.295	1.166	1.338	1.395	1.314	1.428	1.496	1.449	1.520
	FSSE	0.024	0.123	0.141	0.024	0.092	0.109	0.023	0.056	0.071	0.024	0.127	0.135	0.024	0.103	0.102	0.023	0.060	0.072
	RMSE	0.025	0.149	0.148	0.024	0.105	0.114	0.024	0.064	0.076	0.024	0.184	0.140	0.024	0.134	0.105	0.024	0.079	0.075
	$\bar{\nu}$	4.997	4.903	4.826	5.030	4.914	4.824	5.089	4.939	4.844	6.033	4.904	4.970	6.104	4.903	4.966	6.337	4.927	4.984
T=2,500	FSSE	0.675	0.771	0.922	0.806	0.784	0.929	1.028	0.802	0.961	1.022	0.730	0.997	1.285	0.746	1.009	3.133	0.764	1.024
	RMSE	0.674	0.776	0.937	0.805	0.788	0.945	1.031	0.803	0.972	1.021	1.317	1.433	1.288	1.326	1.444	3.147	1.317	1.441
	$\bar{\sigma}$	0.997	0.967	0.902	0.999	0.965	0.898	1.003	0.968	0.888	0.997	0.918	0.852	0.999	0.913	0.849	1.004	0.916	0.842
	FSSE	0.093	0.104	0.173	0.125	0.133	0.195	0.162	0.165	0.222	0.092	0.095	0.153	0.125	0.125	0.172	0.163	0.156	0.199
	RMSE	0.093	0.109	0.198	0.125	0.137	0.219	0.162	0.168	0.248	0.092	0.125	0.213	0.124	0.152	0.229	0.163	0.177	0.254
	$\bar{m}_0$	1.298	1.222	1.332	1.399	1.357	1.427	1.499	1.471	1.518	1.299	1.171	1.327	1.399	1.327	1.419	1.499	1.453	1.512
	FSSE	0.016	0.102	0.115	0.016	0.058	0.076	0.016	0.040	0.056	0.016	0.109	0.105	0.016	0.066	0.075	0.016	0.042	0.055
	RMSE	0.016	0.129	0.119	0.016	0.072	0.080	0.016	0.049	0.059	0.016	0.169	0.108	0.016	0.098	0.077	0.016	0.063	0.057
	$\bar{\nu}$	5.090	4.691	4.684	5.108	4.692	4.689	5.136	4.701	4.696	6.146	4.714	4.842	6.184	4.708	4.839	6.244	4.723	4.847
T=5,000	FSSE	0.520	0.578	0.777	0.591	0.590	0.790	0.691	0.603	0.801	0.772	0.533	0.837	0.902	0.546	0.844	1.104	0.563	0.854
	RMSE	0.527	0.655	0.838	0.600	0.665	0.848	0.703	0.672	0.856	0.784	1.392	1.429	0.919	1.402	1.435	1.129	1.395	1.434
	$\bar{\sigma}$	1.002	0.954	0.902	1.003	0.954	0.900	1.005	0.957	0.896	1.002	0.907	0.854	1.003	0.905	0.854	1.005	0.908	0.852
	FSSE	0.066	0.076	0.144	0.088	0.095	0.156	0.113	0.117	0.172	0.065	0.069	0.124	0.088	0.089	0.135	0.113	0.110	0.153
	RMSE	0.065	0.088	0.174	0.088	0.105	0.185	0.113	0.125	0.201	0.065	0.116	0.191	0.088	0.130	0.199	0.113	0.143	0.212
	$\bar{m}_0$	1.298	1.239	1.330	1.399	1.364	1.422	1.499	1.476	1.515	1.299	1.188	1.321	1.399	1.338	1.414	1.499	1.459	1.509
	FSSE	0.012	0.067	0.078	0.012	0.039	0.053	0.012	0.028	0.039	0.012	0.083	0.075	0.012	0.042	0.053	0.012	0.029	0.038
	RMSE	0.012	0.091	0.084	0.012	0.053	0.058	0.012	0.036	0.042	0.012	0.140	0.078	0.012	0.075	0.055	0.012	0.050	0.039
	$\bar{\nu}$	4.993	4.500	4.491	4.996	4.504	4.496	5.005	4.510	4.505	5.997	4.556	4.655	6.007	4.556	4.655	6.027	4.563	4.664
T=10,000	FSSE	0.345	0.427	0.572	0.402	0.433	0.578	0.475	0.442	0.587	0.512	0.385	0.619	0.609	0.392	0.630	0.733	0.399	0.642
	RMSE	0.345	0.657	0.765	0.402	0.658	0.766	0.475	0.659	0.768	0.511	1.494	1.480	0.608	1.497	1.485	0.732	1.492	1.482
	$\bar{\sigma}$	0.996	0.946	0.894	0.995	0.947	0.892	0.994	0.948	0.888	0.996	0.900	0.849	0.995	0.900	0.848	0.994	0.902	0.845
	FSSE	0.048	0.059	0.112	0.065	0.071	0.122	0.082	0.086	0.137	0.048	0.053	0.097	0.064	0.066	0.106	0.082	0.081	0.119
	RMSE	0.048	0.079	0.154	0.065	0.089	0.163	0.083	0.101	0.177	0.048	0.113	0.179	0.065	0.119	0.185	0.082	0.127	0.195

Table 1: Monte Carlo ML and GMM estimation of the Binomial MSM-t model with $k=8$, $\sigma=1$, $\nu=5,6$ and $m_0=1.3,1.4,1.5$. FSSE: finite sample standard error (e.g. FSSE= $[\sum_{s=1}^{S}(m_{0,s}-\bar{m}_0)^2/S]^{1/2}$), RMSE: root mean squared error (e.g. RMSE= $[\sum_{s=1}^{S}(m_{0,s}-m_0)^2/S]^{1/2}$). The entries $\bar{m}_0$, $\bar{\sigma}$ and $\bar{\nu}$ denote the mean of the estimated parameters over S=400 Monte Carlo runs.

		$\nu = 5$						$\nu = 6$					
		GMM1	GMM2	GMM1	GMM2	GMM1	GMM2	GMM1	GMM2	GMM1	GMM2	GMM1	GMM2
		$m_0 = 1.3$		$m_0 = 1.4$		$m_0 = 1.5$		$m_0 = 1.3$		$m_0 = 1.4$		$m_0 = 1.5$	
	$\bar{m}_0$	1.226	1.333	1.360	1.427	1.475	1.519	1.175	1.325	1.330	1.419	1.456	1.513
	FSSE	0.103	0.117	0.058	0.080	0.040	0.057	0.110	0.111	0.067	0.079	0.043	0.056
	RMSE	0.127	0.122	0.071	0.084	0.048	0.060	0.166	0.113	0.097	0.081	0.062	0.057
	$\bar{\nu}$	4.705	4.627	4.711	4.628	4.726	4.642	4.732	4.776	4.727	4.777	4.743	4.794
BMSM	FSSE	0.591	0.705	0.602	0.710	0.610	0.721	0.546	0.761	0.560	0.770	0.572	0.783
	RMSE	0.660	0.797	0.667	0.801	0.668	0.804	1.380	1.440	1.390	1.445	1.381	1.438
	$\bar{\sigma}$	0.951	0.884	0.951	0.874	0.954	0.857	0.899	0.839	0.897	0.831	0.902	0.816
	FSSE	0.116	0.172	0.150	0.198	0.189	0.231	0.107	0.157	0.141	0.180	0.178	0.212
	RMSE	0.126	0.208	0.158	0.234	0.194	0.272	0.147	0.224	0.174	0.247	0.203	0.280
		$\lambda = 0.05$		$\lambda = 0.10$		$\lambda = 0.15$		$\lambda = 0.05$		$\lambda = 0.10$		$\lambda = 0.15$	
	$\bar{\lambda}$	0.030	0.064	0.079	0.113	0.129	0.163	0.023	0.067	0.069	0.116	0.119	0.165
	FSSE	0.022	0.047	0.027	0.048	0.028	0.049	0.019	0.045	0.025	0.047	0.027	0.050
	RMSE	0.030	0.049	0.034	0.050	0.035	0.051	0.033	0.048	0.040	0.049	0.041	0.052
	$\bar{\nu}$	4.540	4.539	4.601	4.563	4.629	4.574	4.730	4.840	4.684	4.868	4.693	4.891
LMSM	FSSE	0.799	0.911	0.775	0.905	0.773	0.900	0.700	0.894	0.680	0.910	0.672	0.904
	RMSE	0.921	1.020	0.871	1.004	0.856	0.995	1.449	1.464	1.481	1.452	1.469	1.430
	$\bar{\sigma}$	0.922	0.844	0.921	0.821	0.920	0.780	0.886	0.822	0.875	0.799	0.873	0.765
	FSSE	0.119	0.224	0.159	0.258	0.192	0.281	0.105	0.184	0.148	0.222	0.179	0.253
	RMSE	0.142	0.273	0.178	0.314	0.207	0.356	0.155	0.256	0.194	0.299	0.219	0.344

Table 2: Monte Carlo GMM estimation of the Binomial and Lognormal MSM-t models with $T = 5,000$, $k = 10$, $\sigma = 1$, $\nu = 5, 6$, $m_0 = 1.3, 1.4, 1.5$ and $\lambda = 0.05, 0.10, 0.15$. FSSE: finite sample standard error (e.g. FSSE= $[\sum_{s=1}^{S}(m_{0,s} - \bar{m}_0)^2/S]^{1/2}$), RMSE: root mean squared error (e.g. RMSE= $[\sum_{s=1}^{S}(m_{0,s} - m_0)^2/S]^{1/2}$). The entries $\bar{m}_0$, $\bar{\sigma}$ and $\bar{\nu}$ denote the mean of the estimated parameters over S=400 Monte Carlo runs.

		$\nu = 5$									$\nu = 6$								
		ML	GMM1	GMM2	ML	GMM1	GMM2	ML	GMM1	GMM2	ML	GMM1	GMM2	ML	GMM1	GMM2	ML	GMM1	GMM2
	h	$m_0 = 1.3$			$m_0 = 1.4$			$m_0 = 1.5$			$m_0 = 1.3$			$m_0 = 1.4$			$m_0 = 1.5$		
MSE	1	0.971	0.988	0.990	0.957	0.978	0.983	0.947	0.974	0.979	0.964	0.987	0.985	0.947	0.972	0.975	0.937	0.966	0.971
		(0.012)	(0.011)	(0.013)	(0.019)	(0.016)	(0.018)	(0.025)	(0.021)	(0.022)	(0.013)	(0.013)	(0.016)	(0.020)	(0.017)	(0.021)	(0.025)	(0.022)	(0.025)
	5	0.981	0.992	0.994	0.974	0.986	0.989	0.971	0.985	0.988	0.976	0.991	0.990	0.969	0.982	0.985	0.966	0.981	0.983
		(0.009)	(0.008)	(0.011)	(0.013)	(0.012)	(0.013)	(0.016)	(0.014)	(0.015)	(0.010)	(0.010)	(0.013)	(0.014)	(0.013)	(0.015)	(0.017)	(0.015)	(0.017)
	10	0.986	0.994	0.996	0.981	0.990	0.992	0.980	0.989	0.992	0.983	0.993	0.993	0.978	0.987	0.989	0.977	0.987	0.989
		(0.008)	(0.007)	(0.009)	(0.011)	(0.010)	(0.011)	(0.013)	(0.012)	(0.013)	(0.009)	(0.008)	(0.011)	(0.012)	(0.011)	(0.013)	(0.014)	(0.013)	(0.014)
	20	0.991	0.996	0.998	0.988	0.993	0.995	0.987	0.993	0.995	0.988	0.995	0.996	0.985	0.991	0.993	0.985	0.991	0.993
		(0.006)	(0.006)	(0.009)	(0.009)	(0.008)	(0.010)	(0.011)	(0.010)	(0.011)	(0.007)	(0.007)	(0.010)	(0.010)	(0.009)	(0.011)	(0.012)	(0.011)	(0.012)
MAE	1	0.955	0.972	0.938	0.923	0.960	0.923	0.892	0.952	0.909	0.952	0.971	0.931	0.918	0.954	0.913	0.885	0.944	0.898
		(0.061)	(0.047)	(0.064)	(0.084)	(0.065)	(0.072)	(0.108)	(0.082)	(0.082)	(0.059)	(0.042)	(0.059)	(0.083)	(0.062)	(0.068)	(0.106)	(0.079)	(0.078)
	5	0.969	0.977	0.942	0.951	0.971	0.931	0.935	0.968	0.922	0.967	0.975	0.936	0.947	0.966	0.923	0.930	0.963	0.914
		(0.056)	(0.046)	(0.064)	(0.079)	(0.061)	(0.072)	(0.102)	(0.076)	(0.081)	(0.055)	(0.041)	(0.059)	(0.077)	(0.059)	(0.068)	(0.101)	(0.074)	(0.077)
	10	0.977	0.980	0.944	0.964	0.976	0.935	0.954	0.976	0.928	0.975	0.978	0.938	0.961	0.972	0.928	0.951	0.972	0.921
		(0.053)	(0.045)	(0.065)	(0.074)	(0.059)	(0.072)	(0.096)	(0.072)	(0.081)	(0.051)	(0.040)	(0.058)	(0.073)	(0.057)	(0.067)	(0.095)	(0.070)	(0.076)
	20	0.984	0.982	0.946	0.976	0.982	0.939	0.972	0.983	0.933	0.983	0.981	0.940	0.975	0.978	0.932	0.969	0.979	0.926
		(0.048)	(0.044)	(0.065)	(0.067)	(0.055)	(0.072)	(0.087)	(0.068)	(0.081)	(0.046)	(0.040)	(0.058)	(0.066)	(0.053)	(0.066)	(0.086)	(0.066)	(0.075)

Table 3: Monte Carlo Assessment of Bayesian vs. Best Linear Forecasts, Binomial MSM Specification for $k = 8$, $\nu = 5, 6$ and $m_0 = 1.3, 1.4, 1.5$. MSE: mean square errors and MAE: mean absolute errors. MSEs and MAEs are given in percentage of the MSEs and MAEs of a naive forecast using the in-sample variance. All entries are averages over 400 Monte Carlo runs (with standard errors given in parentheses). In each run, an overall sample of 10,000 entries has been split into an in-sample period of 5,000 entries for parameter estimation and an out-of-sample period of 5,000 entries for evaluation of forecasting performance. ML stands for parameter estimation based on the maximum likelihood procedure and pertinent inference on the probability of the 2^k states of the model. GMM1 uses parameters estimated by GMM with moment set 1, while GMM2 implements parameters estimated via GMM with moment set 2.

Equity Markets		Bond Markets	Real Estate Markets
Austria	Italy	Australia	Austria
Argentina	Norway	Belgium	Canada
Belgium	Mexico	Canada	France
Canada	New Zealand	Denmark	Germany
Chile	Japan	France	Italy
Denmark	South Africa	Germany	Japan
Finland	Spain	Ireland	New Zealand
France	Sweden	Netherlands	Spain
Germany	Thailand	Sweden	Sweden
Greece	Turkey	United Kingdom	United Kingdom
Hong Kong	United Kingdom	United States	United States
India	United States		South Africa
Ireland			

Table 4: Equity, Bond and Real Estate markets for the empirical analysis. Countries were chosen upon data availability for the sample period 01/1990 to 01/2008. We employ Datastream calculated (total market) stock indices, 10-year benchmark government bond indices and real estate security indices.

		GARCH							FIGARCH								
		Normal innovations			Student-t innovations				Normal innovations				Student-t innovations				
Mkt.		$\bar{\omega}$	$\bar{\beta}$	$\bar{\alpha}$	$\bar{\omega}$	$\bar{\beta}$	$\bar{\alpha}$	$\bar{\nu}$	$\bar{\omega}$	$\bar{\beta}$	$\bar{\delta}$	$\bar{d}$	$\bar{\omega}$	$\bar{\beta}$	$\bar{\delta}$	$\bar{d}$	$\bar{\nu}$
ST	MG	0.081	0.844	0.112	0.061	0.855	0.122	5.800	0.164	0.271	0.065	0.334	0.101	0.497	0.222	0.428	5.783
		(0.022)	(0.014)	(0.007)	(0.017)	(0.013)	(0.013)	(0.297)	(0.033)	(0.079)	(0.068)	(0.029)	(0.026)	(0.037)	(0.038)	(0.028)	(0.296)
	Min	0.006	0.587	0.054	0.003	0.626	0.040	3.183	0.027	-0.938	-0.915	0.001	0.009	0.039	-0.111	0.286	3.282
	Max	0.491	0.939	0.175	0.425	0.957	0.373	9.387	0.683	0.851	0.762	0.832	0.591	0.873	0.905	0.933	9.608
BO	MG	0.005	0.903	0.075	0.004	0.907	0.078	4.760	0.014	0.558	0.216	0.422	0.009	0.614	0.253	0.443	4.824
		(0.002)	(0.014)	(0.010)	(0.001)	(0.011)	(0.011)	(0.244)	(0.008)	(0.084)	(0.064)	(0.044)	(0.002)	(0.029)	(0.021)	(0.029)	(0.227)
	Min	0.012	0.777	0.031	0.001	0.814	0.028	3.013	0.002	-0.208	-0.378	0.200	0.002	0.437	0.168	0.282	3.533
	Max	0.021	0.946	0.154	0.012	0.951	0.159	6.364	0.095	0.786	0.388	0.685	0.028	0.739	0.344	0.577	6.383
RE	MG	0.064	0.877	0.084	0.064	0.782	0.111	4.512	0.132	0.434	0.332	0.229	0.132	0.308	0.104	0.367	4.831
		(0.013)	(0.019)	(0.011)	(0.018)	(0.074)	(0.011)	(0.324)	(0.031)	(0.154)	(0.159)	(0.037)	(0.035)	(0.149)	(0.115)	(0.069)	(0.255)
	Min	0.004	0.716	0.023	0.002	0.014	0.043	2.000	0.027	-0.598	-0.720	0.029	0.017	-0.506	-0.577	0.044	3.295
	Max	0.137	0.969	0.143	0.192	0.956	0.176	6.256	0.364	0.923	0.968	0.425	0.364	0.926	0.599	0.999	6.529

Table 5: GARCH and FIGARCH in-sample estimates. MG: mean group parameter estimates (with standard errors in parentheses) of GARCH, GARCH-t, FIGARCH and FIGARCH-t models for $N = 25$ international stock market indices (ST), $N = 11$ international 10-year government bond indices (BO), $N = 12$ international real estate security indices (RE). Min: minimum estimated parameter value in the cross-section. Max: maximum estimated parameter value in the cross-section.

		BMSM					LMSM				
		Normal innovations		Student-t innovations			Normal innovations		Student-t innovations		
Mkt.		$\bar{m}_0$	$\bar{\sigma}$	$\bar{m}_0$	$\bar{\sigma}$	$\bar{\nu}$	$\bar{\lambda}$	$\bar{\sigma}$	$\bar{\lambda}$	$\bar{\sigma}$	$\bar{\nu}$
ST	MG	1.435	1.273	1.388	0.808	4.821	0.118	1.273	0.107	0.794	4.921
		(0.019)	(0.113)	(0.033)	(0.080)	(0.192)	(0.012)	(0.113)	(0.017)	(0.087)	(0.190)
	Min	1.264	0.681	1.033	0.449	4.050	0.036	0.681	0.001	0.331	4.050
	Max	1.619	2.932	1.681	2.257	6.779	0.261	2.932	0.341	2.217	6.752
BO	MG	1.487	0.418	1.435	0.252	4.580	0.159	0.418	0.140	0.245	4.588
		(0.037)	(0.029)	(0.053)	(0.027)	(0.291)	(0.025)	(0.029)	(0.034)	(0.028)	(0.289)
	Min	1.234	0.293	1.123	0.160	4.050	0.032	0.293	0.008	0.135	4.050
	Max	1.642	0.601	1.693	0.485	7.089	0.289	0.601	0.362	0.482	7.084
RE	MG	1.594	1.376	1.646	0.765	4.900	0.307	1.374	0.268	0.692	5.192
		(0.041)	(0.143)	(0.063)	(0.094)	(0.128)	(0.085)	(0.143)	(0.056)	(0.109)	(0.127)
	Min	1.393	0.682	1.430	0.487	4.365	0.086	0.682	0.137	0.158	4.574
	Max	1.912	2.287	1.945	1.503	5.555	1.137	2.282	0.741	1.416	6.067

Table 6: Binomial and Lognormal MSM in-sample estimates. MG: mean group parameter estimates (with standard errors in parentheses) of the BMSM, BMSM-t, LMSM and LMSM-t models for $N = 25$ international stock market indices (ST), $N = 11$ international 10-year government bond indices (BO), $N = 12$ international real estate security indices (RE). Min: minimum estimated parameter value in the cross-section. Max: maximum estimated parameter value in the cross-section.

		MSE								MAE							
		Normal innovations				Student-t innovations				Normal innovations				Student-t innovations			
Mkt.	h	GARCH	FIGA	BMSM	LMSM	GARCH	FIGA	BMSM	LMSM	GARCH	FIGA	BMSM	LMSM	GARCH	FIGA	BMSM	LMSM
ST	1	0.873	**0.863**	0.948	0.944	0.883	0.870	0.947	0.948	**0.796**	0.797	0.940	0.935	0.803	0.797	0.825	0.819
		(0.018)	(0.019)	(0.007)	(0.008)	(0.022)	(0.024)	(0.008)	(0.008)	(0.004)	(0.004)	(0.001)	(0.001)	(0.005)	(0.005)	(0.003)	(0.003)
	5	0.919	**0.903**	0.972	0.971	0.962	0.923	0.949	0.953	0.825	0.824	0.962	0.962	0.850	0.837	0.826	**0.822**
		(0.012)	(0.012)	(0.004)	(0.004)	(0.014)	(0.012)	(0.008)	(0.008)	(0.004)	(0.004)	(0.001)	(0.001)	(0.005)	(0.004)	(0.003)	(0.003)
	20	0.928	**0.906**	0.972	0.972	0.981	0.912	0.949	0.954	0.867	0.859	0.964	0.964	0.944	0.890	0.826	**0.822**
		(0.010)	(0.012)	(0.004)	(0.004)	(0.013)	(0.011)	(0.008)	(0.008)	(0.003)	(0.003)	(0.001)	(0.001)	(0.004)	(0.004)	(0.003)	(0.003)
	50	0.945	**0.917**	0.973	0.973	1.039	0.918	0.950	0.954	0.911	0.892	0.964	0.964	1.089	0.944	0.827	**0.823**
		(0.008)	(0.011)	(0.004)	(0.004)	(0.015)	(0.010)	(0.008)	(0.008)	(0.003)	(0.003)	(0.001)	(0.001)	(0.004)	(0.003)	(0.003)	(0.003)
	100	0.958	**0.934**	0.973	0.973	1.181	0.940	0.950	0.954	0.943	0.926	0.964	0.964	1.283	1.000	0.827	**0.822**
		(0.007)	(0.009)	(0.004)	(0.004)	(0.032)	(0.008)	(0.008)	(0.008)	(0.002)	(0.002)	(0.001)	(0.001)	(0.006)	(0.003)	(0.003)	(0.003)
BO	1	0.838	**0.830**	1.023	1.018	0.840	0.832	0.981	0.975	0.791	**0.764**	1.011	1.006	0.793	0.768	0.843	0.825
		(0.008)	(0.009)	(0.002)	(0.002)	(0.008)	(0.009)	(0.005)	(0.006)	(0.003)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
	5	0.856	**0.833**	1.047	1.047	0.856	0.837	0.984	0.979	0.818	**0.776**	1.032	1.032	0.822	0.781	0.845	0.827
		(0.007)	(0.008)	(0.003)	(0.003)	(0.007)	(0.008)	(0.005)	(0.006)	(0.003)	(0.003)	(0.001)	(0.001)	(0.003)	(0.003)	(0.003)	(0.003)
	20	0.910	**0.846**	1.049	1.049	0.922	0.852	0.983	0.977	0.894	**0.809**	1.034	1.034	0.923	0.825	0.844	0.826
		(0.005)	(0.008)	(0.003)	(0.003)	(0.004)	(0.007)	(0.005)	(0.006)	(0.002)	(0.003)	(0.001)	(0.001)	(0.002)	(0.003)	(0.003)	(0.003)
	50	0.975	**0.870**	1.050	1.050	1.072	0.884	0.980	0.975	0.974	0.853	1.034	1.034	1.079	0.882	0.842	**0.824**
		(0.002)	(0.007)	(0.003)	(0.003)	(0.005)	(0.006)	(0.006)	(0.006)	(0.001)	(0.002)	(0.001)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)
	100	1.031	**0.906**	1.051	1.051	1.267	0.932	0.977	0.971	1.042	0.907	1.034	1.035	1.245	0.950	0.839	**0.820**
		(0.002)	(0.005)	(0.003)	(0.003)	(0.013)	(0.004)	(0.006)	(0.006)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)	(0.002)	(0.003)	(0.003)
RE	1	0.827	**0.821**	0.929	0.921	0.839	0.832	0.911	0.910	0.706	0.685	0.897	0.886	0.675	**0.654**	0.691	0.680
		(0.026)	(0.027)	(0.011)	(0.013)	(0.025)	(0.027)	(0.016)	(0.017)	(0.005)	(0.005)	(0.002)	(0.002)	(0.005)	(0.005)	(0.004)	(0.005)
	5	0.858	**0.850**	0.964	0.963	0.880	0.860	0.920	0.925	0.736	0.708	0.932	0.930	0.713	**0.681**	0.697	0.687
		(0.021)	(0.022)	(0.006)	(0.006)	(0.019)	(0.021)	(0.015)	(0.015)	(0.004)	(0.005)	(0.001)	(0.001)	(0.005)	(0.005)	(0.004)	(0.005)
	20	0.879	**0.863**	0.966	0.965	0.911	0.876	0.921	0.925	0.800	0.746	0.934	0.932	0.801	0.727	0.697	**0.687**
		(0.018)	(0.020)	(0.006)	(0.006)	(0.014)	(0.018)	(0.015)	(0.015)	(0.004)	(0.004)	(0.001)	(0.001)	(0.004)	(0.005)	(0.004)	(0.005)
	50	0.897	**0.869**	0.966	0.966	0.944	0.890	0.921	0.926	0.865	0.775	0.934	0.932	0.896	0.780	0.696	**0.687**
		(0.014)	(0.018)	(0.006)	(0.006)	(0.008)	(0.015)	(0.015)	(0.015)	(0.003)	(0.004)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.005)
	100	0.926	**0.875**	0.965	0.964	0.999	0.925	0.914	0.919	0.932	0.799	0.933	0.932	0.992	0.856	0.694	**0.684**
		(0.008)	(0.018)	(0.006)	(0.006)	(0.007)	(0.010)	(0.016)	(0.016)	(0.002)	(0.004)	(0.001)	(0.001)	(0.004)	(0.004)	(0.004)	(0.005)

Table 7: Forecasting results of MSM and (FI)GARCH models. The table shows panel MSE and MAE (with standard errors in parentheses) relative to naive forecasts of historical volatility for $N = 25$ international stock market indices (ST), $N = 11$ international 10-year government bond indices (BO), $N = 12$ international real estate security indices (RE) at horizons $h = 1, 5, 20, 50, 100$. Entries in **bold** denote the model with the lowest MSE or MAE at each horizon h.

		MSE								MAE							
		Normal innovations				Student-t innovations				Normal innovations				Student-t innovations			
Mkt.	h	GARCH	FIGA	BMSM	LMSM	GARCH	FIGA	BMSM	LMSM	GARCH	FIGA	BMSM	LMSM	GARCH	FIGA	BMSM	LMSM
	1	20	23	22	22	21	21	11	10	17	15	22	22	16	14	25	25
	5	18	20	14	15	18	19	9	8	15	13	14	14	13	11	25	25
ST	20	17	19	10	10	15	17	9	8	14	11	14	14	9	10	25	25
	50	11	18	10	10	13	13	9	8	10	10	14	14	6	9	25	25
	100	7	17	10	10	9	11	9	7	8	9	14	14	4	5	25	25
	1	10	9	9	9	8	8	6	6	9	9	9	9	8	9	10	10
	5	9	9	5	5	8	8	6	6	9	9	7	7	8	9	10	10
BO	20	9	9	4	4	7	8	6	6	9	9	6	6	7	8	10	10
	50	6	8	4	4	3	7	6	6	5	7	6	6	2	7	10	10
	100	1	5	4	4	0	4	6	6	3	5	6	6	0	5	10	10
	1	10	9	10	10	9	9	7	7	9	10	11	11	9	10	12	11
	5	9	9	10	10	7	9	6	6	7	9	9	9	7	9	12	11
RE	20	7	9	9	9	4	8	6	5	8	9	9	9	5	9	12	11
	50	5	10	9	9	5	9	6	5	5	9	9	9	2	9	12	11
	100	4	9	8	8	2	7	6	5	3	9	9	9	2	8	12	11

Table 8: Average forecasting accuracy of alternative volatility models. The table shows the number of improvements for single models against historical volatility via the Diebold and Mariano (1995) test for $N = 25$ international stock market indices (ST), $N = 11$ international 10-year government bond indices (BO), $N = 12$ international real estate security indices (RE) at horizons $h = 1, 5, 20, 50, 100$.

		MSE									MAE								
		(FI)GARCH			MSM			(FI)GARCH-MSM			(FI)GARCH			MSM			(FI)GARCH-MSM		
	h	CO1	CO2	CO3	CO4	CO5	CO6	CO7	CO8	CO9	CO1	CO2	CO3	CO4	CO5	CO6	CO7	CO8	CO9
ST	1	**0.864** (0.021)	0.868 (0.020)	0.869 (0.020)	0.936 (0.009)	0.932 (0.010)	0.931 (0.010)	0.877 (0.017)	0.878 (0.017)	0.872 (0.018)	0.792 (0.005)	0.785 (0.005)	0.786 (0.005)	0.830 (0.002)	0.818 (0.003)	0.817 (0.003)	0.750 (0.004)	0.751 (0.004)	**0.746** (0.004)
	5	0.909 (0.012)	0.913 (0.012)	0.919 (0.012)	0.946 (0.008)	0.942 (0.009)	0.941 (0.009)	0.906 (0.013)	0.905 (0.013)	**0.905** (0.013)	0.822 (0.004)	0.810 (0.004)	0.814 (0.004)	0.835 (0.002)	0.821 (0.003)	0.821 (0.003)	0.764 (0.004)	**0.764** (0.004)	0.764 (0.004)
	20	**0.905** (0.012)	0.910 (0.012)	0.921 (0.011)	0.947 (0.008)	0.942 (0.009)	0.942 (0.009)	0.913 (0.012)	0.913 (0.012)	0.910 (0.012)	0.859 (0.003)	0.836 (0.003)	0.845 (0.004)	0.836 (0.002)	0.822 (0.003)	0.822 (0.003)	0.781 (0.003)	**0.780** (0.003)	0.783 (0.003)
	50	**0.912** (0.011)	0.914 (0.011)	0.915 (0.011)	0.947 (0.008)	0.943 (0.009)	0.943 (0.009)	0.921 (0.011)	0.920 (0.011)	0.915 (0.012)	0.892 (0.003)	0.857 (0.003)	0.866 (0.003)	0.836 (0.003)	0.822 (0.003)	0.822 (0.003)	0.800 (0.003)	**0.798** (0.003)	0.801 (0.003)
	100	0.926 (0.010)	0.924 (0.011)	0.924 (0.011)	0.948 (0.008)	0.943 (0.009)	0.943 (0.009)	0.931 (0.011)	0.928 (0.011)	**0.922** (0.012)	0.922 (0.003)	0.878 (0.003)	0.887 (0.002)	0.835 (0.003)	0.821 (0.003)	0.821 (0.003)	0.820 (0.003)	**0.811** (0.003)	0.815 (0.003)
BO	1	**0.829** (0.009)	0.829 (0.009)	0.830 (0.009)	0.947 (0.005)	0.949 (0.005)	0.949 (0.005)	0.842 (0.010)	0.845 (0.010)	0.844 (0.010)	0.763 (0.003)	0.766 (0.003)	0.768 (0.003)	0.851 (0.002)	0.850 (0.003)	0.850 (0.003)	0.721 (0.004)	**0.720** (0.004)	0.721 (0.004)
	5	**0.833** (0.008)	0.834 (0.008)	0.835 (0.008)	0.951 (0.005)	0.953 (0.005)	0.953 (0.005)	0.845 (0.010)	0.847 (0.010)	0.847 (0.010)	0.774 (0.003)	0.778 (0.003)	0.780 (0.003)	0.854 (0.002)	0.853 (0.003)	0.853 (0.003)	0.727 (0.003)	**0.726** (0.003)	0.727 (0.003)
	20	**0.844** (0.008)	0.846 (0.008)	0.848 (0.007)	0.951 (0.005)	0.953 (0.005)	0.953 (0.005)	0.852 (0.009)	0.854 (0.009)	0.852 (0.009)	0.808 (0.003)	0.812 (0.003)	0.816 (0.003)	0.853 (0.003)	0.852 (0.003)	0.852 (0.003)	0.741 (0.003)	**0.740** (0.003)	0.742 (0.003)
	50	0.867 (0.007)	0.870 (0.006)	0.874 (0.006)	0.948 (0.005)	0.951 (0.005)	0.951 (0.005)	0.858 (0.009)	0.861 (0.009)	**0.858** (0.009)	0.851 (0.002)	0.855 (0.002)	0.863 (0.002)	0.851 (0.003)	0.850 (0.003)	0.850 (0.003)	0.758 (0.003)	**0.757** (0.003)	0.759 (0.003)
	100	0.899 (0.005)	0.901 (0.005)	0.907 (0.005)	0.945 (0.006)	0.948 (0.006)	0.948 (0.006)	0.866 (0.009)	0.869 (0.009)	**0.866** (0.009)	0.901 (0.002)	0.902 (0.002)	0.915 (0.002)	0.848 (0.003)	0.847 (0.003)	0.847 (0.003)	0.773 (0.003)	**0.772** (0.003)	0.776 (0.003)
RE	1	**0.823** (0.027)	0.823 (0.027)	0.823 (0.027)	0.881 (0.020)	0.880 (0.020)	0.879 (0.020)	0.828 (0.027)	0.829 (0.027)	0.828 (0.027)	0.653 (0.005)	0.655 (0.005)	0.654 (0.005)	0.696 (0.004)	0.686 (0.004)	0.687 (0.004)	0.596 (0.005)	0.594 (0.005)	**0.590** (0.005)
	5	0.850 (0.023)	0.851 (0.023)	0.854 (0.022)	0.896 (0.018)	0.895 (0.018)	0.895 (0.018)	0.850 (0.024)	0.849 (0.024)	**0.849** (0.024)	0.674 (0.005)	0.676 (0.005)	0.675 (0.005)	0.702 (0.004)	0.691 (0.004)	0.692 (0.004)	0.606 (0.005)	0.602 (0.005)	**0.599** (0.005)
	20	0.860 (0.021)	0.860 (0.021)	0.863 (0.021)	0.897 (0.018)	0.895 (0.018)	0.895 (0.018)	0.858 (0.023)	0.857 (0.023)	**0.857** (0.023)	0.703 (0.005)	0.705 (0.005)	0.706 (0.005)	0.702 (0.004)	0.692 (0.004)	0.692 (0.004)	0.613 (0.005)	0.606 (0.005)	**0.605** (0.005)
	50	0.864 (0.019)	0.863 (0.020)	0.865 (0.019)	0.897 (0.018)	0.896 (0.018)	0.896 (0.018)	0.859 (0.023)	0.857 (0.023)	**0.856** (0.023)	0.727 (0.004)	0.728 (0.004)	0.731 (0.004)	0.701 (0.004)	0.691 (0.004)	0.691 (0.004)	0.614 (0.005)	0.607 (0.005)	**0.607** (0.005)
	100	0.868 (0.019)	0.867 (0.019)	0.868 (0.019)	0.890 (0.020)	0.888 (0.020)	0.888 (0.020)	0.851 (0.025)	0.850 (0.025)	**0.849** (0.025)	0.750 (0.004)	0.752 (0.004)	0.756 (0.004)	0.699 (0.004)	0.689 (0.004)	0.689 (0.004)	0.614 (0.005)	0.606 (0.005)	**0.606** (0.005)

Table 9: Results of forecast combinations from MSM and (FI)GARCH models. The table shows panel MSE and MAE (with standard errors in parentheses) relative to naive forecasts for $N = 25$ international stock market indices (ST), $N = 11$ international 10-year government bond indices (BO), $N = 12$ international real estate security indices (RE) at horizons $h = 1, 5, 20, 50, 100$. Entries in **bold** denote the model with the lowest MSE or MAE at each horizon h. The combination models are CO1: FIGARCH+FIGARCH-t, CO2: GARCH+FIGARCH+FIGARCH-t, CO3: GARCH+GARCH-t+FIGARCH+FIGARCH-t, CO4: BMSM+LMSM-t, CO5: BMSM-t+BMSM+LMSM-t, CO6: BMSM-t+BMSM+LMSM-t+LMSM, CO7: FIGARCH+LMSM-t, CO8: BMSM-t+LMSM-t+FIGARCH, CO9: BMSM-t+LMSM-t+FIGARCH+FIGARCH-t.

		MSE									MAE								
		(FI)GARCH			MSM			(FI)GARCH-MSM			(FI)GARCH			MSM			(FI)GARCH-MSM		
	h	CO1	CO2	CO3	CO4	CO5	CO6	CO7	CO8	CO9	CO1	CO2	CO3	CO4	CO5	CO6	CO7	CO8	CO9
	1	23	23	23	20	20	20	23	23	23	15	16	16	25	25	25	24	24	24
	5	19	18	18	11	11	11	22	22	22	14	16	15	25	25	25	24	24	24
ST	20	19	20	19	10	10	10	20	20	20	12	14	14	25	25	25	23	23	24
	50	19	21	21	10	10	10	19	19	20	9	15	15	25	25	25	24	24	24
	100	19	22	21	10	10	10	18	18	20	9	14	14	25	25	25	24	24	24
	1	9	10	10	7	7	7	8	8	8	9	9	9	10	10	10	11	11	11
	5	9	9	9	6	6	6	8	8	9	9	9	9	10	10	10	11	11	11
BO	20	9	9	9	6	6	6	7	7	9	9	9	8	10	10	10	11	11	11
	50	8	9	8	6	6	6	7	7	8	7	7	7	10	10	10	11	11	11
	100	5	6	6	6	6	6	7	7	8	6	5	5	10	10	10	11	11	10
	1	9	9	9	9	9	9	9	9	9	10	10	9	12	12	12	12	12	12
	5	9	9	8	8	6	6	10	10	10	9	9	9	12	12	12	12	12	12
RE	20	9	9	9	8	6	6	8	8	8	9	9	9	12	12	12	12	12	12
	50	9	10	10	8	6	6	8	8	9	9	10	9	12	12	12	12	12	12
	100	9	10	9	8	6	6	7	6	7	9	10	8	12	12	12	12	12	12

Table 10: Average forecasting accuracy of combination models. The table shows the number of improvements for combination models against historical volatility via the Diebold and Mariano (1995) test for $N = 25$ international stock market indices (ST), $N = 11$ international 10-year government bond indices (BO), $N = 12$ international real estate security indices (RE) at horizons $h = 1, 5, 20, 50, 100$.