Comovement and Predictor Variables for Multifractal Volatility*

Laurent E. Calvet, Adlai J. Fisher and Samuel B. Thompson

This draft: September 2003

This paper implements a frequency decomposition of volatility comovements across financial series. We show that volatility components display little correlation with macroeconomic variables. The components of two series of similar frequencies are nonetheless strongly correlated. This leads us to construct a multivariate extension of the Markov-Switching Multifractal (MSM) introduced in Calvet and Fisher (2001, 2002*b*). Multivariate MSM is a stochastic volatility model characterized by a closed-form likelihood. For state spaces of reasonable size, we implement ML estimation and show the usefulness of the approach for out-of-sample forecasting. A particle filter algorithm is developed and shown to be useful for inference in MSM specifications with a large number of volatility states.

Keywords: Multivariate MSM, comovement, maximum likelihood, particle filter, Markov-switching, stochastic volatility, multifrequency volatility decomposition.

JEL Classification: C13, C32, F37, G15.

*Correspondence: Department of Economics, Littauer Center, Harvard University, Cambridge, MA 02138, and NBER, lcalvet@aya.yale.edu; Department of Finance, Sauder School of Business, University of British Columbia, 2053 Main Mall, Vancouver, BC Canada V6T 1Z2, adlai.fisher@sauder.ubc.ca; Department of Economics, Littauer Center, Harvard University, Cambridge, MA 02138, sthompson@harvard.edu. We thank John Campbell and Jim Stock for useful comments. We are very appreciative of financial support provided for this project by the Social Sciences and Humanities Research Council of Canada under grant 410-2002-0641.

1. Introduction

A growing body of research investigates the comovement of volatility in financial series. The motivation underlying this effort is well-known. Joint movements in volatility, quantified for instance by the conditional covariance matrix, play an important role in risk management, portfolio selection and tests of asset pricing. Comovements in volatility also help our understanding of financial markets, and shed light on issues such as contagion and the transmission of shocks through the financial system.[1] This motivation is particularly strong in the exchange rate literature, where first moments of exchange rates are only weakly related to fundamentals at medium frequencies and movements in volatility can be large and persistent (e.g. Meese and Rogoff, 1983; Rogoff, 1999; Sarno and Taylor, 2002; Clarida *et al.*, 2003).[2]

Multivariate GARCH, which was pioneered by Kraft and Engle (1982) and Bollerslev, Engle and Wooldridge (1988), is perhaps the most commonly used class of models. A natural extension of GARCH, these models assume that a vector transform of the covariance matrix can be written as a linear combination of its lagged values and the return innovations. These models have been shown to exhibit some empirical success over competing alternatives (e.g. Andersen, Bollerslev and Lange, 1999). They are however plagued by a number of difficulties. The dimensionality of the parameter space tends to be high, and it is not always straightforward to guarantee that the generated covariance-matrix is positive-definite. This has led authors to consider simplified versions of the model, such as the constant-conditional correlation GARCH (Bollerslev, 1990), the diagonal BEKK representation of Engle and Kroner (1995), or the more recent restrictions of Engle and Mezrich (1996) and Engle (2002).[3]

These difficulties are even more acute with multivariate stochastic volatility processes (Harvey, Ruiz, and Shephard, 1994). Estimation is generally delicate and is either moment-based (e.g. EMM) or requires the computation of a Hermite polynomial expansion (Ait-Sahalia, 2003). This is an unfortunate situation, since stochastic volatility models are the backbone of modern option pricing (e.g. Hull and White, 1987).

This paper proposes a different approach based on an earlier advance in univariate time series, the Markov Switching Multifractal (MSM) process introduced in Calvet and Fisher (2001, 2002*b*). This earlier research uses Markov-switching to develop the first time-stationary formulation of multifractal diffusions, and also provides a weakly convergent sequence of discrete filters. In this framework, total volatility is the multiplicative product of a large but finite number of random components, each first order Markov with

[1] See for instance Engle, Ito and Lin (1990) and Edwards and Susmel (2003).

[2] See Lyons (1995, 2001) and Andersen and Bollerslev (1998) for stronger evidence at high frequency.

[3] Researchers have also explored factor models (e.g. Engle, 1987; Diebold and Nerlove, 1989; Engle, Ng and Rotshchild, 1990), or proposed estimation methods other than maximum likelihood (Ledoit, Santa-Clara and Wolf, 2003).

an identical marginal distributions. The components are differentiated only by their respective time scales, governed by a unique switching probability for each frequency. The specification is completed by assuming that the progression of switching probabilities is approximately geometric. This parsimonious model delivers long-memory features in volatility, substantial outliers, and a decomposition into components with heterogenous decay rates. MSM compares favorably to earlier specifications both in- and out-of-sample. Univariate multifractal forecasts slightly improve on GARCH(1,1) at daily and weekly intervals, and provide considerable gains in accuracy at horizons of 10 to 50 days (Calvet and Fisher, 2002*b*).

In this paper, we investigate how the MSM volatility components of three exchange rates relate to other macroeconomic and financial variates. We consider the German Mark, the British Pound and the Japanese Yen, all versus the US Dollar over the period 1973-2002. At monthly frequencies, we find little evidence of a correlation between univariate MSM components and macroeconomic variables such as GDP, inflation, money supply or interest rates. There is evidence, however, of strong correlation between the volatility components of different exchange rates. More specifically, we find that components of similar frequencies in two different series tend to be strongly correlated. In contrast, components of different frequencies display less correlation, both within and across series.

This leads us to construct a multivariate stochastic volatility model exhibiting these features. We consider vectors M_{1t}, ..., $M_{\bar{k}t} \in \mathbb{R}^2$, which are first-order Markov and are characterized by the transition probabilities $\gamma_1, ..., \gamma_{\bar{k}}$. The return vector x_t is specified as

$$x_t = (M_{1t} * ... * M_{\bar{k}t})^{1/2} * \varepsilon_t, \qquad (1.1)$$

where $*$ denotes the Hadamard product[4] and the column vectors $\varepsilon_t \in \mathbb{R}^2$ are IID Gaussian $\mathcal{N}(0, \Sigma)$. This specification offers several advantages. It is relatively parsimonious, as the number of parameters is independent of $\bar{k}$. There is no issue of positive semi-definiteness. The likelihood function can be written in closed-form and ML estimation can be implemented for state spaces of reasonable size.

To accommodate larger state spaces, we develop a particle filter algorithm that permits convenient inference and forecasting using simulation methods. The good performance of the method is checked using a Monte Carlo experiment. The algorithm broadens the range of computationally tractable models to include cases where the number of volatility components (state variables) is very large, and to cases where the state variables are drawn from continuous rather than binomial distributions. This innovation thus opens econometric research on multifractal processes to a much wider range of specifications in both the multivariate and univariate cases.

[4] For any $x = (x_1, .., x_n) \in \mathbb{R}^n$ and $y = (y_1, ..., y_n) \in \mathbb{R}^n$, the Hadamard product is $x * y = (x_1 y_1, ..., x_n y_n)$.

Several applications of the model are considered. First, we investigate whether the multivariate model can help improve the out-of-sample forecasts provided by the univariate model. Second, we show that the model can be used to forecast the conditional covariance matrix and correlations of the currencies. Other application might use a frequency decomposition of volatility shocks to analyze the transmission volatility across currency markets.

The rest of the paper is organized as follows. Section 2 reviews univariate MSM and relates the volatility components to other financial and macroeconomic variates. Section 3 introduces the multivariate model, develops likelihood estimation, and proposes a particle filter algorithm. Empirical results are discussed in Section 4. We conclude in Section 5.

2. Comovement of Univariate Volatility Components

2.1. The Univariate Stochastic Volatility Model

We begin by reviewing the Markov-Switching Multifractal (MSM), a discrete-time Markov process with multi-frequency stochastic volatility. Consider an economic series X_t defined in discrete time on the regular grid $t = 0, 1, 2, ..., \infty$. In applications, X_t will be the log-price of a financial asset or exchange rate. We consider an economy with $\overline{k}$ components $M_{1,t}, M_{2,t}, ..., M_{\overline{k},t}$, which decay at heterogeneous frequencies $\gamma_1, .., \gamma_{\overline{k}}$. The notation MSM($\overline{k}$) will refer to versions of the model with $\overline{k}$ frequencies.

The innovations $x_t \equiv X_t - X_{t-1}$ are specified as

$$x_t = (M_{1,t} M_{2,t} ... M_{\overline{k},t})^{1/2} \varepsilon_t, \tag{2.1}$$

where the random variables ε_t are IID standard Gaussians $\mathcal{N}(0, \sigma^2)$. The random *multipliers* or *volatility components* $M_{k,t}$ are persistent, non-negative and satisfy $\mathbb{E}(M_{k,t}) = 1$. We assume for simplicity that the multipliers $M_{1,t}, M_{2,t} ... M_{\overline{k},t}$ at a given time t are statistically independent. The parameter σ is then equal to the unconditional standard deviation of the innovation x_t.

Equation (2.1) defines a return process with stochastic volatility $\sigma_t = \sigma(M_{1,t} M_{2,t} ... M_{\overline{k},t})^{1/2}$. We conveniently stack the period t volatility components into the $\overline{k} \times 1$ row vector

$$M_t = \left(M_{1,t}, M_{2,t}, ..., M_{\overline{k},t} \right).$$

The vector M_t is first-order Markov, which permits maximum likelihood estimation. It is then natural to call M_t the *volatility state vector*. The econometrician observes the returns x_t but not the vector M_t itself. The latent state M_t is inferred recursively by Bayesian updating.

Each $M_{k,t}$ follows a process that is identical except for time scale. Assume that the volatility state vectors have been constructed up to date $t-1$. For each $k \in \{1,..,\bar{k}\}$, the next period multiplier $M_{k,t}$ is drawn from a fixed distribution M with probability γ_k, and is otherwise equal to its current value: $M_{k,t} = M_{k,t-1}$. The switching events and new draws from M are assumed to be independent across k and t. The volatility components $M_{k,t}$ thus differ in their transition probabilities γ_k but not in their marginal distribution M.

The transition probabilities are specified as $\gamma_k = 1 - (1-\gamma_1)^{\left(b^{k-1}\right)}$, where $\gamma_1 \in (0,1)$ and $b \in (1,\infty)$. This specification is introduced in Calvet and Fisher (2001) in connection with the discretization of a Poisson arrival process. Consider a process with a small parameter γ_1. For small values of k, the transition probabilities of low frequency components grow approximately at geometric rate b :

$$\gamma_k \sim \gamma_1 b^{k-1}.$$

In empirical applications, it is numerically convenient to use $(\gamma_{\bar{k}}, b)$ in order to specify the set of transition probabilities.

The multifractal construction imposes only minimal restrictions on the marginal distribution of the multipliers: $M \geq 0$ and $\mathbb{E}(M) = 1$. While this allows flexible parametric or even nonparametric specifications of M, the paper focuses on the parsimonious setup in which M is drawn from a binomial random variable taking values m_0 or $2-m_0$ with equal probability. The full parameter vector is then

$$\psi \equiv (m_0, \sigma, b, \gamma_{\bar{k}}) \in \mathbb{R}^4,$$

where m_0 characterizes the distribution of the multipliers, σ is the unconditional standard deviation of returns, and b and $\gamma_{\bar{k}}$ define the set of switching probabilities.

2.2. Properties

The multiplicative structure (2.1) is appealing to model the high variability and high volatility persistence exhibited by financial time series. When a low level multiplier changes, volatility varies very discontinuously and has strong degree of persistence. In addition, high frequency multipliers introduce substantial outliers.

The MSM construction permits low frequency regime shifts, and thus long volatility cycles in sample paths. In exchange rate series, the duration of the most persistent component, $1/\gamma_1$, is typically of the same order as the length of the data. Estimated processes thus tend to generate volatility cycles with periods proportional to the sample size, a property also apparent in the sample paths of long memory processes. Long memory is often defined by a hyperbolic decline in the autocovariance function as the lag goes to infinity. Fractionally integrated processes generate such patterns by assuming that an innovation linearly affects future periods at a hyperbolically declining weight. As a result, fractional integration tends to produce smooth volatility processes. By contrast, our approach generates long cycles with a switching mechanism that also gives abrupt volatility changes. As shown in Calvet and Fisher (2002*b*), MSM mimics the hyperbolic autocovariograms exhibited by many financial series for a large range of intermediate lags, and is thus consistent with a large body of empirical evidence (e.g., Dacorogna *et al.*, 1993; Ding, Granger and Engle, 1993).[5] The combination of long-memory behavior with sudden volatility movements in MSM has a natural appeal for financial econometrics.

Another interesting property of MSM is that when $\bar{k} \to \infty$, the limiting continuous time process lies outside the class of Itô diffusions. The sample paths are continuous but exhibit a high degree of heterogeneity in local behavior, which is characterized by a continuum of local Hölder exponents in any finite time interval. We refer the reader to Calvet and Fisher (2001, 2002*a*) for a full development of the continuous-time limit.

2.3. Comovement of Volatility Components

The empirical analysis uses daily exchange rate data for the Deutsche Mark (DM), Japanese Yen (JA) and British Pound (UK), all against the US Dollar. The data consists of daily prices reported at noon by the Federal Reserve Bank of New York.[6] The series begin on 1 June 1974, shortly after the demise of the fixed exchange rate system. Since the Deutsche Mark was replaced by the Euro at the beginning of 1999, we chose to end all three series on 31 December 1998. Overall, the dataset contains 6,169 observations for each currency.

We estimate the model by maximum likelihood and report the results in Table 1 for all currencies. The columns correspond to the number of frequencies $\bar{k}$ varying from 1 to 8. The first column is thus a standard Markov-switching model with only two possible values for volatility. As $\bar{k}$ increases, the number of states increases at the rate $2^{\bar{k}}$. The multiplier value $\hat{m}_0$ tends to decline with $\bar{k}$. This is because with a larger number of volatility components, less variability is required in each individual component to generate the same aggregate amount of stochastic volatility. The estimates of $\hat{\sigma}$ show variability of a different type across $\bar{k}$, with no particular pattern of increasing or decreasing. When $\bar{k} = 1$, the parameter $\hat{\gamma}_{\bar{k}}$ indicates a switch in the single multiplier once every few weeks. As $\bar{k}$ increases, the switching probability of the highest frequency multiplier increases until for large values of $\bar{k}$ a switch occurs almost every day. At the same time, the estimated value $\hat{b}$ decreases steadily with $\bar{k}$. In the DM series with

[5] This result complements earlier research that has emphasized the difficulty of distinguishing between long memory and structural change in finite samples (e.g., Bhattacharya *et al.*, 1983; Hidalgo and Robinson, 1996; Diebold and Inoue, 2001).

[6] More specifically, the data consist of buying rates for wire transfers at 12:00 PM Eastern time.

$\bar{k} = 8$, a switch in the lowest frequency multiplier occurs approximately once every seven years, or about one fourth the sample size. Thus, as $\bar{k}$ increases, the range of frequencies spreads out while the spacing between frequencies becomes tighter.

We use the ML parameter estimates to compute the smoothed state probabilities (Kim, 1993), and therefore the conditional expectation of the multipliers $\hat{M}_{k,t}$ of each currency. We then compute the correlations of these components. In Table 2, we see that different components of the DM exchange rate exhibit limited correlation. We note that some correlation in the smoothed beliefs across multipliers is consistent with our assumption of mutual independence in the construction of the process. Consider the case when the volatility of a mid-range frequency multiplier switches to the high state. The econometrician's beliefs about the volatility state of this multiplier are likely to change, but because of imperfect information his beliefs about the state of nearby (in frequency space) components may change in a similar fashion. The correlation of volatility components within a series for the DM is qualitatively similar to unreported results from the UK and JA series.

We also attempted in unreported work to relate volatility components to macroeconomic variables. The variables of interest included monetary aggregates (M1, M2 and M3), short and long interest rates, producer price index, consumer price index, wages and the growth rate of industrial production. Using IMF data for the 1973-2000 period, we computed the correlation between monthly volatility and the macro variables of each country, their difference and the absolute value of their difference. We used several measures of volatility, including the absolute value of the monthly return, the realized monthly volatility, and MSM volatility components. We also ran regressions of volatility on lagged values. We found no robust link between volatility and macroeconomic fundamentals. These results are consistent with the findings of Andersen and Bollerslev (1998*a*), who show that the effect of macro announcements induces volatility shocks that are not unusually large compared to daily volatility. It is thus unsurprising that little impact is found at lower frequencies.

More positive results are obtained, however, when comparing the comovements of volatility across currencies. In Table 2, we report the correlation of volatility components across time series. We see that the correlation between the multipliers $\hat{M}^{\alpha}_{l,t}$ and $\hat{M}^{\beta}_{k,t}$ of two currencies α and β tends to be high when k and l are close, and is low otherwise. This suggests that the components of the series are correlated when the corresponding frequencies are close. We note that the univariate MSM series investigated in Table 2 typically have distinct frequencies, which could in principle complicate the analysis. We confirm the results of Table 2 by estimating a simple bivariate model in which the series are assumed to be statistically independent but are restricted to have identical frequency parameters b and $\gamma_{\bar{k}}$. It is convenient to call this specification the *combined univariate*. In Table 3, we report the estimation results for (DM,JA) and (DM,UK).

We then report in Table 4 the corresponding correlation in inferred multipliers. We note that as in Table 2, multipliers of identical frequencies tend to be most correlated. This suggests to consider a multivariate model in which the vectors $(\hat{M}^{\alpha}_{k,t}, \hat{M}^{\beta}_{k,t})$ have correlated components but are independent across k. The development of this process is presented in the next section.

The findings of Table 2 and 4 also suggest that the volatility comovements of major currencies tend to be positively correlated, and that this correlation is stronger when one considers components of similar frequencies. We also observe that correlation tends to be slightly higher at low than at high frequencies. These findings suggest that the volatility dynamics tend to be determined in the short run by country-specific events, such as innovations to equity returns or investor demand, while long run movements in volatility are more affected by general worldwide conditions. We also observe that volatility components of similar frequencies display much stronger correlations than coarser measures of volatility, such as the absolute value or the square of returns. From a methodological standpoint, this analysis thus confirms the benefits of the frequency decomposition of volatility provided by MSM.

3. A Multivariate Multifrequency Model

3.1. Stochastic Volatility

We consider two financial series α and β defined on the regular grid $t = 0, 1, 2, ..., \infty$. Their log-returns[7] x^{α}_t and x^{β}_t in period t are stacked into the column vector

$$x_t = \begin{bmatrix} x^{\alpha}_t \\ x^{\beta}_t \end{bmatrix} \in \mathbb{R}^2.$$

Volatility is stochastic and hit by shocks with heterogeneous frequencies. For every $k \in \{1, ..., \overline{k}\}$, we consider the first-order Markov column vector

$$M_{k,t} = \begin{bmatrix} M^{\alpha}_{k,t} \\ M^{\beta}_{k,t} \end{bmatrix} \in \mathbb{R}^2.$$

The period-t volatility components $M_{k,t}$ are stacked into the $2 \times \bar{k}$ matrix

$$M_t = (M_{1,t}, M_{2,t}, ..., M_{\overline{k},t}).$$

Each vector $M_{k,t}$ follows a process that is identical except for time scale. Assume that the volatility state vectors have been constructed up to date $t-1$. For each $k \in \{1, .., \overline{k}\}$, the next period vector multiplier $M_{k,t}$ is drawn from a fixed distribution

[7] If X^{α}_t denote the value of the exchange rate at date t, the log-return is $x^{\alpha}_t = \ln(X^{\alpha}_t / X^{\alpha}_{t-1})$.

M with probability γ_k, and is otherwise equal to its current value: $M_{k,t} = M_{k,t-1}$. The dynamics of $M_{k,t}$ can be summarized as:

$$\begin{array}{ll} M_{k,t} \text{ drawn from distribution } M & \text{with probability } \gamma_k \\ M_{k,t} = M_{k,t-1} & \text{with probability } 1-\gamma_k. \end{array}$$

The switching events and new draws from M are assumed to be independent across k and t. The volatility components $M_{k,t}$ thus differ in their transition probabilities γ_k but not in their marginal distribution M. These features greatly contribute to the parsimony of the model.

In applications, M_{kt} is drawn from a bivariate binomial distribution $M = (M^\alpha, M^\beta)'$. The first element of M can take values m_0^α and $m_1^\alpha = 2 - m_0^\alpha$, while the second component is either m_0^β or $m_1^\beta = 2 - m_0^\beta$. The random vector M can thus have four possible values, whose probabilities are parameterized by the matrix $(p_{i,j}) = (\mathbb{P}\{M = (m_i^\alpha, m_j^\beta)\})$. The conditions $\mathbb{P}(M^\alpha = m_0^\alpha) = 1/2$ and $\mathbb{P}(M^\beta = m_0^\beta) = 1/2$ impose that

$$\begin{bmatrix} p_{00} & p_{01} \\ p_{10} & p_{11} \end{bmatrix} = \begin{bmatrix} p & 1/2-p \\ 1/2-p & p \end{bmatrix}$$

for some $p \in [0, 1/2]$. We easily show that the correlation coefficient $\rho_m = Corr(M^\alpha; M^\beta)$ satisfies $\rho_m = 4p - 1$. In empirical applications, it is convenient to report the correlation coefficient ρ_m instead of the probability p. To obtain a parsimonious model, we thus assume that the distribution of volatility components is the same at all frequencies.

The random *volatility components* $M_{k,t}$ are persistent, non-negative and satisfy $\mathbb{E}(M_{k,t}) = 1$. Consistent with previous notation, let $*$ denote element by element multiplication, and $g(M_t)$ the 2×1 vector $M_{1,t} * M_{2,t} * ... * M_{\overline{k},t}$. The return vector x_t is specified as

$$x_t = [g(M_t)]^{1/2} * \varepsilon_t, \tag{3.1}$$

where the column vectors $\varepsilon_t \in \mathbb{R}^2$ are IID Gaussian $\mathcal{N}(0, \Sigma)$. The covariance matrix Σ can be written as

$$\Sigma = \begin{bmatrix} \sigma_\alpha^2 & \rho_\varepsilon \sigma_\alpha \sigma_\beta \\ \rho_\varepsilon \sigma_\alpha \sigma_\beta & \sigma_\beta^2 \end{bmatrix}.$$

The specification implies that the volatility of the series are correlated, and permits correlation in returns when $\rho_\varepsilon \neq 0$.

The univariate series satisfy

$$\begin{aligned} x_t^\alpha &= (M_{1,t}^\alpha M_{2,t}^\alpha ... M_{\overline{k},t}^\alpha)^{1/2} \varepsilon_t^\alpha \\ x_t^\beta &= (M_{1,t}^\beta M_{2,t}^\beta ... M_{\overline{k},t}^\beta)^{1/2} \varepsilon_t^\beta. \end{aligned}$$

The dynamics of the univariate series thus coincide with univariate MSM. We consider for simplicity that the multipliers $M_{1,t}, M_{2,t} ... M_{\overline{k},t}$ at a given time t are statistically independent. The unconditional standard deviation of each univariate series $i \in \{\alpha, \beta\}$ is then equal to σ_i.

As in the univariate case, the transition probabilities $\left(\gamma_1, \gamma_2, ..., \gamma_{\bar{k}}\right)$ are defined as

$$\gamma_k = 1 - (1 - \gamma_1)^{\left(b^{k-1}\right)},$$

where $\gamma_1 \in (0, 1)$ and $b \in (1, \infty)$. The bivariate process is thus specified by eight parameters

$$\psi \equiv (\sigma_\alpha, \sigma_\beta, m_0^\alpha, m_0^\beta, \rho_m, b, \gamma_{\bar{k}}, \rho_\varepsilon),$$

where σ_α and σ_β are the unconditional standard deviations of the series, m_0^α and m_0^β determine the levels of the volatility components, ρ_m their correlation, b and $\gamma_{\bar{k}}$ their transition probabilities, and ρ_ε the correlation of the Gaussian innovations. A key restriction is that the two series are assumed to have identical switching probabilities. Another is that the volatility components of the two series have identical correlations ρ_m at all frequencies. The empirical evidence of Tables 2 and 4 suggest that this is not an entirely unreasonable specification. A possible variant of the model would consider $\rho_1, ..., \rho_{\bar{k}}$, where a functional form involving two parameters would specify correlations as one goes from low to high frequencies.

3.2. Properties

The return series have unconditional means equal to zero: $\mathbb{E}\, x_t = 0$. Their correlation satisfies[8]

$$Corr(x_t^\alpha; x_t^\beta) = \rho_\varepsilon \prod_{k=1}^{\bar{k}} \mathbb{E}[(M_{k,t}^\alpha M_{k,t}^\beta)^{1/2}] \leq \rho_\varepsilon.$$

The upper bound ρ_ε is reached when the multipliers of both series are perfectly correlated. On the other hand when $\rho_\varepsilon < 1$, uncorrelated changes in volatility across series represent additional sources of noise that reduce the correlation of the asset returns.[9]

We now examine the conditional moments of the bivariate model. Returns are unpredictable: $\mathbb{E}_{t-1} x_t = 0$. Their comovement is quantified by the conditional covariance $Cov_t(x_{t+n}^\alpha; x_{t+n}^\beta) = \rho_\varepsilon \sigma_\alpha \sigma_\beta \prod_{k=1}^{\bar{k}} \mathbb{E}_t[(M_{k,t+n}^\alpha M_{k,t+n}^\beta)^{1/2}]$, and the conditional correlation

$$Corr_t(x_{t+n}^\alpha; x_{t+n}^\beta) = \rho_\varepsilon \prod_{k=1}^{\bar{k}} \frac{\mathbb{E}_t[(M_{k,t+n}^\alpha M_{k,t+n}^\beta)^{1/2}]}{[(\mathbb{E}_t M_{k,t+n}^\alpha)(\mathbb{E}_t M_{k,t+n}^\beta)]^{1/2}} \leq \rho_\varepsilon. \tag{3.2}$$

We observe that these quantities fluctuate through time with the multipliers. Thus while the construction assumes constant correlation between the Gaussian innovations and

[8] Note that the inequality stems from the Cauchy-Schwarz inequality.

[9] The correlation can in fact be arbitrary small for any levels of ρ_ε. For instance if $M_{k,t}^\alpha$ and $M_{k,t}^\beta$ are independent, the unconditional correlation $\rho_\varepsilon(\mathbb{E}\sqrt{M})^{2\bar{k}}$ is arbitrary small as M becomes more widely distributed.

between volatility components, the conditional correlation of returns is time-varying. We show in the Appendix that the conditional correlation of returns (3.2) is high when both currencies are volatile, or more specifically when their volatility components are simultaneously high.

Comovement in volatility can be similarly investigated. We show in the Appendix that when $\rho_m > 0$ and $\bar{k}$ is sufficiently large, the conditional correlation of squared returns is

$$Corr_t(|x^{\alpha}_{t+n}|^2; |x^{\beta}_{t+n}|^2) \sim C_{\varepsilon} \prod_{k=1}^{\bar{k}} \frac{\mathbb{E}_t(M^{\alpha}_{k,t+n} M^{\beta}_{k,t+n})}{\{\mathbb{E}_t[(M^{\alpha}_{k,t+n})^2]\mathbb{E}_t[(M^{\beta}_{k,t+n})^2]\}^{\frac{1}{2}}},$$

where $C_{\varepsilon} = \mathbb{E}[(\varepsilon^{\alpha}_1 \varepsilon^{\beta}_1)^2]/\{\mathbb{E}[(\varepsilon^{\alpha}_{t+n})^4]\mathbb{E}[(\varepsilon^{\beta}_{t+n})^4]\}^{\frac{1}{2}} \leq 1$. We note that the conditional correlation is positive even when the Gaussian noises are independent. Consistent with previous intuition, correlation between squared returns is high in periods of high volatility.

3.3. Likelihood Inference

Since the multiplier M has a discrete distribution, there exist a finite number of volatility states. Standard filtering methods then provide the likelihood function in closed-form.

The vector $M_t = (M_{1,t}, M_{2,t}, ..., M_{\overline{k},t})$ has $d = 4^{\overline{k}}$ possible values $m^1, ..., m^d \in \mathbb{R}^{\overline{k}}_{+}$. The dynamics of the Markov chain M_t are characterized by the transition matrix $A = (a_{i,j})_{1 \leq i,j \leq d}$ with components $a_{ij} = \mathbb{P}(M_{t+1} = m^j | M_t = m^i)$. Let $X_t \equiv \{x_s\}^t_{s=1}$ denote the set of past observations. By Bayes' rule, the conditional probability $\Pi^j_{t+1} = \mathbb{P}\left(M_{t+1} = m^j \,|X_t, x_{t+1}\right)$ satisfies

$$\Pi^j_{t+1} = \frac{f_{x_{t+1}}\left(x_{t+1}|M_{t+1} = m^j\right) \sum_{i=1}^{d} \Pi^i_t a_{ij}}{f_{x_{t+1}}\left(x_{t+1}\,|X_t\right)},$$

Let $f\left(x\right)$ the vector with elements $f^i\left(x\right) \equiv f_{x_{t+1}}\left(x_{t+1}|M_{t+1} = m^i\right)$. Note that each density $f^i(x)$ is a bivariate normals with zero mean and covariance matrices that differ across states. We stack these conditional probabilities into the row vector $\Pi_t = \left(\Pi^1_t, .., \Pi^d_t\right) \in \mathbb{R}^d_{+}$. Let $\iota = (1, ., 1) \in \mathbb{R}^d$. The updated probabilities are written in matrix notation as

$$\Pi_{t+1} = \frac{f\left(x_{t+1}\right) * \Pi_t A}{\left[f\left(x_{t+1}\right) * \Pi_t A\right] \iota'}.$$

This formula expresses the conditional probability distribution Π_{t+1} as a function of the observation x_{t+1} and the probability distribution Π_t calculated in period t. These results imply that Π_t can be computed recursively. In empirical applications, we choose the initial vector Π_0 to be equal to the ergodic distribution of the Markov process.

Since the multipliers $(M_{1,1},..,M_{\overline{k},1})$ are independent, the components of Π_0 are uniquely determined by $\Pi_0^j = \prod_{l=1}^{\overline{k}} \mathbb{P}(M = m_l^j)$ for all j.

The log likelihood function is

$$\ln L\,(x_1, ..., x_T; \psi) = \sum_{t=1}^{T} \ln[\omega(x_t) \cdot (\Pi_{t-1} A)].$$

For a fixed $\bar{k}$, the maximum likelihood estimator (ML) is consistent and asymptotically efficient as $T \to \infty$.

3.4. Particle Filter

Multivariate MSM is a parsimonious process that potentially involves very large state spaces. The volatility state vector M_t has $4^{\overline{k}}$ possible values, and the transition matrix A contains $4^{\overline{k}} \times 4^{\overline{k}}$ elements. For instance if $\overline{k} = 8$, the matrix A contains $2^{32} \approx 4 \times 10^9$ elements. More generally when $\overline{k}$ is large, it is computationally expensive to carry out volatility forecasts using the transition matrix. Following the recent literature on Markov chains,[10] we propose a simulation-based inference methodology.

Volatility forecasting involves computing features of the n-period ahead return distribution conditional on past values $X_t = (x_t, ..., x_1)$. We achieve this objective by designing a particle filter, a computational algorithm that generates draws from the distribution $M_t \mid X_t$ given a sample $M_{t-1}^{(1)}, \ldots, M_{t-1}^{(B)}$ from $M_{t-1} \mid X_{t-1}$. We begin the algorithm by drawing $M_0^{(1)}, \ldots, M_0^{(B)}$ from the ergodic distribution Π_0. The particle filter then allows us to recursively draw a sample $\{M_t^{(b)}\}_{b=1}^B$ from $M_t \mid X_t$ given the sample $\{M_{t-1}^{(b)}\}_{b=1}^B$ previously drawn from $M_{t-1} \mid X_{t-1}$. For each $M_t^{(b)}$, we can then simulate the multifractal process forward n steps to obtain a draw from $(M_{t+n}, x_{t+n}) \mid X_t$. Any feature of the forecast distribution may then be approximated. For example, the mean volatility forecast for x_{t+n} is $\mathbb{E}\left(x_{t+n}x_{t+n}' | X_t\right) \approx \frac{1}{B}\sum_{b=1}^B x_{t+n}^{(b)} x_{t+n}^{(b)\prime}$.

The particle filter starts with Bayes theorem: $\mathbb{P}\left(M_t|X_t\right) \propto f_{x_t}\left(x_t|M_t\right)\mathbb{P}\left(M_t|X_{t-1}\right)$, or equivalently

$$\mathbb{P}\left(M_t|X_t\right) \propto f_{x_t}\left(x_t|M_t\right) \sum_{M_{t-1}} \mathbb{P}\left(M_t|M_{t-1}\right) \mathbb{P}\left(M_{t-1}|X_{t-1}\right).$$

The normal density $f_{x_t}\left(x_t|M_t\right)$ is easy to compute, but the sum is computationally expensive for large values of $\overline{k}$. The draws $M_{t-1}^{(1)}, ..., M_{t-1}^{(B)}$ from $M_{t-1} \mid X_{t-1}$ imply the Monte Carlo approximation:

$$\mathbb{P}\left(M_t|X_t\right) \propto f_{x_t}\left(x_t|M_t\right) \frac{1}{B} \sum_{b=1}^{B} \mathbb{P}(M_t|M_{t-1}^{(b)}).$$

[10] See Chib, Nardari and Shephard (2002), Gordon, Salmond and Smith (1994), Jacquier, Polson and Rossi (1994), Kitagawa (1996) and Pitt and Shephard (1999).

As shown in the Appendix, we can use the following importance sampler to simulate the distribution of $M_t \mid X_t$.

1. Draw $M_t^{(1)}$, ..., $M_t^{(B)}$ from the distribution $\frac{1}{B}\sum_{b=1}^{B}\mathbb{P}(M_t|M_{t-1}^{(b)})$ by first drawing a random number j from 1 to B, and then simulating the series forward one period to draw M_t from $\mathbb{P}(M_t|M_{t-1}^{(j)})$.

2. For each $M_t^{(b)}$ compute the multivariate normal density $f_{x_t}(x_t|M_t^{(b)})$.

3. Draw a random number j from 1 to B with probability

$$\mathbb{P}(j=b) \equiv \frac{f_{x_t}(x_t|M_t^{(b)})}{\sum_{j=1}^{B} f_{x_t}(x_t|M_t^{(j)})}.$$

   The vector $M_t^{(j)}$ is the draw from $\mathbb{P}(M_t|X_t)$. Repeat B times to obtain B draws from the conditional distribution.

Likelihood function. We can use the particle filter to compute the likelihood function. Each one-step ahead density can be expressed as a sum over the unobserved volatility states: $f(x_t|X_{t-1}) = \sum_{M_t} f(x_t|M_t)\,\mathbb{P}(M_t|X_{t-1})$. Given simulated draws $M_t^{(b)}$ from $M_t \mid X_{t-1}$, the Monte Carlo approximation to the conditional density is thus

$$\widehat{f}(x_t|X_{t-1}) \equiv \frac{1}{B}\sum_{b=1}^{B} f(x_t|M_t^{(b)}).$$

We infer that the log-likelihood is approximately

$$\ln L\,(x_1,...,x_T;\psi) \approx \sum_{t=1}^{T} \ln \hat{f}(x_t|X_{t-1}).$$

In principle we can carry out simulated maximum likelihood based on this function.

Applications. The particle filter methodology extends the range of computationally feasible multifractal specifications. In previous work with univariate processes, Calvet and Fisher (2002*b*) report an approximate computational upper bound of ten binomial state variables, or 2^{10} states. While this produced good results in the univariate case, multivariate work requires a correspondingly larger number of state variables.

The development of the particle filter permits modelling extensions in two directions.[11] First, we may investigate models with a larger number of state variables. This

[11] Early drafts of Calvet and Fisher (2002*b*) experimented with other computational methods including SMM and EMM.

allows further development of inference and forecasting with multivariate MSM. Second, the particle filter also permits implementation of specifications where the state vector M is drawn from a continuous rather than a binomial distribution. Since earlier work (Calvet and Fisher, 2002*a*) has suggested that exchange rates are best fit by multipliers drawn from a lognormal distribution, it will be interesting in future work to revisit the lognormal specification and compare its out-of-sample performance to the results obtained here and in earlier work with the binomial.

Monte Carlo Simulations. Table 5 presents a Monte Carlo assessment of the importance sampling methodology. For $\bar{k} = 8$, we compute the log-likelihood of the univariate Deutsche Mark series using the particle filter and the ML estimates of Table 1. We compare the result with the exact value obtained in Table 1 by Bayesian updating. In order to assess the variation due to randomness, we compute the simulated likelihood value 1,000 times with $B = 1,000$ volatility draws. The particle filter simulations use the same Deutsche Mark series but are mutually independent. The table presents the average and standard deviation across simulations, along with some quantiles. Similar comparisons are also carried out for the forecasted volatility and kurtosis.

Table 5 reveals that the particle filter estimate of the log-likelihood is fairly precise. The standard deviation is small, but the average across simulations, $-5,406.2$, is below the true value of $-5,393.7$. Even the 99% quantile of the estimates is below the true value. The particle filter estimates of the forecast variance and kurtosis are accurate and approximately unbiased. Judging from the simulation bias and standard error, the particle filter becomes less accurate as the forecast horizon lengthens. These results overall confirm that the particle filter produces reasonable estimates of the likelihood and moments of the series. In future work, we will present a Monte Carlo simulation of the parameter estimates obtained by simulated maximum likelihood.

4. Empirical Results

4.1. Maximum Likelihood Estimates

We report in Table 6 the results of the bivariate ML estimation for (DM,JA) and (DM,UK). As in the univariate case, estimates of m_0 are declining with the number of frequencies, while the standard deviation $\hat{\sigma}$ is variable but displays no apparent trend. For each bivariate series, the estimates of m_0^α, m_0^β and σ appear to be quite close to the values reported in the univariate case (Table 1), especially for larger values of $\bar{k}$. We observe that $\gamma_{\bar{k}}$ increases relatively quickly with $\bar{k}$. The estimates of $\gamma_{\bar{k}}$ and b, which now jointly affect both series, are also broadly comparable to the values obtained in the univariate case.

The correlation between Gaussian innovations ρ_ε is positive and relatively constant across values of $\bar{k}$. The correlation between volatility components ρ_m is rather large and slightly increases with $\bar{k}$. We note that the Gaussian innovations and volatility components of the DM are most correlated with the UK. This result is not particularly surprising since Germany and the UK are geographically close and have markets that are open at the same time.[12] The looser economic and financial connections between Germany and Japan are consistent with the lower estimates of ρ_ε and ρ_m.

The likelihood functions sharply increase with the number of frequencies. For instance with (DM,UK), the log-likelihood increases by more than 200 when $\bar{k}$ goes from 2 to 5. Since the various models are non-nested and specified by the same number of parameters, these results indicate a very substantial increase of fit in-sample.

We can also compare the goodness of fit to the independent case. Consider for instance the (DM,UK) case. The independence assumption implies that the log-likelihood of the bivariate series is the sum of the log-likelihoods obtained in the univariate estimation (Table 1), which is equal to -10,220.86. In contrast, the bivariate estimate implies the likelihood function equal to -8,191.33, implying a gain in log-likelihood of 2,029.53. Note that this comparison requires no adjustment since the independent and the bivariate models are specified by the same number of parameters. The bivariate model thus implies considerable gains in accuracy in-sample as compared to independent univariate models.

4.2. Volatility Forecasts

We report in Table 7 the results of the univariate forecasts for horizons ranging from 1 to 50 days. The univariate is known to be a good benchmark. Previous work (Calvet and Fisher, 2002b) has shown that MSM substantially outperforms the out-of-sample volatility predictions of GARCH(1,1) and Markov-switching GARCH. We estimate MSM on the beginning of the series, and use the last twelve years of data (or approximately half the sample) for our out-of-sample forecasting comparison. Table 7 reports summary forecasting results and significance tests for horizons of 1, 5, 10, 20 and 50 days.

We compute the coefficients α and β from the Mincer-Zarnowitz OLS regressions of realized volatility on the n-period volatility forecast produced by MSM: $s_{t,n}^2 = \alpha + \beta \mathbb{E}_t x_{t,n}^2 + u_t$. More specifically, $x_{t,n} = \ln(X_{t+n}/X_t)$ is the n-period return and the dependent variable $s_{t,n}^2 = \sum_{s=t+1}^{t+n} x_s^2$ is the sum of squared daily returns, as in Andersen and Bollerslev (1998). These regressions are common in the financial econometrics literature,[13] and unbiased forecasts would imply $\alpha = 0$ and $\beta = 1$. We adjust the standard errors of α and β for parameter uncertainty as in West and McCracken

[12] The British Pound was also briefly pegged with other European currencies, including the Mark, when it participated in Exchange Rate Mechanism from October 1990 to September 1992.

[13] See for instance Pagan and Schwert (1990), West and Cho (1995), and Andersen, Bollerslev, and Meddahi (2002).

(1998), and for HAC effects using the weighting and lag selection methodology of Newey and West (1987, 1994). The MSM results show that for each currency, the estimated intercept $\hat{\alpha}$ is slightly positive and the slope $\hat{\beta}$ is slightly lower than 1. These small biases, however, are not statistically significant. In particular, the hypotheses $\alpha = 0$ and $\beta = 1$ are accepted at the 5% confidence level for most values of $\bar{k}$ and the forecast horizon. Because the average size of returns increases with the sampling interval, the estimated intercepts α are larger at longer horizons.

Table 7 also reports the forecast mean squared errors MSE and R^2 at various horizons.[14] We observe that the precision of the out-of-sample volatility forecasts is generally higher for specifications with more frequencies. This finding confirms the benefit of including components of heterogeneous frequencies into the model. The precision of the forecasts is also larger at longer horizons, which stems from the fact that the dependent variable $s^2_{t,n} = \sum_{s=t+1}^{t+n} x_s^2$ becomes a more accurate estimate of volatility as the horizon increases. These results contrast with the poor performance of traditional models such as GARCH(1,1) and Markov-switching GARCH at longer horizons. We also observe that while Table 7 only reports the DM results, analogous results are obtained for other currencies.

We investigate in Table 8 how these results compare to the forecasts obtained from the bivariate models. The challenge is to see whether we can improve the univariate forecasts obtained for the DM by using the bivariate model and thus bringing in information from another currency market. Since volatility is correlated across markets, this will presumably help detect the latent volatility states and thus help refine our forecasts. On the negative side, however, we anticipate that the additional number of parameters used in the bivariate specification make it more difficult to beat the univariate out-of-sample forecasts. Furthermore, we impose that frequencies be the same for both currencies in the bivariate specification, implying that the estimation may fit less precisely the range and spacing of frequencies in the currency of primary interest.

Tables $8a-b$ show that despite these potential difficulties, the DM volatility forecasts obtained from the bivariate model tend to improve on the univariate predictions. More specifically, we find that improvements in accuracy are obtained with UK and JA data. With UK and JA, however, the intercepts α and the slopes β of the Mincer-Zarnowitz regressions are closer to their ideal values than in the univariate case for given values of $\bar{k}$. The reported values are R^2 are also slightly superior. Overall, the multivariate model provides forecasts that are slightly better than the univariate ones.

[14] The multifractal yields a higher R^2 for n-day returns than for daily returns. This stems from the fact that our measure of n-day volatility is a sum of daily squared returns $\sum_{s=t+1}^{t+n} x_s^2$. As in Andersen and Bollerslev (1998), reduced noise in the volatility measure leads to an increase in explanatory power.

4.3. Forecasts of the Conditional Covariance Matrix

We now investigate how the bivariate model helps forecast the conditional covariance matrix of two series. The covariance matrix plays an important role in risk management, portfolio selection and tests of asset pricing. It can only be provided by a bivariate model. [To be completed]

5. Conclusion

This paper implements a multifractal frequency decomposition of exchange rate volatility and investigates whether it can be useful in developing improved forecasting models for variances and covariances. We find no robust relation between the smoothed beliefs about exchange rate volatility components and macroeconomic variates. We do, however, find that across currencies, the inferred states of volatility components of similar frequencies are highly correlated. This motivates our development of a multivariate Markov-switching multifractal process.

We develop a filter for the multivariate model that permits closed form expressions for the likelihood. This algorithm is used for estimation and forecasting of variances and covariances of three exchange rate series. Early results indicate some success, including improved variance forecasts from a bivariate relative to a univariate model for the DM series.

For larger problems the analytical updating algorithm suffers from reduced computational tractability. We therefore develop a particle filter that is particularly suitable for multivariate models with high-dimensional state spaces. We find that this tool shows promise in extending the range of computationally accessible problems, and are implementing the method empirically in ongoing work.

6. Appendix

6.1. Conditional Moments

Conditional Correlation in Returns. The ratios on the right-hand side of (3.2) can be rewritten as

$$\frac{\mathbb{E}(\sqrt{M_{1,t}^{\alpha}M_{1,t}^{\beta}}) + (1-\gamma_k)\left(\sqrt{M_{k,t}^{\alpha}M_{k,t}^{\beta}} - \mathbb{E}\sqrt{M_{1,t}^{\alpha}M_{1,t}^{\beta}}\right)}{\{[1+(1-\gamma_k)(M_{k,t}^{\alpha}-1)][1+(1-\gamma_k)(M_{k,t}^{\beta}-1)]\}^{1/2}}. \tag{6.1}$$

We observe that the correlation between the two returns is large when their volatility components are both large. For instance when $M_{1,t}^{\alpha}$ and $M_{1,t}^{\beta} \to \infty$, the ratio (6.1) is close to 1, which is its upper bound.

Conditional Correlation of Volatility. The conditional covariance of squared returns $Cov_t[(x_{t+n}^{\alpha})^2;(x_{t+n}^{\beta})^2]$ is equal to $\mathbb{E}[(\varepsilon_1^{\alpha}\varepsilon_1^{\beta})^2]\Pi_{k=1}^{\bar{k}}\mathbb{E}_t(M_{k,t+n}^{\alpha}M_{k,t+n}^{\beta}) - \sigma_{\alpha}^2\sigma_{\beta}^2\ \Pi_{k=1}^{\bar{k}}$ $\mathbb{E}_t M_{k,t+n}^{\alpha}\mathbb{E}_t M_{k,t+n}^{\beta}$, which can be approximated by $\mathbb{E}[(\varepsilon_1^{\alpha}\varepsilon_1^{\beta})^2]\Pi_{k=1}^{\bar{k}}\mathbb{E}_t(M_{k,t+n}^{\alpha}M_{k,t+n}^{\beta})$ [when the correlation between the multipliers is large or $\bar{k}$ is sufficiently large]. Their conditional variance of squared returns $Var_t[(x_{t+n}^{\alpha})^2]$ is $\mathbb{E}[(\varepsilon_{t+n}^{\alpha})^4]\Pi_{k=1}^{\bar{k}}\mathbb{E}_t[(M_{k,t+n}^{\alpha})^2] - \sigma_{\alpha}^4$ $\left(\Pi_{k=1}^{\bar{k}}\mathbb{E}_t M_{k,t+n}^{\alpha}\right)^2$.

6.2. Particle Filter

How do we sample from $\mathbb{P}(M_t|X_t)$? Consider a random variable $Y(M_t)$ with conditional expectation $\mathbb{E}[Y(M_t)|X_t]$. The conditional probability $\mathbb{P}(M_t|X_t)$ is unknown, but we can easily simulate from the mixture distribution $g(M_t) \equiv \frac{1}{B}\sum_{b=1}^{B}\mathbb{P}(M_t|M_{t-1}^{(b)})$. We rewrite the conditional expectation

$$\mathbb{E}[Y(M_t)|X_t] = \sum_{M_t} Y(M_t)\,\frac{\mathbb{P}(M_t|X_t)}{g(M_t)}g(M_t).$$

Given draws $M_t^{(1)}$, ..., $M_t^{(B)}$ from $g(M_t)$, the Monte Carlo approximation to this integral is

$$\mathbb{E}[Y(M_t)|X_t] = \sum_{b=1}^{B}\mu_b Y(M_t^{(b)}) \qquad \text{with } \mu_b = \frac{\mathbb{P}(M_t^{(b)}|X_t)}{Bg(M_t^{(b)})}$$

This approximation is valid for any Y, which suggests that we can approximate $\mathbb{P}(M_t|X_t)$ with a discrete distribution where M_t takes on the value $M_t^{(b)}$ with probability μ_b. This approximation to $\mathbb{P}(M_t|X_t)$ is called an importance sampler.

The probabilities μ_b may be simplified

$$\mu_b = \frac{f(x_t|M_t^{(b)})\,\frac{1}{B}\sum_{j=1}^{B}\mathbb{P}(M_t^{(b)}|M_{t-1}^{(j)})}{f(X_t)\,B\frac{1}{B}\sum_{j=1}^{B}\mathbb{P}(M_t^{(b)}|M_{t-1}^{(j)})} = \frac{f(x_t|M_t^{(b)})}{f(X_t)\,B}.$$

$f(x_t|M_t^{(b)})$ is easy to calculate since, conditional on $M_t^{(b)}$, x_t is distributed multivariate normal with zero mean and covariance

$$S_t^{(b)} = \begin{bmatrix} (\sigma_t^{\alpha,(b)})^2 & \rho_\varepsilon \sigma_t^{\alpha,(b)} \sigma_t^{\beta,(b)} \\ \rho_\varepsilon \sigma_t^{\alpha,(b)} \sigma_\beta & (\sigma_t^{\beta,(b)})^2 \end{bmatrix},$$

with $\sigma_t^{\alpha,(b)} = \sigma_\alpha (M_{1,t}^{\alpha,(b)} M_{2,t}^{\alpha,(b)} \cdots M_{\overline{k},t}^{\alpha,(b)})^{1/2}$ and $\sigma_t^{\beta,(b)} = \sigma_\beta (M_{1,t}^{\beta,(b)} M_{2,t}^{\beta,(b)} \cdots M_{\overline{k},t}^{\beta,(b)})^{1/2}$. We do not need to calculate $p\,(X_t)$, since the μ_b must sum to one. Thus we approximate the probabilities with $\mu_b \approx \widehat{\mu}_b \equiv \frac{f(x_t|M_t^{(b)})}{\sum_{j=1}^B f\left(x_t|M_t^{(j)}\right)}$.

6.3. A Multivariate Approximation for Univariate Forecasting

The univariate series have state spaces S^α and S^β of dimension 2^k. The bivariate series has a state space $S^\alpha \times S^\beta$ of dimension 4^k. Our belief over the bivariate state space is given by the vector Π_t. We denote by Π_{t+1}^β the belief vector from the univariate series β.

The conditional probability $\Pi_{t+1}^\alpha(m^\alpha) = \mathbb{P}(M_{t+1}^\alpha = m^\alpha | \Pi_{t+1}^\beta, \Pi_t^\alpha, x_{t+1})$ can be recursively computed as follows:

$$\begin{aligned} &\Pi_{t+1}^\alpha(m^\alpha) \\ \propto \; &\sum_{n^\beta \in S^\beta} \Pi_{t+1}^\beta(n^\beta) f(x_{t+1}|M_{t+1}^\alpha = m^\alpha, M_{t+1}^\beta = n^\beta) \sum_{m \in S^\alpha} \frac{\Pi_t^\alpha(m) \mathbb{P}(M_{t+1}^\alpha = m^\alpha, M_{t+1}^\beta = n^\beta | M_t^\alpha = m)}{\mathbb{P}(M_{t+1}^\beta = n^\beta | M_t^\alpha = m)}. \end{aligned}$$

I like this rule for the following reasons:

- In order for ρ_ε to show up, we need to use both x_{t+1}^β and Π_{t+1}^β.
- In order for ρ_m to show up, it is important to have an expression of the type $\mathbb{P}(M_{t+1}^\alpha = m^\alpha, M_{t+1}^\beta = n^\beta | M_t^\alpha = m)$.
- The computation involves the $2^k \times 4^k$ matrix $\mathbb{P}(M_{t+1}^\alpha = m^\alpha, M_{t+1}^\beta = n^\beta | M_t^\alpha = m)$, instead of the $4^k \times 4^k$ transition matrix of the bivariate process.
- If M_{t+1}^α and M_{t+1}^β are independent, the rule reduces to univariate updating.

Heuristic Derivation. The conditional probability $\Pi_{t+1}^\alpha(m^\alpha) = \mathbb{P}(M_{t+1}^\alpha = m^\alpha | \Pi_{t+1}^\beta, \Pi_t^\alpha, x_{t+1})$ is approximately equal to

$$\sum_{n^\beta \in S^\beta} \Pi_{t+1}^\beta(n^\beta) \mathbb{P}(M_{t+1}^\alpha = m^\alpha | M_{t+1}^\beta = n^\beta, \Pi_t^\alpha, x_{t+1}).$$

Bayes' theorem implies $\mathbb{P}(M^{\alpha}_{t+1} = m^{\alpha}|M^{\beta}_{t+1} = n^{\beta}, \Pi^{\alpha}_{t}, x_{t+1}) \propto f(x_{t+1}|M^{\alpha}_{t+1} = m^{\alpha}, M^{\beta}_{t+1} = n^{\beta})\mathbb{P}(M^{\alpha}_{t+1} = m^{\alpha}|M^{\beta}_{t+1} = n^{\beta}, \Pi^{\alpha}_{t})$. Thus,

$$\Pi^{\alpha}_{t+1}(m^{\alpha}) \propto \sum_{n^{\beta}\in S^{\beta}} \Pi^{\beta}_{t+1}(n^{\beta}) f(x_{t+1}|M^{\alpha}_{t+1} = m^{\alpha}, M^{\beta}_{t+1} = n^{\beta})\mathbb{P}(M^{\alpha}_{t+1} = m^{\alpha}|M^{\beta}_{t+1} = n^{\beta}, \Pi^{\alpha}_{t}).$$

We note that $\mathbb{P}(M^{\alpha}_{t+1} = m^{\alpha}|M^{\beta}_{t+1} = n^{\beta}, \Pi^{\alpha}_{t}) = \sum_{m\in S^{\alpha}} \Pi^{\alpha}_{t}(m)\mathbb{P}(M^{\alpha}_{t+1} = m^{\alpha}|M^{\beta}_{t+1} = n^{\beta}, M^{\alpha}_{t} = m)$, or

$$\mathbb{P}(M^{\alpha}_{t+1} = m^{\alpha}|M^{\beta}_{t+1} = n^{\beta}, \Pi^{\alpha}_{t}) = \sum_{m} \frac{\Pi^{\alpha}_{t}(m)\mathbb{P}(M^{\alpha}_{t+1} = m^{\alpha}, M^{\beta}_{t+1} = n^{\beta}|M^{\alpha}_{t} = m)}{\mathbb{P}(M^{\beta}_{t+1} = n^{\beta}|M^{\alpha}_{t} = m)}.$$

We conclude that $\Pi^{\alpha}_{t+1}(m^{\alpha})$ satisfies the above updating rule.

References

[1] Ait-Sahalia, Y. (2003), "Closed-Form Likelihood Expansions for Multivariate Diffusions," Working Paper, Princeton University.

[2] Andersen, T., and Bollerslev, T. (1998*a*), Deutsche-Mark Dollar Volatility: Intraday Activity Patterns, Macroeconomic Announcements, and Longer-Run Dependencies, *Journal of Finance* **53**, 219-265.

[3] Andersen, T.G., and Bollerslev, T. (1998*b*), Answering the Skeptics: Yes, Standard Volatility Models Do Provide Accurate Forecasts, *International Economic Review* **39**, 885-905.

[4] Andersen, T., Bollerslev, T., and Lange, S. (1999), Forecasting Financial Market Volatility: Sample Frequency vis-à-vis Forecast Horizon, *Journal of Empirical Finance* **6**, 457-477.

[5] Andersen, T. G., Bollerslev, T., and Meddahi, N. (2002), "Analytic Evaluation of Volatility Forecasts," CIRANO Working Paper 2002s-90.

[6] Andersen, T., Bollerslev, T., Diebold, F., and Labys, P. (2003), Modeling and Forecasting Realized Volatility, *Econometrica* **71**, 529-626.

[7] Bhattacharya, R. N., Gupta, V. K., and Waymire, E. (1983), The Hurst Effect under Trends, *Journal of Applied Probability* **20**, 649-662.

[8] Bollerslev, T. (1986), Generalized Autoregressive Conditional Heteroskedasticity, *Journal of Econometrics* **31**, 307-327.

[9] Bollerslev, T. (1990), Modelling the Coherence in Short-Run Nominal Exchange Rates: A Multivariate Generalized ARCH Approach, *Review of Economics and Statistics* **72**, 498-505.

[10] Bollerslev, T., Engle, R., and Wooldridge, J. (1988), A Capital Asset Pricing Model with Time Varying Covariances, *Journal of Political Economy* **96**, 116-131.

[11] Calvet, L., and Fisher, A. (2001), Forecasting Multifractal Volatility, *Journal of Econometrics* **105**, 27-58.

[12] Calvet, L., and Fisher, A. (2002*a*), Multifractality in Asset Returns: Theory and Evidence, *Review of Economics and Statistics* **84**.

[13] Calvet, L., and Fisher, A. (2002*b*), "Regime-Switching and the Estimation of Multifractal Processes," NYU and UBC Working Papers.

[14] Calvet, L., Fisher, A., and Mandelbrot, B. B. (1997), Cowles Foundation Discussion Paper No. 1164-1166, Yale University. Papers available from http://www.ssrn.com.

[15] Chib, S., Nardari, F., and Shephard, N. (2002), Markov Chain Monte Carlo Methods for Stochastic Volatility Models, *Journal of Econometrics* **108**, 281-316.

[16] Clarida, R., Sarno, L., Taylor, M., and Valente, G. (2003), The Out-of-Sample Success of Term Structure Models as Exchange Rate Predictors: A Step Beyond, *Journal of International Economics* **60**, 61-83.

[17] Dacorogna, M., Müller, U., Nagler, R., Olsen, R., and Pictet, O. (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.

[18] Diebold. F. X., and Inoue, A. (2001), Long Memory and Regime Switching, *Journal of Econometrics* **105***,* 27-58.

[19] Diebold, F., and Nerlove, M. (1989), The Dynamics of Exchange Rate Volatility: A Multivariate Latent Factor ARCH Model, *Journal of Applied Econometrics* **4**, 1-21.

[20] Diebold, F., and Rudebusch, G. (1996), Measuring Business Cycles: A Modern Perspective, *Review of Economics and Statistics* **78**, 67-77.

[21] 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.

[22] Edwards, S., and Susmel, S. (2003), Interest-Rate Volatility in Emerging Markets, *Review of Economics and Statistics* **85**, 328-348.

[23] Engle, R. (1987), "Multivariate GARCH with Factor Structures - Cointegration in Variance," Working Paper, Department of Economics, University of California - San Diego.

[24] Engle, R. (2002), Dynamic Conditional Correlation: A Simple Class of Multivariate Geneneralized Autoregressive Conditional Heteroskedasticity Models, *Journal of Business and Economic Statistics* **20**, 339-350.

[25] Engle, R., and Kroner, K. (1995), Multivariate Simultaneous Generalized ARCH, *Econometric Theory* **11**, 122-150.

[26] Engle, R., and Mezrich, J. (1996), "GARCH for Groups," *Risk* **9**, 36-40.

[27] Engle, R., Ito, T., and Lin, W.-L. (1990), Meteor Showers or Heat Waves? Heteroskedastic Intra-Daily Volatility in the Foreign Exchange Market, *Econometrica* **58**, 525-542.

[28] Engle, R., Ng, V., and Rothschild, M. (1990), Asset Pricing with a Factor ARCH Covariance Structure: Empirical Estimates for Treasury Bills, *Journal of Econometrics* **45**, 213-238.

[29] Ghysels, E., Harvey, A. C., and Renault, E. (1996), Stochastic Volatility, in: G. S. Maddala and C. R. Rao Eds., *Handbook of Statistics*, Vol. 14, 119-191. Amsterdam: North-Holland.

[30] Gordon, N., Salmond, D., and Smith, A. (1993), A Novel Approach to Non-Linear and Non-Gaussian Bayesian State Estimation, *IEEE Proceedings F* **140**, 107-113.

[31] Hamilton, J. D. (1989), A New Approach to the Economic Analysis of Nonstationary Time Series and the Business Cycle, *Econometrica* **57**, 357-84.

[32] Hamilton, J. D. (1990), Analysis of Time Series Subject to Change in Regime, *Journal of Econometrics* **45**, 39-70.

[33] Hamilton, J. D. and R. Susmel (1994), Autoregressive Conditional Heteroskedasticity and Changes in Regime, *Journal of Econometrics* **64**, 307-333.

[34] Harvey, A., Ruiz, E., and Shephard, N. (1994), Multivariate Stochastic Variance Models, *Review of Economic Studies* **61**, 247-264.

[35] Hidalgo, J., and Robinson, P. M. (1996), Testing for Structural Change in a Long-Memory Environment, *Journal of Econometrics* **70**, 159-174.

[36] Hull, J., and White, A. (1987), The Pricing of Options on Assets with Stochastic Volatilities, *Journal of Finance* **42**, 281-300.

[37] Ito, T., Engle, R., and Lin, W.-L. (1992), Where does the Meteor Shower Come From? The Role of Stochastic Policy Coordination, *Journal of International Economics* **32**, 221-240.

[38] Jacquier, E., Polson, N., and Rossi, P. (1994), Bayesian Analysis of Stochastic Volatility Models, *Journal of Business and Economic Statistics* **12**, 371-417.

[39] Kim, C. J. (1993), Dynamic Linear Models with Markov-Switching, *Journal of Econometrics* **60**, 1-22.

[40] Kim, C. J., and Nelson, C. R. (1998), Business Cycle Turning Points, A New Coincident Index, and Tests of Duration Dependence Based on a Dynamic Factor Model with Regime Switching, *Review of Economics and Statistics* **80**, 188-201.

[41] Kitagawa, G. (1996), Monte Carlo Filter and Smoother for Gaussian Nonlinear State Space Models, *Journal of Computational and Graphical Statistics* **5**, 1-25.

[42] Kraft, D., and Engle, R. (1982), "Autoregressive Conditional Heteroskedasticity in Multiple Time Series," unpublished manuscript, Department of Economics, UCSD.

[43] Ledoit, O., Santa-Clara, P., and Wolf, M. (2003), Flexible Multivariate GARCH Modeling with an Application to International Stock Markets, *Review of Economics and Statistics* **85**, 735-747.

[44] Lyons, R. (1995), Tests of Microstructural Hypotheses in the Foreign Exchange Market, *Journal of Financial Economics* **39**, 321-351.

[45] Lyons, R. (2001), *The Microstructure Approach to Exchange Rates*, MIT Press.

[46] Meese, R., and Rogoff, K. (1983), Empirical Exchange Rate Models of the Seventies: Do They Fit Out of Sample?, *Journal of International Economics* **14**, 3-24.

[47] Melvin, M., and Melvin, B. P. (2003), The Global Transmission of Volatility in the Foreign Exchange Market, *Review of Economics and Statistics* **85**, 670-679.

[48] Newey, W., and West, K. (1987), A Simple, Positive Semi-Definite, Heteroskedasticity and Autocorrelation Consistent Covariance Matrix, *Econometrica* **55**, 703-708.

[49] Newey, W., and West, K. (1994), Automatic Lag Selection in Covariance Matrix Estimation, *Review of Economic Studies* **61**, 631-654.

[50] Pagan, A., and Schwert, W. (1990), Alternative Models for Conditional Stock Volatility, *Journal of Econometrics* **45**, 267-290.

[51] Pitt, M., and Shephard, N. (1999), Filtering via Simulation: Auxiliary Particle Filter, *Journal of the American Statistical Association* **94**, 590-599.

[52] Rogoff, K. (1999), Perspectives on Exchange Rate Volatility, in *International Capital Flows,* M. Feldstein Ed., 441-53, University of Chicago Press.

[53] Sarno, L., and Taylor, M. (2002), *The Economics of Exchange Rates*, Cambridge University Press.

[54] West, K., and Cho, D. (1995), The Predictive Ability of Several Models of Exchange Rate Volatility, *Journal of Econometrics* **69**, 367-391.

[55] West, K., and McCracken, M. (1998), Regression-Based Tests of Predictive Ability, *International Economic Review* **39**, 817-840.

TABLE 1. – UNIVARIATE MLE

	$\bar{k}=1$	2	3	4	5	6	7	8
				Deutsche Mark				
$\hat{m}_0$	1.644	1.583	1.546	1.485	1.456	1.406	1.373	1.346
	(0.014)	(0.017)	(0.014)	(0.013)	(0.014)	(0.013)	(0.012)	(0.011)
$\hat{\sigma}$	0.667	0.626	0.594	0.568	0.521	0.537	0.540	0.541
	(0.011)	(0.020)	(0.015)	(0.018)	(0.024)	(0.022)	(0.020)	(0.025)
$\hat{\gamma}_{\bar{k}}$	0.072	0.097	0.709	0.728	0.760	0.885	0.956	0.987
	(0.011)	(0.018)	(0.148)	(0.098)	(0.103)	(0.122)	(0.058)	(0.033)
$\hat{b}$	-	9.95	24.03	11.06	8.63	5.59	4.38	3.56
		(4.04)	(8.17)	(2.05)	(1.63)	(0.88)	(0.56)	(0.43)
ln L	-5601.73	-5470.68	-5422.59	-5405.44	-5397.80	-5395.55	-5393.16	-5393.72
				Japanese Yen				
$\hat{m}_0$	1.794	1.767	1.673	1.636	1.620	1.549	1.549	1.500
	(0.011)	(0.003)	(0.009)	(0.011)	(0.011)	(0.012)	(0.011)	(0.009)
$\hat{\sigma}$	0.636	0.542	0.567	0.456	0.684	0.656	0.527	0.506
	(0.011)	(0.009)	(0.017)	(0.012)	(0.021)	(0.024)	(0.020)	(0.017)
$\hat{\gamma}_{\bar{k}}$	0.197	0.285	0.404	0.713	0.791	0.943	0.942	0.999
	(0.022)	(0.030)	(0.085)	(0.089)	(0.087)	(0.053)	(0.049)	(0.002)
$\hat{b}$	-	962.82	17.09	20.95	20.70	10.43	10.40	8.17
		(1062.77)	(3.47)	(4.53)	(4.38)	(1.78)	(1.80)	(1.51)
ln L	-5387.12	-5111.36	-4997.46	-4958.58	-4938.52	-4929.90	-4930.49	-4925.71
				British Pound				
$\hat{m}_0$	1.745	1.697	1.675	1.626	1.592	1.552	1.517	1.470
	(0.013)	(0.012)	(0.014)	(0.012)	(0.002)	(0.013)	(0.016)	(0.011)
$\hat{\sigma}$	0.619	0.585	0.492	0.463	0.393	0.490	0.396	0.393
	(0.010)	(0.016)	(0.016)	(0.015)	(0.013)	(0.023)	(0.017)	(0.017)
$\hat{\gamma}_{\bar{k}}$	0.131	0.247	0.312	0.678	0.711	0.793	0.802	0.956
	(0.018)	(0.035)	(0.055)	(0.091)	(0.063)	(0.080)	(0.081)	(0.053)
$\hat{b}$	-	25.03	17.16	13.32	10.76	8.72	6.58	5.09
		(6.94)	(3.65)	(2.29)	(0.80)	(1.35)	(1.02)	(0.68)
ln L	-5219.33	-4996.72	-4899.76	-4851.44	-4823.06	-4811.97	-4807.47	-4805.59

Notes: This table shows maximum likelihood estimation results for the binomial multifractal model. Columns correspond to the number of frequencies $\bar{k}$ in the estimated model. Asymptotic standard errors are in parenthesis.

TABLE 2. – Correlation of Univariate Volatility Components

	DM1	DM2	DM3	DM4	DM5	DM6	DM7	DM8	$\lvert x_{\text{DM}}\rvert$	x^2_{DM}
DM1	1.000	0.738	0.310	0.136	0.073	0.055	0.028	0.016	0.201	0.126
DM2	0.738	1.000	0.584	0.325	0.155	0.099	0.048	0.028	0.271	0.181
DM3	0.310	0.584	1.000	0.627	0.322	0.186	0.089	0.053	0.320	0.241
DM4	0.136	0.325	0.627	1.000	0.749	0.461	0.224	0.134	0.383	0.297
DM5	0.073	0.155	0.322	0.749	1.000	0.814	0.452	0.284	0.457	0.359
DM6	0.055	0.099	0.186	0.461	0.814	1.000	0.782	0.565	0.629	0.496
DM7	0.028	0.048	0.089	0.224	0.452	0.782	1.000	0.927	0.853	0.673
DM8	0.016	0.028	0.053	0.134	0.284	0.565	0.927	1.000	0.880	0.692
$\lvert x_{\text{DM}}\rvert$	0.201	0.271	0.320	0.383	0.457	0.629	0.853	0.880	1.000	0.866
x^2_{DM}	0.126	0.181	0.241	0.297	0.359	0.496	0.673	0.692	0.866	1.000
JA1	0.887	0.455	0.194	0.067	0.044	0.036	0.017	0.010	0.139	0.081
JA2	0.888	0.460	0.202	0.070	0.045	0.037	0.018	0.010	0.140	0.081
JA3	0.889	0.460	0.204	0.070	0.045	0.037	0.018	0.010	0.140	0.081
JA4	0.739	0.341	0.199	0.072	0.024	0.028	0.015	0.009	0.116	0.075
JA5	0.160	0.156	0.173	0.203	0.137	0.080	0.041	0.025	0.089	0.071
JA6	0.098	0.100	0.158	0.289	0.371	0.306	0.175	0.112	0.208	0.180
JA7	0.033	0.031	0.041	0.093	0.180	0.306	0.369	0.328	0.337	0.299
JA8	0.015	0.014	0.018	0.045	0.096	0.206	0.360	0.397	0.375	0.333
$\lvert x_{\text{JA}}\rvert$	0.216	0.120	0.126	0.154	0.190	0.263	0.356	0.364	0.415	0.384
x^2_{JA}	0.107	0.052	0.090	0.120	0.154	0.217	0.287	0.286	0.331	0.355
UK1	0.974	0.635	0.256	0.113	0.061	0.047	0.023	0.013	0.177	0.107
UK2	0.727	0.619	0.287	0.189	0.097	0.063	0.031	0.018	0.165	0.096
UK3	0.644	0.639	0.251	0.179	0.090	0.059	0.028	0.016	0.170	0.110
UK4	0.344	0.525	0.416	0.284	0.137	0.081	0.037	0.021	0.192	0.143
UK5	0.018	0.073	0.476	0.589	0.404	0.238	0.115	0.069	0.227	0.186
UK6	0.148	0.167	0.228	0.421	0.543	0.470	0.275	0.177	0.312	0.258
UK7	0.081	0.086	0.095	0.178	0.305	0.453	0.465	0.379	0.424	0.365
UK8	0.030	0.033	0.037	0.072	0.139	0.281	0.470	0.503	0.488	0.418
$\lvert x_{\text{UK}}\rvert$	0.213	0.231	0.232	0.267	0.285	0.357	0.466	0.476	0.604	0.557
x^2_{UK}	0.113	0.144	0.182	0.223	0.245	0.305	0.395	0.404	0.543	0.606

Notes: This table shows correlations from a frequency decomposition of the univariate multifractal exchange rate models with eight components. For each series, the smoothed probabilities of different volatility states are calculated. These are used to construct eight sub-series giving the smoothed probability that a given multiplier is in the high volatility state. These series are labeled from one to eight, with the first corresponding to the lowest frequency. The table then shows correlations of these series within and across currency models.

TABLE 3. – Combined Univariate MLE

	$\bar{k}=1$	2	3	4	5	6	7	8
				DM and JA				
$\hat{m}_0^{\mathrm{DM}}$	1.668	1.636	1.538	1.488	1.470	1.415	1.410	1.376
	(0.013)	(0.016)	(0.014)	(0.015)	(0.013)	(0.013)	(0.070)	(0.012)
$\hat{m}_0^{\mathrm{JA}}$	1.778	1.751	1.673	1.624	1.619	1.549	1.499	1.463
	(0.014)	(0.010)	(0.010)	(0.017)	(0.011)	(0.011)	(0.056)	(0.012)
$\hat{\sigma}_{\mathrm{DM}}$	0.664	0.574	0.597	0.571	0.507	0.546	0.449	0.462
	(0.011)	(0.013)	(0.015)	(0.026)	(0.021)	(0.024)	(0.079)	(0.000)
$\hat{\sigma}_{\mathrm{JA}}$	0.626	0.547	0.564	0.482	0.677	0.646	0.614	0.547
	(0.013)	(0.010)	(0.016)	(0.037)	(0.029)	(0.025)	(0.034)	(0.002)
$\hat{\gamma}_{\bar{k}}$	0.124	0.206	0.462	0.735	0.740	0.882	0.987	0.993
	(0.012)	(0.019)	(0.109)	(0.070)	(0.064)	(0.047)	(0.011)	(0.001)
$\hat{b}$	-	84.77	17.05	13.36	13.74	7.42	6.86	5.17
		(40.30)	(3.66)	(2.77)	(2.09)	(0.70)	(0.85)	(0.00)
ln L	-11003.28	-10604.23	-10421.73	-10369.40	-10345.89	-10332.45	-10322.84	-10321.88
				DM and UK				
$\hat{m}_0^{\mathrm{DM}}$	1.657	1.597	1.533	1.487	1.460	1.413	1.407	1.377
	(0.013)	(0.017)	(0.014)	(0.014)	(0.013)	(0.014)	(0.013)	(0.013)
$\hat{m}_0^{\mathrm{UK}}$	1.730	1.688	1.672	1.623	1.591	1.540	1.506	1.467
	(0.013)	(0.012)	(0.013)	(0.012)	(0.011)	(0.012)	(0.013)	(0.011)
$\hat{\sigma}_{\mathrm{DM}}$	0.665	0.624	0.598	0.569	0.526	0.544	0.457	0.459
	(0.011)	(0.022)	(0.015)	(0.022)	(0.022)	(0.024)	(0.019)	(0.018)
$\hat{\sigma}_{\mathrm{UK}}$	0.618	0.587	0.497	0.467	0.393	0.476	0.405	0.398
	(0.010)	(0.016)	(0.015)	(0.014)	(0.012)	(0.021)	(0.022)	(0.018)
$\hat{\gamma}_{\bar{k}}$	0.096	0.171	0.310	0.705	0.725	0.777	0.832	0.954
	(0.010)	(0.023)	(0.063)	(0.055)	(0.063)	(0.070)	(0.076)	(0.040)
$\hat{b}$	-	16.61	13.98	12.22	9.90	6.55	5.77	4.68
		(3.56)	(2.54)	(1.45)	(1.16)	(0.67)	(0.54)	(0.43)
ln L	-10825.35	-10474.47	-10325.40	-10258.07	-10222.33	-10213.72	-10204.79	-10200.43

Notes: This table shows maximum likelihood estimation results for the combined univariate model. This involves joint estimation of two univariate models where the paramters b and $\gamma_{\bar{k}}$ are required to be identical across currencies. This also corresponds to a restricted bivariate model where all correlations are zero. Columns correspond to the number of frequencies $\bar{k}$ in the estimated model. Asymptotic standard errors are in parenthesis.

TABLE 4. – Correlation of Combined Univariate Volatility Components

	DM1	DM2	DM3	DM4	DM5	DM6	DM7	DM8	$\lvert x_{\mathrm{DM}}\rvert$	x^2_{DM}
JA1	0.901	0.895	0.383	0.123	-0.002	0.027	0.015	0.007	0.139	0.081
JA2	0.901	0.895	0.378	0.126	-0.003	0.027	0.015	0.007	0.139	0.081
JA3	0.803	0.792	0.169	0.133	0.003	0.018	0.012	0.006	0.103	0.062
JA4	0.411	0.415	0.263	0.166	0.059	0.028	0.019	0.010	0.098	0.077
JA5	0.065	0.066	0.112	0.191	0.334	0.237	0.107	0.056	0.151	0.122
JA6	0.024	0.024	0.043	0.094	0.261	0.361	0.256	0.154	0.221	0.195
JA7	0.010	0.010	0.014	0.030	0.100	0.229	0.372	0.336	0.340	0.305
JA8	0.005	0.005	0.006	0.014	0.052	0.137	0.333	0.396	0.374	0.335
$\lvert x_{\mathrm{JA}}\rvert$	0.222	0.221	0.090	0.111	0.146	0.211	0.340	0.362	0.415	0.384
x^2_{JA}	0.110	0.111	0.031	0.086	0.119	0.175	0.275	0.283	0.331	0.355
UK1	0.978	0.979	0.603	0.162	0.040	0.022	0.022	0.009	0.176	0.107
UK2	0.717	0.739	0.620	0.204	0.126	0.050	0.035	0.015	0.170	0.101
UK3	0.624	0.618	0.596	0.143	0.113	0.053	0.034	0.013	0.169	0.115
UK4	0.231	0.240	0.451	0.501	0.330	0.142	0.070	0.029	0.215	0.159
UK5	0.037	0.046	0.067	0.434	0.589	0.368	0.184	0.082	0.246	0.201
UK6	0.130	0.138	0.141	0.199	0.402	0.534	0.388	0.200	0.322	0.268
UK7	0.073	0.077	0.079	0.085	0.177	0.336	0.489	0.401	0.433	0.372
UK8	0.028	0.030	0.032	0.034	0.076	0.168	0.387	0.503	0.489	0.419
$\lvert x_{\mathrm{UK}}\rvert$	0.205	0.206	0.212	0.207	0.247	0.286	0.413	0.475	0.604	0.557
x^2_{UK}	0.109	0.108	0.130	0.168	0.210	0.247	0.351	0.401	0.543	0.606

Notes: This table shows correlations from a frequency decomposition of the combined univariate multifractal exchange rate models with eight components. For each series, the smoothed probabilities of different volatility states are calculated. These are used to construct eight sub-series giving the smoothed probability that a given multiplier is in the high volatility state. These series are labeled from one to eight, with the first corresponding to the lowest frequency. The table then shows correlations of these series across currencies.

TABLE 5. – Evaluation of Particle filter

		$\mathrm{E}\left(x^2_{t+n}\|x_t,...,x_1\right)$				Conditional kurtosis			
	$\ln L$	$n=1$	5	20	50	$n=1$	5	20	50
True value	-5393.7	0.304	0.317	0.337	0.347	5.105	5.481	5.892	6.225
Simulation average	-5406.2	0.305	0.318	0.341	0.350	5.069	5.436	5.838	6.113
Std. Deviation	6.6279	0.023	0.028	0.037	0.046	0.205	0.233	0.288	0.318
1% quantile	-5423.7	0.260	0.263	0.271	0.266	4.683	4.979	5.266	5.482
25% quantile	-5410.2	0.289	0.299	0.313	0.313	4.927	5.279	5.622	5.891
50% quantile	-5405.7	0.303	0.316	0.337	0.345	5.037	5.416	5.810	6.095
75% quantile	-5401.9	0.320	0.336	0.367	0.385	5.173	5.570	6.015	6.333
99% quantile	-5396.2	0.345	0.367	0.407	0.430	5.465	5.848	6.362	6.672

Notes: This table compares values generated by the particle filter with their true values generated by exact Bayesian updating. $\ln L$ is the value of the log-likelihood function for the Deutsche Mark series with $k = 8$ evaluated at the maximum likelihood estimates in Table 1. E $x^2_{t+n}|x_t,...,x_1$ is the forecasted variance of the series, and conditional kurtosis is E $x^4_{t+n}|x_t,...,x_1$ /(E $x^2_{t+n}|x_t,...,x_1$)^2. For each quantity, the table provides the true value along with the average, standard deviation, and quantiles over 1000 simulations. In each simulation the data set is the same, but the random draws used to calculate the particle filter are independent across simulations. The particle filter is calculated from $B = 1000$ random draws.

TABLE 6. – Bivariate MLE

	$\bar{k}=1$	2	3	4	5	6	7	8
				DM and JA				
$\hat{m}_0^{DM}$	1.638	1.581	1.538	1.482	1.459	0.000	0.000	0.000
	(0.011)	(0.014)	(0.002)	(0.014)	(0.016)	(0.000)	(0.000)	(0.000)
$\hat{m}_0^{JA}$	1.727	1.694	1.661	1.605	1.578	0.000	0.000	0.000
	(0.011)	(0.010)	(0.000)	(0.014)	(0.012)	(0.000)	(0.000)	(0.000)
$\hat{\sigma}_{DM}$	0.666	0.615	0.566	0.559	0.609	0.000	0.000	0.000
	(0.010)	(0.012)	(0.020)	(0.023)	(0.017)	(0.000)	(0.000)	(0.000)
$\hat{\sigma}_{JA}$	0.694	0.662	0.588	0.596	0.678	0.000	0.000	0.000
	(0.012)	(0.015)	(0.015)	(0.035)	(0.017)	(0.000)	(0.000)	(0.000)
$\hat{\gamma}_{\bar{k}}$	0.125	0.202	0.433	0.703	0.746	0.000	0.000	0.000
	(0.013)	(0.019)	(0.028)	(0.061)	(0.044)	(0.000)	(0.000)	(0.000)
$\hat{b}$	-	12.22	13.93	10.39	8.49	0.00	0.00	0.00
		(1.18)	(2.95)	(1.63)	(0.05)	(0.00)	(0.00)	(0.00)
$\hat{\rho}_{\varepsilon}$	0.639	0.646	0.641	0.645	0.647	0.000	0.000	0.000
	(0.008)	(0.008)	(0.024)	(0.021)	(0.009)	(0.000)	(0.000)	(0.000)
$\hat{\rho}_m$	0.472	0.506	0.575	0.628	0.629	0.000	0.000	0.000
	(0.052)	(0.049)	(0.083)	(0.026)	(0.034)	(0.000)	(0.000)	(0.000)
$\ln L$	-9562.64	-9140.91	-8996.07	-8920.86	-8892.74	0.00	0.00	0.00
				DM and UK				
$\hat{m}_0^{DM}$	1.675	1.581	1.552	1.499	1.499	0.000	0.000	0.000
	(0.012)	(0.012)	(0.015)	(0.006)	(0.013)	(0.000)	(0.000)	(0.000)
$\hat{m}_0^{UK}$	1.754	1.676	1.647	1.594	1.595	0.000	0.000	0.000
	(0.010)	(0.011)	(0.011)	(0.034)	(0.011)	(0.000)	(0.000)	(0.000)
$\hat{\sigma}_{DM}$	0.665	0.684	0.591	0.562	0.459	0.000	0.000	0.000
	(0.010)	(0.015)	(0.015)	(0.007)	(0.010)	(0.000)	(0.000)	(0.000)
$\hat{\sigma}_{UK}$	0.647	0.672	0.576	0.543	0.853	0.000	0.000	0.000
	(0.010)	(0.016)	(0.017)	(0.016)	(0.018)	(0.000)	(0.000)	(0.000)
$\hat{\gamma}_{\bar{k}}$	0.249	0.396	0.595	0.792	0.799	0.000	0.000	0.000
	(0.022)	(0.043)	(0.072)	(0.099)	(0.011)	(0.000)	(0.000)	(0.000)
$\hat{b}$	-	11.63	12.61	9.65	10.04	0.00	0.00	0.00
		(2.05)	(2.09)	(1.24)	(0.05)	(0.00)	(0.00)	(0.00)
$\hat{\rho}_{\varepsilon}$	0.725	0.731	0.725	0.727	0.728	0.000	0.000	0.000
	(0.007)	(0.007)	(0.007)	(0.007)	(0.013)	(0.000)	(0.000)	(0.000)
$\hat{\rho}_m$	0.787	0.819	0.846	0.848	0.844	0.000	0.000	0.000
	(0.032)	(0.030)	(0.028)	(0.045)	(0.032)	(0.000)	(0.000)	(0.000)
$\ln L$	-8723.58	-8392.52	-8246.48	-8187.11	-8191.33	0.00	0.00	0.00

Notes: This table shows maximum likelihood estimation results for the bivariate multifractal model. Columns correspond to the number of frequencies $\bar{k}$ in the estimated model. Asymptotic standard errors are in parenthesis.

TABLE 7. – Univariate Forecast Results, DM

	$\bar{k}=1$	2	3	4	5	6	7	8
				Mincer-Zarnowitz Alpha				
1	0.13	0.10	0.03	0.11	0.10	0.11	0.08	0.07
	(0.11)	(0.09)	(0.10)	(0.07)	(0.07)	(0.07)	(0.08)	(0.08)
5	0.50	0.37	0.03	0.47	0.36	0.45	0.29	0.26
	(0.52)	(0.44)	(0.46)	(0.31)	(0.32)	(0.31)	(0.34)	(0.34)
10	0.57	0.31	-0.21	0.76	0.55	0.75	0.48	0.44
	(1.16)	(1.04)	(1.02)	(0.70)	(0.72)	(0.68)	(0.74)	(0.76)
20	0.58	0.39	-0.55	1.47	1.23	1.71	1.31	1.23
	(2.94)	(2.70)	(2.56)	(1.71)	(1.75)	(1.66)	(1.77)	(1.80)
50	-5.96	5.21	0.45	4.49	4.38	5.64	5.23	5.10
	(12.67)	(9.07)	(9.13)	(5.94)	(5.85)	(5.49)	(5.68)	(5.81)
				Mincer-Zarnowitz Beta				
1	0.69	0.70	0.88	0.69	0.73	0.70	0.77	0.79
	(0.22)	(0.17)	(0.19)	(0.13)	(0.13)	(0.13)	(0.14)	(0.15)
5	0.76	0.75	0.93	0.74	0.78	0.74	0.81	0.83
	(0.21)	(0.16)	(0.18)	(0.12)	(0.12)	(0.11)	(0.13)	(0.13)
10	0.86	0.84	0.99	0.77	0.81	0.77	0.83	0.85
	(0.25)	(0.20)	(0.20)	(0.13)	(0.14)	(0.13)	(0.14)	(0.15)
20	0.95	0.87	1.01	0.78	0.80	0.75	0.80	0.82
	(0.32)	(0.26)	(0.26)	(0.17)	(0.17)	(0.16)	(0.17)	(0.18)
50	1.34	0.71	0.95	0.77	0.76	0.70	0.73	0.75
	(0.59)	(0.37)	(0.39)	(0.24)	(0.24)	(0.22)	(0.23)	(0.24)
				Forecast MSE				
1	0.73	0.73	0.72	0.72	0.72	0.72	0.71	0.71
5	5.09	5.00	4.79	4.77	4.71	4.76	4.68	4.67
10	13.03	12.84	12.20	12.06	11.90	12.03	11.86	11.83
20	37.43	37.43	35.17	34.43	34.19	34.84	34.30	34.15
50	151.43	158.25	146.38	140.56	140.09	143.80	141.82	140.91
				Restricted R^2				
1	0.019	0.028	0.041	0.039	0.044	0.042	0.047	0.048
5	0.069	0.086	0.125	0.129	0.139	0.131	0.144	0.146
10	0.112	0.124	0.168	0.178	0.188	0.179	0.191	0.193
20	0.114	0.114	0.167	0.185	0.191	0.175	0.188	0.192
50	0.090	0.049	0.121	0.156	0.159	0.136	0.148	0.154

Notes: The first two panels of this table show show coefficients from the Mincer-Zarnowitz regressions $x^2_{t+h} = \alpha + \beta E_t[x^2_{t+h}] + u_t$. The leftmost column corresponds to the horizon h of the forecast. Asymptotic standard errors in parenthesis are corrected for heteroskedasticity and autocorrelation using the method of Newey and West (1987,1994) and for parameter uncertainty using the method of West and McCracken (1998). The third panel gives the mean square error of the forecast, and the final panel reports the forecast R^2 which is is one less the MSE divided by the sum of squared demeaned squared returns in the out of sample period.

TABLE 8a. – Bivariate Forecast Results, DM with JA

	$\bar{k}=1$	2	3	4	5	6	7	8
				Mincer-Zarnowitz Alpha				
1	0.08	0.08	0.08	0.05	0.05	0.00	0.00	0.00
	(0.12)	(0.10)	(0.09)	(0.08)	(0.08)	(0.00)	(0.00)	(0.00)
5	0.15	0.19	0.17	0.12	0.12	0.00	0.00	0.00
	(0.60)	(0.49)	(0.44)	(0.38)	(0.37)	(0.00)	(0.00)	(0.00)
10	-0.66	-0.07	-0.04	0.11	0.13	0.00	0.00	0.00
	(1.47)	(1.17)	(1.00)	(0.84)	(0.82)	(0.00)	(0.00)	(0.00)
20	-3.85	-0.04	-0.19	0.42	0.49	0.00	0.00	0.00
	(4.33)	(2.87)	(2.44)	(1.98)	(1.96)	(0.00)	(0.00)	(0.00)
50	-29.98	2.75	-1.44	2.22	3.03	0.00	0.00	0.00
	(23.10)	(9.05)	(8.52)	(6.43)	(6.40)	(0.00)	(0.00)	(0.00)
				Mincer-Zarnowitz Beta				
1	0.83	0.77	0.81	0.84	0.85	0.00	0.00	0.00
	(0.25)	(0.19)	(0.18)	(0.16)	(0.16)	(0.00)	(0.00)	(0.00)
5	0.97	0.87	0.92	0.89	0.91	0.00	0.00	0.00
	(0.27)	(0.20)	(0.18)	(0.15)	(0.15)	(0.00)	(0.00)	(0.00)
10	1.21	0.97	1.02	0.92	0.94	0.00	0.00	0.00
	(0.34)	(0.24)	(0.22)	(0.16)	(0.16)	(0.00)	(0.00)	(0.00)
20	1.53	0.98	1.06	0.90	0.92	0.00	0.00	0.00
	(0.51)	(0.30)	(0.27)	(0.20)	(0.20)	(0.00)	(0.00)	(0.00)
50	2.56	0.89	1.16	0.86	0.86	0.00	0.00	0.00
	(1.13)	(0.40)	(0.41)	(0.27)	(0.27)	(0.00)	(0.00)	(0.00)
				Forecast MSE				
1	0.73	0.72	0.72	0.71	0.71	0.00	0.00	0.00
5	5.04	4.92	4.80	4.67	4.64	0.00	0.00	0.00
10	13.05	12.64	12.20	11.83	11.74	0.00	0.00	0.00
20	38.22	36.52	35.07	33.79	33.59	0.00	0.00	0.00
50	158.99	149.44	144.83	136.30	137.31	0.00	0.00	0.00
				Restricted R^2				
1	0.025	0.033	0.039	0.048	0.050	0.000	0.000	0.000
5	0.079	0.100	0.123	0.147	0.152	0.000	0.000	0.000
10	0.110	0.138	0.168	0.194	0.200	0.000	0.000	0.000
20	0.095	0.136	0.170	0.200	0.205	0.000	0.000	0.000
50	0.045	0.102	0.130	0.181	0.175	0.000	0.000	0.000

Notes: The first two panels of this table show show coefficients from the Mincer-Zarnowitz regressions $x^2_{t+h} = \alpha + \beta E_t[x^2_{t+h}] + u_t$. The leftmost column corresponds to the horizon h of the forecast. Asymptotic standard errors in parenthesis are corrected for heteroskedasticity and autocorrelation using the method of Newey and West (1987,1994) and for parameter uncertainty using the method of West and McCracken (1998). The third panel gives the mean square error of the forecast, and the final panel reports the forecast R^2 which is is one less the MSE divided by the sum of squared demeaned squared returns in the out of sample period.

TABLE 8b. – Bivariate Forecast Results, DM with UK

	$\bar{k}=1$	2	3	4	5	6	7	8
				Mincer-Zarnowitz Alpha				
1	0.09	0.07	0.10	0.06	0.06	0.00	0.00	0.00
	(0.14)	(0.10)	(0.08)	(0.08)	(0.08)	(0.00)	(0.00)	(0.00)
5	-0.06	0.06	0.29	0.13	0.14	0.00	0.00	0.00
	(0.81)	(0.50)	(0.37)	(0.39)	(0.39)	(0.00)	(0.00)	(0.00)
10	-2.17	-0.38	0.27	0.03	0.08	0.00	0.00	0.00
	(2.32)	(1.16)	(0.84)	(0.88)	(0.87)	(0.00)	(0.00)	(0.00)
20	-11.05	-1.11	0.42	0.09	0.26	0.00	0.00	0.00
	(8.08)	(2.86)	(2.08)	(2.12)	(2.09)	(0.00)	(0.00)	(0.00)
50	-69.76	-6.25	1.18	1.00	1.74	0.00	0.00	0.00
	(47.33)	(10.74)	(7.37)	(7.09)	(6.89)	(0.00)	(0.00)	(0.00)
				Mincer-Zarnowitz Beta				
1	0.84	0.81	0.72	0.81	0.81	0.00	0.00	0.00
	(0.31)	(0.19)	(0.15)	(0.16)	(0.16)	(0.00)	(0.00)	(0.00)
5	1.13	0.95	0.81	0.89	0.88	0.00	0.00	0.00
	(0.39)	(0.21)	(0.14)	(0.15)	(0.15)	(0.00)	(0.00)	(0.00)
10	1.67	1.07	0.88	0.94	0.92	0.00	0.00	0.00
	(0.58)	(0.25)	(0.17)	(0.17)	(0.17)	(0.00)	(0.00)	(0.00)
20	2.57	1.14	0.91	0.95	0.93	0.00	0.00	0.00
	(1.04)	(0.31)	(0.21)	(0.22)	(0.21)	(0.00)	(0.00)	(0.00)
50	4.82	1.35	0.93	0.94	0.91	0.00	0.00	0.00
	(2.47)	(0.50)	(0.32)	(0.30)	(0.29)	(0.00)	(0.00)	(0.00)
				Forecast MSE				
1	0.73	0.72	0.72	0.72	0.72	0.00	0.00	0.00
5	5.20	4.87	4.76	4.71	4.71	0.00	0.00	0.00
10	13.86	12.43	12.00	11.92	11.93	0.00	0.00	0.00
20	41.24	35.71	34.30	33.96	34.03	0.00	0.00	0.00
50	171.97	145.69	139.34	136.90	137.28	0.00	0.00	0.00
				Restricted R^2				
1	0.018	0.035	0.037	0.044	0.044	0.000	0.000	0.000
5	0.049	0.110	0.130	0.139	0.139	0.000	0.000	0.000
10	0.055	0.152	0.181	0.187	0.187	0.000	0.000	0.000
20	0.024	0.155	0.188	0.196	0.195	0.000	0.000	0.000
50	-0.033	0.125	0.163	0.178	0.175	0.000	0.000	0.000

Notes: The first two panels of this table show show coefficients from the Mincer-Zarnowitz regressions $x^2_{t+h} = \alpha + \beta E_t[x^2_{t+h}] + u_t$. The leftmost column corresponds to the horizon h of the forecast. Asymptotic standard errors in parenthesis are corrected for heteroskedasticity and autocorrelation using the method of Newey and West (1987,1994) and for parameter uncertainty using the method of West and McCracken (1998). The third panel gives the mean square error of the forecast, and the final panel reports the forecast R^2 which is is one less the MSE divided by the sum of squared demeaned squared returns in the out of sample period.