

# 金融市场的多分形波动率测度、模型及其 SPA 检验<sup>①</sup>

魏 宇

(西南交通大学经济管理学院, 成都 610031)

**摘要:** 提出一种新的金融市场波动率的测度方法——多分形波动率 (multifractal volatility) 测度, 并以上证综指在长达 8 年左右时间内的高频数据样本为例, 构造了多分形波动率的 ARFMA 动力学模型. 同时, 运用最近提出的 SPA (superior predictive ability) 检验法, 实证对比了多分形波动率模型与现有的如实现波动率 (realized volatility) 模型、GARCH 模型以及随机波动 (stochastic volatility, SV) 模型对市场波动预测能力的优劣. 实证结果显示, 在某些损失函数标准下, 文中提出的多分形波动率测度及其动力学模型具有比现有其它模型更优的波动率刻画能力和预测精度.

**关键词:** 多分形波动率; 实现波动率; 波动率测度; 预测; SPA 检验

**中图分类号:** F224 F830 **文献标识码:** A **文章编号:** 1007-9807(2009)05-0088-12

## 0 引 言

对金融资产收益的波动率 (volatility) 描述是现代金融理论的核心内容, 有关波动率大小的测度 (measurement) 及其动力学机制 (dynamics) 的刻画, 对于资产定价理论的检验、最优资产组合的选择、衍生产品套期保值策略的设计以及金融风险的管理而言, 都具有极其重要的理论和实际意义.

1999 年, 分形 (fractal) 理论之父 Mandelbrot 在《科学美国人》(Scientific American) 上撰文指出, 多分形 (multifractal)<sup>②</sup> 理论是一种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 (unifractal) 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<sup>[1]</sup>. 此后, 众多国内外学者运用多分形理论对各种不同类型金融市场的波动现象进行了实证研究, 取得了许多有价值的研究成果<sup>[2-13]</sup>. 其中, 文献

[2~7] 还实证检验了中国股票市场的多分形波动特征.

但需要指出的是, 目前绝大多数的多分形研究还停留在对市场多分形波动特征的实证检验层面之上. 因此, 如何进一步挖掘多分形波动分析中产生的对市场波动特征描述的有益统计信息, 进而为更加准确的市场波动率测度和建模提供依据, 仍然是目前该领域研究中亟待解决的难点问题之一.

同时, 主流金融理论中的各种波动率描述方法及其模型, 无论是早期的 ARCH/GARCH 族模型、随机波动模型 (stochastic volatility, SV), 还是最近非常流行的实现波动率模型 (realized volatility, RV), 都是直接利用股价或者收益率序列本身来直接描述市场的波动率大小. 但是, 由于金融市场的波动率本身是无法直接加以观察的, 因此, 现有主流波动率研究中的众多统计推论也许并不具有广泛的代表性<sup>[9-11]</sup>.

① 收稿日期: 2006-05-22; 修订日期: 2009-04-28.

基金项目: 国家自然科学基金资助项目 (70501025; 70771097; 70771095); 教育部新世纪优秀人才支持计划资助项目 (NCET-08-0826); 教育部长江学者和创新团队发展计划项目 (RT0806).

作者简介: 魏宇 (1975-), 男, 四川攀枝花人, 博士, 教授, 博士生导师. Email: weiyu@126.com

② 国内也有一些学者将 multifractal 翻译为“多标度分形”或者“多重分形”

基于上述考虑,在前期相关研究的基础上<sup>[3-5]</sup>,首先利用高频数据的多分形谱(multifractal spectrum)分析,提炼出对金融市场日波动率(daily volatility)描述的间接统计信息,显然这些间接的定量统计信息,是与上面讨论的主流金融学中各种对波动率的直接描述方法相异的.在此基础上,本文提出了一种基于高频股价数据多分形谱的市场波动率测度方法—多分形波动率(multifractal volatility, MFV)测度,并相应构造了其 ARFMA 动力学模型.同时,为了验证这种新的波动率测度方法及其模型的可靠性和实用性,还实证计算了 MFV 模型、RV 模型、GARCH 模型以及 SV 模型对未来市场波动率的样本外(out-of-sample)预测值,并进一步运用 Hansen and Lunde 提出的 SPA (superior predictive ability) 检验法<sup>[14]</sup>,实证对比了这些模型对市场波动率预测的精确程度.实证结果显示,在经异方差调整后的损失函数标准下,本文提出的多分形波动率测度及其 ARFMA 模型具有比现有其它波动率模型更加优越的波动率刻画能力和预测精度.

## 1 数据样本说明以及日收益率和高频收益率的计算方法

本文研究的数据样本为上证综指(SSEC)从1999年1月19日到2006年10月10日的每5分钟高频数据(共 $N = 1\,849$ 个交易日),记为 $I_{t,d}$ , $t = 1, 2, \dots, N$ ,  $d = 0, 1, 2, \dots, 48$ 其中 $I_{t,0}$ 表示第 $t$ 天的开盘价, $I_{t,48}$ 表示第 $t$ 天的收盘价,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9:30分开盘,到11:30分中午休市,然后13:00开盘,到15:00全天收盘,每天共有4个小时(即240min)连续竞价交易时间,因此,采用每5min记录一个数据的方法每天可以产生48个高频股价记录(不包括 $I_{t,0}$ ),样本总体的高频数据量为88752个.文中的日收益率(daily return) $R_t$ 利用相邻两个交易日的收盘价计算如下

$$R_t = 100(\ln I_{t,48} - \ln I_{t-1,48}), \quad t = 2, 3, \dots, N \quad (1)$$

同理,定义第 $t$ 天的(每5min)高频收益率(high-frequency return) $R_{t,d}$ 为

$$R_{t,d} = 100(\ln I_{t,d} - \ln I_{t,d-1}), \quad d = 1, 2, \dots, 48 \quad (2)$$

## 2 基于高频数据的实现波动率 RV 模型与多分形波动率 MFV 模型

### 2.1 实现波动率 RV 的估计方法及其动力学模型

Andersen and Bollerslev 首次指出<sup>[15]</sup>,传统上用日收益率的平方(squared daily return)作为日波动率(daily volatility)的测度将会面临非常严重的测量误差和噪声(measurement error and noise),而使用基于交易日内高频收益数据(intradaily high-frequency return)的实现波动率(RV)作为日波动率的测度,将大大降低这些误差和噪声对真实潜在波动率过程(underlying volatility process)的影响.最近,Andersen等的研究进一步指出<sup>[16]</sup>,从根本上讲,由于潜在真实的市场波动率是不可观测的(unobservable),因此目前公认的方法是用基于高频收益数据的实现波动率来作为市场真实波动率的代理变量(proxy)和测度基准(benchmark).

根据 Andersen and Bollerslev 的定义,对第 $t$ 天的实现波动率的估计表示为第 $t$ 天内的高频收益平方和<sup>[15]</sup>,即

$$RV_t' = \sum_{d=1}^{48} R_{t,d}^2 \quad (3)$$

但最近 Hansen and Lunde 的研究又指出<sup>[17]</sup>,由于股票市场并不象外汇市场那样在24h连续进行交易,因此,能观察和记录到的高频股价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(即股票市场从收盘到第二天开盘的所谓“close-to-open”波动率).因此为了使实现波动率的估计更加准确地刻画全天的市场波动率大小,采用 Hansen and Lunde 的建议,用某种尺度参数(scale parameter) $\gamma$ 来对 $RV_t'$ 进行尺度变换<sup>[17]</sup>,即对第 $t$ 天的实现波动率估计为

$$RV_t = \gamma \cdot RV_t' \quad (4)$$

其中

$$Y = \frac{N^{-1} \sum_{t=1}^N R_t^2}{N^{-1} \sum_{t=1}^N RV_t'} \quad (5)$$

Andersen等的研究又发现<sup>[18]</sup>,在取自然对数以后,对数实现波动率(以下简称为lnRV)的波动特征可以用一种高斯动力学过程(Gaussian dynamic process)来描述,同时lnRV展现出明显的长期记忆性(long memory)特性.为此,Andersen等又建议采用自回归分整移动平均过程ARFMA来描述lnRV的上述动力学特性<sup>[18]</sup>.

考虑到不同滞后阶数的ARFMA(p,d,q)模型对lnRV的估计结果非常接近,同时结合模型估计的AIC(Akaike's Information Criterion)大小比较,这里采用ARFMA(1,d,1)模型来为lnRV建模.

ARFMA(p,d,q)模型的一般形式为

$$\Phi(L)(1-L)^d(Y-\mu) = \Theta(L)\varepsilon_t \quad (6)$$

其中 $\Phi(L) = 1 - \phi_1L - \phi_2L^2 - \dots - \phi_pL^p$ ,  
 $\Theta(L) = 1 + \theta_1L + \theta_2L^2 + \dots + \theta_qL^q$ 分别为自回归滞后p阶算子以及移动平均滞后q阶算子,L为滞后算子, $(1-L)^d$ 为分数差分算子, $\mu$ 是Y的均值,同时这里假定 $\varepsilon_t \sim ND(0, \sigma_\varepsilon^2)$ .

### 2.2 多分形波动率MFV的估计方法及其动力学模型

与2.1节中对实现波动率(RV)的估计方法不同,本节提出的多分形波动率测度(MFV)的构建并非基于高频收益率 $R_{i,d}$ ,而是基于高频股价数据 $I_{i,d}$ 本身.即首先通过高频股价数据来计算一天当中的市场波动多分形谱(multifractal spectrum,记为 $f(\alpha)$ ),再从多分形谱 $f(\alpha)$ 中计算MFV,具体计算过程如下:

**步骤1** 参照文献[4]的方法,即用“数盒子”(box-counting)的方法来计算每天市场波动的多分形谱.假定整个交易日的的时间长度为标准化的1,则无重复均匀覆盖这48个高频股价数据的“盒子长度”( $\delta < 1$ )可以分别取1/48, 1/24, 1/16, 1/12, 1/8, 1/6, 1/4, 1/3, 1/2和1.举例来说,当取无重复覆盖的盒子长度为1/8时,就表明可以用8个这样的盒子无重复的覆盖住每天产生的这48个交易数据,其中每个盒子中的数据个数为 $48/8 = 6$ 个.

**步骤2** 当取盒子长度为 $\delta$ 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 $I(t), t = 1, 2, \dots, 48$ ),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

$$P_i(\delta) = \frac{\sum_{j=1}^n I(i_j)}{\sum_{t=1}^n I(t)}, \quad i = 1, 2, \dots, m \quad (7)$$

其中, $I(i_j)$ 表示第i个盒子中的第j个指数.根据文献[4]中的相关定义,则有以下幂律关系(power law)存在

$$P_i(\delta) \sim \delta^\alpha \quad (8)$$

$$N_\alpha(\delta) \sim \delta^{-f(\alpha)} \quad (9)$$

其中, $N_\alpha(\delta)$ 表示具有相同Hölder指数 $\alpha$ 的长度为 $\delta$ 的盒子个数.然后,通过以下的“分割函数”- $S_q(\delta)$ 来计算多分形谱 $f(\alpha)$ ,其中

$$S_q(\delta) = \sum_{i=1}^m P_i^q(\delta) \quad (10)$$

同样,根据文献[4]中的相关结论, $S_q(\delta)$ 同样满足以下形式的幂律关系

$$S_q(\delta) \sim \delta^{q\tau(q)} \quad (11)$$

可以看到,当q取正数时,q越大,则 $S_q(\delta)$ 将主要反映的是那些具有大的概率测度的盒子的信息;反之,当q取负数时,q越小,则 $S_q(\delta)$ 将主要反映的是那些具有小的概率测度的盒子的信息.在实际计算时,q的取值范围大小以 $\alpha$ 和 $f(\alpha)$ 达到饱和值为准<sup>[12,13]</sup>. $\tau(q)$ 的值可以通过求取在双对数坐标轴 $\ln S_q(\delta) - \ln \delta$ 上的直线斜率得出,并且通过Legendre变换可以得到<sup>[12,13]</sup>

$$\alpha = \frac{d\tau(q)}{dq} \quad (12)$$

$$f(\alpha) = \alpha q - \tau(q) \quad (13)$$

运用上面的方法,以2004年2月4日和2004年2月5日的连续两个交易日的上证综指为例,计算了其对应的价格波动多分形谱,如图1所示.其中,图1的(a)、(b)两图分别表示的是两天当中的上证综指的高频价格走势,而(c)、(d)两图分别是两天的上证综指的多分形谱 $f(\alpha)$ ,其分布的Hölder指数 $\alpha$ 的最大和最小值分别用 $\alpha_{max}$ 和 $\alpha_{min}$ 表示.

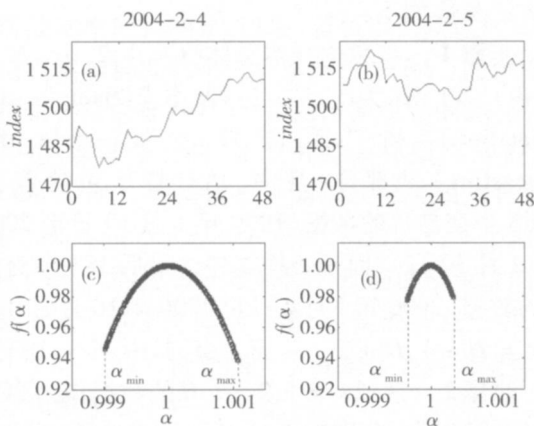


图 1 两天当中上证综指的高频价格走势及多分形谱分布

Fig. 1 Multifractal spectra of SSEC in two days

图 1 中多分形谱表现出的显著弓形表明, 上证综指的价格波动确实展现出明显的多分形特征. 同时, 可以看到, 不同的高频价格波动形式和波动幅度对应着不同的多分形谱形状和大小. 因为每个指数价格盒子的概率测度  $P_i(\delta) \sim \delta^\alpha$ , 所以这个概率测度的大小反映的就是盒子当中指数价格的高低水平. 同时注意到  $\delta < 1$ , 所以  $\alpha_{max}$  和  $\alpha_{min}$  分别代表最小概率测度和最大概率测度盒子的测度值. 显然,  $\alpha_{max}$  表示的是一天当中市场价格波动的相对最低位, 而  $\alpha_{min}$  则表示的是一天当中市场价格波动的相对最高位. 因此,  $\Delta\alpha = \alpha_{max} - \alpha_{min}$  越大 (即多分形谱的宽度越大), 则表明当天价格走势的分布越不均匀, 即当天价格波动的幅度越大, 波动越剧烈, 即  $\Delta\alpha$  是一种能够测度市场日波动率大小的定量描述指标. 同样, 根据 2.1 节的讨论, 为了使多分形波动率 MFV 的估计更加准确地刻画全天的市场波动率大小, 这里仍然采用如公式 (5) 所定义的尺度参数 (scale parameter)  $\gamma$  来对第  $t$  天的多分形波动率测度进行尺度变换, 不同的是将公式 (5) 中的分母变换为  $\Delta\alpha$  的平均值. 下面, 正式定义第  $t$  天的多分形波动率 (multifractal volatility) 为

$$MFV_t = \gamma \cdot \Delta\alpha_t \quad (14)$$

图 2 是样本区间内的日收益率  $R_t$ 、实现波动率估计 RV 以及多分形波动率 MFV 的分布状况. 从图中不难看出, 直观上来讲, 无论是实现波动率 RV 还是多分形波动率 MFV, 都很好地刻画了日收益率的波动状况, 并且提出的 MFV 与作为波动

率测度基准的 RV 具有非常相似的波动率刻画.

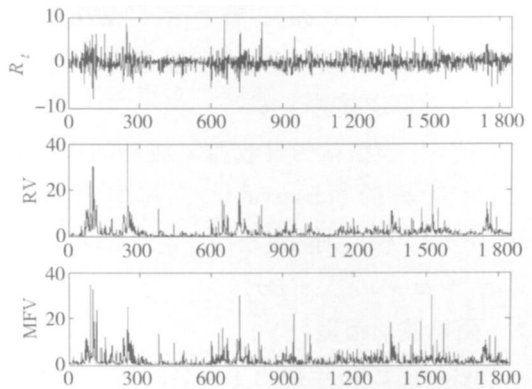


图 2 上证综指收益率及 RV 和 MFV 波动率估计

Fig. 2 Returns, RV and MFV estimates of SSEC

进一步的考虑, 由于 ARFMA( $p, d, q$ ) 对不同类型时间序列的动力学特征都具有很强的刻画能力, 同时很多常见的时间序列模型如自回归模型 AR( $p$ )、自回归移动平均模型 ARMA( $p, q$ ) 以及自回归单整移动平均模型 ARIMA( $p, d, q$ ) 等都可以视为 ARFMA( $p, d, q$ ) 模型的特例, 因此, 与对数实现波动率 (lnRV) 的建模一样, 假定多分形波动率的对数形式 (以下简称为 lnMFV) 仍然服从一个 ARFMA(1,  $d$ , 1) 过程, 具体形式如 2.1 节中的公式 (6) 所示.

### 3 基于日数据的 GARCH 模型与随机波动率 SV 模型

在金融计量研究当中, 除了 2.1 节介绍的实现波动率模型以外, 还有 2 类普遍使用的市场波动率模型, 即 Bollerslev 的广义自回归条件异方差模型 (GARCH) 以及 Taylor 的随机波动模型 (stochastic volatility model, SV)<sup>[19, 20]</sup>. 这两类波动率模型的构建并不依赖于高频数据, 而主要是基于日收益数据, 下面分别简要介绍之:

#### 3.1 广义自回归条件异方差模型 (GARCH)

GARCH 模型是目前金融计量研究当中运用最为广泛的波动率模型之一, 该模型假定金融资产 (日) 收益率满足以下的离散形式

$$R_t = \mu_t + \varepsilon_t = \mu_t + \sigma_t z_t \quad (15)$$

其中  $\mu_t$  是收益波动的条件均值 (conditional

mean),  $\sigma_t^2$  是条件方差 (conditional variance), 而假定新生量 (innovation)  $z_t$  满足:  $z_t \sim NID(0, 1)$ <sup>③</sup>. 同时由于收益率的条件均值一般很小, 因此在实证研究当中假定其等于零.

GARCH 模型认为日收益率波动的条件方差  $\sigma_t^2$  是可观测的 (observable), 其中最常见的是 GARCH(1, 1) 模型则假定条件方差满足以下形式

$$\sigma_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + \beta \sigma_{t-1}^2 \quad (16)$$

### 3.2 随机波动模型 (SV)

Taylor 在 1986 年提出了著名的随机波动模型 (SV)<sup>[20]</sup>, 与 3.1 节讨论的 GARCH 模型不同的是, SV 模型假定金融收益的条件方差  $\sigma_t^2$  是不可观测的 (unobservable), 且其服从以下的随机过程

$$\sigma_t^2 = \sigma^{*2} \exp(h_t) \quad (17)$$

其中, 不可观测的对数波动率 (log-volatility)  $h_t$  则满足

$$h_t = \phi h_{t-1} + \sigma_\eta \eta_{t-1} \quad (18)$$

且假定  $\eta_t \sim NID(0, 1)$ ,  $h_1 \sim NID(0, \sigma_\eta^2 / (1 - \phi^2))$ .

由于在 SV 模型当中, 条件方差是一个不可观测的变量, 很难计算出其精确的似然函数, 因此对 SV 模型进行参数估计存在着较大困难. 近年来, 众多学者提出了很多不同类型的估计方法, 逐步克服了 SV 模型参数估计的困难, 使得 SV 模型逐渐成为了一种与 GARCH 模型一样被广泛使用的波动率模型. 有关 SV 模型及其多变量形式扩展的技术细节可以参见 Shephard 的深入研究<sup>[21]</sup>, 另外, 对 SV 模型的实证估计方法中涉及的模拟极大似然估计以及 Kalman 滤波等技术可以参见 Sandmann and Koopman 的研究<sup>[22]</sup>.

## 4 波动率预测方法及 SPA 检验

### 4.1 波动率预测方法说明

在本文的实证研究当中, 对上面讨论的 4 种波动率模型进行了所谓的“样本外预测能力检验” (tests for out-of-sample predicting ability). 预

测方法具体如下:

步骤 1 将数据样本总体 ( $t = 1, 2, \dots, N = 1849$ ) 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  $H = 1666$  个交易日的数据 (1999 年 1 月 19 日至 2005 年 12 月 30 日), 而预测样本包含最后 183 个交易日的数据 (2006 年 1 月 4 日至 2006 年 10 月 10 日, 即  $t = H + 1, H + 2, \dots, H + M$ , 其中  $M = 183$ ).

步骤 2 选取  $t = 1, 2, \dots, H$  的数据作为第一次的估计样本, 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 1 天的波动率预测, 记为  $\hat{\sigma}_{H+1}^2$ . 也就是说,  $\hat{\sigma}_{H+1}^2$  是在前面 1666 个样本数据的模型估计基础上对第 1667 天的市场波动率预测.

步骤 3 保持估计样本的时间区间长度不变 ( $H = 1666$ ), 将估计样本时间区间向后平行移动 1 天, 即第 2 次选取的是  $t = 2, 3, \dots, H + 1$  的数据样本作为新的估计样本, 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 1 天的市场波动率预测, 记为  $\hat{\sigma}_{H+2}^2$ .

步骤 4 同理, 不断重复步骤 3 可以得到  $\hat{\sigma}_{H+3}^2, \hat{\sigma}_{H+4}^2, \hat{\sigma}_{H+5}^2, \dots$  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  $t = N = H + M = 1849$  d 的市场波动率预测  $\hat{\sigma}_{H+M}^2$ .

简言之, 对前面讨论的每一种波动率模型, 分别重复进行了 183 次的模型估计, 从而得到了每个模型的 183 个未来 1 天的市场波动率预测值, 记为  $\hat{\sigma}_m^2, m = H + 1, H + 2, \dots, H + M$ . 同时, 记预测样本区间的实现波动率估计为  $RV_m, m = H + 1, H + 2, \dots, H + M$ . 这里对  $RV_m$  的估计方法来自于 2.1 节的说明, 并以此作为真实市场波动率的代理 (proxy), 用以衡量各类波动率模型的预测精度.

### 4.2 SPA 检验说明

有了以上所讨论的 4 类波动率模型及其对市

③ 限于篇幅, 只讨论了假定新生量  $z_t$  服从正态分布的情况, 当然还可以推广到假定其服从更具胖尾特征的  $t$  分布或广义误差分布 (GED) 的情况.

场波动率的预测  $\hat{\sigma}_m^2$  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) 究竟有多大了. 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 and Lunde 建议, 可以尽可能多地采用不同形式的损失函数来作为预测模型精度的判断标准<sup>[14]</sup>. 基于这样的考虑, 在实证研究当中, 采用了 6 种不同的损失函数来分别作为各类波动率模型预测精度的评判标准.

这 6 种损失函数分别标记为  $L_i, i = 1, 2, \dots, 6$  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形式.  $L_3$  和  $L_4$  分别是经异方差调整的 MSE 和 MAE (heteroskedastic adjusted MSE and MAE), 而限于篇幅, 对  $L_5$  和  $L_6$  的具体含义讨论可以参考 Hansen and Lunde 的深入讨论<sup>[14]</sup>. 各损失函数的具体定义为

$$L_1 : \text{MSE} = M^{-1} \sum_{m=H+1}^{H+M} (RV_m - \hat{\sigma}_m^2)^2 \quad (19)$$

$$L_2 : \text{MAE} = M^{-1} \sum_{m=H+1}^{H+M} |RV_m - \hat{\sigma}_m^2| \quad (20)$$

$$L_3 : \text{HMSE} = M^{-1} \sum_{m=H+1}^{H+M} (1 - RV_m / \hat{\sigma}_m^2)^2 \quad (21)$$

$$L_4 : \text{HMAE} = M^{-1} \sum_{m=H+1}^{H+M} |1 - RV_m / \hat{\sigma}_m^2| \quad (22)$$

$$L_5 : \text{QLKE} = M^{-1} \sum_{m=H+1}^{H+M} (\ln(\hat{\sigma}_m^2) + RV_m / \hat{\sigma}_m^2) \quad (23)$$

$$L_6 : \text{R}^2\text{LOG} = M^{-1} \sum_{m=H+1}^{H+M} [\ln(RV_m / \hat{\sigma}_m^2)]^2 \quad (24)$$

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, 那么我们只能判断: “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<sup>④</sup>.

为了解决这一问题, Hansen and Lunde 提出了一种所谓的“高级预测能力检验法” (superior predictive ability SPA)<sup>[14]</sup>. 他们的研究证明, 因为采用了所谓的“自举法” (bootstrap), SPA 检验比类似的 White<sup>[23]</sup> 提出的 reality check (RC) 检验法具有更加优异的模型判别能力, 且 SPA 检验的结论具有更好的稳健性 (robustness). 也就是说, 与基于一个单一样本的其他检验法相比, SPA 得到的检验结论更加可靠, 且其得到的结论可以推广到其它类似的数据样本当中去.

SPA 检验的实现过程如下: 首先, 假定有  $J+1$  种不同的波动率模型, 记为  $M_k, k = 0, 1, \dots, J$ . 每种波动率模型  $M_k$  得到的未来 1 天的波动率预测记为  $\hat{\sigma}_{m,k}^2$  其中  $m = H+1, H+2, \dots, H+M$ . 对每一个预测值, 都可以计算公式 (19) — (24) 所定义的 6 种损失函数值, 记为  $L_{i,m,k}$ , 其中  $i = 1, 2, \dots, 6$  下面, 用  $M_0$  表示作为 SPA 检验的基础模型 (base model) 即用该模型作为与其它模型的预测表现进行对比检验的基础, 因此, 对于其它的  $k = 1, 2, \dots, J$  种波动率模型, 可以计算其相对于基础模型  $M_0$  的“相对损失函数值” (relative loss function), 记为

$$X_{k,m} = L_{i,0,m} - L_{i,k,m} \quad (25)$$

现在感兴趣的问题是: 是否在  $k = 1, 2, \dots, J$  的模型  $M_k$  当中有比基础模型 ( $M_0$ ) 表现更加优异的模型呢? 为了得到这一问题的答案, 首先可以定义这样的零假设  $H_0$ : “与其它模型  $M_k$  相比, 基础模型  $M_0$  是表现最好的预测模型.” 这一零假设可以用数学表达式表示为

$$\max_k \lambda_k = E(X_{k,m}) \leq 0, k = 1, 2, \dots, J \quad (26)$$

Hansen and Lunde 证明了这一假设检验的检验统计量可以表示为<sup>[14]</sup>

$$T = \max_k \frac{\sqrt{M} \bar{X}_k}{\hat{\omega}_{kk}}; k = 1, 2, \dots, J \quad (27)$$

其中

④ 比方说, 数据样本中的少数奇异点 (outliers) 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, 进而可能会导致对波动率模型优劣的错误判断. 深入讨论可以参见 Hansen and Lunde 的研究<sup>[17]</sup>.

$$\bar{X}_k = M^{-1} \sum_{m=H+1}^{H+M} X_{k,m}, \hat{\omega}_{kk}^2 = \text{var}(\sqrt{M}\bar{X}_k) \quad (28)$$

为了获得公式 (27) 的  $T$  检验量的分布状况及其  $p$  值, Hansen and Lunde 建议可以采用一种所谓的“自举法”(Bootstrap procedure) 来取得<sup>[14]</sup>. 首先, 需要获得一个长度为  $M$  的  $X_{k,m}$  新样本. 要获得这样一个样本, 则先要从  $\{X_{k,m}\}$  的集合当中随机抽取一个新的子样本 (New subsample), 而该子样本的长度则来自一个服从均值为  $q$  的几何分布 (geometric distribution) 的随机数, 同时控制这些子样本的组合长度为所要求的  $M$ .

重复这样的 bootstrap 过程  $B$  次, 可以获得  $B$  个长度为  $M$  的  $X_{k,m}$  新样本, 记为  $X_{k,m}^i, i = 1, 2, \dots, B$ . 在后面的实证研究当中, 选取  $q = 0.5$  和  $B = 1000$  次作为这一 bootstrap 过程的控制参数. 对每一个 bootstrap 样本的均值表示为

$$\bar{X}_k^i = M^{-1} \sum_{m=1}^M X_{k,m}^i; i = 1, 2, \dots, B \quad (29)$$

而所有  $B$  个 bootstrap 样本均值的方差估计表示为

$$\hat{\omega}_{kk} = B^{-1} \sum_{i=1}^B (\bar{X}_k^i - \bar{\bar{X}}_k)^2; \bar{\bar{X}}_k = B^{-1} \sum_{i=1}^B \bar{X}_k^i \quad (30)$$

其次, 定义  $\bar{Z}_k^i$  为

$$\bar{Z}_k^i = (\bar{X}_k^i - \bar{\bar{X}}_k) \times I\{\bar{X}_k^i > -A_k\} \quad (31)$$

表 1 相关序列的描述性统计

Table 1 Descriptive statistics of relevant time series

	Daily return		Realized volatility		Multifractal volatility	
	$R_t$	$R_t^2$	RV	hRV	MFV	hMFV
Mean	0.02354	2.003	1.017	-0.501	2.003	0.127
Std	1.414	5.366	1.453	0.981	2.895	1.041
Skewness	0.580**	8.676***	5.245**	0.222**	4.777***	0.127*
Kurtosis	5.175**	103.118***	42.825**	0.132	34.169***	0.248*
JB	2.166**	842.415**	149.767**	16.568**	96.978**	9.712**
$Q(20)$	30.094*	233.483**	4.298.619**	9.468.857**	1.385.762**	4.032.765**

说明: \*, \*\*, \*\*\* 分别代表在 10%、5% 和 1% 水平下显著, 其中 Std 为标准差, 峰态系数 Kurtosis 为超额峰态, JB 为 Jarque-Bera 统计量,  $Q(20)$  为 Ljung-Box  $Q(20)$  统计量.

从表 1 中可以看到, 无论是收益率序列本身, 还是各种波动率的测度及其对数序列, 都表现出较为明显的“有偏”(skewed) 和“尖峰胖尾”(leptokurtic and fat tailed) 形态, 它们的分布都显著拒绝正态分布假设, 且在滞后 20 期的范围之内, 都具有明显的自相关特征.

其中

$$A_k = \frac{1}{4} M^{-4} \hat{\omega}_{kk} \quad (32)$$

而  $I(\cdot)$  是一个指示函数 (indicator function), 即当  $\{\cdot\}$  中的条件成立时, 其取值为 1, 否则取值为 0. 最后, 可以得到如下的实证统计量

$$T^i = \max \frac{\sqrt{M} \bar{Z}_k^i}{\hat{\omega}_{kk}}; i = 1, 2, \dots, B \quad (33)$$

Hansen and Lunde 的研究表明<sup>[14]</sup>, 在式 (26) 所示的零假设条件下, 公式 (33) 所示的实证统计量收敛于公式 (27) 所定义的统计检验  $T$ . 因此, 该统计检验  $T$  的  $p$  值可以直接从下式得出:

$$p = B^{-1} \sum_{i=1}^B I\{T^i > T\} \quad (34)$$

SPA 检验的  $p$  值越大 (越接近于 1), 则表明越不能拒绝公式 (26) 所定义的零假设  $H_0$ : “与其它模型  $M_k$  相比, 基础模型  $M_0$  是表现最优的预测模型.” 即说明该基础模型的预测精度越高.

## 5 实证结果

### 5.1 相关序列的描述性统计

表 1 是对日收益率  $R_t$ 、收益率平方  $R_t^2$ 、实现波动率估计 RV、对数 RV (hRV) 以及多分形波动率 MFV 和对数 MFV (hMFV) 序列的描述性统计结果.

### 5.2 波动率模型参数估计结果

表 2 是在样本总体基础上 (1999 年 1 月 19 日—2006 年 10 月 10 日) 的各类波动率模型参数估计结果. 其中, ARFMA 模型的定义如第 2 节所示, 标准 GARCH 模型的定义如 3.1 节所示, 而随机波动模型 (SV) 的定义如 3.2 节所示.

表 2 基于样本总体的各类波动率模型参数估计结果 ⑤

Table 2 Estimation results of volatility models based on observation population

Model	ARFMA			GARCH		SV
	lnMFV	lnRV				
$\phi_1$	0.191 (0.120)	0.365 (0.488)	$\alpha$	0.068 (0.026)	$\sigma^{*2}$	1.344 (0.049)
$\theta_1$	-0.373 (0.133)	-0.346 (0.483)	$\beta$	0.851 (0.029)	$\phi$	0.935 (0.223)
$d$	0.461 (0.037)	0.487 (0.017)	$\omega$	0.123 (0.026)	$\sigma_{\eta}^2$	0.087 (0.103)
$\sigma_{\varepsilon}^2$	0.842	0.425				
lnL	-2305.79	-1720.84		-3116.15		-3042.76
AIC	4619.58	3449.67		6238.30		6098.01
$Q(20)$	22.638 [0.307]	23.504 [0.265]		26.379 [0.154]		30.489 [0.0623]

说明: 表中数字为各类波动率模型的参数估计结果, 数字下方的圆括号中是估计的标准误差; lnL 是模型估计的对数极大似然函数值, AIC 是赤池信息,  $Q(20)$  是 Ljung-Box  $Q(20)$  统计量, 其下面的方括号中是对应检验的  $p$  值。

从不同模型估计的 lnL, AIC 和  $Q(20)$  检验结果的比较中, 可以清楚地看到, 论文采用的 ARFMA(1, d, 1) 模型, 对 lnMFV 的波动特征具有较好的刻画能力, 这也证明了在 2.2 节中对 lnMFV 的动力学模型假定是合理的。

### 5.3 波动率模型预测结果

各类波动率模型对未来 1 天的市场波动率预测方法如 4.1 节所示。图 3 表示了各种波动模型在预测样本区间内 ( $m = 1, 2, \dots, 183$ ) 的波动率预测结果。其中, 图 3 的 (a) 图表示的是 ARFMA-lnRV 模型和 ARFMA-lnMFV 模型的预测结果 (分别用实线和虚线表示), 而对实际市场波动率基准 RV 的估计则用实心的小方块表示。类似地, 图 3 的 (b) 图表示的是 SV 以及 GARCH 模型的预测结果。

表 3 是一般的单一样本预测检验结果, 其中 6 种损失函数  $L_i$  的定义如 4.2 节所示。表中用粗体加下划线表示的是在某一损失函数  $L_i$  的判断准则下, 得到的最小损失函数值。

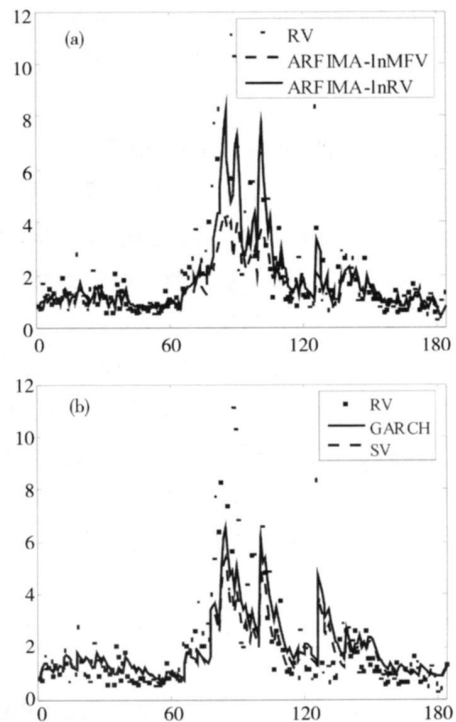


图 3 不同波动率模型在预测样本区间的预测结果

Fig. 3 Forecasting results of different volatility models during forecasting period

⑤ 文中对各类波动率模型的估计是在 OX 编程语言环境下实现的, 关于 OX 语言可以参见 Doomik, J. A. (2002), Object-Oriented Matrix Programming Using OX, 3rd ed. London: Timberlake Consultants Press and Oxford. www.nuff.ox.ac.uk/users/Doomik.



表3 单一样本的各类波动率模型预测损失函数值

Table 3 Loss function values for different volatility forecasting models in a single sample

Model	Loss function $L_i$					
	MSE	MAE	HMSE	HMAE	QLKE	R <sup>2</sup> LOG
ARFMA-lmFV	4.441	0.966	<b>0.201</b>	<b>0.365</b>	2.869	0.280
ARFMA-lnRV	<b>2.914</b>	<b>0.853</b>	0.243	0.381	<b>2.787</b>	<b>0.223</b>
SV	3.768	0.953	0.351	0.443	2.944	0.274
GARCH	3.217	0.946	0.503	0.503	3.917	0.281

说明:表中用粗体加下划线表示的是在某一损失函数判断准则下的最小损失值。

从表3的实证结果可以看到,对所考察的这个单一样本来讲,ARFMA-lnRV和本文提出的ARFMA-lmFV模型是较为合适的波动率模型选择。具体来讲,在6种损失函数标准下,ARFMA-lnRV模型有4种获得了最小损失函数值,而ARFMA-lmFV也有2种得到了最好的预测精度(即HMSE和HMAE:经异方差调整的MSE和MAE)。但是正如在4.2节中所讨论的那样,如果要得到更加稳健和适用范围更广的结论,则必须对这一预测结果进行进一步的SPA检验。

#### 5.4 波动率模型的SPA检验结果

表4是SPA检验结果<sup>⑥</sup>,表4的第1列表示的是6种损失函数 $L_i$ ,第2列是被选作基础模型(Base model  $M_0$ )的模型名称,表中数字为SPA检验的 $p$ 值。 $p$ 值越大(越接近于1),则表明越不能拒绝公式(26)所定义的零假设 $H_0$ :“与其它模型 $M_k$ 相比,基础模型 $M_0$ 是表现最优的预测模型。”这也就是说,在某一损失函数 $L_i$ 判断标准下,如果基础模型 $M_0$ 相对于其它模型的SPA检验 $p$ 值越大,则表明该模型的预测精度越高。反之,如果 $p$ 值很小(一般为小于10%),则有理由相信,基础模型 $M_0$ 的表现要劣于所考察的对比模型 $M_k$ 。从表中的实证结果可以看出:

1) 以上证综指为例,总体来说,基于高频收益数据的ARFMA-lnRV模型是一种预测精度相当高的波动率模型,但在HMSE和HMAE这两种经过异方差调整的损失函数标准下,提出的基于高频价格数据的多分形波动率模型—

ARFMA-lmFV却是预测精度最高的波动率模型。具体表现为,在这两种损失函数标准下,作为基础模型 $M_0$ 的ARFMA-lmFV的SPA检验的 $p$ 值都近似为1(表中用粗体加下划线表示之)。

2) 基于日数据(Daily data)的SV模型似乎比GARCH模型的预测表现稍好一些,在6种损失函数标准下,SV模型有4种优于GARCH模型,这也与目前国内学者的一些研究结果类似。

3) 总体来讲,基于高频数据的波动率模型(RV和MFV模型)比基于日数据的波动率模型(GARCH和SV)具有更好的波动率预测表现。这也证明了与日数据相比,高频数据当中确实蕴含着更加丰富的市场波动信息。

4) 表4中SPA检验得到的结论与表3得到的结论基本一致,因此这些结论对中国股票市场的实际波动特征而言,具有相当的普遍性。

## 6 结束语

自从分形理论的创始人Mandelbrot首次提出multifractal是一种刻画金融市场异常波动特征的有力工具以来<sup>[1]</sup>,许多学者都运用多分形理论对各种不同类型金融市场的波动现象进行了实证分析。但是,如何充分运用多分形分析所产生的丰富统计信息来帮助我们对金融市场波动规律的深入理解和把握,仍然是这一领域研究中的难点之一。

本文在前期研究的基础之上<sup>[3-5]</sup>,提出了一种新的基于高频股价数据的市场波动率测度方法

⑥ 作者感谢斯坦福大学经济系的P. R. Hansen博士授权使用并提供了基于OX编程语言的SPA检验程序。

——多分形波动率 (MFV) 测度, 并相应构造了其 ARFMA 动力学模型. 与主流的 ARCH /GARCH 族模型、SV 模型以及 RV 模型对市场波动率的直接刻画方式不同, 多分形波动率的测度方法是运用多分形分析中产生的对市场波动刻画的间接统计信息, 从多分形谱的参数中提炼出市场波动率的定量测度数据. 进一步, 为了验证这一新的波动

率测度方法及其模型在实际金融市场中的有效性和实用性, 进一步运用 Hansen and Lunde 提出的对波动率模型优劣判断的 SPA 检验法<sup>[14]</sup>, 实证对比了其与现有的其它波动率模型, 如基于高频收益数据的 RV 模型以及基于日数据的 SV 和 GARCH 模型对中国股市波动率的刻画和预测能力.

表 4 不同波动率模型预测结果的 SPA 检验结果

Table 4 SPA results for different volatility forecasting results

$L_i$	Basemodel $M_0$	Alternative models $M_k$			
		ARFMA-hMFV	ARFMA-hRV	SV	GARCH
MSE	ARFMA-hMFV	—	0.031	0.040	0.041
	ARFMA-hRV	0.969	—	0.967	0.929
	SV	0.960	0.033	—	0.045
	GARCH	0.959	0.071	0.955	—
MAE	ARFMA-hMFV	—	0.043	0.358	0.372
	ARFMA-hRV	0.957	—	0.970	0.984
	SV	0.642	0.030	—	0.408
	GARCH	0.628	0.016	0.592	—
HMSE	ARFMA-hMFV	—	<b>0.955</b>	<b>0.990</b>	<b>0.992</b>
	ARFMA-hRV	0.045	—	0.972	0.988
	SV	0.010	0.028	—	0.992
	GARCH	0.008	0.012	0.008	—
HMAE	ARFMA-hMFV	—	<b>0.854</b>	<b>0.995</b>	<b>0.999</b>
	ARFMA-hRV	0.146	—	0.980	1.000
	SV	0.005	0.020	—	0.999
	GARCH	0.001	0.000	0.001	—
QLIKE	ARFMA-hMFV	—	0.005	0.017	0.020
	ARFMA-hRV	0.995	—	0.954	0.817
	SV	0.983	0.046	—	0.064
	GARCH	0.980	0.183	0.980	—
$R^2$ LOG	ARFMA-hMFV	—	0.019	0.415	0.533
	ARFMA-hRV	0.981	—	0.989	0.985
	SV	0.585	0.011	—	0.664
	GARCH	0.467	0.015	0.336	—

说明: 表中数字为经过 1000 次 bootstrap 过程后的 SPA 检验的  $p$  值,  $p$  值越大, 表明与所考察的其它模型  $M_k$  相比, 基础模型  $M_0$  的表现越好.

实证结果显示, 在某些损失函数标准下, 提出的多分形波动率测度及其模型对市场波动特征具有优异的刻画能力和预测精度. 同时, 就中国股市而言, 基于高频数据的波动率模型 (RV 和 MFV

模型) 比基于日收益数据的 SV 和 GARCH 模型具有更高的波动率预测精度. 因此, 如何充分利用高频数据中蕴含的丰富的市场波动信息来指导金融学研究工作, 将是一个非常有意义的发展

方向.

论文提出的这一新的市场波动率测度方法及其模型,丰富了现有对市场波动率的研究手段,为进一步探究金融市场的真实波动特征提供了一种新的视角和实证方法.但还需指出的是,虽然在某些损失函数标准下,本文提出的MFV模型具有最

高的波动率预测精度,但在其它损失函数标准下,该模型的表现并不出色.因此,如何进一步调整MFV的ARFMA( $p, d, q$ )模型参数,以改善其对市场波动状况的描述能力,并进一步寻找更加适合中国股市实际风险特征的波动率模型,仍然是下一步研究的主要方向.

## 参考文献:

- [1] Mandelbrot B B. A multifractal walk down Wall Street [J]. *Scientific American*, 1999, 298(1): 70—73.
- [2] 何建敏, 常松. 中国股票市场多重分形游走及其预测 [J]. *中国管理科学*, 2002, 11(3): 11—17.  
He Jianming, Chang Song. Chinese stock markets multifractal walk and forecasting [J]. *Chinese Management Science*, 2002, 11(3): 11—17. (in Chinese)
- [3] 魏宇, 黄登仕. 金融市场多尺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)
- [4] 魏宇, 黄登仕. 基于多尺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)
- [5] Wei Y, Huang D-S. Multifractal analysis of SSEC in Chinese stock market: A different empirical result from Heng Seng index [J]. *Physica A*, 2005, 355(5): 497—508.
- [6] 张永东, 毕香秋. 中国股票市场多尺度行为的实证分析 [J]. *预测*, 2002, 18(4): 56—59.  
Zhang Yongdong, Bi Xiangqiu. Empirical analysis of multiscale behavior of Chinese stock market [J]. *Forecast*, 2002, 18(4): 56—59. (in Chinese)
- [7] 朱林, 常松, 何建敏. 我国股票市场多仿射特性研究 [J]. *管理工程学报*, 2002, 22(3): 86—89.  
Zhu Lin, Chang Song, He Jianmin. Multi-affine property study of our country stock market [J]. *Journal of Management Engineering in China*, 2002, 22(3): 86—89. (in Chinese)
- [8] Plerou V, Gopkrishnan P, Amaral A, et al. Scaling of the distribution of price fluctuations of individual companies [J]. *Phys Rev E*, 1999, 60(3): 6519—6529.
- [9] Gopkrishnan P, Plerou V, Amaral A, et al. Scaling of the distributions of fluctuations of financial market indices [J]. *Phys Rev E*, 1999, 60(3): 5305—5316.
- [10] Schmitt F, Schertzer D, Lovejoy S. Multifractal fluctuations in finance [J]. *Int J Theor Appl Fin*, 2000, 3(3): 361—364.
- [11] Bacry E, Delour J, Muzy F. Modeling financial time series using multifractal random walks [J]. *Physica A*, 2001, 299(5): 84—92.
- [12] 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.
- [13] 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.
- [14] 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(1): 873—889.
- [15] Andersen T G, Bollerslev T. Answering the skeptics: Yes, standard volatility models do provide accurate forecasts [J]. *Int*

- temational Economic Review, 1998, 39(3): 885—905.
- [16] 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.
- [17] Hansen P R, Lunde A. Consistent ranking of volatility models[J]. *Journal of Econometrics*, 2006, 131(2): 97—121.
- [18] 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.
- [19] Bollerslev T. Generalized autoregressive conditional heteroskedasticity[J]. *Journal of Econometrics*, 1986, 31(6): 307—327.
- [20] Taylor S J. *Modeling Financial Time Series*[M]. Chichester: John Wiley and Sons, 1986.
- [21] Shephard N. Statistical aspects of ARCH and stochastic volatility[A]. In: Cox D. R., Ingle D. V., Bamdorff Nielsen O. E. (Eds.), *Time Series Models in Econometrics, Finance and Other Fields*, Number 65 in Monographs on Statistics and Applied Probability[M]. London: Chapman and Hall, 1996, 1—67.
- [22] Sandmann G, Koopman S J. Estimation of stochastic volatility models via Monte Carlo maximum likelihood[J]. *Journal of Econometrics*, 1988, 87(5): 271—301.
- [23] White H. A reality check for data snooping[J]. *Econometrica*, 2000, 68(3): 1097—1126.

## Multifractal volatility measure, its model and SPA test in financial market

WEI Yu

School of Economics & Management, Southwest Jiaotong University, Chengdu 610031, China

**Abstract** In this paper, a new volatility measure, multifractal volatility, is constructed. Based on about 8 years' high-frequency data of SSE, we choose ARFMA model as the dynamic model of multifractal volatility and use SPA by Hansen and Lunde (2005) to test the predicting performance of the multifractal model and other popular models, such as Realized volatility model, GARCH and Stochastic volatility model. The empirical results show that for some kinds of loss functions, multifractal volatility measure and its ARFMA model have better predicting performance than other existing volatility models.

**Key words** multifractal volatility; realized volatility; volatility measures; predicting SPA test