

多分形波动率预测模型及其 MCS 检验^①

魏 宇, 马 锋, 黄登仕

(西南交通大学经济管理学院, 成都 610031)

摘要: 以上证综指的 5 min 高频数据为例, 在已有的多分形波动率(multifractal volatility)测度方法基础上, 提出了新的波动率测度方法及模型. 运用滚动时间窗的样本外预测技术以及比 SPA 检验更具优势的“模型信度设定检验”(model confidence set, MCS), 对比了新的波动率测度模型和主流的 GARCH 族以及已实现波动率(realized volatility)模型的预测精度. 实证结果显示: 不论是短记忆模型还是长记忆模型, 多分形波动率模型的预测精度明显优于 GARCH 族模型, 且长记忆模型的预测能力要好于短记忆模型. 同时, 在多数损失函数下, 新提出的多分形波动率测度方法及其动力学模型的预测效果都是最优的.

关键词: 多分形波动率; GARCH 族模型; 已实现波动率模型; 滚动预测; MCS 检验

中图分类号: C93; F224 文献标识码: A 文章编号: 1007-9807(2015)08-0061-12

0 引 言

在金融市场中, 分形(fractal)被认为是个“典型事实”(stylized fact), 因多数学者已经发现金融市场存在显著且复杂的分形特征^[1-4]. 有趣的是, 分形不仅存在于成熟股票市场, 而且还存在于新兴股票市场, 但成熟市场的分形强度要弱于新兴市场^[5]. 值得一提的是, 对中国股票市场的分形特征研究, 所得结论都基本一致, 即不论是深市还是沪市, 都存在显著的分形特征. 早在 1997 年, “分形理论之父”Mandelbrot 就在其著名论著《Fractals and Scaling in Finance》中指出了分形理论在金融领域的广阔前景^[6]. 之后, Mandelbrot 又在世界权威的科学杂志《Scientific American》上撰文指出, 多分形理论是定量刻画金融市场各种复杂波动特征的有力工具, 且与一般的单分形描述相比, 多分形理论的工具和方法在金融市场中具有更强的实用性^[7].

另一方面, 对波动率的精确预测是控制金融

市场风险的有效手段, 且 Poon 和 Granger^[8]发现其在资产组合、期权定价以及市场监管方面等扮演着至关重要的角色. 早期的波动率预测研究多集中在 GARCH 族模型, 随着 GARCH 族模型的扩展, 它变得更为复杂, 且在波动率预测方面取得了一定的成效. 然而, GARCH 族模型并不能非常好地解释和预测金融市场中那些令人捉摸不透的复杂波动行为. 考虑到多分形是被用来刻画复杂对象非均匀和各向异性特征的有力工具, 因而近年来有学者基于分形的思想, 提出了多分形波动率(multifractal volatility)测度方法, 并且进行了相关的研究探讨, 发现了基于多分形波动率的模型在预测能力方面明显好于 GARCH 族模型和随机波动模型(SV)^[9]. 让人感兴趣的是, 多分形波动率预测模型不仅仅用于波动率的预测, 还有学者发现可将其运用到权证定价^[10]和风险管理等领域^[11-14]. 由此可见, 对多分形波动率的研究具有较为重要的理论和实践意义.

① 收稿日期: 2012-12-10; 修订日期: 2013-09-02.

基金项目: 国家自然科学基金资助项目(71071131; 71371157); 高等学校博士学科点专项科研基金资助项目(20120184110020); 教育部人文社科基金规划资助项目(14YJC790073); 四川省科技青年基金资助项目(2015JQ0010).

作者简介: 魏宇(1975—), 男, 四川攀枝花人, 博士, 教授. Email: weiyusy@126.com

在多分形波动率的测度方法上,本文基于已有研究^[12,15-16]提出了新的测度方法,而这种测度方法与现有研究的不同之处在于以下两点:第1,本文的测度方法中利用了每天多分形谱奇异指数 α 的标准差信息,而非极差 $\Delta\alpha$ 。这一做法主要是考虑到极差 $\Delta\alpha$ 不能充分反映价格波动中不同概率变化的全部信息;其次,本文利用“已实现波动率测度”(realized volatility measure, RV)而非“平方收益率”(squared returns)来构建多分形波动率的尺度参数。因为 Andersen 和 Bollerslev^[17]和 Barndorff-Nielsen 和 Shephard^[18]发现已实现波动率比日收益率的平方具有更少的交易噪音。进一步, Ashely 等^[19], Lo 和 Mackinlay^[20]及 Foster 等^[21]发现样本内拟合优度的检验往往会受到数据挖掘偏误的影响,而样本外的滚动预测方法则可以规避此类偏误所造成的问题。因此,本文结合样本内估计和样本外的滚动时间窗预测法,对现有各种多分形波动率模型的预测能力进行了评价。为了保证论文结论的稳健性,还利用了 Hansen 等^[22]最近提出的“模型信度设定检验法”(model confidence set, MCS)对比了新旧多分形波动率模型和主流的 GARCH 族以及已实现波动率(RV)模型的预测精度。相比 Hansen 在 2005 年提出的“高级预测能力检验法”(superior predictive ability, SPA)^[23], MCS 检验法的条件更为宽松,与 SPA 检验主要的区别是它不需要选择基础模型,因此其检验结果更具稳健性。由于以往关于多分形波动率模型的检验多用 SPA 方法,因而本文采用的 MCS 方法丰富了对波动率模型预测能力的检验手段。

1 波动率测度方法及其计量模型

1.1 已实现波动率测度(realized volatility, RV)及其计量模型

Andersen 和 Bollerslev^[17]发现传统研究中多采用日收益率的平方作为日波动率的代理变量将面临非常严重的测量误差和噪声,而基于交易日内高频收益数据的已实现波动率则可以有效地降低这些误差和噪声对真实潜在的波动率过程的影响。Andersen 等^[24]的最近研究表明,由于潜在真

实的波动率是不可观察的,因而学术界普遍利用已实现波动率作为市场波动率的代理变量。根据 Andersen 和 Bollerslev^[17]对已实现波动率的定义,对第 t 天的已实现波动率估计表示为第 t 天内高频数据的收益率平方和。

首先,将第 t 天的日内高频价格数据表示为 $P_{t,m}$,其中 $m = 1, 2, \dots, n$,而 P_t 为第 t 天的收盘价,则第 t 天的第 m 个高频收益率表示为

$$R_{t,m} = 100(\ln P_{t,m} - \ln P_{t,m-1}) \quad (1)$$

第 t 天的日收益率表示为

$$R_t = 100(\ln P_t - \ln P_{t-1}) \quad (2)$$

第 t 天的已实现波动率测度(RV_t^*)则定义为

$$RV_t^* = \sum_{m=1}^n R_{t,m}^2 \quad (3)$$

但最近 Hansen 和 Lunde^[25]的研究又指出,由于股票现货与期货市场并不像外汇市场那样在 24 h 连续进行交易,因此,能观察和记录到的高频价格数据只能反映有交易时段的(active)市场波动状况,而无法包含无交易时段的(inactive)市场波动信息(即市场从第 1 天收盘到第 2 天开盘的所谓“close-to-open”波动率)。因此为了使已实现波动率的估计能更加准确地刻画真实市场波动幅度的大小,这里采用 Hansen 和 Lunde^[25]的建议,用某种尺度参数(scale parameter) λ_0 来对 RV_t^* 进行尺度变换,即对第 t 天的已实现波动率估计为

$$RV_t = \lambda_0 RV_t^* \quad (4)$$

其中,在时间段 N 内的尺度参数 λ_0 为

$$\lambda_0 = \frac{N^{-1} \sum_{t=1}^N R_t^2}{N^{-1} \sum_{t=1}^N RV_t^*} \quad (5)$$

其次,Andersen 等^[26]的研究进一步发现,在对已实现波动率取对数后(标记为 $\ln RV$),其波动特征可以用高斯动力学过程来描述,且 $\ln RV$ 序列展现出非常明显的长期记忆性特征。这一结论在魏宇^[9]、王鹏和魏宇^[14]、王鹏和王建琼^[15]对我国股市的实证研究中同样得到了证实。因此,Andersen 等^[26]还建议采用“自回归分整移动平均(ARFIMA)模型”来描述 $\ln RV$ 的上述动力学特征。由于已实现波动测度对抽样频率敏感(频率越高,噪声影响越大),因此需要确定抽样频

率. Andersen 和 Bollerslev^[17] 证实了外汇交易波动能很好地用 288 个每 5 min 取样的数据的变化平方和来估计, 并进一步发现 5 min 的取样频率可在兼顾精确性和减少微观噪声之间取得平衡. 另外, 我国学者^[27-28] 以上证高频数据为例, 也证实了 5 min 抽样频率数据的合理性. 为此, 本文仍然以 5 min 高频数据为样本, 对已实现波动率 (RV) 及后面提及的多分形波动率 (MFV) 进行建模和分析.

根据 AIC、SC 法则、极大似然值以及模型“简约型”的要求, 本文对 $LnRV$ 建立了模型 ARFIMA(1, d , 1). ARFIMA(p , d , q) 模型的一般形式表示为

$$\varphi(L) (1-L)^d (Y-u) = \Theta(L) \varepsilon_t \quad (6)$$

$$\varphi(L) = 1 - \varphi_1 L - \varphi_2 L^2 - \dots - \varphi_p L^p \quad (7)$$

$$\Theta(L) = 1 + \theta_1 L + \theta_2 L^2 + \dots + \theta_q L^q \quad (8)$$

其中, 式(7)、式(8) 分别为自回归滞后 p 阶算子和移动平均滞后 q 阶算子 L 为滞后算子, $(1-L)^d$ 为分数差分算子, μ 是 Y 的均值, 同时这里假定 $\varepsilon_t \sim NID(0, \sigma_\varepsilon^2)$. 本文对利用 $LnRV$ 构建的 ARFIMA 模型标记为 ARFIMA- $LnRV$. 由于模型 ARFIMA- $LnRV$ 着重刻画波动的长记忆性, 基于稳健性考虑, 在本文的实证研究部分也考虑了具有短记忆特征的 ARMA(1, 1)- $LnRV$ 模型.

1.2 多分形波动率测度 (multifractal volatility, MFV) 及其计量模型

本文对多分形波动率测度 (MFV) 的计算方法采用常用的“数盒子”法, 具体计算步骤如下: 首先, 通过一天当中的高频价格数据来计算当天的市场波动多分形谱 (记为 $f(\alpha)$) , 其次再从多分形谱 $f(\alpha)$ 中计算 MFV.

第 1 步, 假设每个交易日的的时间长度为标准化的 1, 则无重复均匀覆盖每天 48 个高频价格数据^②的“盒子长度”分别取: $1/48, 1/24, 1/16, 1/12, 1/8, 1/4, 1/3, 1/2$ 和 1 (对于不同的市场, “盒子长度”可能不一样).

第 2 步, 当取盒子长度为 δ 时, 假定覆盖每天 48 个高频股价数据需要 m 个盒子 (这里为了公式表述的清晰, 另记一天当中的高频股价数据为

$I(t) \quad t = 1, 2, \dots, 48$), 且每个盒子内有 n 个数据记录, 那么定义在第 i 个盒子上的指数概率测度为

$$P_i(\delta) = \frac{\sum_{j=1}^n I(i_j)}{\sum_{t=1}^{48} I(t)}, \quad i = 1, 2, \dots, m \quad (9)$$

其中 $I(i_j)$ 表示第 i 个盒子中的第 j 个指数. 根据文献 [27] 等可知, 有如下幂律关系存在

$$P_i(\delta) \sim \delta^\alpha \quad (10)$$

$$N_\alpha(\delta) \sim \delta^{-f(\alpha)} \quad (11)$$

其中 $N_\alpha(\delta)$ 表示具有相同的奇异指数 α 的长度为 δ 的盒子个数, 而多分形谱 $f(\alpha)$ 为 α 子集的分形维数. 由文献 [29] 可知, $f(\alpha)$ 和 α 可以通过满足幂律关系的“配分函数” $S_q(\delta)$ 来计算, 即

$$S_q(\delta) = \sum_{i=1}^m P_i^q(\delta) \sim \delta^{\tau(q)} \quad (12)$$

可以发现, 当 q 取正数时, q 值越大, 则 $S_q(\delta)$ 将主要反映的是那些具有大的概率测度盒子的信息; 反之, 当 q 取负数时, q 值越小, 则 $S_q(\delta)$ 将主要反映的是那些具有小概率测度盒子的信息. $\tau(q)$ 的值可以通过求取在对数坐标轴上 $\ln S_q(\delta) - \ln(\delta)$ 的直线斜率得出, 同时通过 Legendre 变换可以得到^[30-31]

$$\alpha = \frac{d\tau(q)}{dq} \quad (13)$$

$$f(\alpha) = \alpha q - \tau(q) \quad (14)$$

通过以上步骤, 可以计算出每天的多分形谱极差 $\Delta\alpha = \alpha_{\max} - \alpha_{\min}$, 其中 α_{\max} 和 α_{\min} 分别代表最小概率测度和最大概率测度盒子的测度值. 因而 $\Delta\alpha$ 越大, 则表明当天价格走势的分布越不均匀, 当天价格的波动幅度越大. 换句话说, $\Delta\alpha$ 是能够测度市场日波动率大小的定量描述指标. 基于这一方法, 有学者构建了多分形波动率的测度指标 (以下记为 MFV_t)^[12]

$$MFV_t = \lambda_1 \Delta\alpha_t \quad (15)$$

$$\lambda_1 = \frac{N^{-1} \sum_{t=1}^N R_t^2}{N^{-1} \sum_{t=1}^N \Delta\alpha_t} \quad (16)$$

② 我国股市每天共有 4 h (即 240 min) 连续竞价交易时间, 因此, 采用每 5 min 记录一个数据的方法每天可以产生 48 个高频股价记录. 对不同股市的计算方法同理.

进一步有学者认为^[14-15],极差 $\Delta\alpha$ 仅运用了一天中 α 值的两个记录(即最大值和最小值),并不能充分反映价格波动中不同概率变化的全部信息,即不能充分反映价格波动的真实大小.他们的研究发现 $\Delta\alpha$ 的分布特征极易受到极端值的影响,因而他们在构建多分形波动率时利用的是 α 值的标准差(记为 S_α).为了清晰起见,本文将这种多分形波动率测度记为 $MFVW_t$,其定义如下

$$MFVW_t = \lambda_2 S_\alpha \tag{17}$$

$$\lambda_2 = \frac{N^{-1} \sum_{t=1}^N R_t^2}{N^{-1} \sum_{t=1}^N S_{\alpha,t}} \tag{18}$$

最近,由于研究发现已实现波动率比日收益率的平方具有更少的噪音,Chen和Wu^[16]利用已实现波动率构建了新的多分形波动率测度,记为 MV_t ,其定义为

$$MV_t = \frac{\Delta \alpha_t}{E(\Delta \alpha_t)} E(RV_t) \tag{19}$$

从式(19)可以看出,该多分形波动率测度与式(15)定义的 MFV_t 测度方法的不同在于其对尺度参数的构建依据为已实现波动率,而非收益率的平方.

结合上述已有研究,本文提出新的多分形波动率测度方法.该测度方法中不仅包含了多分形谱指数 α 的标准差信息,并且在尺度参数的构建上采用了更具优势的已实现波动率测度.换句话说,相比已有的多分形波动率测度,本文新提出的多分形波动率所包含的信息含量更为丰富,因此相比其他3种多分形波动率模型,在理论上应是最优的.这里将该测度方法标记为 MVM_t ,其定义如下

$$MVM_t = \lambda_3 S_{\alpha,t} \tag{20}$$

$$\lambda_3 = \frac{N^{-1} \sum_{t=1}^N RV_t}{N^{-1} \sum_{t=1}^N S_{\alpha,t}} \tag{21}$$

从式(21)中不难发现,本文以已实现波动率和多分形谱指数 α 的标准差 S_α 作为尺度参数的计算基准,而这两个指标可从不同的角度反应出日内(intraday)股票价格的变化特征.比方说,假如利用日收益率的平方作为尺度参数的计算基

准,即使日内价格波动较大,也不能在尺度参数中得以表现,而已实现波动率可很好地解决该问题.值得一提的是,相比日收益的平方,已实现波动率具有更小的交易噪声.另外,由于不同的 α 代表了价格变化中不同概率子集的测度值,而具有极差意义的 $\Delta\alpha$ ($\Delta\alpha = \alpha_{\max} - \alpha_{\min}$) 只运用了一天中不同的 α 值的两个记录(即 α_{\max} 和 α_{\min}),因此 $\Delta\alpha$ 并不能充分反映价格波动中不同概率变化的全部信息.进一步,与极差统计量本身所具有的特点一样, $\Delta\alpha$ 不仅忽视了 α 值的分布特征,且极易受到极端值(extreme value)的影响,因此在衡量当天价格波动程度时不具稳健性.且与极差统计量相比,标准差是更为常用的衡量数据变化特征的统计指标,该指标本身具有信息全面性、稳健性等优点,而这些优点正是极差统计量所欠缺的.由此可见,本文提出的新的多分形波动率测度方法同时具备了这两种参数的优势,因此相比已有的多分形波动率测度而言,其在理论上应是更具优势的.

进一步,为了使研究结论更加稳健,本文既采用了长记忆的 ARFIMA 模型,又使用了短记忆的 ARMA 模型来刻画多分形波动率的动力学特征,依次表示为 ARFIMA-LnMFV($LnMFVW$ 、 $LnMV$ 以及 $LnMVM$) 和 ARMA-LnMFV($LnMFVW$ 、 $LnMV$ 以及 $LnMVM$).另外,由于 GARCH 族模型是目前学术界和实务界运用较为广泛的波动率预测模型,为了更好地展示和突出本文提出的新的多分形波动率模型预测效果,进一步引入了 GARCH 族模型中常见的短记忆模型: GARCH、GJR-GARCH 和长记忆模型 FIEGARCH、HYGARCH 与之进行对比研究.最后,除了 GARCH 族模型之外,本文在对比模型中也加入了被众多实证研究证明表现优异的高频数据波动率模型——ARFIMA-LnRV 和 ARMA-LnRV^[32-33].综上所述,本文对比的长记忆模型和短记忆模型各有7个,共计14种不同类型的波动率模型.

2 波动率预测方法及 MCS 检验

2.1 波动率预测方法说明

在本文的实证研究中,对上面讨论的14种模型进行了滚动时间窗(rolling time windows)的

“样本外预测能力检验”(test for out-of-sample predicting ability). 本文的预测方法具体如下:

1) 第 1 步 将数据样本总体($t = 1, 2, \dots, N$, $N = 2019$) 划分为“估计样本”(sample for estimation) 和“预测样本”(sample for forecasting) 两部分. 其中, 估计样本包含 $H = 1819$ 个交易日的数据, 而预测样本包含最后 200 个交易日的数据(即 $t = H + 1, H + 2, \dots, H + M$, 其中 $M = 200$)^③.

2) 第 2 步 选取 $t = 1, 2, \dots, H$ 的数据作为第 1 次的估计样本, 分别对上述各种波动率模型的参数进行估计, 然后在此估计基础之上, 运用递推法获得未来一天的波动率预测, 记为 \hat{h}_{H+1}^2 . 也就是说 \hat{h}_{H+1}^2 是在前面 1819 个样本数据的模型估计基础上对第 1820 天的市场波动率预测.

3) 第 3 步 保持估计样本的时间区间长度不变($H = 1819$) 将估计时间区间向后平行移动 1 天, 即第 2 次选取的是 $t = 2, 3, \dots, H + 1$ 的数据作为新的估计样本, 然后重新估计上述各类波动率模型的参数, 并在此新的估计模型基础上获得未来一天的市场波动率预测值, 记为 \hat{h}_{H+2}^2 .

4) 第 4 步 同理, 不断重复步骤 3, 可以得到 $\hat{h}_{H+3}^2, \hat{h}_{H+4}^2, \dots$, 直到最后一次的估计样本区间为 $t = M, M + 1, \dots, H + M - 1$, 以获得对最后一天, 即第 $t = N = H + M$ 天的市场波动率预测 \hat{h}_{H+M}^2 .

简言之, 对前面所讨论的每一种波动率模型, 都分别重复进行了 200 次的模型参数估计(共计 $14 \times 200 = 2800$ 次不同的估计), 从而每个模型都获得了 200 个未来一天的样本外市场波动率估计, 记为 \hat{h}_m^2 , $m = H + 1, H + 2, \dots, H + M$. 同时, 记预测样本区间的已实现波动率估计为 RV_m , $m = H + 1, H + 2, \dots, H + M$. 这里对 RV_m 的估计方法来自于 1.1 节说明, 并以此作为真实市场波动率的代理变量, 用以衡量各类波动率模型的预测精度.

2.2 模型信度设定(model confidence set, MCS)

有了以上所讨论的 14 种波动率模型及其对市场波动的预测值 \hat{h}_m^2 以后, 就可以比较这些预测值与真实市场波动率估计基准—— RV_m 的偏差

(或损失) 究竟有多大. 然而需要说明的是, 到目前为止, 学术界还不清楚用哪一种损失函数(loss function) 作为衡量预测偏差的标准最为合理. 因此, Hansen 和 Lunde^[25] 建议, 可以尽可能多地采用不同形式的损失函数作为预测模型精度的判断标准. 基于这样的考虑, 在本文的实证研究中采用了 6 种不同的损失函数分别作为各类波动率模型预测精度的评判标准.

这 6 种损失函数分别记为 L_i ($i = 1, 2, \dots, 6$), 其中 L_1 和 L_2 分别称为平均误差平方(mean squared error, MSE) 和平均绝对误差(mean absolute error, MAE), 它们是此类判断中最常用的两类损失函数. L_3 和 L_4 分别是经异方差调整的 MSE 和 MAE, 而限于篇幅, 对 L_5 和 L_6 的具体含义可以参考文献[25]的讨论. 各损失函数的具体定义如下:

$$L_1: MSE = M^{-1} \sum_{m=H+1}^M (RV_m - \hat{h}_m^2)^2 \quad (22)$$

$$L_2: MAE = M^{-1} \sum_{m=H+1}^M |RV_m - \hat{h}_m^2| \quad (23)$$

$$L_3: HMSE = M^{-1} \sum_{m=H+1}^M \left(1 - \frac{\hat{h}_m^2}{RV_m}\right)^2 \quad (24)$$

$$L_4: HMAE = M^{-1} \sum_{m=H+1}^M \left|1 - \frac{\hat{h}_m^2}{RV_m}\right| \quad (25)$$

$$L_5: QLIKE = M^{-1} \sum_{m=H+1}^M \left(\ln \hat{h}_m^2 + \frac{RV_m}{\hat{h}_m^2}\right) \quad (26)$$

$$L_6: R^2 LOG = M^{-1} \sum_{m=H+1}^M \left(\frac{\ln(RV_m)}{\hat{h}_m^2}\right)^2 \quad (27)$$

需要指出的是, 如果在一次实证研究中发现: 采用某种 L_i 作为判断标准, 得到了模型甲比模型乙的预测损失值小的话, 那么只能判断“在这样一个特定的数据样本中, 采用这一特定的损失函数 L_i 时, 模型甲比模型乙的预测精确度高”. 很明显, 这一判断是不稳健的, 且无法推广到其它类似的数据样本或者其它的损失函数判断标准. 比方说, 数据样本中的少数奇异点往往会严重影响损失函数的计算结果, 从而引起损失函数值的异常增加, 进而可能导致对波动率模型优劣的错误

③ 选择不同长度的估计样本和预测样本的实证结果基本一致, 如果需要, 作者可以提供不同样本长度上的模型检验结果.

判断^[25]。

为了解决这一问题, Hansen 和 Lunde^[23] 提出了所谓的“高级预测能力检验法”(superior predictive ability SPA), 然而 SPA 检验需要选择基准模型, 会导致产生“与对照组的多重比较”(multiple comparisons with controlling set) 问题。为了克服这一缺陷, Hansen 等^[22] 最近又提出了名为“模型信度设定检验法”(model confidence set MCS) 的模型比较方法。与 SPA 检验相比, MCS 检验具有很多明显的优势。比方说, 由于 SPA 检验的零假设是复合假设(composite hypothesis), 它会影响到 SPA 检验统计量的渐进分布, 从而会产生“厌恶参数”(nuisance parameters) 的问题, 进而难以控制第一类错误(或称“拒真”) 概率。而 MCS 检验则不用进行此类复合假设检验, 从而可以大大减小第一类错误发生的概率。本文采用这种新的 MCS 方法来比较众多波动率模型对我国股市波动的预测精度, 从而其结论应该具有更好的稳健性。

MCS 检验的实现过程如下: 首先, 假定存在 m_0 个波动率预测模型(就本文而言 $m_0 = 14$) 这些模型都在一个集合中, 该集合记为 M_0 , 则有 $M_0 = \{1, 2, \dots, m_0\}$ 。每个预测模型都可获得 M 个未来一天的样本外市场波动率估计值 \hat{h}_m^2 ($m = H+1, H+2, \dots, H+M$)。对每个预测值, 都可按照本文给出的 6 种损失函数 (L_1, L_2, \dots, L_6) 计算出相应的损失函数值, 记为 $L_{i,j,m}$ ($i = 1, 2, 3, 4, 5, 6; j = 1, 2, \dots, \mu, \nu, \dots, m_0; m = H+1, H+2, \dots, H+M$)。因此, 对于 M_0 中任意两个波动率预测模型 u, v ($u, v \in M_0$), 都可以计算出相对应的波动率预测值的相对损失函数值, 记为 $d_{i,\mu,\nu}$, 其定义如下

$$d_{i,\mu,\nu} = L_{i,\mu,m} - L_{i,\nu,m} \quad (28)$$

其次, 定义“高级对象的集合”(the set of superior objects) 为 M^* , 则 M^* 的表达式为

$$M^* \equiv \{u \in M_0 : E(d_{i,\mu,\nu}) \leq 0 \text{ 对所有 } v \in M_0\} \quad (29)$$

MCS 检验过程是在集合 M_0 中进行一系列的显著性检验, 剔除集合 M_0 中预测能力较差的模型。因此, 在每一次检验中, 零假设都是两个模型具有相同的预测能力, 即为

$$H_{0,M} : E(d_{i,\mu,\nu}) = 0, \quad (30)$$

对所有 $u, v \in M \subset M_0$

MCS 检验的过程是依据等价检验(equivalence test) δ_M 和剔除准则(elimination rule) e_M 。其中, 对于任何 $M \subset M_0$, 等价检验 δ_M 都是用来检验零假设, 而剔除准则 e_M 是用来剔除拒绝零假设的模型。因此, MCS 算法步骤为: 第 1 步, 设定 $M = M_0$; 第 2 步, 在显著水平 α 下, 利用等价检验 δ_M 检验零假设 $H_{0,M}$; 第 3 步, 如果接受零假设, 则定义 $\hat{M}_{1-\alpha}^* = M$, 否则利用剔除准则 e_M 将拒绝零假设的模型从 M 中剔除, 这一过程一直持续到不再出现拒绝零假设的情况, 最后得到 MCS 检验下的幸存模型(surviving objects)。集合 $\hat{M}_{1-\alpha}^*$ 包含了在 $1 - \alpha$ 的置信水平下的最优预测模型。如果对于给定的模型 k ($k \in M_0$), 该模型属于集合 $\hat{M}_{1-\alpha}^*$ 的条件是其 MCS 检验的 p 值大于显著水平 α 。换句话说, p 值越大, 表明该模型得到的波动率预测结果越精确。

相比 SPA 而言, MCS 的检验统计量更为复杂。考虑到篇幅, 本文仅介绍较为常用的范围统计量(range statistic) 和半二次方统计量(semi-quadratic statistic)^④, 其定义分别为

$$T_R = \max_{u, v \in M} \frac{|\bar{d}_{i,\mu,\nu}|}{\sqrt{\text{var}(d_{i,\mu,\nu})}},$$

$$T_{SQ} = \max_{u, v \in M} \frac{(\bar{d}_{i,\mu,\nu})^2}{\text{var}(\bar{d}_{i,\mu,\nu})}, \quad (31)$$

$$\bar{d}_{i,\mu,\nu} = \frac{1}{M} \sum_{m=H+1}^{H+M} d_{i,\mu,\nu}$$

其中, $\bar{d}_{i,\mu,\nu}$ 表示模型 u 和 v 波动率预测值的相对损失函数值的平均值。如果统计量 T_R, T_{SQ} 大于给定的临界值, 表明拒绝零假设。由于统计量 T_R 和 T_{SQ} 的渐近分布依赖于“厌恶参数”(nuisance parameters), 因此它们的真实分布非常复杂。然而, 统计量 T_R, T_{SQ} 及相应的 p 值可以通过采用所

④ 除以上两种方法外, 还有最大偏差(maximum deviation), 偏差(deviation from common average), 最大 t 值(Max t) 统计量等, 具体内容可以参考 Hansen 和 Lunde^[22] 的相关论文。

谓的 bootstrap(“自举法”)来获得^⑤。

3 实证结果

3.1 数据样本说明及统计描述

本文研究的数据样本为上证综指(SSEC)从2004-01-01至2012-04-30的每5min高频数据(共 $N = 2\,019$ 个交易日),数据来源于“中国经济研究中心(CCER)-色诺芬股票市场高频数据库”。上海证券交易所每个交易日有4h(即240min)连续竞价交易时间,因此,采用每5min记录1个数据的方法每天可以产生48个高频股价记录,样本总体的高频数据量为96 912个。

表1为论文研究中各时间序列的描述性统计结果。从表1可以看到,无论是已实现波动率本身,还是各种波动率的测度及其对数序列,都表现出显著的“有偏”(skewed)和“尖峰”(leptokurtic)形态,且在滞后5、10和20期内都具有明显的自相关特征(Q 统计量都很显著)。本文还利用R/S分析方法对各波动率序列进行了长记忆性检验,其Hurst指数计算结果也表明各波动率序列存在明显的长记忆性特征。因此,可以认为已实现波动率和各种多分形波动率及其对数序列都存在较为显著的长记忆(long memory)特征。另外,表中ADF单位根检验结果表

明,各序列都显著拒绝了存在单位根的原假设。因此,可以认为各序列都是平稳(stationary)时间序列,进而可以直接作下一步的分析和计量建模。

3.2 基于已实现波动率(RV)的各模型预测精度检验

本文检验了7种长记忆性模型:ARFIMA-LnMFV、ARFIMA-LnMV、ARFIMA-LnMFVW、ARFIMA-LnMVM、ARFIMA-LnRV、HYGARCH及FIGARCH(依次记为LM1到LM7)和7种短记忆性模型:ARMA-LnMFV、ARMA-LnMV、ARMA-LnMFVW、ARMA-LnMVM、ARMA-LnRV、GARCH、GJR-GARCH(依次记为SM1到SM7)对中国股市波动率的预测能力。

为了清晰起见,图1仅表示4种短记忆模型在预测样本区间的波动率预测结果(图2表示长记忆的),图中对实际市场波动率的估计RV用空心的五角星表示。为了得到MCS检验中的各统计量及 p 值,这里选取 $d = 2$ (block length)和模拟次数 $B = 10\,000$ 次作为Bootstrap过程的控制参数。参照Hansen和Lunde^[22]的做法,MCS检验的显著性水平 α 取值为0.1,则 p 值小于0.1的波动率预测模型是样本外预测能力差的模型,将在MCS检验过程中被剔除,而 p 值大于0.1的波动率预测模型是样本外预测能力较好的模型,在MCS检验中能幸存下来。

表1 已实现波动率RV序列和多分形波动率序列的描述性统计

Table 1 Description statistic for realized volatility and multifractal volatility series

参数	已实现波动率相关序列		多分形波动率相关序列							
	RV_i	$LnRV_i$	MVF_i	$LnMVF_i$	$MFVW_i$	$LnMFVW_i$	MV_i	$LnMV_i$	MVM_i	$LnMVM_i$
平均值	3.176	0.069	3.176	0.635	3.176	0.595	3.176	0.635	3.176	0.595
标准差	3.972	0.915	3.949	1.004	4.24	1.032	3.947	1.004	4.24	1.032
偏度	4.416	0.407	3.246	0.234	3.602	0.29	3.246	0.234	3.602	0.29
峰态系数	37.393*	2.721*	17.728*	2.562*	21.636*	2.622*	17.728*	2.562*	21.636*	2.622*
J-B 统计量	106.063*	62.254*	21.792*	34.550*	33.584*	40.234*	21.792*	34.550*	33.584*	40.234*
ADF	-4.915*	-4.849*	-7.729*	-6.113*	-13.643*	-6.166*	-7.729*	-6.113*	-13.643*	-6.166*
Q 统计量	$Q(5)$	2 535.37*	4 948.10*	968.68*	1 700.40*	880.77*	1 666.60*	968.68*	1 700.40*	880.77*
	$Q(10)$	3 739.00*	8 611.00*	1 528.50*	2 957.00*	1 353.60*	2 890.50*	1 528.50*	2 957.00*	1 353.60*
	$Q(20)$	6 316.00*	14 720.00*	2 579.90*	4 932.90*	2 270.80*	4 822.90*	2 579.90*	4 932.90*	2 270.80*
H	0.890 7	0.941 7	0.887 1	0.926 6	0.887 1	0.926 6	0.887 1	0.926 6	0.887 1	0.926 6

注: *表示在1%的水平上显著;峰态系数为超额峰态; $Q(n)$ 为滞后阶数为 n 的Ljung-Box Q 统计量,ADF为Augmented Dickey-Fuller单位根检验。 H 是Hurst指数值,当 $1 > H > 0.5$ 时,表明序列存在长记忆性。

⑤ 关于Bootstrap的实现过程可以参考Hansen和Lunde^[22]的相关论文。

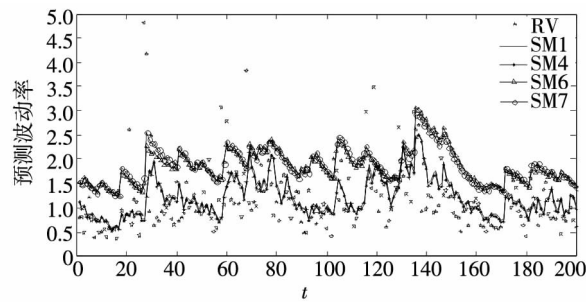


图1 短记忆模型的在预测样本区间的预测结果

Fig. 1 Forecasting results of short memory volatility models during forecasting period

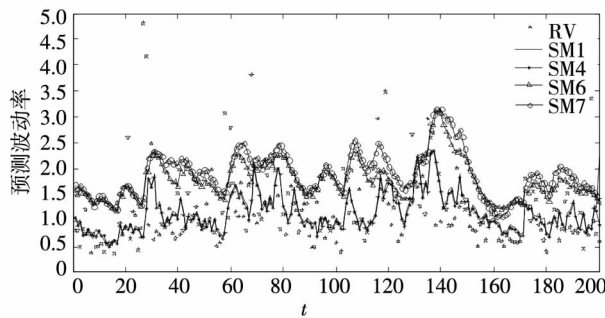


图2 长记忆模型的在预测样本区间的预测结果

Fig. 2 Forecasting results of long memory volatility models during forecasting period

表2是MCS检验结果表的第1列是本文所涉及的全部波动率模型(共14个)第1行是各损失函数 L_i 表中数字是在各损失函数下统计量 T_R 和 T_{SQ} 对应的 p 值。 p 值大于0.1(用加粗和下划线指示的数字)表示的是在MCS检验过程中的幸存模型,即预测能力较好的模型。 p 值越大,表明该模型得到的波动率预测精度越高。从表2的实证结果可以发现:

1) 不论是在何种损失函数下的MCS检验,GARCH族模型(SM6对应的GARCH、SM7对应的GJR-GARCH、LM6对应的HYGARCH以及LM7对应的FIEGARCH)几乎都会被剔除(p 值小于0.1)。也就是说,GARCH族模型的预测效果明显不如多分形波动率模型;

2) 对比表中短记忆模型(SM)和长记忆模型(LM)的检验结果可以看出,总的来说,在不同损失函数和统计检验量下,长记忆模型中的幸存模型数量要稍多于短记忆类模型。特别是在损失函数 R^2LOG 下,这一现象更为明显。进一步,通过对比长记忆和短记忆模型中幸存模型对应的 p 值也可发现,长记忆模型的 p 值普遍高于短记忆模型。因此,认为在对中国股市的波动率预测中,长记忆模型的表现要优于短记忆模型,即长记忆模型是更为合适的选择;

3) 基于已实现波动率测度RV的短记忆模型

ARMA-LnRV(模型SM5)和长记忆模型ARFIMA-LnRV(模型LM5)的波动率预测效果也较为突出。这一结果一方面可能来源于RV测度方法本身的科学性,另一方面也可能是因为所采用的波动率预测基准就是RV测度本身;

4) 在所有14种对比模型中,只有本文新提出的长记忆多分形波动率模型ARFIMA-LnMVM(即模型LM4)在所有损失函数和检验统计标准下幸存。特别是在MAE、HMSE及HMAE标准下,其他13种模型都被剔除,而本文提出的模型均得到了最大的 p 值($p = 1$)。这一结果有力地证明了多分形波动率测度方法及其模型在我国股市中具有良好的实际应用价值,且本文提出的新的多分形波动率测度方法及其ARFIMA-LnMVM模型的优势更为显著。

3.3 基于已实现双幂次变差(BPV)的各模型预测精度检验

近年来,基于金融高频数据的波动率研究成为金融研究领域的热点,因此已实现波动率也获得了不断的改进和发展。尤其是Barndorff-Nielsen等^[34-35]提出的(realized bipower variation,记为BPV)广泛应用于高频数据的波动率建模中。该方法不但具备已实现波动率的各种优点,而且具有更好的稳健性。最近,有研究^[36]将该方法应用于

中国股票市场的实证研究,并进一步探讨其在我国市场中的有效性.为此,本文又以已实现双幂次变差(BPV)为基准波动率(即真实波动率的代理变量),再一次对上述 14 种模型进行了 MCS 检验,实证结果与以已实现波动率作为真实波动率的替代变量基本一致(如表 3 所示).另外,通过对

比表 2 和表 3,不难发现以 BPV 作为基准波动率时,本文新提出的多分形波动率模型预测精度明显提高.总而言之,不论是以 RV 还是 BPV 作为真实波动率的替代变量,本文提出的多分形波动率测度方法及其 ARFIMA-LnMVM 模型的优势都是非常明显的.

表 2 不同波动率模型预测结果的 MCS 检验结果(基于 RV)

Table 2 Based on RV, MCS results for different volatility forecasting results

损失函数	MSE		MAE		HMSE		HMAE		QLIKE		R ² LOG	
	T _R	T _{SQ}	T _R	T _{SQ}	T _R	T _{SQ}	T _R	T _{SQ}	T _R	T _{SQ}	T _R	T _{SQ}
SM1	0.267 9	0.338 1	0.000 3	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.880 9	0.901 5	0.028 0	0.042 1
SM2	0.267 9	0.374 8	0.000 1	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.880 9	0.861 6	0.035 4	0.078 6
SM3	0.628 9	0.448 5	0.011 8	0.004 0	0.021 0	0.000 0	0.000 0	0.000 0	0.394 0	0.514 2	0.142 7	0.333 9
SM4	0.267 9	0.386 1	0.001 7	0.002 0	0.000 0	0.000 0	0.000 0	0.000 0	0.394 0	0.515 7	0.063 9	0.143 3
SM5	0.718 2	0.569 1	0.016 7	0.005 5	0.000 0	0.000 0	0.000 0	0.000 0	0.880 9	0.861 6	0.138 6	0.171 5
SM6	0.000 0	0.000 1	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.011 4	0.024 4	0.000 0	0.000 0
SM7	0.000 1	0.000 4	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.013 4	0.088 3	0.000 0	0.000 0
LM1	0.718 2	0.569 1	0.023 9	0.036 5	0.059 6	0.000 5	0.007 9	0.000 1	0.880 9	0.861 6	0.221 4	0.552 8
LM2	0.875 5	0.827 0	0.016 7	0.012 5	0.028 9	0.000 0	0.002 2	0.000 0	0.880 9	0.901 5	0.333 5	0.650 1
LM3	0.875 5	0.827 0	0.055 6	0.080 3	0.059 6	0.000 5	0.007 9	0.000 1	0.822 8	0.682 2	0.768 7	0.799 4
LM4	0.718 2	0.825 2	1.000 0	1.000 0	1.000 0	1.000 0	1.000 0	1.000 0	0.394 0	0.515 7	1.000 0	1.000 0
LM5	1.000 0	1.000 0	0.055 6	0.080 3	0.059 6	0.000 5	0.007 9	0.000 1	1.000 0	1.000 0	0.768 7	0.799 4
LM6	0.000 1	0.000 2	0.000 1	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.013 4	0.041 8	0.000 0	0.000 0
LM7	0.267 9	0.001 1	0.000 1	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.011 4	0.013 3	0.000 0	0.000 0

注:表中数字是进行了 10 000 次 Bootstrap 模拟的 MCS 检验 p 值. p 值大于 0.1(加粗和下划线的数字)表示的是在 MCS 检验过程中的幸存模型,即预测能力较好的模型.7 种长记忆性模型分别为 ARFIMA-LnMFV、ARFIMA-LnMV、ARFIMA-LnMFVW、ARFIMA-LnMVM、ARFIMA-LnRV、HYGARCH 及 FIEGARCH(依次记为 LM1 到 LM7);7 种短记忆性模型分别为 ARMA-LnMFV、ARMA-LnMV、ARMA-LnMFVW、ARMA-LnMVM、ARMA-LnRV、GARCH、GJR-GARCH(依次记为 SM1 到 SM7).

表 3 不同波动率模型预测结果的 MCS 检验结果(基于 BPV)

Table 3 Based on BPV, MCS results for different volatility forecasting results

损失函数	MSE		MAE		HMSE		HMAE		QLIKE		R ² LOG	
	T _R	T _{SQ}	T _R	T _{SQ}	T _R	T _{SQ}	T _R	T _{SQ}	T _R	T _{SQ}	T _R	T _{SQ}
SM1	0.074 9	0.129 9	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.861 6	0.749 5	0.001 2	0.002 3
SM2	0.076 0	0.174 5	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.861 6	0.749 5	0.001 2	0.004 4
SM3	0.333 0	0.301 3	0.000 0	0.001 8	0.018 5	0.000 0	0.000 4	0.000 0	0.861 6	0.746 3	0.039 9	0.043 0
SM4	0.080 4	0.224 8	0.000 0	0.000 3	0.000 0	0.000 0	0.000 0	0.000 0	0.861 6	0.746 3	0.003 0	0.014 9
SM5	0.398 9	0.462 6	0.000 8	0.000 1	0.000 0	0.000 0	0.000 0	0.000 0	0.861 6	0.749 5	0.039 9	0.026 6
SM6	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.861 6	0.024 1	0.001 2	0.000 0
SM7	0.000 0	0.000 3	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.861 6	0.097 7	0.001 2	0.000 0
LM1	0.398 9	0.776 9	0.010 4	0.004 8	0.028 5	0.000 2	0.003 6	0.000 0	0.861 6	0.749 5	0.064 5	0.115 4
LM2	0.802 0	0.959 6	0.001 5	0.000 5	0.019 2	0.000 0	0.000 0	0.000 0	0.861 6	0.793 0	0.064 5	0.115 4
LM3	0.963 3	0.959 6	0.010 4	0.0069	0.028 5	0.000 2	0.003 6	0.000 0	0.861 6	0.749 5	0.204 1	0.230 3
LM4	0.963 3	0.959 6	1.000 0	1.000 0	1.000 0	1.000 0	1.000 0	1.000 0	0.861 6	0.746 3	1.000 0	1.000 0
LM5	1.000 0	1.000 0	0.010 4	0.006 9	0.028 5	0.000 2	0.003 6	0.000 0	1.000 0	1.000 0	0.204 1	0.230 3
LM6	0.000 1	0.000 1	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.861 6	0.041 2	0.001 2	0.000 0
LM7	0.074 9	0.000 7	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.000 0	0.861 6	0.011 1	0.001 2	0.000 0

注:表中数字是进行了 10 000 次 Bootstrap 模拟的 MCS 检验 p 值. p 值大于 0.1(加粗和下划线的数字)表示的是在 MCS 检验过程中的幸存模型,即预测能力较好的模型.7 种长记忆性模型分别为 ARFIMA-LnMFV、ARFIMA-LnMV、ARFIMA-LnMFVW、ARFIMA-LnMVM、ARFIMA-LnRV、HYGARCH 及 FIEGARCH(依次记为 LM1 到 LM7);7 种短记忆性模型分别为 ARMA-LnMFV、ARMA-LnMV、ARMA-LnMFVW、ARMA-LnMVM、ARMA-LnRV、GARCH、GJR-GARCH(依次记为 SM1 到 SM7).

4 结束语

多分形波动率测度是种全新的金融市场波动率测度方法. 近年来, 各国金融市场都展现出复杂的分形特征^[37], 因此, 本文在现有多分形波动率测度(MFV、MV及MFVW)的基础上, 提出了新的测度方法(MVM)及其计量模型. 进一步, 运用MCS检验, 实证对比了各类多分形波动率模型和主流的GARCH族以及已实现波动率(realized volatility)模型对中国股市波动的刻画和预测能力. 主要实证结果显示: 首先, 多分形波动率模型对中国股市波动率的预测效果明显好于GARCH族模型; 其次, 无论是GARCH族模型、RV模型还是各类多分形波动率模型, 其长记忆形式的预测能力要好于短记忆模型; 最后, 在所有14种对比模型中, 只有新提出的长记忆多分形波动率模型ARFIMA-LnMVM(即模型LM4)在所有损失函数和检验统计标准下幸存. 特别是在MAE、HMSE及HMAE标准下, 其他13种模型都被剔除, 而本文提出的模型均得到了最大的 p 值($p = 1$). 本文

又以相比已实现波动率更具优势和稳健性的已实现双幂次变差为基准波动率, 再次对本文涉及的14种模型进行MCS检验. 实证结果与上述结论基本一致, 且本文提出的多分形波动率模型预测精度明显提高. 简言之, 从总体预测能力上来看, 本文提出的多分形波动率模型具有更好的波动率预测精度.

然而需要指出的是, 虽然多分形波动率模型的优越性在中国股市已经得到了不少实证检验^[9-15], 但有学者发现上证指数真实时间序列的奇异性强度与随机序列的奇异性强度相差无几, 因此上证指数本身并不具有多分形特性, 从而用价格序列来构建多分形的概率测度方法也许并不合适^[38-39]. 但无论如何, 多分形波动率为刻画市场波动的复杂特征提供了可能的新视角, 并且它可以灵活描述很多金融市场的“典型事实”(stylized fact), 如波动的长记忆性和条件收益的胖尾分布特征等. 从本文的实证结果来看, 对金融市场多分形波动特征的研究及进一步探讨, 对于更加深入和全面的市场波动率刻画以及金融风险定量描述和预测都具有较强的实际应用价值.

参考文献:

- [1]Cao G X, Xu L B, Cao J. Multifractal detrended cross-correlations between the Chinese exchange market and stock market [J]. *Physica A*, 2012, 391(20): 4855 - 4866.
- [2]Wang Y D, Wei Y, Wu C F. Cross-correlations between Chinese A-share and B-share markets [J]. *Physica A*, 2010, 389(23): 5468 - 5478.
- [3]Wang Y D, Liu L, Gu R B. Analysis of efficiency for Shenzhen stock market based on multifractal detrended fluctuation analysis [J]. *International Review of Financial Analysis*, 2009, 18(5): 271 - 276.
- [4]Jiang Z Q, Zhou W X. Multifractal analysis of Chinese stock volatilities based on the partition function approach [J]. *Physica A*, 2008, 387(19): 4881 - 4888.
- [5]Zunino L, Tabak B M, Figliola A, et al. A multifractal approach for stock market inefficiency [J]. *Physica A*, 2008, 387(26): 6558 - 6566.
- [6]Mandelbrot B B. *Fractals and Scaling in Finance* [M]. New York: Springer, 1997.
- [7]Mandelbrot B B. A multifractal walk down Wall Street [J]. *Scientific American*, 1999, 280(2): 70 - 73.
- [8]Poon S H, Granger C W J. Forecasting volatility in financial markets: A review [J]. *Journal of Economic Literature*, 2003, 41(2): 478 - 539.
- [9]魏宇. 金融市场的多分形波动率测度、模型及其SPA检验 [J]. *管理科学学报*, 2009, 12(5): 88 - 99.
Wei Yu. Multifractal volatility measures, its model and SPA test in financial market [J]. *Journal of Management Sciences in China*, 2009, 12(5): 88 - 99. (in Chinese)
- [10]王鹏, 魏宇, 王建琼. 基于多分形波动率测度的权证定价方法研究 [J]. *管理科学*, 2009, 12(2): 106 - 113.
Wang Peng, Wei Yu, Wang Jianqiong. Warrant pricing based on multifractal volatility measurement [J]. *Management Sciences in China*, 2009, 12(2): 106 - 113. (in Chinese)

- [11] 魏 宇, 黄登仕. 金融市场多标度分形现象及与风险管理的关系 [J]. 管理科学学报, 2003, 6(1): 87-91.
Wei Yu, Huang Dengshi. Multifractal phenomenon and financial risk management [J]. Journal of Management Sciences in China, 2003, 6(1): 87-91. (in Chinese)
- [12] 魏 宇. 多分形波动率测度的 VaR 计算模型 [J]. 系统工程理论与实践, 2009, 29(9): 7-15.
Wei Yu. VaR model based on multifractal volatility measurement [J]. Systems Engineering-Theory & Practice, 2009, 29(9): 7-15. (in Chinese)
- [13] 魏 宇. 基于多分形理论的动态 VaR 预测模型研究 [J]. 中国管理科学, 2012, 20(5): 7-18.
Wei Yu. Forecasting model for dynamic value-at-risk based on multifractal theories [J]. Chinese Journal of Management Science, 2012, 20(5): 7-18. (in Chinese)
- [14] 王 鹏, 魏 宇. 基于多分形波动率测度的 ES 风险度量 [J]. 系统管理学报, 2012, 21(2): 192-200.
Wang Peng, Wei Yu. Excepted shortfall estimation based on multifractal volatility [J]. Journal of Systems & Management, 2012, 21(2): 192-200. (in Chinese)
- [15] 王 鹏, 王建琼. 中国股票市场的多分形波动率测度及其有效性研究 [J]. 中国管理科学, 2008, 16(6): 9-15.
Wang Peng, Wang Jianqiong. Multifractal volatility measure of China's Stock Market and its validity [J]. Chinese Journal of Management Science, 2008, 16(6): 9-15. (in Chinese)
- [16] Chen H T, Wu C F. Forecasting volatility in Shanghai and Shenzhen markets based on multifractal analysis [J]. Physica A, 2011, 390(16): 2926-2935.
- [17] Andersen T G, Bollerslev T. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts [J]. International Economic Review, 1998, 39(4): 885-905.
- [18] Barndorff-Nielsen O E, Shephard N. Non-Gaussian OU based models and some of their uses in financial economics [J]. Journal of the Royal Statistical Society, 2001, 63(2): 167-241.
- [19] Ashely R, Granger C W J, Schmalensee R. Advertising and aggregate consumption: An analysis of causality [J]. Econometrica, 1980, 48(5): 1149-1167.
- [20] Lo A W, MacKinlay A C. Data-snooping biases in tests of financial asset pricing models [J]. Review of Finance Studies, 1990, 3(3): 431-467.
- [21] Foser F D, Smith T, Whaley R E. Assessing goodness-of-fit of asset pricing models: the distribution of the maximal R^2 [J]. Journal of Fiance, 1997, 52(2): 591-607.
- [22] Hansen P R, Lunde A, James M N. The model confidence set [J]. Econometrica, 2011, 79(2): 453-497.
- [23] Hansen P R, Lunde A. A test for superior predictive ability [J]. Journal of Business and Economic Statistics, 2005, 23(4): 365-380.
- [24] Andersen T G, Bollerslev T, Meddahi N. Correcting the errors: Volatility forecast evaluation using high frequency data and realized volatilities [J]. Econometrica, 2005, 73(3): 279-296.
- [25] Hansen P R, Lunde A. A forecast comparison of volatility models: Does anything beat a GARCH(1,1)? [J]. Journal of Applied Econometrics, 2005, 20(7): 873-889.
- [26] Andersen T G, Bollerslev T, Diebold F X, et al. The distribution of realized stock return volatility [J]. Journal of Financial Economics, 2001, 61(1): 43-76.
- [27] 唐 勇. 金融资产跳跃检验方法实证比较 [J]. 中国管理科学, 2012, 20(5): 290-299.
Tang Yong. The empirical comparison among jump tests of financial asset based on high frequency data [J]. Chinese Journal of Management Science, 2012, 20(5): 290-299. (in Chinese)
- [28] 叶五一, 缪柏其. 已实现波动率与日内价差条件下的 CVaR 估计 [J]. 管理科学学报, 2012, 15(8): 60-71.
Ye Wuyi, Miao baiqi. Estimating of CVaR with consideration of realized volatility and price range [J]. Journal of Management Sciences in China, 2012, 15(8): 60-71. (in Chinese)
- [29] 魏 宇, 黄登仕. 基于多标度分形理论的金融风险测度指标研究 [J]. 管理科学学报, 2005, 8(4): 50-59.
Wei Yu, Huang Dengshi. Study on financial risk measure based on multifractal theory [J]. Journal of Management Sciences in China, 2005, 8(4): 50-59. (in Chinese)
- [30] Sun X, Chen H, Wu Z, et al. Multifractal analysis of Hang Seng index in Hong Kong stock market [J]. Physica A, 2001, 291(7): 553-562.

- [31] Sun X, Chen H, Yuan Y, et al. Predictability of multifractal analysis of Hang Seng stock index in Hong Kong [J]. *Physica A*, 2001, 301(12): 473–482.
- [32] Andersen T G, Bollerslev T, Diebold F X, et al. Modeling and forecasting realized volatility [J]. *Econometrica*, 2003, 71(2): 579–625.
- [33] Koopman S H, Jungbacker B, Hol E. Forecasting daily variability of the S&P 100 stock index using historical, realised and implied volatility measurements [J]. *Journal of Empirical Finance*, 2005, 12(3): 445–475.
- [34] Barndorff-Nielsen O E, Shephard N. Power and bipower variation with stochastic volatility and jumps [J]. *Journal of Financial Econometrics*, 2004, 2(1): 1–37.
- [35] Barndorff-Nielsen O E, Shephard N. Econometrics of testing for jumps in financial economics using bipower variation [J]. *Journal of Financial Econometrics*, 2006, 4(1): 1–30.
- [36] 李胜歌, 张世英. 已实现双幂次变差与多幂次变差的有效性分析 [J]. *系统工程学报*, 2007, 22(3): 280–286.
Li Shengge, Zhang Shiying. Analysis of the efficiency of realized bipower variation and realized multipower variation [J]. *Journal of Systems Engineering*, 2007, 22(3): 280–286. (in Chinese)
- [37] 周炜星. 金融物理学导论 [M]. 上海, 上海财经大学出版社, 2007.
Zhou Weixing. *A Guide to Econophysics* [M]. Shanghai: Shanghai University of Finance and Economic Press, 2007. (in Chinese)
- [38] 周炜星. 上证指数高频数据的多重分形错觉 [J]. *管理科学学报*, 2010, 13(3): 81–86.
Zhou Weixing. Illusionary multifractal in high-frequency data of Shanghai Stock Exchange Composite Index [J]. *Journal of Management Sciences in China*, 2010, 13(3): 81–86. (in Chinese)
- [39] Jiang Z Q, Zhou W X. Multifractal in stock indexes: Fact or fiction? [J]. *Physica A*, 2008, 387(14): 3605–3614.

Multi-fractal volatility forecasting model and its MCS test

WEI Yu, MA Feng, HUANG Deng-shi

School of Economics & Management, Southwest Jiao Tong University, Chengdu 610031, China

Abstract: This paper introduces a new volatility measure and constructs its model based on multifractal volatility method. Taking 5-minute high frequency data of the Shanghai Composite Index as an example, and applying the out-of-sample rolling time window forecasting combined with Model Confidence Set which is proved superior to SPA test, this paper compares the empirical performance of the new model and those of the GARCH-type and Realized volatility (RV) models. The empirical results show that the forecasting accuracy of the multifractal volatility measure model in the short term as well as in the long term are better than the GARCH-type and RV models. Moreover, the forecasting models in the long term perform better than those in the short term. The performance in most loss function of the new method based on multifractal volatility measure is superior to other forecasting models.

Key words: multifractal volatility; GARCH-type models; realized volatility model; rolling forecasting; MCS test