

基于随机波动性模型的中国股市波动性估计

王春峰, 蒋祥林, 李刚

(天津大学管理学院、天津大学金融工程研究中心, 天津 300072)

摘要:采用动态随机波动性模型实证研究了中国股票市场的波动性. 通过基于马尔可夫链蒙特卡罗(MCMC)模拟的贝叶斯分析方法, 较好地估计了随机波动性模型中的参数与波动性序列. 基于中国股市数据进行的实证结果表明, 与ARCH类模型相比, 随机波动性模型能更好地描述股票市场回报的异方差和波动性的序列相关性.

关键词:随机波动性模型; 贝叶斯分析; 马尔可夫链蒙特卡罗(MCMC); ARCH模型

中图分类号: F830.91 **文献标识码:** A **文章编号:** 1007-9807(2003)04-0063-10

0 引言

波动性是金融市场最为重要特性之一. 一方面它与市场的不确定性和风险直接相关, 是体现金融市场质量和效率的最简洁和最有效的指标之一; 另一方面波动性对企业的投资与财务杠杆决策, 消费者的消费行为和模式, 经济周期及相关宏观经济变量等都具有重要影响. 此外波动性也是证券组合理论、资本资产定价模型(CAPM)、套利定价模型(APT)及期权定价公式的核心变量. 因此, 波动性的估计和预测一直是金融经济学研究的持续热点^[1].

针对股票市场波动性表现出的时变特点与“集聚效应(clustering effect)”, Engle、Bollerslev等提出了ARCH类模型^[2,3]. ARCH类模型采用时变的条件方差来捕捉价格波动的时变性和序列相关性, 但由于ARCH模型将条件方差定义为过去观测值的平方项和前期条件方差的确定性函数, 条件方差的估计与过去观测值直接相关, 因此当存在异常观测值时, 估计的波动性序列不是很稳定. 另外, ARCH模型对于长期波动性的预测能力较差.

为此, 一些学者提出了随机波动性模型^[4-7]. 随机波动性模型与金融资产定价理论中扩散过程有着直接的联系, 假定时变方差遵循某种不可观测的随机过程, 并不依赖于过去观测值, 因此其估计的波动性序列比ARCH模型更加稳定. 此外, 随机波动性模型比ARCH模型可更好地描述金融数据自身固有的尖峰性. 大量实证研究表明^[8-10], 随机波动性模型比ARCH模型能更好地描述金融数据的特性, 特别是对两个模型预测的均方误差和SC信息指标的比较以及模型的残差分析发现, 随机波动性模型比GARCH模型具有更好的预测能力, 尤其是对长期波动性的预测.

随机波动性模型只是由于其估计的困难性才一直没有受到重视. 从计量经济学的角度看, 随机波动性模型在实践上的困难在于不能直接得到模型的似然函数与模型的无条件矩的解释形式, 这是因为下文式(1)中均值方程的方差是一个潜在的、非直接观察到的变量并且模型是非正态的, 似然函数和无条件矩只能通过高维积分得到. 近年来, 在随机波动性的模型估计方面已经取得了极大的进展. 估计技术可以分为三大类: 第1类方法是基于传统的参数估计方法, 用近似的方法或者模拟

收稿日期: 2001-10-11; 修订日期: 2002-12-12.

基金项目: 国家杰出青年科学基金项目(70225002); 国家自然科学基金资助项目(70041039); 教育部跨世纪优秀人才基金资助项目; 教育部优秀青年教师教学科研奖励基金资助项目.

作者简介: 王春峰(1966—), 男, 河北省人, 博士, 教授, 博士生导师.

的方法构造模型的似然函数和无条件矩.前者包括拟最大似然估计(QML)、广义矩估计(GMM)等,后者包括模拟最大似然估计(SML)、模拟矩估计(SMM)等.第2类方法通过引入一个辅助模型(如ARMA、GARCH模型)或者半参数方法间接地估计随机波动性模型,包括非直接推断(indirect inference)、有效矩估计(EMM)等.第3类方法是基于贝叶斯原理的参数后验分布分析,但因为高维积分的原因,参数的后验均值和标准差的计算极为困难.随着马尔可夫链蒙特卡罗(MCMC)模拟技术的发展和计算机计算能力的提高,后验分布计算上的困难得以克服.然而,MCMC方法对计算能力的要求,使用非常规的软件,使得研究者和从业人员实施起来极为困难.金融时间序列研究与应用的焦点是波动性序列本身,在以上估计方法中,有些能有效地估计模型的参数,但不适合估计模型的潜在波动性序列.作者在一个较为成熟的软件包WINBUGS(Spiegelhalter等,1996)的基础上,利用SV模型层次结构的特点,采用贝叶斯分析中的增参数技术,构造出了SV模型的后验分布,使用Gibbs抽样对SV模型参数和波动性的后验分布进行数值积分,使得基于MCMC的SV模型的参数和波动性序列估计极为方便.

本文的目的在于引入动态随机波动性模型来描述我国股票市场的波动性,并与传统的ARCH类模型进行对比分析.

1 随机波动性模型

实证研究表明,资产价格变化的方差随时间而变化,其价格变化的分布类似于对数正态分布.假定 t 时资产价格为 P_t , 则资产价格的对数回报 $r_t = 100\ln(P_t/P_{t-1})$ 可表示为

$$r_t \sim N(0, \sigma_t^2)$$

大量实证研究还发现在 r_t 过程中存在序列相关性^[11-15], 为此人们引入了不同的模型来描述这种相关性:ARCH类模型假定 r_t 是资产价格过去观测值的非线性函数,而随机波动性模型则假定 r_t 是一个单独的随机过程,可能包含资产价格的过去值,但存在一个不同的、独立的误差结构.由于可以在随机过程中包含ARCH类模型,因

此把 r_t 表示为一个特定的随机过程有很大的吸引力.

最初的随机波动性模型将 r_t 表示为具有独立同分布的对数正态过程^[4],但大量研究(如Wiggins关于期权定价的研究^[16]、Karpoff关于量价关系的研究^[17])指出,波动性存在持续的均值回复现象,为了刻画这种现象,Danielsson^[6]等人假定对数波动性服从一阶自回归过程.动态随机波动性模型结构为

$$\begin{aligned} r_t &= \mu_t, t = 1, \dots, T \quad \mu_t \sim N(0, 1) \\ \ln(\sigma_t^2) &= \alpha + \beta \ln(\sigma_{t-1}^2) + v_t, v_t \sim N(0, 1) \end{aligned} \quad (1)$$

上述结构包括了均值方程和波动性方程.其中, T 是观测值的数目.要估计的参数向量可以表示为 $\theta = (\alpha, \beta, \nu)$. 记 $R_T = \{r_t\}_{t=1}^T$, r_t 是潜在变量,记 $\sigma_T = \{\sigma_t\}_{t=1}^T$.

对于股票数据,通常认为市场价格向下的冲击比向上的冲击会导致更高的波动性,人们将这种现象称为波动性的非对称性,或“杠杆效应”.为了测试这种现象,Engle和Victor给出了一种不对称的消息冲击曲线^[22](asymmetric news impact curve),见图1所示.为了使GARCH模型能够描述波动性的这种非对称性,Nelson于1991年提出了EGARCH模型,Nelson指出GARCH模型中对系数参数的非负性约束太强,过度地限制了条件方差的动态性,而在EGARCH模型中这些参数不受此约束. EGARCH模型是在波动性方程中包含了资产价格变化冲击的滞后项,本文将这种方式引入随机波动性模型描述波动性的非对称性,模型可以表示为

$$\begin{aligned} r_t &= \mu_t, t = 1, \dots, T \quad \mu_t \sim N(0, 1) \\ \ln(\sigma_t^2) &= \alpha + \beta_1 \mu_{t-1} / \sigma_{t-1} + \beta_2 |\mu_{t-1} / \sigma_{t-1}| + \ln(\sigma_{t-1}^2) + v_t, v_t \sim N(0, 1) \end{aligned} \quad (2)$$

模型的左侧是条件方差的对数,这说明“杠杆效应(leverage effect)”是指数形式的,而不是二次的,同时这也保证了条件方差的非负性.如果 $\beta_2 > 0$,则冲击是非对称的;如果 $\beta_2 < 0$,则样本数据具有“杠杆效应”.将式(1)对应的模型简称为SV模型,式(2)对应的模型简称为SV- μ 模型.

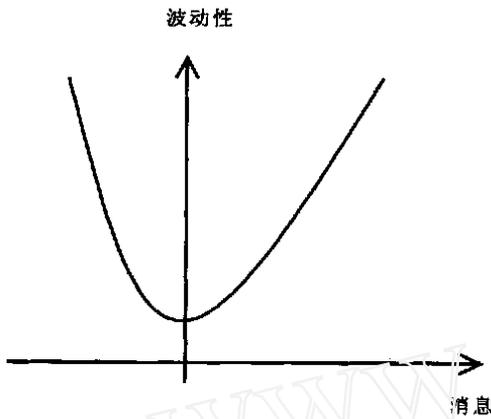


图 1 消息冲击曲线

2 随机波动性模型的估计方法:基于 MCMC 的贝叶斯分析

2.1 随机波动性模型的参数与波动性的贝叶斯推断

首先分析 SV 模型. 波动性模型中的参数定义为 $\theta = (\alpha, \beta, \nu)$. 采用贝叶斯分析中的增参数方法, 将潜在的波动序列 $\sigma_t = \{ \sigma_i \}_{i=0}^T$ 增加为要估计的参数, 因此, 参数个数总计为 $(4 + T)$. 可将模型看成由层次结构的条件分布组成: $f(R_T | \sigma_T, \theta), f(\sigma_T | \theta), f(\theta)$. $f(R_T | \sigma_T, \theta)$ 的含义是指在波动性过程和参数给定情况下, 样本的条件密度; $f(\sigma_T | \theta)$ 是对波动性过程的一个先验的信息, 采用了模型(1)的对数正态的 AR(1) 过程; 而 $f(\theta)$ 反映了波动性模型参数的先验分布, θ 的先验信息选择要保证 $\ln(\sigma_t^2)$ 是平稳过程. $f(\sigma_T, \theta)$ 为 σ_T, θ 的联合先验分布, 则

$$f(\sigma_T, \theta) = f(\theta) f(\sigma_T | \theta) \quad (3)$$

其中

$$f(\sigma_T, \theta) = f(\alpha, \beta, \nu) f(\sigma_T | \alpha, \beta, \nu) = f(\alpha) f(\beta) f(\nu) f(\sigma_0) \prod_{t=1}^T f(\sigma_t | \sigma_{t-1}, \alpha, \beta, \nu) = f(\alpha) f(\beta) f(\nu) f(\sigma_0) \times \prod_{t=1}^T \left[\frac{\sqrt{2}}{\nu \sqrt{\sigma_t}} \exp \left\{ - \frac{1}{2\nu^2} (\ln \sigma_t^2 - \mu_{t-1} \nu / \sigma_{t-1} - 2 / \mu_{t-1} / \sigma_{t-1} - \ln \sigma_{t-1}^2)^2 \right\} \right] \quad (9)$$

$$f(\sigma_T | \theta) = f(\sigma_0) \prod_{t=1}^T f(\sigma_t | \sigma_{t-1}, \theta) = f(\sigma_0) \prod_{t=1}^T \left[\frac{\sqrt{2}}{\nu \sqrt{\sigma_t}} \exp \left\{ - \frac{1}{2\nu^2} (\ln \sigma_t^2 - \mu_{t-1} \nu / \sigma_{t-1} - 2 / \mu_{t-1} / \sigma_{t-1} - \ln \sigma_{t-1}^2)^2 \right\} \right] \quad (4)$$

给定 σ_T 下, 样本 R_T 的条件分布为

$$f(R_T | \sigma_T) = f(r_T | \sigma_T) = \prod_{t=1}^T \left[\frac{1}{\sigma_t \sqrt{2}} \exp \left\{ - \frac{r_t^2}{2\sigma_t^2} \right\} \right] \quad (5)$$

因此, 根据贝叶斯原理, 参数 θ 、波动性序列 σ_T 的联合后验分布与条件分布和先验分布的乘积成正比, 即

$$f(\theta, \sigma_T | R_T) = K \times f(\theta) f(\sigma_0) \prod_{t=1}^T f(\sigma_t | \sigma_{t-1}, \theta) \times \prod_{t=1}^T f(r_t | \sigma_t) \quad (6)$$

其中, K 为归一化参数. 样本 R_T 与 σ_T 的联合分布为

$$f(R_T, \sigma_T) = f(R_T | \sigma_T) f(\sigma_T) \quad (7)$$

样本似然函数

$$f(R_T) = \int f(R_T, \sigma_T) d\sigma_T \quad (8)$$

以与 Kim 等(1998) 相同的方式设定参数的先验分布, 假定参数的先验分布是独立的. 假定 $\alpha = 2^* - 1$, 其中 $*$ 服从贝塔分布 $*$ \sim Bet(20, 1.5), 这样 α 的先验均值是 0.86; β 服从正态分布 $\beta \sim N(0, 100)$, 假定 $\ln(\nu)$ 的先验分布为正态分布 $\ln(\nu) \sim N(-2.49, 0.73)$, 其先验均值和先验标准差分别为 0.2 和 0.125.

只需对上述的过程做一定的改变, 就能建立起 SV- Γ 模型的后验分布. 假定先验分布 $\sigma_1 \sim N(0, 100)$, $\sigma_2 \sim N(0, 100)$; $\sigma_T, \sigma_1, \sigma_2$ 的联合先验分布为

在构造公式(3)、(9)的密度时, 隐含的假定 σ 和 r 在 0 时刻的值存在, 初始的密度也可以假定为是长期的稳态密度, 或者设置为 0, 这不会使估计结果产生很大的差异.

其中: $\mu_{t-1} = r_{t-1} / r_{t-1}$. 这样可以建立起参数和波动性序列的后验分布 $f(\theta, \sigma^2, \tau | R_T)$.

将 $g(\theta, \sigma^2, R_T)$ 定义为模型参数、波动性序列和样本数据的函数, 则 $g(\theta, \sigma^2, R_T)$ 的后验均值为

$$E\{g(\theta, \sigma^2, R_T)\} = \int g(\theta, \sigma^2, R_T) f(\theta, \sigma^2, \tau | R_T) d\theta d\sigma^2 d\tau \quad (10)$$

选择 $g(\theta, \sigma^2, R_T)$ 可以计算出模型参数和波动性的后验均值、标准差和样本对数似然函数等统计值.

从以上的分析可知, 对 SV 模型后验分布的计算将是一个高维的积分问题, 例如, 对于样本量为 1 000 的 SV 模型, 参数和波动性的后验均值的计算将是一个 1 004 维的积分. 对于这样的高维积分问题, 用数值分析的方法是可能实现的, 而传统的模拟方法的精度与效率是极差的. 针对这种情况, 本文采用了基于马尔可夫链蒙特卡罗 (Markov chain Monte Carlo, MCMC) 模拟的方法计算 SV 模型参数和波动性的后验分布.

2.2 基于 MCMC 的后验分布的模拟

MCMC 方法最初应用于计算物理 (Metropolis 等, 1953), Hastings (1970) 的工作使其更为一般化, 但主要应用于统计问题. Gelfand 和 Smith (1990) 的研究显示出 MCMC 方法在贝叶斯计算上有着巨大的潜力, 他们将马尔可夫链的方法与 Tanner 和 Wong (1987) 的增参数方法相结合, 证明了对含有潜在动态变量的计量模型的估计是极为成功. Gibbs 抽样和 Hastings-metropolis 算法是两种最为常用的 MCMC 方法. 本文使用了 Gibbs 抽样的方法模拟 SV 模型的后验分布.

Gibbs 抽样的方法已经成功地应用于大量的计量模型, 其内在的思想是通过在参数向量 θ 的状态空间 Θ 上构造出一个具有唯一的固定分布 $f(\theta | R_T)$ 的马尔可夫链 $\{\theta^{(k)}\}_{k=1}^N$, 取已经进入收敛的状态的 $\{\theta^{(k)}\}_{k=m}^N$ 来近似计算 $E(g(\theta))$, 即

$$E(g(\theta | R_T)) = \frac{1}{N - m + 1} \sum_{k=m}^N g(\theta^{(k)} | R_T) \quad (11)$$

其中: $E(g(\theta))$ 可定义为参数 g 基于样本 R_T 的后验均值和后验标准差等.

Gibbs 抽样首先将参数 θ 分成 B 块: $\theta = (\theta^{(1)}, \dots, \theta^{(B)})$.

在应用中, 应该通过适当的分块使得从每一个条件分布 $f(\theta^{(b)} | \theta^{(1)}, \dots, \theta^{(b-1)}, \theta^{(b+1)}, \dots, \theta^{(B)}, R_T)$ 中的抽样成为可能. 下面写出分块 Gibbs 抽样的具体步骤:

- 1) 确定初始点 $\theta^{(0)} = (\theta^{(0)}_{(1)}, \dots, \theta^{(0)}_{(B)})$, 设 $i = 0$,
- 2) 按以下的方式从条件分布中抽样

$$\theta^{(i+1)}_{(1)} \sim f(\theta^{(i+1)}_{(1)} | \theta^{(i)}_{(2)}, \dots, \theta^{(i)}_{(B)}; R_T)$$

.....

$$\theta^{(i+1)}_{(b)} \sim f(\theta^{(i+1)}_{(b)} | \theta^{(i+1)}_{(1)}, \theta^{(i+1)}_{(2)}, \dots, \theta^{(i+1)}_{(b-1)}, \dots, \theta^{(i)}_{(B)}; R_T)$$

.....

$$\theta^{(i)}_{(B)} \sim f(\theta^{(i)}_{(B)} | \theta^{(i+1)}_{(1)}, \theta^{(i+1)}_{(2)}, \dots, \theta^{(i+1)}_{(B-1)}; R_T)$$

这样就定义了一个从 $\theta^{(i)}$ 到 $\theta^{(i+1)}$ 的转移过程.

- 3) $i = i + 1$, 转到第 2) 步.

上述的算法产生了一个向量序列 $\theta^{(0)}, \theta^{(1)}, \dots, \theta^{(i)}, \dots$ 作为马尔可夫链的一个实现过程, 从 $\theta^{(i)}$ 到 $\theta^{(i+1)}$ 的转移密度是

$$K_G(\theta^{(i)}, \theta^{(i+1)}) = \prod_{b=1}^B f(\theta^{(i+1)}_{(b)} | \theta^{(i+1)}_{(1)}, \theta^{(i+1)}_{(2)}, \dots, \theta^{(i+1)}_{(b-1)}, \dots, \theta^{(i)}_{(B)}) \quad (12)$$

在本文中, 上述的后验分布的构造和 MCMC 后验分布模拟都是基于一个较为成熟的软件包 WINBUGS (Spiegelhalter 等, 1996) 上实现的, 基于 WINBUGS 不仅能估计出模型的基本的参数 θ , 并且可以得到波动性序列 σ^2_t 、对数似然函数值 $\ln(L)$ 等一系列统计值.

3 实证结果分析

3.1 数据特性描述

在实证研究中使用了上证综合指数 1991 年 1 月—2001 年 12 月和深证成分指数 1991 年 4 月—2001 年 12 月的收盘价的日数据, 样本量分别是 2 726 和 2 692. 两种股票价格指数都以连续复利的形式转化为价格回报序列: $r_t = 100 \ln(p_t / p_{t-1})$. 回报序列见图 1, 统计特性总结为表 1, 回报和回报平方序列自相关特性总结为表 2.

据表 1 所示, 上证与深证指数的回报序列都

有较大的正的偏度,这与文献中所报告的大多数的指数回报序列具有负向偏度不同;两种序列的峰度统计量都大于 3,这表明回报序列的分布具有尖峰厚尾的特性;正态分布统计量很大,拒绝正

态分布的假定;ARCH 效应测试表明,回报序列存在显著的 ARCH 效应.测试自相关程度的 Box-Pierce Q 统计量表明,回报序列不存在明显的自相关,而回报平方序列存在着显著的自相关.

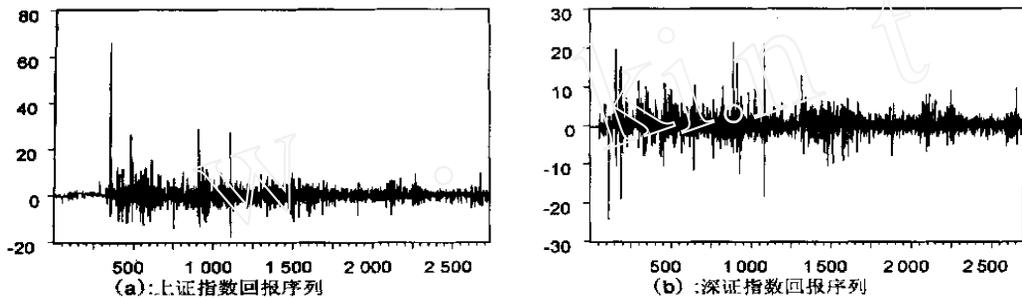


图 2 股指回报序列

表 1 回报序列的统计特性

	标准偏差	偏度 (Skew)	峰度 (Kurtosis)	正态统计量 (J - B)	ARCH 效应
上证	3.058	1.350	20.939	35.389 (0.000)	166.226 (0.000)
深证	2.600	0.504	16.868	12.694 (0.000)	81.943 (0.000)

注:() 内为显著概率水平.

表 2 回报和回报平方序列的自相关性

	回报序列			回报平方序列		
	$Q(5)$	$Q(10)$	$Q(20)$	$Q(5)$	$Q(10)$	$Q(20)$
上证	31.716 (1.000)	44.319 (1.000)	76.960 (0.000)	599.030 (0.000)	734.423 (0.000)	813.659 (0.000)
深证	33.842 (1.000)	44.237 (1.000)	58.680 (0.000)	118.570 (0.000)	288.542 (0.000)	300.380 (0.000)

注:() 内为显著概率水平.

3.2 SV 模型实证结果分析

在本文的实证研究中,只估计了不具有“杠杆效应”的随机波动性模型.表 3 给出了 SV 模型的 MCMC 后验分布模拟的结果.通过参数后验均值的 MC 误差、标准差与参数后验均值大小的比较,可以对后验均值的 MCMC 模拟的精确性做出评价,虽然上证指数和深证指数分别为 2 729 和 2 696 的高维积分,但 MC 误差、标准差与参数后验均值相比,MC 误差是相当的小,并且标准差远大于 MC 误差.因此, MCMC 后验均值估计的结果是相当精确的.上证指数和深证指数的波动性过程的持续性参数显著的大于 0.92,结果表明波动性冲击是高度持续

的,这与国外文献对 SV 模型和 ARCH 类模型的实证研究的结果是一致的.图 3a、图 4a 分别为上证指数和深证指数回报的波动性序列.

表 4 和图 3—图 6 给出了根据随机波动性模型的均值方程的残差序列 $\mu_t = r_t / \sqrt{h_t}$ 诊断分析结果.从表 4 可以看出,经过 SV 模型拟合后,与原始数据相比,SV 模型的均值方程的残差的偏度和峰度都明显降低,标准偏差接近 1,正态分布统计量也明显降低,残差序列与正态分布的差别大大减小;ARCH 效应测试表明,经 SV 模型拟合后,残差的 ARCH 效应不再显著,自相关程度测试表明,残差序列不存在自相关,残差平方的自相关性也

不再显著.从图3和图5的残差序列、残差序列柱形图、残差 $Q-Q$ 图可以得出同样的结论,即经SV模型拟合后,模型残差序列较原始数据的正态性得到了加强,在 $Q-Q$ 图上,代表残差序列的曲线大部分位于对角线上(代表正态分布);在极端的

情况下,残差序列与正态分布有明显的差别,表明有必要使用极值理论模型或者对假定残差服从后尾分布(如 t 分布)的模型来描述回报序列的波动性特征.总之,SV模型能较好地拟合中国股票市场回报的波动性特征.

表3 SV模型的MCMC后验分布模拟的结果

参数			ν	$\ln(L)$
上证综指	0.035 6	(0.955 0)	0.452 9	- 5 889.504
	(0.011 0)	(0.007 6)	(0.027 0)	
	[2.846E - 4]	[3.048E - 4]	[0.001 6]	[1.023 6]
深证成指	0.071 9	(0.924 8)	0.512 2	- 5 981.256
	(0.015 0)	(0.011 1)	(0.031 6)	
	[4.692E - 4]	[4.87E - 4]	[0.001 9]	[1.307 5]

注:() 中为点估计的标准差,[] 中为点估计的MC样本标准差.估计是基于Gibbs抽样.

表4 SV模型和GARCH(1,1)模型均值方程残差比较

		上证指数		深证指数		
		SV模型	GARCH模型	SV模型	GARCH模型	
标准偏差		0.985 2	0.994 5	0.983 7	0.999	
偏度		0.041 1	0.773 4	- 0.014 2	- 0.313 5	
峰度		2.628	26.651	2.533 6	36.198	
正态统计量(J-B)		16.438 3 (0.000)	32 767.04 (0.000)	24.618 (0.000)	35 643.982 (0.000)	
ARCH效应		1.258 (0.321 2)	1.563 (0.175 3)	0.736 7 (0.472 5)	0.876 (0.437 6)	
量计统Q e e c r e i P - x o B	Residuals	$Q(5)$	12.568 (1.000)	75.928 (0.000)	7.329 (1.000)	46.629 (0.000)
		$Q(10)$	16.965 (1.000)	95.536 (0.000)	13.783 (1.000)	52.774 (0.000)
		$Q(20)$	41.032 (0.003)	229.592 (0.000)	18.628 (0.091)	63.359 (0.000)
	Squared residuals	$Q(5)$	6.126 (0.125 7)	0.500 3 (1.000)	3.527 (0.625 9)	0.238 4 (1.000)
		$Q(10)$	17.283 (0.052 7)	1.594 0 (1.000)	8.451 (0.565 2)	0.506 4 (1.000)
		$Q(20)$	26.924 (0.189 5)	3.236 6 (1.000)	16.528 (0.589 4)	1.178 0 (1.000)

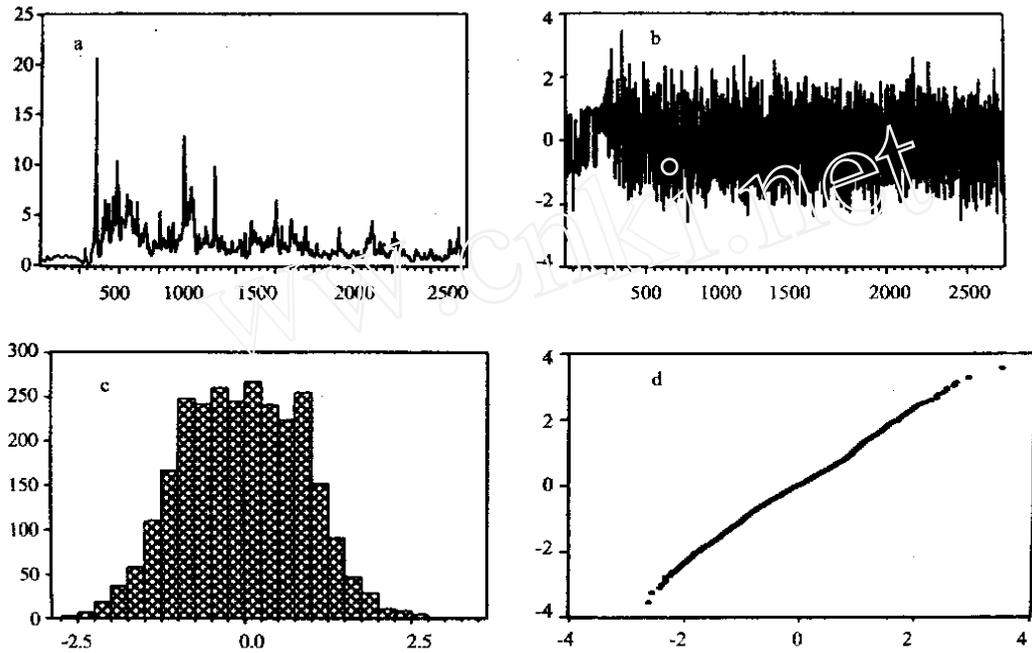
注:() 内为显著概率水平.

3.3 SV模型与GARCH模型的比较

为了说明随机波动性模型的优点,在同样数据集下与ARCH类的估计结果进行了比较.在各种ARCH模型中,分别选择了GARCH(p, q)模型、GR-GARCH和EGARCH等模型进行估计.根据对

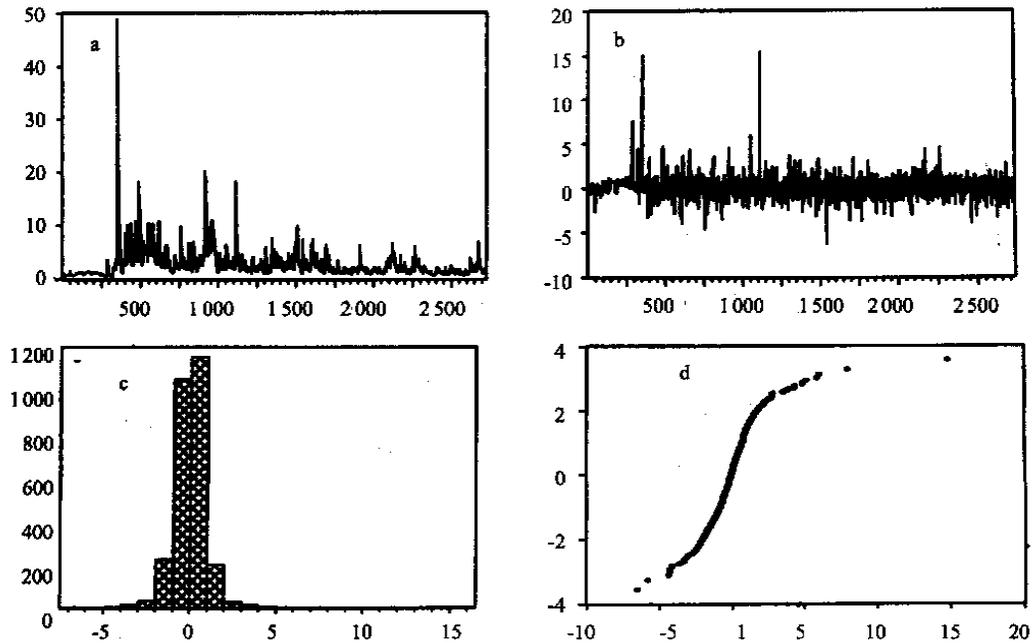
数似然函数和AIC准则,在大部分样本区间中,GARCH(1,1)模型的估计效果优于GR-GARCH、EGARCH等其他模型.因此这里仅给出GARCH(1,1)模型与随机波动性模型的对比结果.由表5给出的从对数似然函数和SC准则两方

面的比较结果可知,随机波动性模型对两种指数的估计结果均优于 GARCH(1,1) 模型.



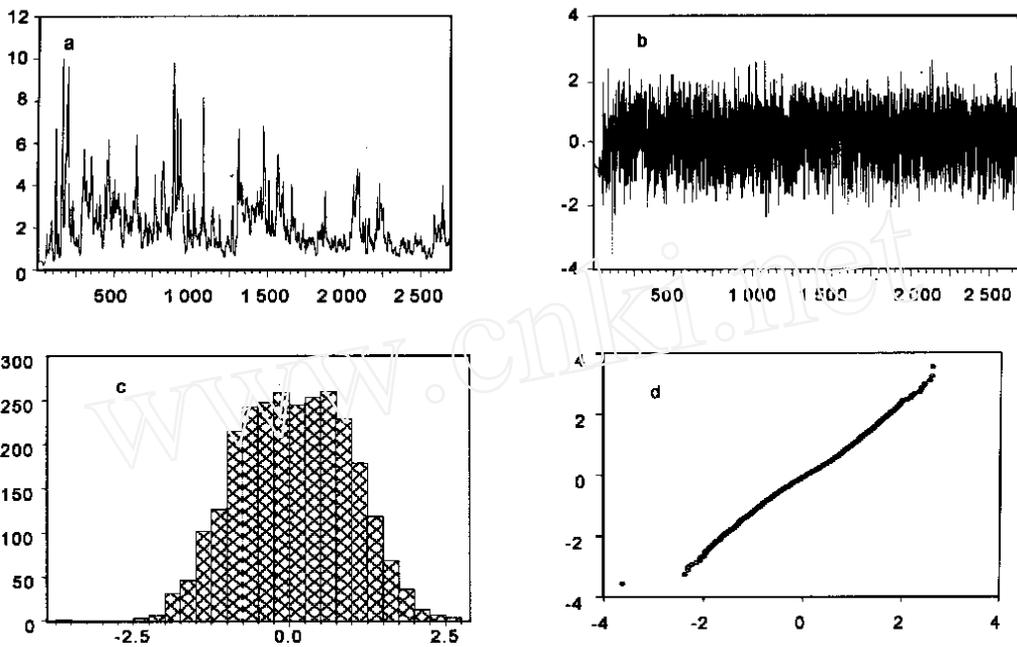
(a:波动性序列, b:残差序列, c:残差序列柱形图, d:残差 Q-Q 图(与正态比较))

图 3 基于 SV 模型的波动性与残差序列(上证指数)



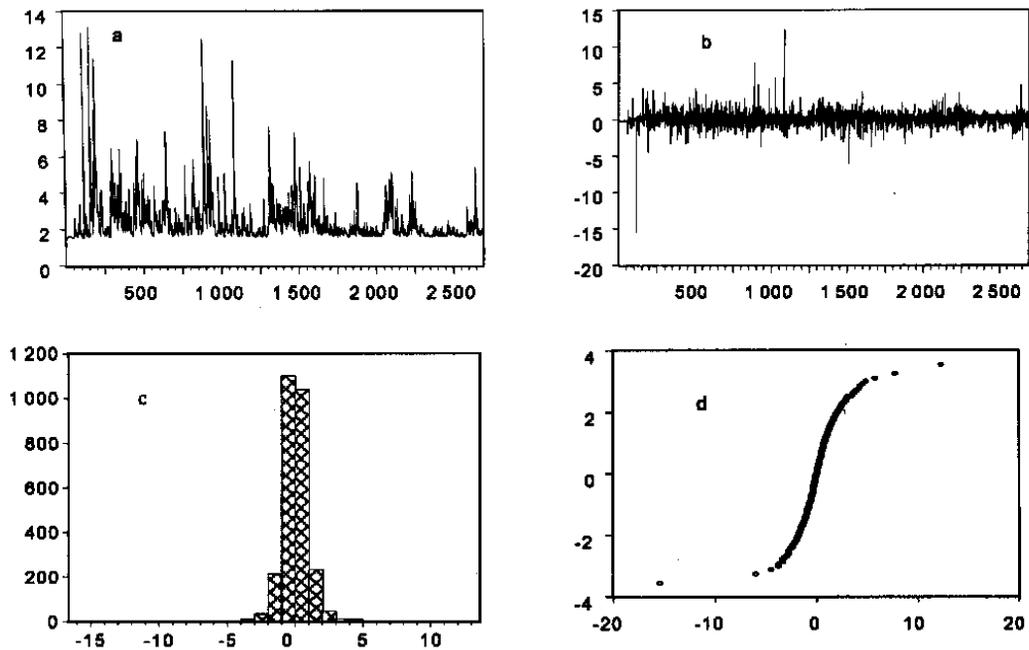
(a:波动性序列, b:残差序列, c:残差序列柱形图, d:残差 Q-Q 图(与正态比较))

图 4 基于 GARCH 模型的波动性与残差序列(深证指数)



(a:波动性序列 ,b:残差序列 ,c:残差序列柱形图 ,d:残差 $Q-Q$ 图(与正态比较))

图5 基于 SV 模型的波动性与残差序列(深证指数)



(a:波动性序列 ,b:残差序列 ,c:残差序列柱形图 ,d:残差 $Q-Q$ 图(与正态比较))

图6 基于 GARCH 模型的波动性与残差序列(上证指数)

通过表 4、图 4 和图 6 可知 ,GARCH 模型均值方程的残差序列的偏度和峰度较原始数据而言已有了一定程度的减小,但仍然拒绝残差序列正态分布的假定.从正态分布统计量和 $Q-Q$ 图来看也

可以得到相同的结论.因此,从残差分析来看,SV 模型较 GARCH 模型能更好地拟合我国股票市场回报的波动性.

在模型的参数已经估计出来的条件下,那么

r_t^2 就是 r_t^2 的条件预期. 因此 r_t^2 在给定的条件下, 均方误差损失函数 $E\{(r_t^2 - \phi_t) | u_{t-1}, u_{t-2}, \dots\}$ 在 $\phi_t = r_t^2$ 时取最小值. 故 $(r_t^2 - \hat{r}_t^2)^2$ 的均值可由 $MSE = T^{-1} \sum_{t=1}^T (r_t^2 - \hat{r}_t^2)^2$ 给出. 均方预测误差是基于实际

样本数据 r_t 的 4 阶矩, 而在某些参数条件下, SV 模型和 GARCH 模型作为数据产生过程时, r_t 的 4 阶矩是不存在的. 因此, 用 MSE 判断模型的预测能力在某些情况下可能是不合理的, 为了全面地反映模型的预测能力, 还计算了另外 2 个指标: $MAE =$

表5 SV模型和GARCH模型比较

	上证指数		深证指数	
	随机波动性	GARCH(1,1)	随机波动性	GARCH(1,1)
对数似然函数	- 5 877.376	- 5 889.504	- 5 967.738	- 5 981.256
SC 准则	4.322 4	4.331 2	4.444 1	4.454 2
MSE	783.628	1 044.187	486.896	699.750
MAE	6.832	9.545	5.710	8.035
LE	1.636	2.060	1.648	2.172

注: SC 代表 Schwarz 准则, 它可以评价模型的估计效果, 其数值越小越好. $SC = (-2 \times \text{对数似然函数} + (\ln T) \times \text{参数个数}) / T$

$T^{-1} \sum_{t=1}^T |r_t^2 - \hat{r}_t^2|$, $|LE| = T^{-1} \sum_{t=1}^T |\ln(r_t^2) - \ln(\hat{r}_t^2)|$. 从表 5 的计算结果看, SV 模型的预测能力要优于 GARCH 模型.

综上所述, 通过 SV 模型与 GARCH 模型的似然函数、SC 准则、残差拟合程度、模型预测能力的综合比较分析, 可以得出结论: 随机波动性模型比 ARCH 类模型能更好地描述中国股票市场的异方差和波动性的序列相关性.

4 结 论

本文采用动态随机波动性模型实证研究了中国股票市场的波动性. 通过使用基于 MCMC 的贝叶斯分析较好地估计了随机波动性模型中的参数. 利用中国股市数据进行的实证结果表明, 与 ARCH 类模型相比, 随机波动性模型能更好地描述股票市场的异方差和波动性的序列相关性.

参 考 文 献:

- [1] 王春峰. 金融市场风险管理[M]. 天津: 天津大学出版社, 2001
- [2] Engle R F. Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation[J]. Econometrica, 1982, 50: 987—1007
- [3] Bollerslev T. Generalized autoregressive conditional heteroscedasticity[J]. Journal of Econometrics, 1986, 31: 307—327
- [4] Clark P. A subordinated stochastic process model with finite variance for speculative process[J]. Econometrica, 1973, 41: 135—155
- [5] Epps T W, Epps M L. The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distribution hypothesis[J]. Econometrica, 1976, 44: 305—321
- [6] Danielsson J. Estimation of dynamic stochastic volatility model for asset price determination by simulated maximum likelihood[D]. Durham: Duke University, 1991
- [7] Taylor S J. Modeling stochastic volatility[J]. Mathematical Finance, 1994, 4(2): 183—204
- [8] Danielsson J. Stochastic volatility in asset prices estimation with simulated maximum likelihood[J]. Journal of Econometrics, 1994, 64: 375—400
- [9] Kim S. Shephard N. Stochastic Volatility: Likelihood Inference and Comparison with ARCH Models[R]. Nuffield College, London: Oxford University, 1994

- [10] Shephard N. Statistical Aspects of ARCH and Stochastic Volatility[R]. Nuffield College, London: Oxford University, 1995
- [11] Bollerslev T. A conditional heteroskedastic time series model for speculative prices and rates of return[J]. The Review of Economics and Statistics, 1987, 59: 542—547
- [12] Baillie R T, Bollerslev T. The message of daily exchange rates: A conditional variance tale[J]. Journal of Business and Economic Statistics, 1989, 7: 297—306
- [13] Diebold F X. Empirical Modelling of Exchange Rate[M]. Berlin: Springer-Verlag, 1987
- [14] Nelson D B. Conditional heteroskedasticity in asset returns: A new approach[J]. Econometrica, 1991, 59: 347—370
- [15] Hsieh D A. Statistical properties of daily exchange rates: 1974—1963[J]. Journal of International Economics, 1988, 24: 129—145
- [16] Wiggins J B. Option values under stochastic volatility: Theory and empirical estimates[J]. Journal of Financial Economics, 1987, 19: 351—372
- [17] Karpoﬀ J M. The relation between price changes and trading volume: A survey[J]. Journal of Financial and Quantitative Analysis, 1987, 22: 109—123
- [18] Hendry D F, Richard J. Likelihood Evaluation for Dynamic Latent Variables Models[R]. ISDS discussion paper, Durham: Duke University, 1990
- [19] Polson J E, Rossi G E. Bayesian analysis of stochastic volatility models[J]. Journal of Business & Economic Statistics, 1994, 12: 371—390
- [20] Polson J E, Rossi G E. Bayesian Analysis of Stochastic Volatility Model with Fat-tails and Correlated errors[R]. www2. bc. edu/jacquier, 2001
- [21] Gilks W R, Richardson S. Spiegelhalter, D.J. Markov Chain Monte Carlo in Practice[M]. London: Chapman & Hall Press, 1998
- [22] Engle Robert F, Victor K Ng. Measuring and testing the impact of news on volatility[J]. Journal of Finance, 1993, 48: 1022—1082

Estimating volatility of Chinese stock market by stochastic volatility model

WANG Chun-feng, JIANG Xiang-lin, LI Gang

Center of Finance Engineering, Management School, Tianjin University, Tianjin 300072, China

Abstract: The volatility of Chinese stock market is investigated using the dynamic version of stochastic volatility model, and Bayesian analysis based on MCMC is introduced to improve the parameters estimation in stochastic volatility model. Empirical results on Chinese stock market indicate that stochastic volatility model outperforms the ARCH model in capturing the heteroskedasticity and serial correlation of volatility of the stock market returns.

Key words: stochastic volatility model; Bayesian analysis; MCMC; ARCH model