# Higher Dimensional Multifractal Processes: A GMM Approach

Ruipeng Liu[†] and Thomas Lux[‡]

[‡]Department of Economics, University of Kiel, Olshausenstraße 40, 24118 Kiel, Germany, and
Institute for the World Economy, Hindenburgufer 66, 24105 Kiel, Germany
[†]School of Accounting, Economics and Finance, Deakin University, Melbourne, VIC 3125, Australia

**Abstract**

Multifractal processes have recently been introduced as a new tool for modeling the stylized facts in financial time series. In this paper, we extend it to a multivariate multifractal model with parsimonious settings. Since there are restrictions when applying likelihood approaches due to the extreme large state spaces, we implement its estimation via Generalized Methods of Moment (GMM). Our Monte Carlo studies demonstrate convincing performances of the GMM estimator; we also present its empirical applications in terms of volatility forecasting and portfolio management.

Keyword: Multivariate, Multifractal, Long memory, GMM estimation, Volatility forecast, Portfolio choice.

JEL Classification: C20, G15

## 1 Introduction

Despite the rich literature that exists on volatility modeling, multifractal (MF) processes have been recently introduced as an alterative formalisation, which conceives volatility as a hierarchical, multiplicative process with heterogeneous components. The essential new feature of MF models is their ability of generating different degrees of long-term dependence in various powers of returns - a feature pervasively found in empirical financial data, cf. Lo (1991), Ding et al (1993), Beran (1994), Lobato and Savin (1998), Zumbach (2004), Liu et. al (2007), Lux and Morales-Arias (2010), etc.[1] Research on multifractal models

[1]Though there are considerable numbers of long memory volatility models, such as ARFIMA (Fractional Integrated Autoregressive Moving Average), and FIGARCH (Fractional

originated from statistical physics (Mandelbrot, 1974), unfortunately, the models used in physics are of a combinatorial nature and suffer from non-stationarity due to the limitation to a bounded interval and the non-convergence of moments in the continuous-time limit. This major weakness was overcome by introducing iterative versions of multifractal processes, cf. Calvet and Fisher (2001, 2004, 2006), Lux (2006, 2007).

So far, available multifractal models are mostly univariate ones. However, for many important questions in empirical research, multivariate settings are preferable, cf. Bollerslev (1990), Liesenfeld and Richard (2003). For instance, it is now well accepted that financial volatilities move together over time across assets and markets. This is particularly important when considering asset allocation, risk measurement and management and portfolio strategies; since the information on the source of long memory in the volatility process is quite silent, the multivariate setting may provide additional insight into the factors responsible for the long term dependence.

The rest of this paper is organized as follows: Section 2 introduces a parsimonious multivariate multifractal model after a brief review of the univariate ones. Section 3 implements its estimation via GMM approach, Monte Carlo studies have been conducted to whose assess its applicability. Empirical applications on volatility forecasting and portfolio management are presented in Section 4. Concluding remarks are provided in Section 5. The appendix provides details of the analytical moments calculation.

# 2 Multivariate multifractal models

Mandelbrot et al. (1997) first introduced the multifractal model of asset returns (MMAR), translating the approach of energy flux cascade from the statistical physics, where "cascades" are typically modeled by multiplicative operations on probability measures, cf. Mandelbrot (1974) and Harte (2001). However, the practical applicability of MMAR suffers from its combinatorial nature, i.e. the non-causal nature of the time transformation and from its non-stationarity due to the inherent restriction to a bounded interval.

## 2.1 Review of univariate multifractal models

These limitations have been overcome by the introduction of iterative versions of the MF processes, one of the most prominent development is the Markov-switching multifractal model (MSM), cf. Calvet and Fisher (2001, 2004) and Lux (2008). In their approach, returns are modeled as:

$$r_t = \sigma \left( \prod_{i=1}^{k} M_t^{(i)} \right)^{1/2} \cdot u_t \tag{1}$$

Integrated General Autoregressive Conditional Heteroscedasticity) models, most of them are only limited on the second moment.

with $u_t$ drawn from a standard Normal distribution $N(0,1)$ and instantaneous volatility being determined by the product of $k$ volatility components or multipliers $M_t^{(1)}$, $M_t^{(2)}$ ..., $M_t^{(k)}$, and a constant scale parameter $\sigma$. Volatility components are renewed at time $t$ with probability $\gamma_i$ depending on its rank '$i$' within the hierarchy of multipliers or remains unchanged with probability $1-\gamma_i$. The transition probabilities are specified as:

$$\gamma_i = 1 - (1-\gamma_1)^{(b^{i-1})}, \qquad \text{for } i = 1, 2, \dots k, \tag{2}$$

with parameters $\gamma_1 \in [0,1]$ and $b \in (1, \infty)$.

This iterative version of the multifractal model preserves the hierarchical structure of MMAR while dispensing with its restriction to a bounded interval. While this model is asymptotically "well-behaved" (i.e. it shares all the convenient properties of Markov-switching processes), it is still capable of capturing some important properties of financial time series, namely, volatility clustering and the power-law behaviour of the autocovariance function of absolute moments:

$$Cov(|r_t|^q, |r_{t+\tau}|^q) \propto \tau^{2d(q)-1}. \tag{3}$$

Note, however, that the power-law behavior of the MSM model holds only approximately in a preasymptotic range. Rather than displaying asymptotic power-law behavior of autocovariance functions in the limit $t \to \infty$ or divergence of the spectral density at zero, the Markov-switching MF model is rather characterized by only 'apparent' long memory with an approximately hyperbolic decline of the autocorrelation of absolute powers over a finite horizon and exponential decline thereafter. In particular, approximately hyperbolic decline as expressed in eq. (3) holds only over an interval $1 \ll \tau \ll b^k$ with $b$ the parameter of the transition probabilities of eq. (2) and $k$ the number of hierarchical cascade levels.

## 2.2 Higher dimensional multifractal models

Let us consider an $N$-dimensional asset returns process evolving in discrete time over the interval $[0,T]$ with equally spaced discrete time points $t = 1, \dots, T$, and $r_t = (r_1, \dots, r_N)'$:

$$r_t = \sigma. * [g(M_t)]^{1/2}. * u_t, \tag{4}$$

$\sigma$, $u_t$ are $N \times 1$ vectors, $.*$ being element by element multiplication, $u_t$ follows the multivariate standard Normal distribution which has the correlation matrix being composed of unknown correlation parameters $\rho$:

$$\begin{bmatrix} 1 & \rho_{12} & \rho_{13} & \cdots \\ \rho_{12} & 1 & \rho_{23} & \cdots \\ \rho_{13} & \rho_{23} & 1 & \cdots \\ \vdots & \vdots & & \ddots \end{bmatrix}$$

$\sigma$ is a vector of constant scale of parameters and can be viewed as unconditional standard deviation. $g(M_t)$ is a $N \times 1$ vector of the products of multifractal volatility components, i.e., $g(M_t) = [g(M_{1,t}), \ldots, g(M_{N,t})]'$:

$$g(M_q, t) = \prod_{i=1}^{j} M_{q,t}^{(i)} .* \prod_{l=j+1}^{k} M_{q,t}^{(l)}, \tag{5}$$

$$M_{1,t}^{(i)} = M_{2,t}^{(i)} = \ldots M_{N,t}^{(i)}, \qquad \text{for } 1 < i \leq j,$$

$q = 1, \ldots, N$. Eq. (5) means that each element of $g(M_t)$ is the instantaneous volatility of univariate multifractal processes; in addition, $N$ time series share a number of $j$ joint cascades that govern the strength of their volatility correlation. Consequently, the larger $j$, the higher the correlation between them. After $j$ joint multipliers, each series has additional independent multifractal components. Instead of introducing additional correlation parameters for the specification of new arrivals at hierarchy level $i$ among different time series, our assumption of joint cascade level simplifies the characterization of new arrivers of volatility components across different assets. This simplification has been demonstrated to be well performed in the bivariate case according to both its simulation and empirical applications, cf. Liu and Lux (2010).

Furthermore, to constrain the space of parameters further, a restriction for the specification of the transition probabilities is imposed:

$$\gamma_i \quad = \quad 2^{-(k-i)}, \qquad\qquad for \quad i = 1, 2 \ldots, j, \ldots, k. \tag{6}$$

Each volatility component is renewed at time $t$ with probability of $\gamma_i$ depending on its rank within the hierarchy of multipliers and remains unchanged with probability of $1-\gamma_i$. Lux (2008) Liu and Lux (2010) report that transition probabilities of form Eq. (6) possesses sufficient flexibility in the remaining parameters so that its empirical performance is relatively little hampered by fixing these parameters.

We specify volatility components for all assets to be random draws from either a binomial distribution or Lognormal distribution. For the binomial case in which we assume two draws, i.e., $m_0 \in (0, 2)$ and alternative $m_1 = 2 - m_0$; for the latter, we assume $logM \sim N(-\lambda, \sigma_m^2)$, and normalize it by assigning constraint $E[M_t^{(i)}] = 1$ for the sake of explosion.

# 3 GMM Estimation

Although the multifractal model is a rather new tool in volatility modelling, various approaches have already been explored to estimate its parameters. The parameters of the combinatorial MMAR have been estimated via an adaptation of the scaling estimator and Legendre transformation approach from statistical physics although this approach has been shown to yield unreliable results, cf. Lux (2004). A broad range of more rigorous estimation methods have been developed for iterative MF processes. Calvet and Fisher (2001, 2004) propose maximum likelihood (ML), whose applicability is, however, confined to the case of discretely distributed multipliers, in addition, it imposes an upper bound computational limit of the cascade level, i.e., it is not feasible to implement when $k > 5$ (bivariate model) using personal computers. Calvet, et. al (2006) have introduced a simulation based ML to estimate the parameters of a bivariate extension of the MSM model. Lux (2008) proposes a Generalized Method of Moments approach, which can be applied not only to discrete but also to continuous distributions of the volatility components.

In this paper, we adopt the GMM (Generalized Method of Moments) approach by Hansen (1982) with analytical solutions of a set of appropriate moment conditions. In the GMM approach, the vector of parameter estimates of the model, say $\beta$, can be obtained as:

$$\widehat{\beta} = \arg\min_{\beta \in \Theta} \bar{M}(\beta)' W \bar{M}(\beta) \tag{7}$$

$\Theta$ is the parameter space, $\bar{M}(\beta)$ is the vector of differences between sample moments and analytical moments, and $W$ a positive definite weighting matrix, which controls the over-identification when applying GMM. Implementing (7), one typically starts with the identity matrix, then the inverse of the covariance matrix obtained from the first round estimation is used as the weighting matrix in the next step, and the procedure will continue until the estimates and weighting matrices converge. Under suitable conditions, $\hat{\beta}$ is consistent and asymptotically converges to $T^{1/2}(\hat{\beta} - \beta_0) \sim N(0, \Xi)$ with covariance matrix $\Xi$. As is well-known, $\hat{\beta}_T$ is consistent and asymptotically Normal if suitable 'regularity conditions' are fulfilled (sets of which are detailed, for example, in Harris and Mátyás (1999)). $\hat{\beta}_T$ then converges to

$$T^{1/2}(\hat{\beta}_T - \beta_0) \sim N(0, \Xi), \tag{8}$$

with covariance matrix $\Xi = (\bar{F}_T' \bar{V}_T^{-1} \bar{F}_T)^{-1}$ in which $\beta_0$ is the true parameter vector, $\hat{V}_T = T var \bar{M}_T(\beta)$ is the covariance matrix of the moment conditions, $\hat{F}_T(\beta) = \frac{\partial \bar{M}_T(\beta)}{\partial \beta}$ is the matrix of first derivatives of the moment conditions, and $\bar{V}_T$ and $\bar{F}_T$ are the constant limiting matrices to which $\hat{V}_T$ and $\hat{F}_T$ converge.

The applicability of GMM for multifractal models has been discussed by Lux (2008). In order to account for the proximity to long memory characterizing multifractal models, it is recommended to use logarithmic differences of absolute returns together with the pertinent analytical moment conditions, i.e. to transform the observed data $r_t$ into $\tau$th differences of the log observations:

$$\begin{aligned} X_{t,\tau} &= \ln|r_t| - \ln|r_{t-\tau}| \\ &= \left(\sigma_1 + 0.5\sum_{i=1}^{k}\varepsilon_t^{(i)} + 0.5\sum_{h=k+1}^{n}\varepsilon_t^{(h)} + ln|u_t|\right) - \left(\sigma_1 + 0.5\sum_{i=1}^{k}\varepsilon_{t-\tau}^{(i)} + \right. \\ &\quad \left. 0.5\sum_{h=k+1}^{n}\varepsilon_{t-\tau}^{(h)} + ln|u_{t-\tau}|\right) \\ &= 0.5\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-\tau}^{(i)}) + 0.5\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-\tau}^{(h)}) + (ln|u_t| - ln|u_{t-\tau}|) \end{aligned} \tag{9}$$

with $\varepsilon_t^{(i)} = \ln\left(M_t^{(i)}\right)$. The variable resulted in Eq. (9) has nonzero autocovariance over a limited number of time lags. In order to exploit the temporal scaling properties of multifractal processes, we select moment conditions for the covariances of different orders over various time lags $\tau$. More precisely, the moment conditions that we consider include two categories: the first set of conditions is obtained by considering some order of log-squared observations, and the second set of moment conditions is derived from the absolute observations. In particular, we select moment conditions for the powers of $X_{t,\tau}$ i.e. moments of the raw observations and square observations:

$$Cov[X_{t+\tau,\tau}, X_{t,\tau}]; \qquad Cov[X^2_{t+\tau,\tau}, X^2_{t,\tau}];$$

$$Cov[X_{t+\tau,\tau}, X^-_{t,\tau}]; \qquad Cov[X^2_{t+\tau,\tau}, (X^-_{t,\tau})^2];$$

$X^-_{t,\tau}$ stands for the time series other than $X_{t,\tau}$, We recognize the transformation in Eq. (9) excluding the scale parameters $\sigma$, while estimating the scale parameters can be pursued by adding additional moment conditions, i.e., the second moment of empirical data by considering each observation's contribution to the standard deviation of the sample returns.

Similar to Andersen and Sorensen (1996), we proceed by conducting Monte Carlo experiments to explore the performance of the GMM estimation regarding to the bivariate cases. Moment conditions for the bivariate binomial and Lognormal models can be found in the Appendix. We start with the bivariate binomial model with number of cascade level $k = 12$, we fixed correlation parameter $\rho = 0.5$; scale parameter (unconditional variance) $\sigma_1 = \sigma_2 = 1$, and choose multipliers from $m_0 = 1.2$ to 1.5 by 0.1 increment with sample sizes $N_1 = 2000$, $N_2 = 5000$, and $N_3 = 10000$. Table 1 shows the statistical result of our GMM estimator for the case of joint multipliers $j = 6$: for the binomial distribution parameter $\hat{m_0}$, not only the bias but also the finite sample standard deviation and root mean squared error show quite encouraging behavior, even in the smaller sample sizes $N = 2000$ and $N = 5000$, the average bias of the Monte Carlo estimates is moderate throughout and practically zero for the larger sample sizes $N = 10000$. It is also interesting to note that our estimates are in harmony with $T^{\frac{1}{2}}$ consistency. All these results can be viewed as a positive signal of the log transformation in practice.

Then, we turn to the MF model with volatility components being continuous distributed, i.e., $-logM \sim N(\lambda, \sigma_m^2)$. Unlike the binomial model, multifractal processes with continuous distribution of volatility components imply an infinite dimension of the transitional matrix, and the exact form of likelihood function can not be identified explicitly. Therefore, the maximum likelihood approach is not applicable to the Lognormal case.[2] GMM provides a solution for estimating multifractal processes with continuous state spaces. Moment conditions for the Lognormal model are given in Appendix B. Note that the admissible parameter space for the location parameter $\lambda$ is $\lambda \in [0, 1)$ where in the borderline case $\lambda = 0$ the volatility process collapses to a constant (as $m_0 = 1$ in binomial model).

In our Monte Carlo studies of the bivariate Lognormal model reported in Table 2, we cover parameter values of $\lambda = 0.10$ to $\lambda = 0.40$ with the increment of 0.1, and use same numbers of joint multiplier cascade levels and the sample sizes as in the binomial case.[3] As can be seen, results are not too different from those obtained with the binomial model: biases are moderate and close to zero again; SD and RMSE are moderate and decreasing with increasing in the sub-sample sizes from 2000 to 10000, somewhat in contrast to the binomial case, we notice a some slight deterioration of efficiency with smaller sample size when increasing $\lambda$, which might be due to increasing $\lambda$ leading to increasing $\sigma_m^2$ by their dependence; by recalling that we the normalization with $E[M_t^{(i)}] = 1$, it implies $exp(-\lambda + 0.5\sigma_m^2) = 1$ and leads to $\sigma_m^2 = 2\lambda$.

When studying higher dimensional multifractal models, for instance trivariate case (when modeling 3 assets), maximum likelihood approach only applies when the number of cascade levels $k < 3$ (binomial model), which certainly does not meet the empirical demands. In contrast, GMM provides a more convenient way to implement estimation of much higher dimensional MF processes because it allows us to treat each pair of time series as a bivariate case, and select the moment conditions of each bivairate time series. We have also conducted Monte carlo studies for tri-variate ML processes. Analogously to the bivariate models, we select moments conditions for a collections of 3 pairs of bivariate data. Table 3 provides the performance our GMM estimator of the trivariate binomial multifractal model with $k = 12$, $j = 4$, and initial parameters used for simulations are $m_0 = 1.3$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\sigma_3 = 1$, $\rho_{12} = 0.3$, $\rho_{23} = 0.5$, $\rho_{13} = 0.7$. Table 4 report the Monte Carlo studies for trivariate Lognormal model with the same design initial parameters (except with $\lambda = 0.2$) as in the Table 3 of the binomial model. We observe again very positive performance similar to the Monte Carlo studies for the bivariate model. [4]

[2]Theoretically, simulation based maximum likelihood could be applicable, however there are very few applications in the existing literature so far.

[3]We have also repeated experiments with respect to different joint multipliers $j$, which provides us similar results for both binomial and Lognormal cases.

[4]We have also studied with various combinations of parameters, and we obtained similar

All in all, the performance from both the binomial and Lognormal Monte Carlo simulation and estimation shows that GMM approach workes quite well for higher dimensional multifractal processes both in the discrete and in the continuous state space.

# 4 Empirical applications

In this section, we present the performance of our new parsimonious multivariate MSM model by reporting the empirical results of volatility forecasts and portfolio management.

We consider daily data for two stock exchange indices: the Dow Jones composite 65 average index and the $NIKKEI$ 225 average index ($DOW/NIK$, $6^{th}$ January 1970 - $30^{th}$ December 2010), two foreign exchange rates, the U.S. Dollar to British Pound, and German Mark to British Pound ($US/DM$, $1^{st}$ March 1973 - $31^{st}$ December 2010); and a bond portfolio of U.S. 1-year and 2-year treasury constant maturity bond rates ($TB1/TB2$, $1^{st}$ June 1976 - $31^{st}$ December 2010), where the first symbol inside the parentheses gives the acronym for the corresponding time series, followed by the starting and ending dates for the sample at hand. Asset return are calculated as the log differences of prices $r_t = 100 \times (\log(p_t) - \log(p_{t-1}))$, with $p_t$ denoting daily price observations.[5]

We separate each time series into two subsets (in-sample data used for estimation, out-of-sample data for forecast assessment). For the in-sample periods we use for $DOW/NIK$: $6^{th}$ January 1970 - $31^{st}$ August 1992; $US/DM$: $1^{st}$ March 1973 - $30^{th}$ April 1994; and $TB1/TB2$: $1^{st}$ June 1976 - $31^{st}$ May 1996. The remaining out-of-sample subsets are for the $DOW/NIK$: $4^{th}$ September 1992 - $30^{th}$ December 2010; $US/DM$: $1^{st}$ May 1994 - $31^{st}$ December 2010 and $TB1/TB2$: $1^{st}$ June 1996 - $31^{st}$ December 2010.

We use the number of cascades $k = 8$ for $DOW/NIK$, $k = 10$ for $US/DM$ and $k = 8$ for $TB1/TB2$ as Liu (2008). Table 5 reports the in sample empirical GMM estimates for $DOW/NIK$ for different choices of joint cascade levels $j$ ranging from 1 to 7; Table 6 provide the empirical estimates for $US/DM$ with $j$ ranging from 1 to 9; Table 7 provide the empirical estimates for $TB1/TB2$ with $j$ ranging from 1 to 7. To specify an optimal choice of joint cascade levels, we proceed by matching the simulated long memory GPH parameter estimates (Geweke and Porter-Hudak(1983)) and empirical GPH estimates, cf. Liu and Lux (2010). for the detailed heuristic model selection scheme. For the $DOW/NIK$ portfolio the preferred model according to the heuristic model selection scheme detailed in Liu and Lux (2010) is $j = 2$, while it is $j = 3$ for $US/DM$ and $j = 5$ for $TB1/TB2$ respectively.

results, we skip these tables to save spaces.

[5]The U.S. one and two-year treasury constant maturity rates have been converted to equivalent bond prices before calculating returns.

### 4.1 Volatility Forecast

Unlike the volatility forecast implied from maximum likelihood approach for MF models (that is, based on the exact identifying the elements of transition matrix, the conditional probabilities can be used to obtain the multi-step forecasts according to Bayes rule), in this section, we assess the applicability of our multivariate MF by its volatility forecast computed on the base of GMM parameter estimates. We construct the best linear forecast with the auxiliary of the generalized Durbin-Levinson algorithm. As outlined in Brockwell and Davis (1991). Assuming $X_t = \{X_{t1}, \ldots, X_{tm}\}$ being an m-variate stationary time series with mean zero and covariance function given by the $m \times m$ matrix $\Gamma(\cdot)$, the best linear forecasts are obtained

$$\hat{X}_{n+1} = \Theta_{n1} X_n + \cdots \Theta_{nn} X_1 \tag{10}$$

where the $m \times m$ matrix $\Theta_{n1}, \ldots, \Theta_{nn}$ are any solution of

$$\sum_{j=1}^{n} \Theta_{nj} \Gamma(i-j) = \Gamma(i), \qquad i = 1, \ldots, n. \tag{11}$$

The coefficients of $\Theta_{n1}, \ldots, \Theta_{nn}$ can be computed recursively using the multivariate version of generalized Durbin-Levinson algorithm, see Brockwell and Davis (1991 chapter 14) for details.

We first report the performance of volatility forecast by using simulation data, we adopt the traditional criteria of relative mean squared error (RMSE) and relative mean absolute error (RMAE), i.e., mean squared error and mean absolute error divided by the pertinent MSE and MAE of the naive predictor using historical volatility (the sample mean of squared returns over the in-sample period). Before applying empirical data, we first assess its applicability by using simulated ones. We have conducted 400 simulations and estimations with each simulation of 10000 realizations, first 5000 observations are used for estimation, the remaining 5000 observations used for out of sample forecast assessment. The first two columns of Table 8 report the RMSE and RMAE for simulated data from the bivariate MF model with the forecast horizons are $1, 5, 10, 20, 50, 100$ days, and we observe the very successful results that all forecasts based on the multivariate MF model outperforms the ones based on naive predictor, particularly at short-time horizons.

We also present the volatility forecast based on the univariate MF model which is given in Table 9 for comparison reason. For the results from bivariate MF model, the upper two panels of Table 10 report the RMSE and RMAE of the out of sample forecasts regarding to our empirical data at various horizons of $1, 5, 10, 20, 50, 100$ days. We find results based on the bivariate model give better forecasts than ones from the univariate model, indeed, multivariate models provide more information (e.g. correlation parameters etc.) used for voaltility forecasting, Table 10 report the successful out of sample forecasts of equal-weighted portfolio (EW) according their RMSE and RMAE.

## 4.2 Optimal Portfolio Choice

In mean-variance analysis, the maximum expected return strategy leads to a portfolio allocation on the efficient frontier. Consider an investor who has a one-step ahead horizon and constructs a dynamically rebalanced portfolio that maximizes the conditional expected return subject to achieving a target conditional volatility. Computing the time-varying weights of this portfolio requires one-step ahead forecasts of the conditional mean and the conditional variance-covariance matrix. Let $r_{t+1}$ denote the $N \times 1$ vector of risky asset returns, $\mu_{t+1|t} = E_t[r_{t+1}]$ is the conditional expectation of $r_{t+1}$, and $\Sigma_{t+1}|t = E_t[(r_{t+1} - \mu_{t+1|t})(r_{t+1} - \mu_{t+1|t})]$ is the conditional variance-covariance matrix of $r_{t+1}$. At each period $t$, the investor solves the following utility ($U(\cdot)$) maximization problem:

$$\max_{w_t}\{E[U(W_{t+1})] = \mu_{p,t+1} - \frac{c}{2}\sigma^2_{p,t+1}\} \tag{12}$$

with

$$\mu_{p,t+1} = w_t^{'}\mu_{t+1|t} + (1 - w_t^{'})r_f; \qquad \text{and} \sigma^2_{p,t+1} = w_t^{'}\Sigma_{t+1|t}w_t, \tag{13}$$

where $w_t$ is a vector of portfolio weights, $c$ the coefficient of the absolute risk aversion, and $r_f$ the riskless return. Thus, the optimal portfolio weight is given:

$$w_t = \frac{1}{c}\Sigma^{-1}_{t+1|t}(\mu_{p,t+1} - r_f) \tag{14}$$

One commonly used performance measure is the Sharpe ratio $SR$. We compute and compare the *ex post* Sharpe ratios $SR = (\mu_p - r_f)/\sigma_p$. However, it does not take into account time-varying conditional volatility because the sample standard deviation (SD) overestimates the conditional risk when following dynamic strategies. Consequently, the realized Sharpe ratio underestimates the performance of dynamic strategies. In addition, the Sharpe ratio cannot quantify the economic gains of the dynamic strategies over the buy-and-hold strategies. One alternative performance measure that is directly related to the Sharpe ratio, but also quantifies the outperformance is the $M2$ measure developed by Modigliani and Modigliani (1997).[6] The $M2$ measure is the abnormal return that the dynamic strategy would have earned if it had the same risk as the benchmark. It is defined as

$$M_2 = \frac{\sigma_b}{\sigma_p}(\mu_p - r_f)(\mu_b - r_f) \tag{15}$$

$\sigma_b$ and $\mu_b$ are the benchmark portfolio mean and standard deviation respectively, as in the subsection of vlatility forecast, we use the naive predictor of historical volatility (the sample mean of squared returns over the in-sample period). $M_2$ in Eq. (15) is also directly related to the Sharpe ratio

[6]Graham and Harvey (1997) introduced a similar measure called $GH2$, which includes the correlation between the risk-free asset and other assets. As the risk-free rate we use is constant, the two measures $GH2$ and the $M2$ actually are identical in our study.

as $M_2 = \sigma_b(SR_p - SR_b)$, i.e. a measure of the difference of Sharpe ratios of portfolio and benchmark.

The third column of Table 8 reports the $M_2$ for simulated data from the bivariate MF model with the forecast horizons are $1, 5, 10, 20, 50, 100$ days. The lower panel of Table 10 report the out of sample forecasts of $M_2$ regarding to our empirical data at various time horizons. As can be seen, $M_2$ forecasts from the bivariate MF model are very positive (in particular for the equal-weighted portfolios), except with the failure of U.S one-year treasury rate.

# 5 Concluding remarks

In this paper we have developed a parsimonious multivariate multifractal model extending the univariate Markov-switching multifractal model. Since there are limitations of maximum likelihood (ML) approach due to its high dimensional structure with extreme large state spaces, we implement its estimation via alternative Generalized Method Moment (GMM). The moments conditions have been employed through the log transformation of observations. Our Monte Carlo experiments indicate the successful performance of our GMM estimator, i.e., it does not pose computational restrictions on the choice of the number of cascade levels with GMM, compared to a maximum of about 5 cascade levels in ML estimation (bivariate case); GMM also applies the multifractal models with continuous distributed volatility components; Furthermore, empirically speaking, GMM is much faster compared to the very time-consuming ML approaches. In the last part of this paper, we applied the model to volatility forecast and portfolio management with empirical financial time series of stock exchange indices, foreign exchange rates and U.S Bond maturity rates. We demonstrate the applicability of the multivariate multifractal model.

## References


[1] Andersen, T. G. and B. E. Sorensen (1996), GMM Estimation of A Stochastic Volatility Model: A Monte Carlo Study. *Journal of Business and Economic Statistics* **14**, 328-352.

[2] Beran, J. (1994) *Statistics for Long-Memory Processes*, New York: Chapman & Hall.

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

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

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

[6] Calvet, L. and A. Fisher (2002), Multifractality in Asset Returns: Theory and Evidence, *Review of Economics* & *Statistics*, **84**, 381-406.

[7] Calvet, L. and A. Fisher (2004), Regime Switching and the Estimation of Multifractal Processes, *Journal of Financial Econometrics*, **2**, 49-83.

[8] Calvet, L., A. Fisher, and S. Thompson, (2006), Volatility Comovement: A Multifrequency Approach, *Journal of Econometrics*, **131**, 179-215.

[9] Ding, Z., R. Engle, and C. Granger, (1993), A Long Memory Property of Stock Market Returns and a New Model, *Journal of Empirical Finance*, **1**, 83-106.

[10] Geweke, J. and S. Porter-Hudak, (1983), The Estimation and Application of Long Memory Time Series Models, *Journal of Time Series*, **4**, 221-238.

[11] Graham, J. R., and C. R. Harvey, (1997), Grading the Performance of Market-Timing Newsletters, *Financial Analysts Journal*, **53**, 5468.

[12] Granger, C. W. J. and R. Joyeux, (1980), A Introduction to Long-memory Time Series Models and Fractional Differencing, *Journal of Time Series Analysis*, **1**, 15-29.

[13] Han, Y. (2006), Asset Allocation with a High Dimensional Latent Factor Stochastic Volatility Model, *The Review of Financial Studies* **19**, 237 - 270.

[14] Hansen, L.P (1982), Large Sample Properties of Generalized Method of Moments Estimators, *Econometrica*, **50**, 1029-1054.

[15] Harris, D. and Mátyás, L. (1999), Generalized Method of Moments Estimation. Cambridge, University Press.

[16] Harte, D. (2001), Multifractals: Theory and Applications, London: Chapman and Hall.

[17] Idier, J. (2008), *Long Term vs. Short Term Comovements in Stock Markets. The Use of Markov-Switching Multifractal Models.*, Manuskript, Banque de France.

[18] Liesenfeld, R. and J. Richard (2003), Univariate and Multivariate Stochastic Volatility Models: Estimation and Diagnostics, *Journal of Empirical Finance* **10**, 505-531.

[19] Liu, R. (2008), Multi-variate Multifractal Models: Estimation of Parameters and Applications to Risk Management, University of Kiel, Ph.D thesis, Chapter 3.

[20] Liu, R., and T. Lux (2010), Flexible and Robust Modelling of Volatility Comovements: A Comparison of Two Multifractal Models, *Manuskript*,

[21] Lo, A. W. (1991), Long-Term Memory in Stock Market Prices, *Econometrica*, **59**, 1279-1313.

[22] Lobato, I. N. and N. E. Savin, (1998), Real and Spurious Long-Memory Properties of Stock Market Data, *Journal of Business* & *Economics Statistics*, **16**, 261-283.

[23] Lux, T. (1996), Long-term Stochastic Dependence in Financial Prices: Evidence from the German Stock Market, *Applied Economics Letters*, **3**, 701-706.

[24] Lux, T. (2004), Detecting Multifractal Properties in Asset Returns: The Failure of the "Scaling Estimator" *International Journal of Modern Physics*, **15**, 481-491.

[25] Lux, T. and T. Kaizoji (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.

[26] 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.

[27] Lux, T. and L. Morales-Arias (2010), Forecasting Volatility under Fractality, Regime-Switching, Long Memory and Student-t Innovations, *Computational Statistics & Data Analysis* (in press).

[28] Mandelbrot, B. (1974), Intermittent Turbulence in Self Similar Cascades: Divergence of High Moments and Dimension of Carrier, *Journal of Fluid Mechanics*, **62**, 331-358.

[29] Mandelbrot, B. Fisher, A. and Calvet, L. (1997), *Cowles Foundation Discussion Papers 1164–1166* available at http://www.ssrn.com Yale University.

[30] Modigliani, F., and L. Modigliani, 1997, Risk-Adjusted Performance, *Journal of Portfolio Management*, **23**, 4554.

[31] Zumbach, G. (2004), Volatility Processes and Volatility Forecast with Long Memory, *Quantitative Finance*, **4**, 70-86.

# Appendix: Moment Conditions

Recall the model from Section 3. Let $\varepsilon_t^{(\cdot)} = ln(|M_t^{(\cdot)}|)$, and we compute the first log difference:

$$\begin{aligned} X_{t,1} &= ln(|r_{1,t}|) - ln(|r_{1,t-1}|) \\ &= \frac{1}{2}\sum_{i=1}^{k}\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right) + \frac{1}{2}\sum_{l=k+1}^{n}\left(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}\right) + (ln|u_{2,t}| - ln|u_{1,t-1}|) \end{aligned}$$

$$\begin{aligned} Y_{t,1} &= ln(|r_{2,t}|) - ln(|r_{2,t-1}|) \\ &= \frac{1}{2}\sum_{i=1}^{k}\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right) + \frac{1}{2}\sum_{h=k+1}^{n}\left(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}\right) + (ln|u_{2,t}| - ln|u_{2,t-1}|) \end{aligned}$$

# A Binomial case

$$\begin{aligned} & cov(X_{t,1}, Y_{t,1}) \\ &= E\left[(X_{t,1} - E[X_{t,1}]) \cdot (Y_{t,1} - E[Y_{t,1}])\right] = E[X_{t,1} \cdot Y_{t,1}] \\ &= E\left\{\left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) + (ln|u_{1,t}| - ln|u_{1,t-1}|)\right] \cdot \right. \\ & \left. \left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) + (ln|u_{2,t}| - ln|u_{2,t-1}|)\right]\right\} \\ &= \tfrac{1}{4}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] - 2E[u_t]^2 + 2E[ln|u_{1,t}| \cdot ln|u_{2,t}|]. \end{aligned} \tag{A1}$$

We firstly consider $E[(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})^2]$, the only one non-zero contribution is $[ln(m_0) - ln(2-m_0)]^2$, and it occurs when new draws take place in cascade level $i$ between $t$ and $t-1$, whose probability by definition is $\frac{1}{2}\frac{1}{2^{k-i}}$. Summing up we get:

$$\begin{aligned} cov(X_{t,1}, Y_{t,1}) = 0.25 \cdot [ln(m_0) - ln(2-m_0)]^2 \cdot \sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} - 2E[u_t]^2 \\ +2E[ln|u_{1,t}| \cdot ln|u_{2,t}|]. \end{aligned}$$

$cov(X_{t+1,1}, Y_{t,1})$

$$= E\left\{\left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)} - \varepsilon_{t}^{(i)}) + \tfrac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(l)} - \varepsilon_{t}^{(l)}) + (ln|u_{1,t+1}| - ln|u_{1,t}|)\right] \cdot \left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_{t}^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{h=k+1}^{n}(\varepsilon_{t}^{(h)} - \varepsilon_{t-1}^{(h)}) + (ln|u_{2,t}| - ln|u_{2,t-1}|)\right]\right\}$$

$$= \tfrac{1}{4}E\left[\sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)} - \varepsilon_{t}^{(i)}\right) \cdot \sum_{i=1}^{k}\left(\varepsilon_{t}^{(i)} - \varepsilon_{t-1}^{(i)}\right)\right] + E[u_t]^2 - E[ln|u_{1,t}| \cdot ln|u_{2,t}|]. \tag{A2}$$

For $(\varepsilon_{t+1}^{(i)} - \varepsilon_{t}^{(i)})(\varepsilon_{t}^{(i)} - \varepsilon_{t-1}^{(i)})$, the non-zero value only occurs in case of two changes of the multiplier from time $t+1$ to time $t-1$, the probability of this occurrence is $(\frac{1}{2}\frac{1}{2^{k-i}})^2$. So, we have the result:

$$cov[X_{t+1,1}, Y_{t,1}]$$
$$= 0.25 \cdot [2ln(m_0) \cdot ln(2-m_0) - (ln(m_0))^2 - (ln(2-m_0))^2] \cdot \sum_{i=1}^{k}(\tfrac{1}{2}\tfrac{1}{2^{k-i}})^2$$
$$+E[u_t]^2 - E[ln|u_{1,t}| \cdot ln|u_{2,t}|].$$

Then, we look at the moment condition for one single time series:

$cov[X_{t+1,1}, X_{t,1}]$

$$= E\left\{\left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)} - \varepsilon_{t}^{(i)}) + \tfrac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_{t+1}^{(l)} - \varepsilon_{t}^{(l)}) + (ln|u_{1,t+1}| - ln|u_{1,t}|)\right] \cdot \left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_{t}^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_{t}^{(l)} - \varepsilon_{t-1}^{(l)}) + (ln|u_{1,t}| - ln|u_{1,t-1}|)\right]\right\}$$

$$= \tfrac{1}{4}E\left[\sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)} - \varepsilon_{t}^{(i)}\right) \cdot \sum_{i=1}^{k}\left(\varepsilon_{t}^{(i)} - \varepsilon_{t-1}^{(i)}\right)\right] + \tfrac{1}{4}E\left[\sum_{l=k+1}^{n}\left(\varepsilon_{t+1}^{(l)} - \varepsilon_{t}^{(l)}\right) \cdot \sum_{l=k+1}^{n}\left(\varepsilon_{t}^{(l)} - \varepsilon_{t-1}^{(l)}\right)\right]$$
$$+E\left[ln|u_t|\right]^2 - E\left[ln|u_t|^2\right]. \tag{A3}$$

The first component is identical to the one of the case of $cov[X_{t+1,1}, Y_{t,1}]$, and the second component can be derived in the same way. Adding together we arrive at:

$$\begin{aligned}
&cov[X_{t+1,1}, X_{t,1}] \\
&= 0.25 \cdot \left[2ln(m_0) \cdot ln(2-m_0) - (ln(m_0))^2 - (ln(2-m_0))^2\right] \cdot \sum_{i=1}^{k} (\tfrac{1}{2}\tfrac{1}{2^{k-i}})^2 \\
&+0.25 \cdot \left[2ln(m_0) \cdot ln(2-m_0) - (ln(m_0))^2 - (ln(2-m_0))^2\right] \cdot \sum_{i=k+1}^{n} (\tfrac{1}{2}\tfrac{1}{2^{n-i}})^2 \\
&+E[ln|u_t|]^2 - E[ln|u_t|^2].
\end{aligned} \tag{A4}$$

By our assumption of both time series having the same number of cascade levels, the moments for the two individual time series are identical for the same length of time lags.

Then, let's turn to the squared observations:

$$\begin{aligned}
&E[X_{t,1}^2 \cdot Y_{t,1}^2] \\
&= E\left\{\left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) + (ln|u_{1,t}| - ln|u_{1,t-1}|)\right]^2 \cdot \right. \\
&\qquad \left. \left[\tfrac{1}{2}\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) + (ln|u_{2,t}| - ln|u_{2,t-1}|)\right]^2\right\} \\
&= \tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^4\right] + \tfrac{1}{16}E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right] \\
&+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right] \\
&+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right] \\
&+\tfrac{1}{4}\left\{2E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] + 2E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right]\right\} \cdot \left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2\right) \\
&+2E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 8E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|] \\
&+4E[ln|u_{1,t}| \cdot (ln|u_{2,t}|]^2 + 2E[(ln|u_t|)^2]^2.
\end{aligned}$$

By examining each component in the expression above combining with the calculations of the previous moments, it is not difficult to find the solution:

$E[X_{t,1}^2 \cdot Y_{t,1}^2]$

$= [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \frac{1}{16} \sum_{i=1}^{k} \frac{1}{2} \frac{1}{2^{k-i}} + [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \frac{1}{16} \sum_{i=k+1}^{n} \frac{1}{2} \frac{1}{2^{n-i}} \sum_{i=k+1}^{n} \frac{1}{2} \frac{1}{2^{n-i}}$

$+2\,[ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \frac{1}{16} \sum_{i=1}^{k} \frac{1}{2} \frac{1}{2^{k-i}} \sum_{i=k+1}^{n} \frac{1}{2} \frac{1}{2^{n-i}}$

$+ \left(E[ln|u_t|^2] - E[ln|u_t|]^2\right) \cdot [ln\,(m_0) - ln\,(2-m_0)]^2 \cdot \left( \sum_{i=1}^{k} \frac{1}{2} \frac{1}{2^{k-i}} + \sum_{i=k+1}^{n} \frac{1}{2} \frac{1}{2^{n-i}} \right)$

$+2E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 8E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|]$

$+4E[ln|u_{1,t}| \cdot ln|u_{2,t}|]^2 + 2E[(ln|u_t|)^2]^2.$

(A5)

$E[X_{t+1,1}^2 \cdot Y_{t,1}^2]$

$= E\left\{ \left[ \frac{1}{2} \sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) + \frac{1}{2} \sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) + (ln|u_{1,t+1}| - ln|u_{1,t}|) \right]^2 \cdot \right.$

$\left. \left[ \frac{1}{2} \sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \frac{1}{2} \sum_{h=k+1}^{n} (\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) + (ln|u_{2,t}| - ln|u_{2,t-1}|) \right]^2 \right\}$

$= \frac{1}{16} E\left[ \left( \sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \left( \sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \right]$

$+ \frac{1}{16} E\left[ \left( \sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) \right)^2 \left( \sum_{h=k+1}^{n} (\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) \right)^2 \right]$

$+ \frac{1}{16} E\left[ \left( \sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \left( \sum_{h=k+1}^{n} (\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)}) \right)^2 \right]$

$+ \frac{1}{16} E\left[ \left( \sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \left( \sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) \right)^2 \right]$

$+ \frac{1}{4} \left\{ 2E\left[ \left( \sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \right] + 2E\left[ \left( \sum_{l=k+1}^{n} (\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) \right)^2 \right] \right\} \cdot \left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2\right)$

$+E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 4E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|] + 4E[ln|u_{1,t}| \cdot ln|u_{2,t}|]E[ln|u_t|]^2$

$+3E[ln|u_t|^2]^2 - 4E[ln|u_t|^2]E[ln|u_t|]^2.$

Until now, the only unfamiliar component is the first term:

$E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^{2}\cdot\left(\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^{2}\right]$, there are three different forms to be considered:

(1) $\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)^{2}\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)^{2}$, which have non-zero value only if $\varepsilon_{t+1}^{(i)} \neq \varepsilon_{t}^{(i)} \neq \varepsilon_{t-1}^{(i)}$. and this possibility is $(\frac{1}{2}\frac{1}{2^{k-i}})^{2}$, combining with the non-zero expectation value,
we have $\left(\sum_{i=1}^{k}(\frac{1}{2}\frac{1}{2^{k-i}})^{2}\right)\cdot[ln\,(m_0)-ln\,(2-m_0)]^{4}$.

(2) $\left(\varepsilon_{t+1}^{(j)}-\varepsilon_{t}^{(j)}\right)^{2}\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)^{2}$, which are non-zero for $i \neq j$, $\varepsilon_{t+1}^{(j)} \neq \varepsilon_{t}^{(j)}$ and $\varepsilon_{t}^{(i)} \neq \varepsilon_{t-1}^{(i)}$, the probability of its occurrence is $\sum_{i=1}^{k}\left(\frac{1}{2^{k-i}}\sum_{j=1,j\neq i}^{k}\frac{1}{2^{k-j}}\right)$.

Putting together these two possible forms we get
$[ln\,(m_0)-ln\,(2-m_0)]^{4}\cdot\left(\sum_{i=1}^{k}\frac{1}{2}\frac{1}{2^{k-i}}\sum_{j=1}^{k}\frac{1}{2}\frac{1}{2^{k-j}}\right)$.

(3) Form $\left(\varepsilon_{t+1}^{(j)}-\varepsilon_{t}^{(j)}\right)\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\left(\varepsilon_{t}^{(j)}-\varepsilon_{t-1}^{(j)}\right)\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)$, which for $i \neq j$ and $\varepsilon_{t+1}^{(n)} \neq \varepsilon_{t}^{(n)} \neq \varepsilon_{t-1}^{(n)}$, $n=i,j$ are non-zero, and which implies
$2\left\{\sum_{i=1}^{k}\left((\frac{1}{2^{k-i}})^{2}\sum_{j=1,j\neq i}^{k}(\frac{1}{2^{k-j}})^{2}\right)\right\}\cdot[ln\,(m_0)-ln\,(2-m_0)]^{4}$.

Then we have the solution for the first component in the above moment condition:

$$E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^{2}\cdot\left(\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^{2}\right]$$
$$=[ln\,(m_0)-ln\,(2-m_0)]^{4}\left[\sum_{i=1}^{k}\frac{1}{2}\frac{1}{2^{k-i}}\sum_{j=1}^{k}\frac{1}{2}\frac{1}{2^{k-j}}+2\sum_{i=1}^{k}(\frac{1}{2}\frac{1}{2^{k-i}})^{2}\sum_{j=1,j\neq i}^{k}(\frac{1}{2}\frac{1}{2^{k-j}})^{2}\right]$$

The other components can be solved by recalling previous calculations. All in all, we finally arrive at:

$$E[X_{t+1,1}^2 \cdot Y_{t,1}^2]$$
$$= [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{16} \left[ \sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} \sum_{j=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-j}} + 2\sum_{i=1}^{k} (\tfrac{1}{2}\tfrac{1}{2^{k-i}})^2 \sum_{j=1, j\neq i}^{k} (\tfrac{1}{2}\tfrac{1}{2^{k-j}})^2 \right]$$
$$+ [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{16} \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{n-i}} \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{n-i}}$$
$$+ \tfrac{1}{8}\,[ln\,(m_0) - ln\,(2-m_0)]^4 \sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{n-i}}$$
$$+ \left(E[ln|u_t|^2] - E[ln|u_t|]^2\right) \cdot [ln\,(m_0) - ln\,(2-m_0)]^2 \cdot \left( \sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} + \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{n-i}} \right)$$
$$+ E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 4E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|] + 4E[ln|u_{1,t}| \cdot ln|u_{2,t}|]E[ln|u_t|]^2$$
$$+ 3E[ln|u_t|^2]^2 - 4E[ln|u_t|^2]E[ln|u_t|]^2. \tag{A6}$$

$$E[X_{t+1,1}^2 \cdot X_{t,1}^2]$$
$$= E\left\{ \left[ \tfrac{1}{2}\sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) + \tfrac{1}{2}\sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) + (ln|u_{1,t+1}| - ln|u_{1,t}|) \right]^2 \cdot \right.$$
$$\left. \left[ \tfrac{1}{2}\sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) + \tfrac{1}{2}\sum_{l=k+1}^{n} (\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) + (ln|u_{1,t}| - ln|u_{1,t-1}|) \right]^2 \right\}$$

$$= \tfrac{1}{16}E\left[ \left( \sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \cdot \left( \sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \right] + \tfrac{1}{16}E\left[ \left( \sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \cdot \left( \sum_{l=k+1}^{n} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \right]$$
$$+ \tfrac{1}{16}E\left[ \left( \sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \left( \sum_{l=k+1}^{n} (\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}) \right)^2 \right] + \tfrac{1}{16}E\left[ \left( \sum_{i=1}^{k} (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}) \right)^2 \left( \sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) \right)^2 \right]$$
$$+ \tfrac{1}{4}\left\{ 2E\left[ \left( \sum_{i=1}^{k} (\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \right)^2 \right] + 2E\left[ \left( \sum_{l=k+1}^{n} (\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}) \right)^2 \right] \right\} \cdot \left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2\right)$$
$$+ \tfrac{1}{16} \cdot 4E\left[ \sum_{i=1}^{k} \left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right) \sum_{i=1}^{k} \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right) \right] E\left[ \sum_{l=k+1}^{n} \left(\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}\right) \sum_{l=k+1}^{n} \left(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}\right) \right]$$
$$+ \tfrac{1}{4} \cdot 4E\left[ \sum_{i=1}^{k} \left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right) \sum_{i=1}^{k} \left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right) \right] \cdot \left(E[ln|u_t|]^2 - E[ln|u_t|^2]\right)$$
$$+ \tfrac{1}{4} \cdot 4E\left[ \sum_{l=k+1}^{n} \left(\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}\right) \sum_{l=k+1}^{n} \left(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}\right) \right] \cdot \left(E[ln|u_t|]^2 - E[ln|u_t|^2]\right)$$
$$+ 3E[ln|u_t|^2]^2 + E[ln|u_t|^4] - 4E[ln|u_t|^3]E[ln|u_t|].$$

(A7)

The first and second term are the same as the first one in the case $E[X^2_{t+1,1}, Y^2_{t,1}]$, and the rest are our familiars. Adding together, we have the result:

$$
\begin{aligned}
&E[X^2_{t+1,1} \cdot X^2_{t,1}] \\
&= [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{16} \left[ \sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} \sum_{j=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-j}} + 2\sum_{i=1}^{k} (\tfrac{1}{2}\tfrac{1}{2^{k-i}})^2 \sum_{j=1, j\neq i}^{k} (\tfrac{1}{2}\tfrac{1}{2^{k-j}})^2 \right] \\
&+ [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{16} \left[ \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{k-i}} \sum_{j=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{k-j}} + 2\sum_{i=k+1}^{n} (\tfrac{1}{2}\tfrac{1}{2^{k-i}})^2 \sum_{j=k+1, j\neq i}^{n} (\tfrac{1}{2}\tfrac{1}{2^{k-j}})^2 \right] \\
&+ [ln\,(m_0) - ln\,(2-m_0)]^4 \cdot \tfrac{1}{8} \sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{n-i}} \\
&+ \left(E[ln|u_t|^2] - E[ln|u_t|]^2\right) \cdot [ln\,(m_0) - ln\,(2-m_0)]^2 \cdot \left( \sum_{i=1}^{k} \tfrac{1}{2}\tfrac{1}{2^{k-i}} + \sum_{i=k+1}^{n} \tfrac{1}{2}\tfrac{1}{2^{n-i}} \right) \\
&+0.25 \left[ 2ln\,(m_0)\, ln\,(2-m_0) - (ln\,(m_0))^2 - (ln\,(2-m_0))^2 \right]^2 \left( \sum_{i=1}^{k} \left(\tfrac{1}{2}\tfrac{1}{2^{k-i}}\right)^2 \sum_{i=k+1}^{n} \left(\tfrac{1}{2}\tfrac{1}{2^{n-i}}\right)^2 \right) \\
&+2 \left[ 2ln\,(m_0)\, ln\,(2-m_0) - (ln\,(m_0))^2 - (ln\,(2-m_0))^2 \right] \cdot \\
&\left( \sum_{i=1}^{k} \left(\tfrac{1}{2}\tfrac{1}{2^{k-i}}\right)^2 + \sum_{i=k+1}^{n} \left(\tfrac{1}{2}\tfrac{1}{2^{n-i}}\right)^2 \right) \cdot \left(E[ln|u_t|]^2 - E[ln|u_t|^2]\right) \\
&+3E[ln|u_t|^2]^2 + E[ln|u_t|^4] - 4E[ln|u_t|^3]E[ln|u_t|].
\end{aligned}
$$

(A8)

# B Lognormal case

$$cov(X_{t,1}, Y_{t,1}) = E[(X_{t,1} - E[X_{t,1}]) \cdot (Y_{t,1} - E[Y_{t,1}] = E[X_{t,1} \cdot Y_{t,1}]$$

$$= \tfrac{1}{4} E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] + 2E[ln|u_{1,t}| \cdot ln|u_{2,t}|] - 2E[u_t]^2$$

$$= 0.5\sigma_\varepsilon^2 \sum_{i=1}^{k} \tfrac{1}{2^{k-i}} + 2E[ln|u_{1,t}| \cdot ln|u_{2,t}|] - 2E[u_t]^2. \tag{B1}$$

Because the non-zero outcomes occur when $\varepsilon_t^{(i)} \neq \varepsilon_{t-1}^{(i)}$, which implies:

$$E\left[(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})^2\right] = 2(E[(\varepsilon_t^{(i)})^2] - E[\varepsilon_t^{(i)}]^2) = 2\sigma_\varepsilon^2$$

$$cov(X_{t+1,1}, Y_{t,1}) = \tfrac{1}{4} E\left[\sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right) \cdot \sum_{i=1}^{k}\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)\right] + E[u_t]^2 - E[ln|u_{1,t}| \cdot ln|u_{2,t}|]$$

$$= -0.25\sigma_\varepsilon^2 \sum_{i=1}^{k}\left(\tfrac{1}{2^{k-i}}\right)^2 + E[u_t]^2 - E[ln|u_{1,t}| \cdot ln|u_{2,t}|]. \tag{B2}$$

Because the non-zero outcomes occur when $\varepsilon_{t+1}^{(i)} \neq \varepsilon_t^{(i)} \neq \varepsilon_{t-1}^{(i)}$, which implies:

$$E\left[(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}) \cdot (\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right] = E[\varepsilon_t^{(i)}]^2 - E[(\varepsilon_t^{(i)})^2] = -\sigma_\varepsilon^2$$

$$cov(X_{t+1,1}, X_{t,1})$$

$$= \tfrac{1}{4} E\left[\sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right) \cdot \sum_{i=1}^{k}\left(\varepsilon_{t+1}^{(i)} - \varepsilon_t^{(i)}\right)\right] + \tfrac{1}{4} E\left[\sum_{l=k+1}^{n}\left(\varepsilon_{t+1}^{(l)} - \varepsilon_t^{(l)}\right) \cdot \sum_{l=k+1}^{n}\left(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)}\right)\right]$$

$$+E\left[ln|u_t|\right]^2 - E\left[ln|u_t|^2\right]$$

$$= -0.25\sigma_\varepsilon^2 \left[\sum_{i=1}^{k}\left(\tfrac{1}{2^{k-i}}\right)^2 \sum_{i=k+1}^{n}\left(\tfrac{1}{2^{n-i}}\right)^2\right] + E\left[ln|u_t|\right]^2 - E\left[ln|u_t|^2\right]. \tag{B3}$$

$E[X_{t,1}^2 \cdot Y_{t,1}^2]$

$$= \tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^4\right] + \tfrac{1}{16}E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right]$$
$$+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\left(\sum_{h=k+1}^{n}(\varepsilon_t^{(h)} - \varepsilon_{t-1}^{(h)})\right)^2\right]$$
$$+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right]$$
$$+\tfrac{1}{4}\left\{2E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^2\right] + 2E\left[\left(\sum_{l=k+1}^{n}(\varepsilon_t^{(l)} - \varepsilon_{t-1}^{(l)})\right)^2\right]\right\} \cdot \left(2E[ln|u_t|^2] - 2E[ln|u_t|]^2\right)$$
$$+2E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 8E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|]$$
$$+4E[ln|u_{1,t}| \cdot (ln|u_{2,t}|]^2 + 2E[(ln|u_t|)^2]^2$$
$$= 0.75\sigma_\varepsilon^4 \sum_{i=1}^{k} \tfrac{1}{2^{k-i}} + 0.25\sigma_\varepsilon^4 \sum_{l=k+1}^{n} \tfrac{1}{2^{n-l}} \sum_{h=k+1}^{n} \tfrac{1}{2^{n-h}}$$
$$+0.25\sigma_\varepsilon^4 \sum_{i=1}^{k} \tfrac{1}{2^{k-i}} \sum_{h=k+1}^{n} \tfrac{1}{2^{n-h}} + 0.25\sigma_\varepsilon^4 \sum_{i=1}^{k} \tfrac{1}{2^{k-i}} \sum_{l=k+1}^{n} \tfrac{1}{2^{n-l}}$$
$$+2\sigma_\varepsilon^2\left(E[ln|u_t|^2] - E[ln|u_t|]^2\right) \cdot \left(\sum_{i=1}^{k} \tfrac{1}{2^{k-i}} + \sum_{l=k+1}^{n} \tfrac{1}{2^{n-l}}\right)$$
$$+2E[(ln|u_{1,t}|)^2 \cdot (ln|u_{2,t}|)^2] - 8E[(ln|u_{1,t}|)^2 \cdot ln|u_{2,t}|] \cdot E[ln|u_t|]$$
$$+4E[ln|u_{1,t}| \cdot (ln|u_{2,t}|]^2 + 2E[(ln|u_t|)^2]^2. \tag{B4}$$

For the first term $E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^4\right]$, let's begin with $E\left[\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)^4\right]$, the non-zero value implies:

$$E\left[\left(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)}\right)^4\right] = 2E[\varepsilon_t^{(i)}]^4 + 6E[(\varepsilon_t^{(i)})^2]^2 - 8E[(\varepsilon_t^{(i)})^3]E[\varepsilon_t^{(i)}] = 12\sigma_\varepsilon^4.$$

This occurs with probability $2^{\frac{1}{k-i}}$. Then we have the solution:

$$E\left[\left(\sum_{i=1}^{k}(\varepsilon_t^{(i)} - \varepsilon_{t-1}^{(i)})\right)^4\right] = 12\sigma_\varepsilon^4 \cdot \sum_{i=1}^{k} \frac{1}{2^{k-i}}.$$

$E[X^2_{t+1,1}\cdot Y^2_{t,1}]$

$$= \frac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t})\right)^2\cdot\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1})\right)^2\right]$$
$$+\frac{1}{16}E\left[\left(\sum_{l=k+1}^{n}(\varepsilon^{(l)}_{t+1}-\varepsilon^{(l)}_{t})\right)^2\left(\sum_{h=k+1}^{n}(\varepsilon^{(h)}_{t}-\varepsilon^{(h)}_{t-1})\right)^2\right]$$
$$+\frac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t})\right)^2\left(\sum_{h=k+1}^{n}(\varepsilon^{(h)}_{t}-\varepsilon^{(h)}_{t-1})\right)^2\right]$$
$$+\frac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1})\right)^2\left(\sum_{l=k+1}^{n}(\varepsilon^{(l)}_{t+1}-\varepsilon^{(l)}_{t})\right)^2\right]$$
$$+\frac{1}{4}\left\{2E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1})\right)^2\right]+2E\left[\left(\sum_{l=k+1}^{n}(\varepsilon^{(l)}_{t}-\varepsilon^{(l)}_{t-1})\right)^2\right]\right\}\cdot\left(2E[ln|u_t|^2]-2E[ln|u_t|]^2\right)$$
$$+E[(ln|u_{1,t}|)^2\cdot(ln|u_{2,t}|)^2]-4E[(ln|u_{1,t}|)^2\cdot ln|u_{2,t}|]\cdot E[ln|u_t|]+4E[ln|u_{1,t}|\cdot(ln|u_{2,t}|]E[ln|u_t|]^2$$
$$+3E[ln|u_t|^2]^2-4E[ln|u_t|^2]E[ln|u_t|]^2.$$

For the first term $E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t})\right)^2\cdot\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1})\right)^2\right]$, there are three different possible forms:

(1) $\left(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t}\right)^2\left(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1}\right)^2$, has non-zero value only if $\varepsilon^{(i)}_{t+1}\neq\varepsilon^{(i)}_{t}\neq\varepsilon^{(i)}_{t-1}$. then $E\left[\left(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t}\right)^2\left(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1}\right)^2\right]=E[\varepsilon_t^4]+3E[\varepsilon_t^2]^2-4E[\varepsilon_t^3]E[\varepsilon_t]=6\sigma_\varepsilon^4$. ($E[\varepsilon_t^3]=3\lambda\sigma_\varepsilon^2+\lambda^3$ and $E[\varepsilon_t^4]=3\sigma_\varepsilon^4+6\lambda^2\sigma_\varepsilon^2+\lambda^4$,) and the probability of this occurance is $(\frac{1}{2^{k-i}})^2$. Putting together we get $\left[\sum_{i=1}^{k}\left(\frac{1}{2^{k-i}}\right)^2\right]\cdot 6\sigma_\varepsilon^4$

(2) $\left(\varepsilon^{(j)}_{t+1}-\varepsilon^{(j)}_{t}\right)^2\left(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1}\right)^2$, does not equal zero for $i\neq j$, $\varepsilon^{(j)}_{t+1}\neq\varepsilon^{(j)}_{t}$ and $\varepsilon^{(i)}_{t}\neq\varepsilon^{(i)}_{t-1}$. since $E\left[\left(\varepsilon^{(j)}_{t+1}-\varepsilon^{(j)}_{t}\right)^2\left(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1}\right)^2\right]=4E[(\varepsilon^{(i)}_{t})^2]^2-8E[(\varepsilon^{(i)}_{t})^2]E[\varepsilon^{(i)}_{t}]^2+4E[\varepsilon^{(i)}_{t}]^4=4\sigma_\varepsilon^4$, together with the probability, this overall contribution yields:
$\left[\sum_{i=1}^{k}\left(\frac{1}{2^{k-i}}\sum_{j=1,j\neq i}^{k}\frac{1}{2^{k-j}}\right)\right]\cdot 4\sigma_\varepsilon^4$

(3) $\left(\varepsilon_{t+1}^{(j)}-\varepsilon_{t}^{(j)}\right)\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\left(\varepsilon_{t}^{(j)}-\varepsilon_{t-1}^{(j)}\right)\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right)$, which for $i \neq j$ and $\varepsilon_{t+1}^{(n)} \neq \varepsilon_{t}^{(n)} \neq \varepsilon_{t-1}^{(n)}$, $n = i, j$ are non-zero, since $\left(\varepsilon_{t+1}^{(j)}-\varepsilon_{t}^{(j)}\right)\left(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)}\right)\left(\varepsilon_{t}^{(j)}-\varepsilon_{t-1}^{(j)}\right)\left(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)}\right) = 4\sigma_{\varepsilon}^{4}$, we obtain a contribution $2\left[\sum_{i=1}^{k}\left(\frac{1}{2^{k-i}}\right)^{2}\sum_{j=1,j\neq i}^{k}\left(\frac{1}{2^{k-j}}\right)^{2}\right]\cdot\sigma_{\varepsilon}^{4}$.

Combining those three cases, we have the result:

$$E\left[\left(\sum_{i=1}^{k}(\varepsilon_{t+1}^{(i)}-\varepsilon_{t}^{(i)})\right)^{2}\cdot\left(\sum_{i=1}^{k}(\varepsilon_{t}^{(i)}-\varepsilon_{t-1}^{(i)})\right)^{2}\right]$$

$$= 6\sigma_{\varepsilon}^{4}\cdot\sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^{2} + 4\sigma_{\varepsilon}^{4}\cdot\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{j=1,j\neq i}^{k}\tfrac{1}{2^{k-j}} + 2\sigma_{\varepsilon}^{4}\cdot\sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^{2}\sum_{j=1,j\neq i}^{k}(\tfrac{1}{2^{k-j}})^{2}$$

$$E[X_{t+1,1}^{2}\cdot Y_{t,1}^{2}]$$

$$\begin{aligned}
&= \tfrac{1}{16}\left[6\sigma_{\varepsilon}^{4}\cdot\sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^{2} + 4\sigma_{\varepsilon}^{4}\cdot\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{j=1,j\neq i}^{k}\tfrac{1}{2^{k-j}} + 2\sigma_{\varepsilon}^{4}\cdot\sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^{2}\sum_{j=1,j\neq i}^{k}(\tfrac{1}{2^{k-j}})^{2}\right]\\
&+0.25\sigma_{\varepsilon}^{4}\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\sum_{h=k+1}^{n}\tfrac{1}{2^{n-h}}\\
&+0.25\sigma_{\varepsilon}^{4}\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{h=k+1}^{n}\tfrac{1}{2^{n-h}} + 0.25\sigma_{\varepsilon}^{4}\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\\
&+2\sigma_{\varepsilon}^{2}\cdot\left(E[ln|u_t|^{2}] - E[ln|u_t|]^{2}\right)\cdot\left(\sum_{i=1}^{k}\tfrac{1}{2^{k-i}} + \sum_{i=k+1}^{n}\tfrac{1}{2^{n-i}}\right)\\
&+E[(ln|u_{1,t}|)^{2}\cdot(ln|u_{2,t}|)^{2}] - 4E[(ln|u_{1,t}|)^{2}\cdot ln|u_{2,t}|]\cdot E[ln|u_t|] + 4E[ln|u_{1,t}|\cdot(ln|u_{2,t}|]E[ln|u_t|]^{2}\\
&+3E[ln|u_t|^{2}]^{2} - 4E[ln|u_t|^{2}]E[ln|u_t|]^{2}
\end{aligned}\tag{B5}$$

$$E[X^2_{t+1,1}\cdot X^2_{t,1}]$$

$$
\begin{aligned}
&= \tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t})\right)^2\cdot\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1})\right)^2\right]+\tfrac{1}{16}E\left[\left(\sum_{l=k+1}^{n}(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t})\right)^2\cdot\left(\sum_{l=k+1}^{n}(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1})\right)^2\right]\\
&+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t})\right)^2\left(\sum_{l=k+1}^{n}(\varepsilon^{(l)}_{t}-\varepsilon^{(l)}_{t-1})\right)^2\right]+\tfrac{1}{16}E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1})\right)^2\left(\sum_{l=k+1}^{n}(\varepsilon^{(l)}_{t+1}-\varepsilon^{(l)}_{t})\right)^2\right]\\
&+\tfrac{1}{4}\left\{2E\left[\left(\sum_{i=1}^{k}(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t})\right)^2\right]+2E\left[\left(\sum_{l=k+1}^{n}(\varepsilon^{(l)}_{t+1}-\varepsilon^{(l)}_{t})\right)^2\right]\right\}\cdot\left(2E[ln|u_t|^2]-2E[ln|u_t|]^2\right)\\
&+4\cdot\tfrac{1}{16}E\left[\sum_{i=1}^{k}\left(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t}\right)\sum_{i=1}^{k}\left(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1}\right)\right]E\left[\sum_{l=k+1}^{n}\left(\varepsilon^{(l)}_{t+1}-\varepsilon^{(l)}_{t}\right)\sum_{l=k+1}^{n}\left(\varepsilon^{(l)}_{t}-\varepsilon^{(l)}_{t-1}\right)\right]\\
&+4\cdot\tfrac{1}{4}E\left[\sum_{i=1}^{k}\left(\varepsilon^{(i)}_{t+1}-\varepsilon^{(i)}_{t}\right)\sum_{i=1}^{k}\left(\varepsilon^{(i)}_{t}-\varepsilon^{(i)}_{t-1}\right)\right]\cdot\left(E[ln|u_t|]^2-E[ln|u_t|^2]\right)\\
&+4\cdot\tfrac{1}{4}E\left[\sum_{l=k+1}^{n}\left(\varepsilon^{(l)}_{t+1}-\varepsilon^{(l)}_{t}\right)\sum_{l=k+1}^{n}\left(\varepsilon^{(l)}_{t}-\varepsilon^{(l)}_{t-1}\right)\right]\cdot\left(E[ln|u_t|]^2-E[ln|u_t|^2]\right)\\
&+3E[ln|u_t|^2]^2+E[ln|u_t|^4]-4E[ln|u_t|^3]E[ln|u_t|]\\
&= \tfrac{1}{16}\left[6\sigma^4_\varepsilon\cdot\sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^2+4\sigma^4_\varepsilon\cdot\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{j=1,j\neq i}^{k}\tfrac{1}{2^{k-j}}+2\sigma^4_\varepsilon\cdot\sum_{i=1}^{k}(\tfrac{1}{2^{k-i}})^2\sum_{j=1,j\neq i}^{k}(\tfrac{1}{2^{k-j}})^2\right]\\
&+\tfrac{1}{16}\left[6\sigma^4_\varepsilon\cdot\sum_{l=k+1}^{n}(\tfrac{1}{2^{n-l}})^2+4\sigma^4_\varepsilon\cdot\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\sum_{j=k+1,j\neq i}^{n}\tfrac{1}{2^{n-j}}+2\sigma^4_\varepsilon\cdot\sum_{l=k+1}^{n}(\tfrac{1}{2^{n-l}})^2\sum_{j=k+1,j\neq i}^{n}(\tfrac{1}{2^{n-j}})^2\right]\\
&+0.25\sigma^4_\varepsilon\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}+0.25\sigma^4_\varepsilon\sum_{l=k+1}^{n}\tfrac{1}{2^{n-l}}\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}\\
&+2\sigma^2_\varepsilon\left(\sum_{i=1}^{k}\tfrac{1}{2^{k-i}}+\sum_{i=k+1}^{n}\tfrac{1}{2^{n-i}}\right)\cdot\left(E[ln|u_t|^2]-E[ln|u_t|]^2\right)\\
&+0.25\sigma^4_\varepsilon\sum_{i=1}^{k}\left(\tfrac{1}{2^{k-i}}\right)^2\cdot\sum_{i=k+1}^{n}\left(\tfrac{1}{2^{n-i}}\right)^2\\
&-\sigma^2_\varepsilon\cdot\left(\sum_{i=1}^{k}\left(\tfrac{1}{2^{k-i}}\right)^2+\sum_{i=k+1}^{n}\left(\tfrac{1}{2^{n-i}}\right)^2\right)\cdot\left(E[ln|u_t|]^2-E[ln|u_t|^2]\right)\\
&+3E[ln|u_t|^2]^2+E[ln|u_t|^4]-4E[ln|u_t|^3]E[ln|u_t|].
\end{aligned}
\tag{B6}
$$

Because the first term is identical with the first one of case $E[X^2_{t+1,1}, Y^2_{t,1}]$.

Table 1: GMM Estimation of the bivariate MF binomial Model

| | | $\hat{m}_0$ | | | $\hat{\rho}$ | | | $\hat{\sigma}_1$ | | | $\hat{\sigma}_2$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *Bias* | *SD* | *RMSE* | *Bias* | *SD* | *RMSE* | *Bias* | *SD* | *RMSE* | *Bias* | *SD* | *RMSE* |
| $m_0 = 1.20$ | $N_1$ | -0.095 | 0.128 | 0.159 | 0.000 | 0.073 | 0.073 | 0.004 | 0.042 | 0.042 | 0.001 | 0.041 | 0.041 |
| | $N_2$ | -0.071 | 0.122 | 0.141 | 0.002 | 0.047 | 0.047 | 0.002 | 0.028 | 0.028 | 0.000 | 0.027 | 0.027 |
| | $N_3$ | -0.054 | 0.103 | 0.116 | 0.003 | 0.032 | 0.032 | 0.001 | 0.019 | 0.019 | 0.000 | 0.019 | 0.018 |
| $m_0 = 1.30$ | $N_1$ | -0.099 | 0.144 | 0.175 | 0.007 | 0.084 | 0.084 | 0.006 | 0.063 | 0.063 | 0.000 | 0.061 | 0.061 |
| | $N_2$ | -0.045 | 0.107 | 0.116 | 0.002 | 0.052 | 0.052 | 0.004 | 0.042 | 0.042 | 0.000 | 0.041 | 0.041 |
| | $N_3$ | -0.019 | 0.067 | 0.070 | 0.000 | 0.035 | 0.035 | 0.001 | 0.029 | 0.029 | 0.001 | 0.028 | 0.028 |
| $m_0 = 1.40$ | $N_1$ | -0.064 | 0.120 | 0.136 | 0.007 | 0.086 | 0.086 | 0.009 | 0.090 | 0.090 | -0.001 | 0.088 | 0.088 |
| | $N_2$ | -0.015 | 0.059 | 0.060 | -0.004 | 0.052 | 0.052 | 0.005 | 0.060 | 0.060 | 0.000 | 0.058 | 0.058 |
| | $N_3$ | -0.004 | 0.033 | 0.034 | -0.008 | 0.035 | 0.036 | 0.001 | 0.042 | 0.042 | 0.001 | 0.041 | 0.041 |
| $m_0 = 1.50$ | $N_1$ | -0.041 | 0.074 | 0.085 | 0.005 | 0.090 | 0.090 | 0.009 | 0.132 | 0.132 | -0.018 | 0.117 | 0.118 |
| | $N_2$ | -0.005 | 0.040 | 0.040 | -0.016 | 0.054 | 0.057 | 0.004 | 0.082 | 0.082 | -0.005 | 0.084 | 0.084 |
| | $N_3$ | 0.001 | 0.024 | 0.024 | -0.019 | 0.038 | 0.043 | 0.001 | 0.058 | 0.058 | 0.002 | 0.060 | 0.060 |

Note: All simulations are based on the bivariate Multifractal process with the whole number of cascade levels equal to 12, $j = 6$, $\rho = 0.5$, $\sigma_1 = 1$, $\sigma_2 = 1$, and eight moment conditions as in the Appendix A are used. Sample lengths are $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$. Bias denotes the distance between the given and estimated parameter value, SD and RMSE denote the standard deviation and root mean squared error, respectively. For each scenario, 400 Monte Carlo simulations have been carried out.

Table 2: GMM estimation of the bivariate MF (Lognormal) Model

| | | $\hat{\lambda}$ | | | $\hat{\rho}$ | | | $\hat{\sigma_1}$ | | | $\hat{\sigma_2}$ | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | *Bias* | *SD* | *RMSE* | *Bias* | *SD* | *RMSE* | *Bias* | *SD* | *RMSE* | *Bias* | *SD* | *RMSE* |
| $\lambda = 0.10$ | $N_1$ | -0.029 | 0.061 | 0.068 | -0.066 | 0.352 | 0.358 | -0.077 | 0.345 | 0.353 | 0.012 | 0.069 | 0.070 |
| | $N_2$ | 0.017 | 0.045 | 0.049 | 0.049 | 0.272 | 0.276 | 0.051 | 0.267 | 0.272 | -0.007 | 0.050 | 0.051 |
| | $N_3$ | -0.009 | 0.033 | 0.035 | -0.010 | 0.204 | 0.204 | 0.011 | 0.204 | 0.204 | -0.002 | 0.036 | 0.036 |
| $\lambda = 0.20$ | $N_1$ | -0.041 | 0.082 | 0.091 | -0.166 | 0.441 | 0.471 | -0.151 | 0.477 | 0.499 | -0.011 | 0.077 | 0.078 |
| | $N_2$ | 0.018 | 0.047 | 0.050 | -0.108 | 0.362 | 0.377 | 0.087 | 0.396 | 0.405 | 0.006 | 0.041 | 0.041 |
| | $N_3$ | -0.009 | 0.033 | 0.034 | -0.061 | 0.314 | 0.319 | -0.05 | 0.331 | 0.334 | -0.003 | 0.030 | 0.031 |
| $\lambda = 0.30$ | $N_1$ | 0.043 | 0.084 | 0.095 | -0.191 | 0.624 | 0.652 | -0.192 | 0.609 | 0.637 | -0.016 | 0.084 | 0.086 |
| | $N_2$ | 0.022 | 0.048 | 0.052 | 0.141 | 0.505 | 0.524 | -0.122 | 0.544 | 0.557 | 0.011 | 0.042 | 0.043 |
| | $N_3$ | -0.009 | 0.034 | 0.035 | -0.086 | 0.452 | 0.46 | -0.077 | 0.451 | 0.457 | 0.006 | 0.032 | 0.032 |
| $\lambda = 0.40$ | $N_1$ | -0.052 | 0.082 | 0.097 | 0.259 | 0.651 | 0.701 | -0.268 | 0.652 | 0.704 | 0.009 | 0.08 | 0.08 |
| | $N_2$ | 0.025 | 0.046 | 0.053 | 0.177 | 0.603 | 0.627 | 0.189 | 0.533 | 0.565 | -0.006 | 0.047 | 0.048 |
| | $N_3$ | 0.01 | 0.034 | 0.036 | -0.153 | 0.541 | 0.561 | -0.175 | 0.436 | 0.470 | -0.003 | 0.031 | 0.030 |

Note: All simulations are based on the bivariate multi-fractal process with the whole number of cascade levels equal to 12, $j = 6$, $\rho = 0.5$, $\sigma_1 = 1$, $\sigma_2 = 1$, and eight moment conditions as in the Appendix are used. Sample lengths are $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$. Bias denotes the distance between the given and estimated parameter value, SD and RMSE denote standard deviation and root mean squared error, respectively. For each scenario, 400 Monte Carlo simulations have been carried out.

Table 3: GMM estimation for the trivariate multifractal (binomial) model

| $\hat{\theta}$ | Sub-sample Size | *Bias* | *SD* | *RMSE* |
|---|---|---|---|---|
| | $N_1$ | 0.097 | 0.128 | 0.161 |
| $\hat{m_0}$ | $N_2$ | 0.042 | 0.075 | 0.086 |
| | $N_3$ | -0.019 | 0.056 | 0.059 |
| | $N_1$ | 0.011 | 0.078 | 0.079 |
| $\hat{\sigma_1}$ | $N_2$ | -0.001 | 0.055 | 0.055 |
| | $N_3$ | -0.001 | 0.038 | 0.038 |
| | $N_1$ | 0.000 | 0.084 | 0.084 |
| $\hat{\sigma_2}$ | $N_2$ | 0.000 | 0.055 | 0.055 |
| | $N_3$ | -0.004 | 0.039 | 0.039 |
| | $N_1$ | 0.002 | 0.086 | 0.086 |
| $\hat{\sigma_3}$ | $N_2$ | -0.003 | 0.052 | 0.052 |
| | $N_3$ | 0.002 | 0.040 | 0.040 |
| | $N_1$ | 0.011 | 0.133 | 0.133 |
| $\hat{\rho_{12}}$ | $N_2$ | 0.000 | 0.102 | 0.102 |
| | $N_3$ | -0.009 | 0.085 | 0.085 |
| | $N_1$ | 0.014 | 0.124 | 0.124 |
| $\hat{\rho_{23}}$ | $N_2$ | 0.017 | 0.109 | 0.110 |
| | $N_3$ | -0.021 | 0.098 | 0.100 |
| | $N_1$ | -0.006 | 0.089 | 0.089 |
| $\hat{\rho_{13}}$ | $N_2$ | 0.011 | 0.073 | 0.074 |
| | $N_3$ | 0.009 | 0.056 | 0.057 |

Note: Simulations are based on the trivariate binomial multifractal process with $k = 12$, $j = 4$, and initial value $m_0 = 1.3$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\sigma_3 = 1$, $\rho_{12} = 0.3$, $\rho_{23} = 0.5$, $\rho_{13} = 0.7$. Sample lengths are $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$.

Table 4: GMM estimation for the trivariate multifractal (Lognormal) model

| $\hat{\theta}$ | Sub-sample Size | $Bias$ | $SD$ | $RMSE$ |
|---|---|---|---|---|
| | $N_1$ | -0.057 | 0.051 | 0.068 |
| $\hat{\lambda}$ | $N_2$ | 0.012 | 0.031 | 0.033 |
| | $N_3$ | 0.003 | 0.021 | 0.021 |
| | $N_1$ | 0.056 | 0.295 | 0.300 |
| $\hat{\sigma_1}$ | $N_2$ | -0.029 | 0.210 | 0.211 |
| | $N_3$ | -0.027 | 0.154 | 0.156 |
| | $N_1$ | -0.068 | 0.277 | 0.285 |
| $\hat{\sigma_2}$ | $N_2$ | -0.033 | 0.213 | 0.215 |
| | $N_3$ | -0.008 | 0.158 | 0.158 |
| | $N_1$ | -0.055 | 0.283 | 0.288 |
| $\hat{\sigma_3}$ | $N_2$ | -0.034 | 0.200 | 0.203 |
| | $N_3$ | -0.011 | 0.177 | 0.177 |
| | $N_1$ | 0.014 | 0.142 | 0.142 |
| $\hat{\rho_{12}}$ | $N_2$ | -0.018 | 0.101 | 0.102 |
| | $N_3$ | -0.029 | 0.073 | 0.078 |
| | $N_1$ | 0.020 | 0.088 | 0.088 |
| $\hat{\rho_{23}}$ | $N_2$ | -0.013 | 0.056 | 0.058 |
| | $N_3$ | -0.016 | 0.040 | 0.043 |
| | $N_1$ | 0.009 | 0.048 | 0.048 |
| $\hat{\rho_{13}}$ | $N_2$ | 0.016 | 0.027 | 0.031 |
| | $N_3$ | -0.019 | 0.021 | 0.029 |

Note: Simulations are based on the trivariate Lognormal multifractal process with $k = 12$, $j = 4$, and initial value $\lambda = 0.2$, $\sigma_1 = 1$, $\sigma_2 = 1$, $\sigma_3 = 1$, $\rho_{12} = 0.3$, $\rho_{23} = 0.5$, $\rho_{13} = 0.7$. Sample lengths are $N_1 = 2,000$, $N_2 = 5,000$ and $N_3 = 10,000$.

Table 5: GMM estimates of bivariate MF model ($Dow/Nik$) model

| | $j=1$ | $j=2$ | $j=3$ | $j=4$ | $j=5$ | $j=6$ | $j=7$ | $j=8$ | $j=9$ |
|---|---|---|---|---|---|---|---|---|---|
| Binomial model | | | | | | | | | |
| $\hat{m_0}$ | 1.224 | 1.210 | 1.292 | 1.281 | 1.274 | 1.268 | 1.267 | 1.260 | 1.261 |
| | (0.024) | (0.023) | (0.022) | (0.022) | (0.023) | (0.023) | (0.023) | (0.021) | (0.022) |
| $\hat{\sigma_1}$ | 1.032 | 1.110 | 1.105 | 1.106 | 1.101 | 1.110 | 1.109 | 1.107 | 1.107 |
| | (0.014) | (0.010) | (0.019) | (0.019) | (0.018) | (0.014) | (0.016) | (0.015) | (0.012) |
| $\hat{\sigma_2}$ | 1.007 | 0.981 | 0.980 | 1.044 | 1.003 | 1.002 | 0.982 | 0.986 | 0.998 |
| | (0.015) | (0.016) | (0.014) | (0.013) | (0.019) | (0.020) | (0.021) | (0.024) | (0.025) |
| $\hat{\rho}$ | 0.217 | 0.212 | 0.201 | 0.202 | 0.205 | 0.198 | 0.202 | 0.196 | 0.207 |
| | (0.013) | (0.012) | (0.012) | (0.011) | (0.015) | (0.013) | (0.013) | (0.010) | (0.010) |
| Lognormal model | | | | | | | | | |
| $\hat{\lambda}$ | 0.053 | 0.058 | 0.058 | 0.056 | 0.058 | 0.058 | 0.057 | 0.055 | 0.056 |
| | (0.015) | (0.016) | (0.015) | (0.014) | (0.015) | (0.015) | (0.015) | (0.017) | (0.016) |
| $\hat{\sigma_1}$ | 0.972 | 0.971 | 0.970 | 0.994 | 0.982 | 0.978 | 0.972 | 0.997 | 1.004 |
| | (0.027) | (0.025) | (0.025) | (0.026) | (0.026) | (0.026) | (0.026) | (0.027) | (0.028) |
| $\hat{\sigma_2}$ | 0.997 | 1.013 | 0.984 | 0.976 | 0.977 | 0.996 | 0.986 | 0.975 | 0.974 |
| | (0.017) | (0.018) | (0.017) | (0.017) | (0.017) | (0.020) | (0.019) | (0.025) | (0.023) |
| $\hat{\rho}$ | 0.199 | 0.194 | 0.191 | 0.194 | 0.202 | 0.211 | 0.204 | 0.197 | 0.198 |
| | (0.018) | (0.020) | (0.018) | (0.017) | (0.016) | (0.018) | (0.018) | (0.017) | (0.017) |

Note: Each column corresponds to the empirical estimate with different joint numbers of cascade level $j$ ($k = 10$); $\hat{d}$ is the mean of 100 simulated GPH estimators, and numbers in parenthesis are standard errors.

Table 6: GMM estimates of bivariate MF model ($US/DM \sim £$) model

| | $j=1$ | $j=2$ | $j=3$ | $j=4$ | $j=5$ | $j=6$ | $j=7$ | $j=8$ | $j=9$ | $j=10$ | $j=11$ |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Binomial model | | | | | | | | | | | |
| $\hat{m_0}$ | 1.534 | 1.551 | 1.560 | 1.555 | 1.550 | 1.543 | 1.540 | 1.564 | 1.550 | 1.557 | 1.560 |
| | (0.023) | (0.022) | (0.019) | (0.027) | (0.026) | (0.028) | (0.029) | (0.026) | (0.020) | (0.025) | (0.028) |
| $\hat{\sigma_1}$ | 0.643 | 0.591 | 0.605 | 0.652 | 0.635 | 0.609 | 0.637 | 0'495 | 0.607 | 0.613 | 0.614 |
| | (0.028) | (0.028) | (0.021) | (0.024) | (0.025) | (0.022) | (0.024) | (0.021) | (0.029) | (0.029) | (0.024) |
| $\hat{\sigma_2}$ | 0.496 | 0.487 | 0.472 | 0.484 | 0.482 | 0.524 | 0.539 | 0.505 | 0.524 | 0.518 | 0.572 |
| | (0.019) | (0.020) | (0.018) | ( 0.016) | ( 0.018) | (0.015) | (0.017) | (0.017) | (0.018) | (0.017) | (0.015) |
| $\hat{\rho}$ | 0.589 | 0.622 | 0.621 | 0.621 | 0.621 | 0.591 | 0.601 | 0.587 | 0.582 | 0.590 | 0.599 |
| | (0.012) | (0.017) | (0.016) | (0.015) | (0.012) | (0.014) | (0.012) | (0.013) | (0.014) | (0.013) | (0.013) |
| Lognormal model | | | | | | | | | | | |
| $\hat{\lambda}$ | 0.292 | 0.263 | 0.271 | 0.266 | 0.264 | 0.264 | 0.262 | 0.265 | 0.262 | 0.265 | 0.264 |
| | (0.024) | (0.040) | (0.037) | (0.035) | (0.035) | (0.034) | (0.029) | (0.020) | (0.022) | (0.020) | (0.025) |
| $\hat{\sigma_1}$ | 0.602 | 0.597 | 0.603 | 0.595 | 0.605 | 0.608 | 0.610 | 0.608 | 0.611 | 0.597 | 0.601 |
| | (0.018) | (0.019) | (0.024) | (0.016) | (0.013) | (0.013) | (0.009) | (0.015) | (0.012) | (0.016) | (0.014) |
| $\hat{\sigma_2}$ | 0.516 | 0.525 | 0.523 | 0.516 | 0.496 | 0.507 | 0.506 | 0.496 | 0.495 | 0.511 | 0.496 |
| | (0.021) | (0.025) | (0.016) | (0.019) | (0.026) | (0.025) | (0.023) | (0.014) | (0.015) | (0.017) | (0.017) |
| $\hat{\rho}$ | 0.629 | 0.620 | 0.621 | 0.616 | 0.621 | 0.627 | 0.628 | 0.631 | 0.629 | 0.628 | 0.633 |
| | (0.015) | (0.011) | (0.015) | (0.022) | (0.020) | (0.019) | (0.015) | (0.023) | (0.025) | (0.022) | (0.025) |

Note: Each column corresponds to the empirical estimate with different joint numbers of cascade level $j$ ($k = 12$); $\hat{d}$ is the mean of 100 simulated GPH estimators, and numbers in parenthesis are standard errors.

Table 7: GMM estimates of bivariate MF model $TB2/TB1$ model

| | $j=1$ | $j=2$ | $j=3$ | $j=4$ | $j=5$ | $j=6$ | $j=7$ | $j=8$ | $j=9$ |
|---|---|---|---|---|---|---|---|---|---|
| Binomial model | | | | | | | | | |
| $\hat{m_0}$ | 1.725 | 1.725 | 1.726 | 1.728 | 1.727 | 1.753 | 1.740 | 1.732 | 1.724 |
| | (0.025) | (0.020) | (0.024) | (0.020) | (0.028) | (0.023) | (0.023) | (0.027) | (0.030) |
| $\hat{\sigma_1}$ | 0.117 | 0.121 | 0.121 | 0.112 | 0.112 | 0.111 | 0.125 | 0.122 | 0.121 |
| | (0.019) | (0.015) | (0.017) | (0.010) | (0.013) | (0.014) | (0.012) | (0.019) | (0.017) |
| $\hat{\sigma_2}$ | 0.204 | 0.206 | 0.206 | 0.213 | 0.203 | 0.212 | 0.208 | 0.206 | 0.206 |
| | (0.021) | (0.016) | (0.022) | ( 0.024) | (0.023) | (0.021) | (0.026) | (0.024) | (0.025) |
| $\hat{\rho}$ | 0.820 | 0.821 | 0.820 | 0.819 | 0.816 | 0.818 | 0.818 | 0.821 | 0.822 |
| | (0.031) | (0.031) | (0.032) | (0.031) | (0.031) | (0.031) | (0.031 | (0.030) | (0.030)) |
| Lognormal model | | | | | | | | | |
| $\hat{\lambda}$ | 0.426 | 0.417 | 0.418 | 0.379 | 0.389 | 0.395 | 0.420 | 0.411 | 0.412 |
| | (0.023) | (0.021) | (0.022) | (0.022) | (0.024) | (0.025) | (0.030) | (0.028) | (0.028) |
| $\hat{\sigma_1}$ | 0.144 | 0.126 | 0.123 | 0.121 | 0.124 | 0.120 | 0.119 | 0.119 | 0.129 |
| | (0.034) | (0.031) | (0.031) | (0.031) | (0.030) | (0.031) | (0.029) | (0.033) | (0.034) |
| $\hat{\sigma_2}$ | 0.222 | 0.219 | 0.214 | 0.207 | 0.204 | 0.207 | 0.201 | 0.217 | 0.221 |
| | (0.034) | (0.030) | (0.027) | (0.024) | (0.024) | (0.022) | (0.020) | (0.021) | (0.021) |
| $\hat{\rho}$ | 0.918 | 0.917 | 0.912 | 0.906 | 0.893 | 0.871 | 0.901 | 0.881 | 0.895 |
| | (0.021) | (0.021) | (0.020) | (0.018) | (0.022) | (0.024) | (0.030) | (0.027) | (0.026) |

Note: Each column corresponds to the empirical estimate with different joint numbers of cascade level $j$ ($k = 10$); $\hat{d}$ is the mean of 100 simulated GPH estimators, and numbers in parenthesis are standard errors.

Table 8: Best linear forecast for simulated data

| | RMSE | | | AMSE | | | M2 | | |
|---|---|---|---|---|---|---|---|---|---|
| | $asset1$ | $asset2$ | $EW$ | $asset1$ | $asset2$ | $EW$ | $asset1$ | $asset2$ | $EW$ |
| 1 | 0.852 | 0.852 | 0.827 | 0.792 | 0.790 | 0.763 | 8.447 | 8.344 | 10.866 |
| 5 | 0.886 | 0.889 | 0.875 | 0.832 | 0.829 | 0.811 | 8.189 | 8.103 | 11.799 |
| 10 | 0.903 | 0.905 | 0.888 | 0.855 | 0.851 | 0.823 | 8.042 | 7.967 | 11.496 |
| 20 | 0.918 | 0.921 | 0.903 | 0.880 | 0.875 | 0.850 | 7.901 | 7.840 | 11.179 |
| 50 | 0.937 | 0.941 | 0.919 | 0.914 | 0.910 | 0.886 | 7.780 | 7.721 | 10.801 |
| 100 | 0.953 | 0.955 | 0.931 | 0.937 | 0.933 | 0.910 | 7.730 | 7.680 | 10.530 |

Note: This table shows the RMSE and RMAE for simulated data from the bivariate MF model (binomial case with parameters $k = 12$, $j = 6$, $\rho = 0.5$, $\sigma_1 = 1$, $\sigma_2 = 1$), 400 simulations and estimations have be conducted, with each simulation of 10000 realizations, first 5000 observations are used for estimation, the remaining 5000 observations used for out of sample forecast assessment. The forecast horizons are $1, 5, 10, 20, 50, 100$ days. EW denotes equal-weighted portfolio.

Table 9: Univariate MF model volatility forecast

| | | $DOW$ | $NIK$ | $US$ | $DM$ | $TB1$ | $TB2$ |
|---|---|---|---|---|---|---|---|
| RMSE | 1 | 1.000 | 0.789 | 0.883 | 0.924 | 0.857 | 0.784 |
| | 5 | 1.000 | 0.847 | 0.903 | 0.942 | 0.872 | 0.815 |
| | 10 | 1.001 | 0.888 | 0.920 | 0.937 | 0.902 | 0.845 |
| | 20 | 1.001 | 0.941 | 0.949 | 0.963 | 0.914 | 0.851 |
| | 50 | 1.001 | 0.972 | 0.978 | 0.989 | 0.950 | 0.873 |
| | 100 | 1.001 | 0.980 | 0.990 | 0.995 | 0.965 | 0.926 |
| | | $DOW$ | $NIK$ | $US$ | $DM$ | $TB1$ | $TB2$ |
| AMSE | 1 | 0.989 | 1.055 | 0.927 | 0.865 | 0.934 | 0.699 |
| | 5 | 0.989 | 1.045 | 0.925 | 0.870 | 0.968 | 0.720 |
| | 10 | 0.989 | 1.047 | 0.930 | 0.875 | 1.030 | 0.806 |
| | 20 | 0.989 | 1.054 | 0.942 | 0.891 | 1.052 | 0.844 |
| | 50 | 0.989 | 1.036 | 0.955 | 0.911 | 1.071 | 0.876 |
| | 100 | 0.989 | 1.022 | 0.970 | 0.924 | 1.109 | 0.912 |

Note: This table shows the volatility forecast by using the univariate MF model. Stocks are Dow Jones composite 65 average index (DOW) and NIKKEI 225 stock average index (NIK); FXs are foreign exchange rate of U.S. Dollar (US) and German Mark (DM) to British Pound; Bonds are the U.S. 1-year and 2-year treasury constant maturity rate (TB1, TB2 respectively).

Table 10: Bivariate MF model volatility forecast

| | | $DOW$ | $NIK$ | $EW$ | $US$ | $DM$ | $EW$ | $TB1$ | $TB2$ | $EW$ |
|---|---|---|---|---|---|---|---|---|---|---|
| RMSE | 1 | 0.809 | 0.754 | 0.791 | 0.889 | 0.909 | 0.889 | 0.685 | 0.793 | 0.687 |
| | 5 | 0.831 | 0.819 | 0.839 | 0.898 | 0.915 | 0.913 | 0.693 | 0.815 | 0.694 |
| | 10 | 0.863 | 0.852 | 0.860 | 0.914 | 0.927 | 0.921 | 0.712 | 0.835 | 0.701 |
| | 20 | 0.902 | 0.916 | 0.925 | 0.941 | 0.949 | 0.955 | 0.781 | 0.863 | 0.719 |
| | 50 | 0.975 | 0.955 | 0.968 | 0.970 | 0.973 | 0.986 | 0.832 | 0.894 | 0.733 |
| | 100 | 0.994 | 0.970 | 0.973 | 0.983 | 0.999 | 0.996 | 0.908 | 0.929 | 0.756 |
| | | $DOW$ | $NIK$ | $EW$ | $US$ | $DM$ | $EW$ | $TB1$ | $TB2$ | $EW$ |
| AMSE | 1 | 0.903 | 0.904 | 1.057 | 0.869 | 0.868 | 0.864 | 0.716 | 0.618 | 0.766 |
| | 5 | 0.906 | 0.926 | 1.068 | 0.875 | 0.874 | 0.871 | 0.754 | 0.669 | 0.792 |
| | 10 | 0.919 | 0.964 | 1.079 | 0.882 | 0.896 | 0.872 | 0.787 | 0.711 | 0.835 |
| | 20 | 0.934 | 0.997 | 1.092 | 0.892 | 0.912 | 0.884 | 0.811 | 0.766 | 0.878 |
| | 50 | 0.963 | 1.031 | 1.095 | 0.908 | 0.927 | 0.902 | 0.851 | 0.838 | 0.896 |
| | 100 | 0.980 | 1.045 | 1.096 | 0.922 | 0.944 | 0.912 | 0.924 | 0.896 | 0.938 |
| | | $DOW$ | $NIK$ | $EW$ | $US$ | $DM$ | $EW$ | $TB1$ | $TB2$ | $EW$ |
| $M2$ | 1 | 13.192 | 1.939 | 12.039 | 27.122 | 33.825 | 38.171 | -53.057 | 40.812 | 60.736 |
| | 5 | 12.942 | 1.994 | 12.553 | 26.892 | 32.911 | 39.426 | -73.449 | 37.702 | 59.274 |
| | 10 | 12.809 | 2.034 | 12.589 | 26.616 | 32.116 | 39.484 | -84.713 | 35.307 | 55.060 |
| | 20 | 12.685 | 2.084 | 12.657 | 26.239 | 31.065 | 39.687 | -96.659 | 32.254 | 50.128 |
| | 50 | 12.674 | 2.138 | 12.762 | 25.496 | 29.123 | 40.012 | -106.658 | 25.778 | 43.374 |
| | 100 | 12.710 | 2.172 | 12.680 | 24.847 | 27.549 | 39.718 | -119.272 | 21.876 | 38.018 |

Note: This table shows
Stocks are Dow Jones composite 65 average index (DOW) and NIKKEI 225 stock average index (NIK); FXs are foreign exchange rate of U.S. Dollar (US) and German Mark (DM) to British Pound; Bonds are the U.S. 1-year and 2-year treasury constant maturity rate (TB1, TB2 respectively). EW denotes equal-weighted portfolio.