



# 基于 C\_TMPV 的中国股市 高频波动率的跳跃行为研究

杨科, 陈浪南

中山大学 岭南学院, 广州 510275

**摘要:** 运用2000年1月4日至2008年12月31日上证综指每5分钟的高频金融数据, 采用核估计量估计中国股市高频波动率序列, 运用修正的已实现门阀多次幂变差估计中国股市高频波动率的跳跃序列, 实证分析中国股市高频波动率跳跃的各种特征, 并运用ACD模型、ACH模型以及扩展的ACH模型进一步分析中国股市高频波动率跳跃的持续期的特征。研究结果表明, 中国股市高频波动率及其跳跃都具有集聚的特征, 高频波动率发生显著跳跃的比例相当高, 高频波动率跳跃的幅度、强度和跳跃幅度的分布都具有时变性, 而跳跃对高频波动率的贡献却具有相对稳定性; 在样本期, 中国股市高频波动率跳跃表现出较强的正相关性, 且跳跃的持续期存在较强的长记忆性和周日历效应。

**关键词:** 高频波动率; 跳跃; 修正的已实现门阀多次幂变差; C\_TZ 统计量

**中图分类号:** F830.91

**文献标识码:** A

**文章编号:** 1672-0334(2011)02-0103-10

## 1 引言

作为资产价格、组合选择和衍生产品设计的核心, 金融资产收益波动率的估计和预测至关重要。由于金融资产(尤其是股票)的交易价格一般具有时间连续性, 因此金融资产的收益率应该是平稳的。然而, 随着日内高频交易数据的可用性, 许多研究表明, 金融资产的收益率在日内近似连续的时间内有可能出现大幅波动, 这种现象称为跳跃, 跳跃在金融资产收益波动率的估计和预测中具有非常重要的意义。虽然跳跃行为发生的概率一般较小, 然而一旦发生, 却会对股票、债券以及衍生品等市场带来巨大的冲击, 这种冲击比连续性波动率的影响要大得多, 历史上股票市场发生的多次崩盘事件均是以前股票价格向下的大幅度波动开始的。鉴于此, 如何有效地分析中国股市波动率的跳跃性特征, 对于完善中国证券市场监管机制、合理构建投资组合和实行风险管理都具有重大的理论和现实意义。

股票市场波动率跳跃是国外金融研究的热点问题

题, 而中国对波动率跳跃的研究文献很少, 对高频波动率跳跃的研究就更少。本研究试图采用修正的已实现门阀多次幂变差和 C\_TZ 统计量估计中国股市高频波动率跳跃序列, 进而检验和分析高频波动率跳跃的各种特征, 并运用 ACD 模型、ACH 模型以及扩展的 ACH 模型进一步分析中国股市高频波动率跳跃的持续期的特征。

## 2 相关研究评述

国外对波动率跳跃分析主要有两类研究方法, 一是基于高度离散化的低频数据的参数化方法, 二是近些年提出的基于高频数据的非参数化方法。在低频数据的参数化方法方面, Chan 等<sup>[1]</sup>通过设定跳跃幅度服从自回归过程建立自回归跳跃幅度模型(GARJI模型), 并采用极大似然估计方法对IBM等11只股票以及道琼斯指数等3种指数进行实证, 发现跳跃幅度存在自相关关系, 跳跃呈现出聚集的特征; Eraker 等<sup>[2]</sup>用马尔可夫链蒙特卡尔(Markov chain

**收稿日期:** 2010-04-15 **修返日期:** 2010-08-29

**基金项目:** 国家自然科学基金(70673116); 教育部人文社会科学重点研究基地重大项目(05JJD790075); 国家社会科学基金(07BJY167); 中山大学985工程产业与区域发展研究创新基地项目; 广东省普通高校人文社会科学重点研究基地项目; 广东省自然科学基金(9151027501000032)

**作者简介:** 杨科(1983-), 男, 湖南岳阳人, 中山大学岭南学院博士研究生, 研究方向: 金融经济学和货币经济学等。  
E-mail: yangkeda@foxmail.com

Monte Carlo, MCMC)方法估计了同时考虑收益率序列跳跃和波动率序列跳跃的含有同时性跳跃的随机波动率(stochastic volatility with contemporaneous jumps, SVCJ)模型和含有独立性跳跃的随机波动率(stochastic volatility with independent jumps, SVIJ)模型,证实随机波动率过程中跳跃的存在以及在波动率方程中引入跳跃的重要性;Maheu等<sup>[3]</sup>在GARJI模型的基础上,通过在GARCH波动率方程中考虑过去跳跃的影响,进一步扩展GARJI模型,发现过去跳跃对波动率的不对称性影响不明显,而当前的跳跃会产生非对称性。然而,这种参数化方法利用的低频数据过于离散,难以反映资产价格变化过程中包含的全部有用信息,并且参数较难估计。随着高频数据的获取越来越便利以及计算机处理数据能力的飞速发展,国外学者又提出解决跳跃问题的新方法,即基于高频数据的非参数化方法。这种非参数化方法的理论基础是利用高频数据将已实现波动率分解为连续部分和跳跃部分,可以得到波动率跳跃序列,进而可以检验和分析波动率跳跃的特征。Barndor-Nielsen等<sup>[4-6]</sup>在此领域取得一系列的成果,他们以随机过程理论为基础,基于二次幂变差方法寻找到股市波动率中连续变动和跳跃的渐进统计量,证明当不存在跳跃时,二次幂变差是积分波动率的一致估计量,从而可以将跳跃从已实现波动率中分离出来;Bollerslev等<sup>[7]</sup>的研究表明,这种非参数化的方法在测度波动率跳跃成分时利用高频数据所包含的信息,这样即便是很简单的模型也能比那些复杂的参数波动率模型取得更好的效果;Thomakos等<sup>[8]</sup>运用Barndor-Nielsen等<sup>[5-6]</sup>的方法从各种不同的已实现波动率估计值中分离出波动率的跳跃成分,考察雅典股票市场波动率跳跃的特征,发现雅典股票市场波动率跳跃具有较强的相关性,其跳跃的持续期也具有相关性。

中国国内对中国股市波动率跳跃成分的研究非常少。在低频数据的参数化方法方面,潘祺<sup>[9]</sup>以Merton<sup>[10]</sup>的跳跃破产模型对上证综指进行拟合,得到的结论是该模型与传统扩散模型相比更好地描述上证综指收益率的突变行为;童汉飞等<sup>[11]</sup>利用GARCH-跳跃模型对沪深两市A股和B股的跳跃性特征进行建模和分析,发现该模型能有效估计沪深两市收益率和波动率的跳跃性变化,B股市场发生跳跃的概率高于A股市场,而A股跳跃幅度的均值比B股高,并且沪深A股间跳跃的一致性比B股强;杜军等<sup>[12]</sup>利用GARCH-跳跃模型对上证综指和道琼斯指数进行比较研究,发现跳跃幅度具有时变的特征。而基于高频数据的非参数化方法方面,中国国内文献则更少。王春峰等<sup>[13]</sup>基于Barndor-Nielsen等<sup>[5-6]</sup>的方法对上证综指已实现波动率跳跃和连续部分进行研究,采用2000年1月4日至2005年5月31日5分钟间隔高频数据,发现上证综指日间波动率发生显著跳跃的比例很高,并且跳跃幅度具有明显的相关性。然而,王春峰等<sup>[13]</sup>运用的是基于Huang等<sup>[14]</sup>研究中的Z统计量来检验波动率跳跃。

Corsi等<sup>[15]</sup>的研究指出,当波动率跳跃连续出现的频率很高时,基于Z统计量的方法有可能检测不出一些跳跃,并且王春峰等<sup>[13]</sup>没有通过建模具体地分析波动率跳跃持续期的特征。基于此,本研究以Corsi等<sup>[15]</sup>修正的已实现门阀多次幂变差(C-TMPV)为理论基础实证研究中国股市高频波动率跳跃的特征。

### 3 实证模型和方法

#### 3.1 高频波动率的估计

假设在一个给定的交易日 $t$ ,将每日价格运动轨迹分割为 $n_t$ , $\Lambda_t$ 为每个时点对数价格的观察值, $\Lambda_t = \{\tau_0, \dots, \tau_{n_t}\}$ , $p_{t,i}$ 为交易日 $t$ 的第 $i$ 个对数价格观测值, $i = 1, \dots, n_t$ ,令 $r_{t,i}$ 为交易日 $t$ 内第 $i$ 个时间段上的对数收益率, $r_{t,i} = p_{t,i} - p_{t,i-1}$ 。定义已实现高频波动率(RV)为所有可以获得的交易日日内高频收益率平方的总和,即

$$RV_t = \sum_{i=1}^{n_t} r_{t,i}^2 \quad (1)$$

由于市场微观结构动态通过噪声影响价格过程,因此在基于金融高频时间序列对波动率进行研究时不得不考虑微观结构噪声的影响。以(1)式所述的已实现波动率估计量在不考虑微观结构效应的影响时是积分波动率的无偏估计量,但在考虑微观结构效应时就是有偏的。目前国内外的学者对已实现波动率估计量进行偏差校正方面已经做了大量的研究工作,相关文献综述可以参见李胜歌<sup>[16]</sup>的研究。

本研究采用Hansen等<sup>[17]</sup>的核估计量对RV进行偏差校正,即

$$RV_t = \sum_{i=1}^{n_t} r_{t,i}^2 + 2 \sum_{h=1}^q \left(1 - \frac{h}{q+1}\right) \sum_{i=1}^{n_t-h} r_{t,i} r_{t,i+h} \quad (2)$$

其中, $q$ 为一个很小的非负整数, $h$ 为不大于 $q$ 的非负整数。设 $q=1$ ,因此 $h$ 也仅取1。

#### 3.2 高频波动率跳跃的检验和估计

假设股价的对数 $p_t$ 服从连续时间的跳跃扩散过程,即

$$dp_t = \mu_t dt + \sigma_t dW_t + \kappa_t dq_t, \quad 0 \leq t \leq T \quad (3)$$

其中, $\mu_t$ 为连续的局部有界变差过程; $\sigma_t$ 为随机波动率过程; $W_t$ 为标准布朗运动; $q_t$ 为泊松跳跃过程,其跳跃强度为 $\lambda_\kappa$ ; $\kappa_t$ 为跳跃幅度,服从正态分布 $N(\mu_\kappa, \sigma_\kappa)$ 。对数收益率从 $t$ 时到 $(t+1)$ 时的二次变差为

$$[r, r]_{t+1} = \int_t^{t+1} \sigma_s^2 ds + \sum_{t < s \leq t+1} \kappa_s^2 \quad (4)$$

其中, $\sigma_s$ 为瞬时波动率,也即(3)式中的随机波动率 $\sigma_t$ , $\kappa_s$ 为跳跃幅度;等号右边第一部分称为积分波动率,它是由(3)式中的连续部分得到;第二部分代表非连续的跳跃部分。在现实中观察到的股票价格是离散的,具有一定的时间间隔,即使超高频的分笔数据也是如此,因此只能用离散状态下已实现波动率的估计值代替二次变差。Andersen等<sup>[18]</sup>和Barndor-Nielsen等<sup>[4]</sup>指出,已实现波动率满足 $RV_{t+1} \rightarrow \int_t^{t+1} \sigma_s^2 ds + \sum_{t < s \leq t+1} \kappa_s^2$ 。

定义二次幂变差为

$$BV_t = \frac{\pi}{2} \sum_{j=2}^{n_t} |r_{t,j}| |r_{t,j-1}| \quad (5)$$

其中,  $j$  为正整数

Bardor-Nielsen 等<sup>[4]</sup>证明,当每个交易日内的价格观测值个数趋于无穷大时,二次幂变差是波动率连续部分的一个有效估计量,即有  $BV_{t+1} \rightarrow \int_t^{t+1} \sigma_s^2 ds$ 。因此,当不存在跳跃时,已实现波动率与二次幂变差的差为 0。Huang 等<sup>[14]</sup>建议用下面的  $Z$  统计量检验波动率跳跃的存在性,即

$$Z_t = \frac{RV_t - BV_t}{RV_t} \sqrt{\left[ \left( \frac{\pi}{2} \right)^2 + \pi - 5 \right] \frac{1}{n_t} \max \left( 1, \frac{TP_t}{BV_t^2} \right)} \quad (6)$$

在不存在跳跃的条件下,该统计量在  $n_t \rightarrow \infty$  时渐进服从标准正态分布,其中  $TP_t$  是三次幂变差, Bardor-Nielsen 等<sup>[5]</sup>指出

$$TP_t = n_t \vartheta_{\frac{4}{3}}^{-3} \frac{n_t}{n_t - 2} \sum_{j=3}^{n_t} |r_{t,j-2}|^{\frac{4}{3}} |r_{t,j-1}|^{\frac{4}{3}} |r_{t,j}|^{\frac{4}{3}} \rightarrow \int_{t-1}^t \sigma_s^4 ds \quad (7)$$

其中,  $\vartheta_k = 2^{\frac{k}{2}} \frac{\Gamma(\frac{k+1}{2})}{\Gamma(\frac{1}{2})}$ ,  $k = \frac{4}{3}$ ,  $\Gamma(\cdot)$  为 Gamma 函数。

当  $Z$  统计量大于标准正态分布在显著水平  $\alpha$  的临界值时(记为  $\Phi_\alpha$ ),波动率跳跃被认为是显著存在的。基于  $Z$  统计量,可以定义第  $t$  交易日的显著波动率跳跃为

$$Z_{J_t} = I(Z_t > \Phi_\alpha) \cdot (RV_t - BV_t)^+ \quad (8)$$

其中,  $I(\cdot)$  为示性函数,  $(RV_t - BV_t)^+ = \max((RV_t - BV_t), 0)$ 。同理,已实现波动率跳跃对已实现波动率的贡献可以定义为

$$Z_{RJ_t} = \frac{Z_{J_t}}{RV_t} = I(Z_t > \Phi_\alpha) \cdot \left( 1 - \frac{BV_t}{RV_t} \right)^+ \quad (9)$$

然而, Corsi 等<sup>[15]</sup>的研究指出,当高频数据中跳跃连续出现的频率很高时, Huang 等<sup>[14]</sup>的  $Z$  统计量有可能检测不出一些跳跃。因此,他们基于 C\_TMPV 构建一个修正的  $Z$  统计量,记为  $C_{TZ}$  统计量。通过 Monte Carlo 模拟,他们发现当不存在跳跃时,  $C_{TZ}$  统计量的检验能力和  $Z$  统计量的检验能力相当;当存在跳跃时,尤其是跳跃连续出现时,  $C_{TZ}$  统计量的检验能力显著高于传统的  $Z$  统计量。

Corsi 等<sup>[15]</sup>构建的修正的已实现 C\_TMPV 估计量的表达式为

$$C\_TMPV_t^{\gamma_1, \dots, \gamma_M} = \left( \frac{1}{n_t} \right)^{1-\frac{1}{2}(\gamma_1 + \dots + \gamma_M)} \sum_{j=M}^{n_t} \prod_{k=1}^M Z_{\gamma_k}(r_{t,j-k+1}, v_{j-k+1}) \quad (10)$$

其中,  $\gamma_1, \dots, \gamma_M$  为任意正数,  $v_t$  为随机门阀函数,  $v_t = c_v^2 \hat{V}_t$ ,  $c_v$  为一个常数,本研究取  $c_v = 3$ ,  $\hat{V}_t$  是连续时间波动率  $\sigma_t$  的辅助估计量(本研究采用 Fan 等<sup>[19]</sup>的非参

数方法进行估计)。(10)式中函数  $Z_\gamma(x, y)$  定义为

$$Z_\gamma(x, y) = \begin{cases} |x|^\gamma & x^2 \leq y \\ \frac{1}{2N(-c_v) \sqrt{\pi}} \left( \frac{2}{c_v} y \right)^{\frac{\gamma}{2}} \Gamma\left( \frac{\gamma+1}{2}, \frac{c_v^2}{2} \right) & x^2 > y \end{cases}$$

其中,  $N(\cdot)$  为标准正态分布累积密度函数。由(10)式的定义可知,修正的已实现门阀二次幂变差  $C\_TBV_t = \vartheta_1^{-2} C\_TMPV_t^{[1,1]}$ ,修正的门阀三次幂变差  $C\_TTP_t = \vartheta_{\frac{4}{3}}^{-3} C\_TMPV_t^{[\frac{4}{3}, \frac{4}{3}, \frac{4}{3}]}$ 。基于修正的已实现门阀多次幂变差,  $C_{TZ}$  统计量为

$$C_{TZ_t} = \frac{RV_t - C\_TBV_t}{RV_t} \sqrt{\left[ \left( \frac{\pi}{2} \right)^2 + \pi - 5 \right] \frac{1}{n_t} \max \left( 1, \frac{C\_TTP_t}{C\_TBV_t^2} \right)} \quad (11)$$

Corsi 等<sup>[15]</sup>已经证明在不存在跳跃的条件下,  $C_{TZ}$  统计量在  $n_t \rightarrow \infty$  时渐进服从标准正态分布。当  $C_{TZ}$  统计量大于标准正态分布在显著水平  $\alpha$  的临界值时(记为  $\Phi_\alpha$ ),波动率跳跃被认为是显著的,基于  $C_{TZ}$  统计量,可以定义第  $t$  交易日的显著波动率跳跃为

$$C_{TZ_{J_t}} = I(C_{TZ_t} > \Phi_\alpha) \cdot (RV_t - C\_TBV_t)^+ \quad (12)$$

同理,波动率跳跃对波动率的贡献可以定义为

$$C_{TZ_{RJ_t}} = \frac{C_{TZ_{J_t}}}{RV_t} = I(C_{TZ_t} > \Phi_\alpha) \cdot \left( 1 - \frac{C\_TBV_t}{RV_t} \right)^+ \quad (13)$$

### 3.3 高频波动率跳跃的持续期模型

高频波动率跳跃的持续期模型主要考虑跳跃之间的时间间隔,较长的持续期预示着高频波动率较少发生跳跃。假设  $\{t_0, t_1, \dots, t_n, \dots\}$  表示高频波动率随机发生跳跃的交易日,而跳跃的持续期定义为  $d_i$ ,  $d_i = t_i - t_{i-1}$ ,它表示两个相继发生跳跃的交易日之间的天数。Engle 等<sup>[20]</sup>的自回归条件持续期模型(ACD模型)可以表述为

$$d_i = \varphi_i \varepsilon_i$$

$$\varphi_i = \omega_d + \sum_{j=1}^m \alpha_{d,j} d_{i-j} + \sum_{j=1}^q \beta_{d,j} \varphi_{i-j}$$

$$E(d_i | d_{i-1}, \dots, d_1) = \varphi_i(d_{i-1}, \dots, d_1; \theta) = \varphi_i \quad (14)$$

其中,  $\varphi_i$  为高频波动率跳跃的第  $i$  个持续期的条件期望值,  $\varepsilon_i$  为服从一个标准指数分布或者标准化的 Weibull 分布的扰动项,  $\omega_d, \alpha_{d,j}$  和  $\beta_{d,j}$  为模型待估计的参数,  $\theta$  为模型的参数向量,  $\theta = (\omega_d, \alpha_{d,j}, \beta_{d,j})'$ ,  $m$  和  $q$  为滞后期数目。

虽然 ACD 模型能够较好地描述波动率跳跃持续期的动态相关性,但该模型仅当跳跃发生时才会更新持续期的条件期望,而在现实世界中,人们总是偏好于利用新的信息不断更新自己的期望。Hamilton 等<sup>[21]</sup>的自回归条件风险模型(ACH模型)很好地解决了这个问题。本研究分别估计两个不同的 ACH 模型,其中第一个模型是存在跳跃的交易日之间没有信息更新的 ACH(1,1)模型,即



$$h_t = \frac{1}{\varphi_{N(t-1)}}$$

$$\varphi_{N(t)} = \omega + \alpha_1 d_{N(t)-1} + \beta_1 \varphi_{N(t)-1} \quad (15)$$

其中,  $h_t$  为条件风险概率, 表示在信息集  $\Gamma_{t-1}$  下高频波动率发生跳跃的条件概率,  $h_t = \Pr[N(t) \neq N(t-1) | \Gamma_{t-1}]$ ,  $\Gamma_{t-1}$  为到  $(t-1)$  为止能观察到的信息,  $N(t)$  为到时间  $t$  为止高频波动率总共发生跳跃的天数。

第二个模型是扩展的 ACH 模型, 即在 ACH 模型中加入星期一、星期二、星期三和星期四的虚拟变量, 用来捕获波动率跳跃持续期的周日历效应, 该模型可以表述为

$$h_t = \frac{1}{\alpha_1 d_{N(t)-1} + \beta_1 \varphi_{N(t)-1} + \delta' z_{t-1}}$$

$$\delta' z_{t-1} = \delta_0 + \delta_M D_M + \delta_T D_T + \delta_W D_W + \delta_{Th} D_{Th} \quad (16)$$

其中,  $\delta_0, \delta_M, \delta_T, \delta_W$  和  $\delta_{Th}$  为日历效应影响系数;  $\delta$  为影响系数向量,  $\delta = (\delta_0, \delta_M, \delta_T, \delta_W, \delta_{Th})'$ ;  $D_M, D_T, D_W$  和  $D_{Th}$  分别为星期一、星期二、星期三和星期四的虚拟变量;  $z_{t-1}$  为虚拟变量构成的向量,  $z_{t-1} = (1, D_M, D_T, D_W, D_{Th})'$ 。

#### 4 数据和描述性分析

本研究选取沪市 5 分钟间隔上证综指作为高频采样数据进行研究。数据采样时间为 2000 年 1 月 4 日至 2008 年 12 月 31 日, 剔除数据不完整的交易日, 得到 2 085 个交易日, 每个交易日有 48 个数据, 共计 100 080 个上证综指数据。数据来源于中国经济研究中心股票市场高频数据库。

表 1 给出上证综指日对数收益率和日已实现波动率序列(由(2)式估计得到)的描述性统计量, 图 1

给出上证综指日对数收益率和日已实现波动率序列图。从表 1 和图 1 可以看出, 日已实现波动率的均值在时间区间 2006 年 1 月 4 日至 2008 年 12 月 31 日这个近似于一个牛市和熊市的完整周期上达到最大值, 并且标准方差也达到最大值, 说明已实现波动率在这个时间区间上的变动幅度最大; 在熊市区间 2007 年 11 月 1 日至 2008 年 12 月 31 日的日已实现波动率的均值和标准方差都比牛市的大, 意味着已实现波动率在熊市的变化幅度要比牛市的大, 并且从已实现波动率的时序图可以看出, 已实现波动率最大的时点几乎都密集在熊市。在区间 2003 年 3 月 3 日至 2005 年 12 月 30 日已实现波动率的取值相对较小, 在一个很小窄幅区间内变动。

#### 5 实证结果和分析

##### 5.1 高频波动率跳跃的特征分析

图 2 (a) ~ (d) 是高频波动率跳跃(分别由(8)式和(12)式估计得到)和跳跃对高频波动率的贡献(分别由(9)式和(13)式得到)的时间序列图, 表 2 给出各个时间区间的统计特征。本研究统一选取显著水平为 0.100%, 用 Z 统计量对上证综指高频波动率跳跃进行甄别, 共甄别出 310 个交易日高频波动率发生显著跳跃, 而 C\_TZ 统计量共甄别出 616 个交易日高频波动率发生显著跳跃, 说明 C\_TZ 统计量比 Z 统计量能更多的甄别出显著跳跃。从图 2 (a) ~ (d) 可以看出, 使用上证综指的高频数据甄别日间波动率发生显著跳跃的比例(为 0.149 和 0.295)是相当高的, 远远高于对日或者更低频的数据应用参数跳跃扩散模型估计的跳跃频率。出现这一现象一个可能的原

表 1 日对数收益率和日已实现波动率序列的描述性统计量

Table 1 Descriptive Statistics of Daily Log Returns and Realized Volatility

	均值	最大值	最小值	标准方差	偏度系数	峰度系数
日对数收益率						
2000-01-04 - 2003-02-26	4.522e-5	0.094	-0.065	0.014	0.596	9.693
2003-03-03 - 2005-12-30	-3.969e-4	0.079	-0.040	0.013	0.756	5.529
2006-01-04 - 2007-10-31	0.004	0.056	0.093	0.019	-1.139	7.020
2007-11-01 - 2008-12-31	-0.004	0.090	-0.128	0.028	0.044	4.821
2000-01-04 - 2008-12-31	1.041e-4	0.094	-0.128	0.017	-0.128	8.166
日已实现波动率						
2000-01-04 - 2003-02-26	1.566e-4	0.008	1.593e-6	4.880e-4	12.475	186.070
2003-03-03 - 2005-12-30	1.261e-4	0.001	1.186e-5	1.357e-4	3.610	21.331
2006-01-04 - 2007-10-31	3.333e-4	0.004	1.395e-5	4.190e-4	3.623	22.834
2007-11-01 - 2008-12-31	7.192e-4	0.011	6.766e-5	9.779e-4	6.062	52.250
2000-01-04 - 2008-12-31	2.579e-4	0.011	1.593e-6	5.370e-4	9.507	138.490

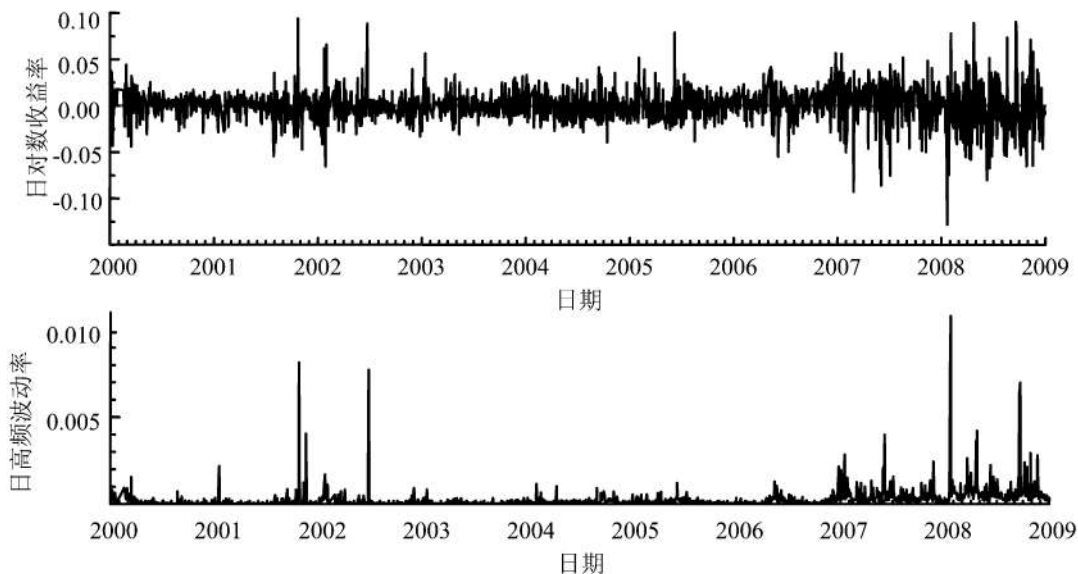


图1 上证综指日对数收益率和日已实现波动率序列  
Figure 1 Daily Log Returns and Realized Volatility for SSE

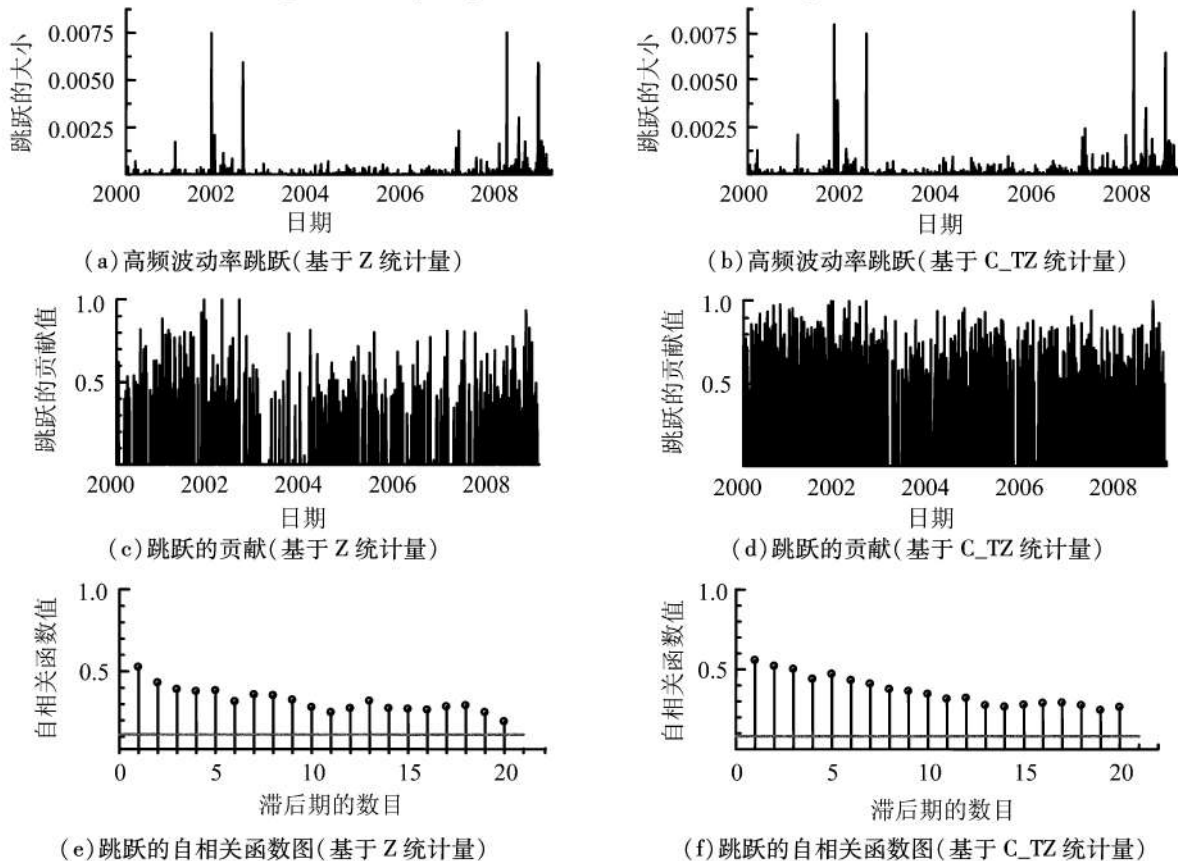


图2 高频波动率跳跃、贡献及显著跳跃的自相关函数图  
Figure 2 Jumps, Contribution and Autocorrelations for Jumps

因是,使用日或者更低频数据时可能会损失一些重要的信息,导致甄别波动率发生跳跃的方法出现严重的误差,而使用高频数据甄别日间波动率是否发生跳跃就可以避免这一问题。中国股市波动率发生跳跃的概率之所以较大,其中的一个重要原因是中国股票市场投机气氛比较严重,投资者较少关注当前收益,而较多的注重资本利得,在市场上容易追涨

杀跌,带来股价和股市波动率的频繁大幅度变动。因此,应加强对投资者的理性投资教育,改善市场投资主体的构成,发挥投资基金的理性投资主导作用。

由图2(a)~(d)和表2可知,在不同的时间区间内,无论是基于Z统计量还是C\_TZ统计量,高频波动率跳跃的幅度是变化不定的,即高频波动率跳跃幅度具有时变性。特别地,在熊市区间内跳跃的幅

表2 高频波动率跳跃及其对波动率的贡献的统计特征  
Table 2 Statistical Characteristics of Jump  
of High-Frequency Volatility and Its Contribution to Volatility

	跳跃			贡献		
	最小值	最大值	均值	最小值	最大值	均值
2000-01-04 – 2003-02-26						
基于 Z 统计量	1.590e-6	0.008	2.356e-4	0.302	0.901	0.523
基于 C_TZ 统计量	1.590e-6	0.008	2.038e-4	0.299	0.934	0.672
2003-03-03 – 2005-12-30						
基于 Z 统计量	9.460e-6	6.807e-4	1.153e-4	0.298	0.815	0.445
基于 C_TZ 统计量	7.840e-6	9.536e-4	1.277e-4	0.205	0.884	0.618
2006-01-04 – 2007-10-31						
基于 Z 统计量	1.560e-5	0.002	2.898e-4	0.291	0.809	0.480
基于 C_TZ 统计量	9.540e-6	0.002	3.166e-4	0.290	0.921	0.635
2007-11-01 – 2008-12-31						
基于 Z 统计量	5.650e-5	0.008	8.061e-4	0.314	0.933	0.482
基于 C_TZ 统计量	6.330e-5	0.009	7.975e-4	0.251	0.937	0.591
2000-01-04 – 2008-12-31						
基于 Z 统计量	1.590e-6	0.008	3.226e-4	0.291	0.933	0.487
基于 C_TZ 统计量	1.590e-6	0.009	2.825e-4	0.205	0.937	0.638

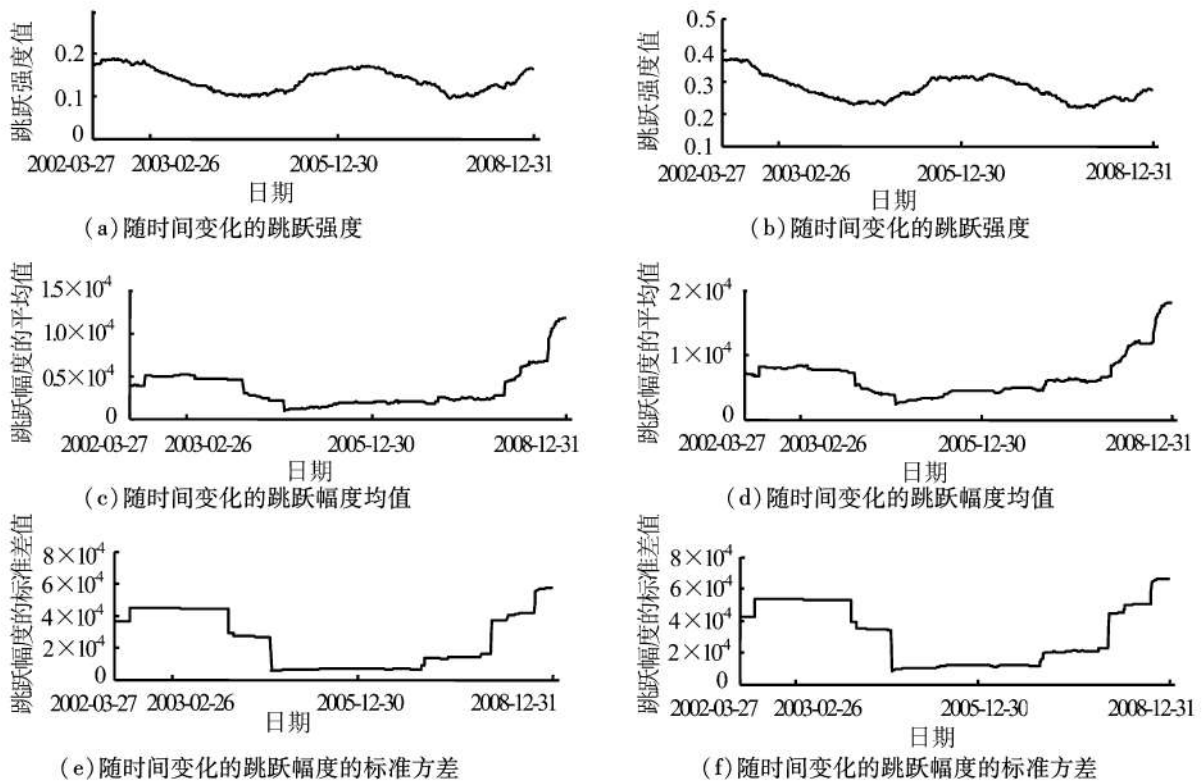
度均值是牛市区间的 2 倍多;无论是基于 Z 统计量还是 C\_TZ 统计量,当跳跃发生时,跳跃对高频波动率的贡献在不同的时间区间内都占相当大的比例,并且跳跃对高频波动率的贡献具有相对稳定性。基于 Z 统计量估计的跳跃对高频波动率的贡献的均值大致为 50% 左右,而本研究基于 C\_TZ 统计量估计的跳跃对高频波动率的贡献的均值却为 60% 左右,进一步说明 C\_TZ 统计量比 Z 统计量能更多地甄别出显著跳跃。从图 2 (a)~(b) 还可以发现,已实现波动率跳跃表现出较明显的聚集特征,意味着高频波动率跳跃存在自相关性。图 2 (e)~(f) 确认了这一点,它清晰地显示,无论是基于 Z 统计量还是 C\_TZ 统计量,显著的高频波动率跳跃都存在较强正相关性,并且自相关函数值递减得很慢,说明显著的高频波动率跳跃具有较强的可预测性。由于股市中跳跃性波动产生的根源在于包括货币政策、财政政策在内的宏观经济政策的变动<sup>[18]</sup>,因此中国股票市场高频波动率跳跃的长期滞后相关性是由于宏观政策对股票市场的波动会产生较长时间的影响,或者说政策对股票市场造成的影响具有长期有效性。

由高频波动率跳跃的聚集性和跳跃幅度的时变性可以推测,高频波动率跳跃强度和跳跃幅度的分布也同样是时变的。本研究以两年的滚动时间窗口估计高频波动率跳跃强度以及跳跃幅度分布的均值和标准方差<sup>[22]</sup>,图 3 给出这些估计值的动态图。

由图 3 可知,上证综指高频波动率跳跃的频率在 2002 年是相当高的,而在 2002 年之后高频波动率跳跃的频率有一个下降趋势,这个下降趋势一直延续到 2004 年,之后跳跃频率有一个上升趋势,在 2006 年初这个上升趋势达到最大值,在此之后的牛市区间上证综指高频波动率跳跃的频率又处于下降趋势,而在熊市区间跳跃的频率又开始不断的增加。但上证综指高频波动率跳跃幅度的分布的均值和标准方差与跳跃频率的动态特征不同,两者在 2004 年时有一个突然的下降,在 2004 年至 2006 年底高频波动率跳跃幅度的分布的均值和标准方差动态变化都不是很大,两者都比较接近于 0,在 2007 年 10 月左右至 2008 年初,即在牛市与熊市转换的时期,跳跃的幅度明显放大,之后跳跃的幅度还有一个上升趋势,意味着在牛市和熊市这个完整周期内,绝大部分巨幅跳跃均集中在熊市。

## 5.2 高频波动率跳跃的持续期分析

图 4 为 4 个不同时间区间以及总体区间高频波动率跳跃持续期的自相关函数图,由图 4 可以很明显的看出各个区间上的自相关特征都不同,在区间 2000 年 1 月 4 日至 2003 年 2 月 26 日,高频波动率跳跃持续期的自相关系数滞后 20 期仍然显著并都表现出正的自相关性,而在牛市和熊市区间自相关系数仅滞后 4 期就不再显著,并且之后自相关性还表现出不太显著的周期性。在总体区间,上证综指高频



注:左边三图是基于 Z 统计量,右边三图是基于 C\_TZ 统计量

图 3 高频波动率跳跃强度、跳跃幅度的均值以及跳跃幅度的标准差的动态图

Figure 3 Jump Intensity, Jump Size's Mean and Jump Size's Standard Deviation for High-Frequency Volatility

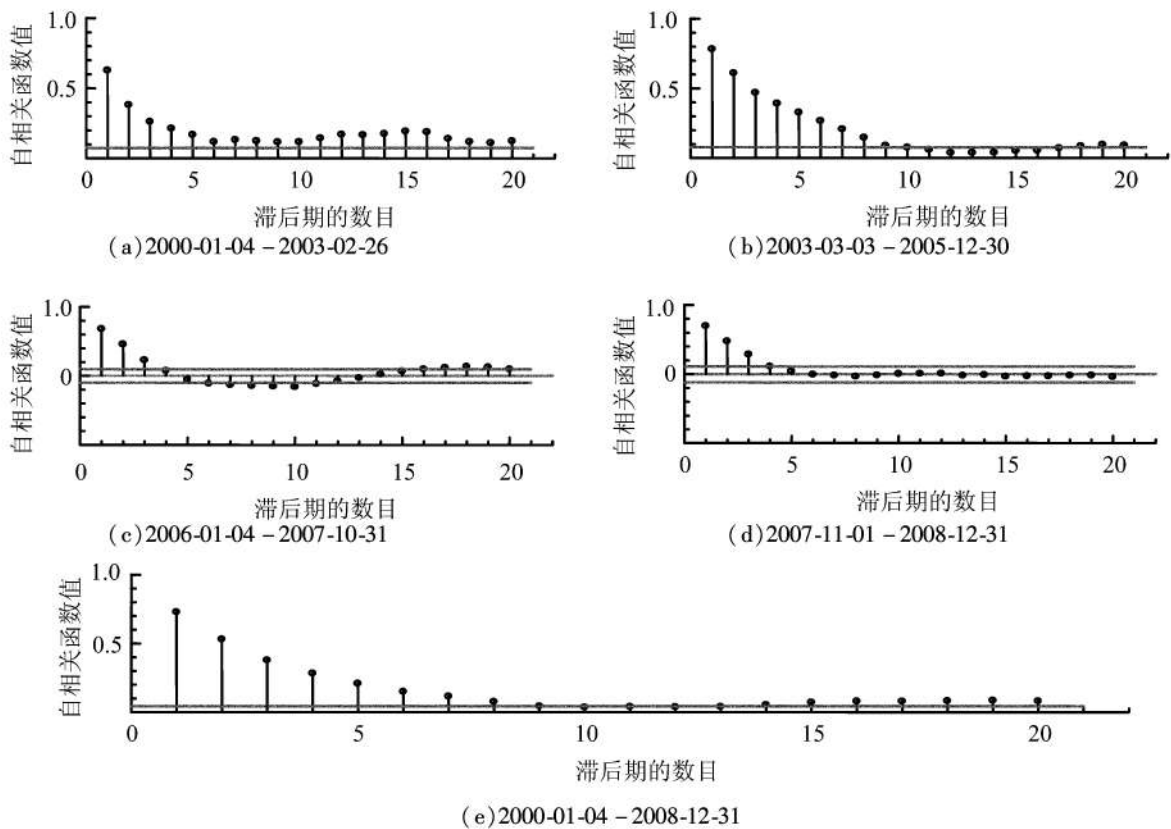


图 4 不同时间区间高频波动率跳跃持续期的自相关特征(基于 C\_TZ 统计量)

Figure 4 Autocorrelations for Duration of High-Frequency Volatility's Jump at Different Time Intervals(Based on C\_TZ Statistic)

表3 模型估计结果  
Table 3 Estimate Results of Models

3种持续期模型在总体区间的估计结果				
	ACD(1,1)模型	ACH(1,1)模型	扩展的ACH模型	
$\omega$	0.151(0.077)[0.049]	0.364(0.096)[0.012]		
$\alpha_1$	0.065(0.021)[0.002]	0.037(0.016)[0.000]	0.037(0.011)[0.000]	
$\beta_1$	0.893(0.037)[0.000]	0.903(0.052)[0.000]	0.892(0.058)[0.000]	
$\delta_0$	-	-	0.825(0.067)[0.001]	
$\delta_M$	-	-	-0.848(0.045)[0.003]	
$\delta_T$	-	-	-0.556(0.097)[0.011]	
$\delta_W$	-	-	0.263(0.069)[0.023]	
$\delta_{Th}$	-	-	-0.004(0.001)[0.000]	
$\log L$	-1 318.877	-1 248.290	-1 240.353	
ACH模型在各个时间区间的估计结果				
	2000-01-04 - 2003-02-26	2003-03-03 - 2005-12-30	牛市区间	熊市区间
$\omega$	0.259(1.392)	0.333(0.548)	1.088(0.871)	1.319(0.635)
$\alpha_1$	0.016*(0.011)	0.045*(0.027)	0.061*(0.038)	0.079*(0.083)
$\beta_1$	0.956*(0.041)	0.883*(0.070)	0.686*(0.092)	0.524*(0.045)

注:表中圆括号内数据为标准误,方括号内数据为p值,\*为在5%的显著水平下显著,logL为对数似然函数值。

波动率跳跃持续期显著的滞后自相关系数较长,说明高频波动率跳跃的持续期序列也具有聚集效应的特征。

表3为3种持续期模型在总体区间以及ACH模型在各个时间区间的估计结果。由表3可知,对于3种持续期模型而言,所有的估计参数都高度显著不为0,且估计的 $\alpha_1$ 和 $\beta_1$ 之和都大于0.900,意味着上证综指高频波动率跳跃持续期在总体时间区间存在较强的长记忆性。对于ACH(1,1)模型而言, $\alpha_1$ 和 $\beta_1$ 的估计值在各个时间区间都显著不为0,并且 $\alpha_1$ 和 $\beta_1$ 的估计值之和在各个时间区间都大于0.500,说明上证综指高频波动率跳跃持续期在各个时间区间都存在长记忆性特征,但在牛熊市区间其长记忆性要比其他区间弱,且牛市比熊市稍微强。由表3还可以推断,如果波动率跳跃持续期的条件期望为其均值1.926时,则在交易日为星期五时,上证综指高频波动率发生跳跃的概率约为 $0.364(\frac{1}{1.926+0.825})$ ,

即大概有 $\frac{1}{3}$ 的概率高频波动率会发生跳跃。然而当交易日为星期一时,高频波动率发生跳跃的概率增加至0.526。当交易日为星期二、星期三和星期四时,这个典型的风险概率分别为0.456、0.332和0.364。因此,在一周内,中国股市高频波动率在星期一发生跳跃的概率最大,而星期三发生跳跃的概率最小,大部分的较大跳跃都集中在星期一,意味着中国股市高

频波动率跳跃的持续期在总体时间区间具有周日历效应。

图5分别给出由ACH(1,1)模型和扩展的ACH模型估计的条件风险概率(见(15)式和(16)式中的 $h_t$ )的时序图,即高频波动率发生跳跃的概率的时序图。比较两图可以发现,加入每天更新信息的扩展的ACH模型估计的高频波动率发生跳跃的概率比ACH(1,1)模型估计有更加不稳定的波动。同时两图都展示了一个比较清晰的时序模式,在2003年至2004年和2007年至2008年两个区间,高频波动率发生跳跃的概率比较小,因此在这两个时间区间高频波动率发生显著跳跃的数量比较少,而在熊市行情中高频波动率发生跳跃的概率要比牛市的高,这与图3(a)和(b)得出的结论一致。中国股票市场在熊市行情中波动率发生跳跃的概率要比牛市高的一个可能的原因是,在熊市行情中,中国政府政策多变,干预市场过多。因此要使中国股票市场长期平稳运行,应该有一个规范化、法律化的市场运行环境。

## 6 结论

本研究采用修正的已实现门阈多次幂变差和C-TZ统计量,利用上证综指2000年1月4日至2008年12月31日的每5分钟的高频数据,估计高频波动率跳跃的时间序列,检验和分析中国股市高频波动率跳跃的各种特征,并运用ACD模型、ACH模型以及扩展的ACH模型进一步分析中国股市高频波动率



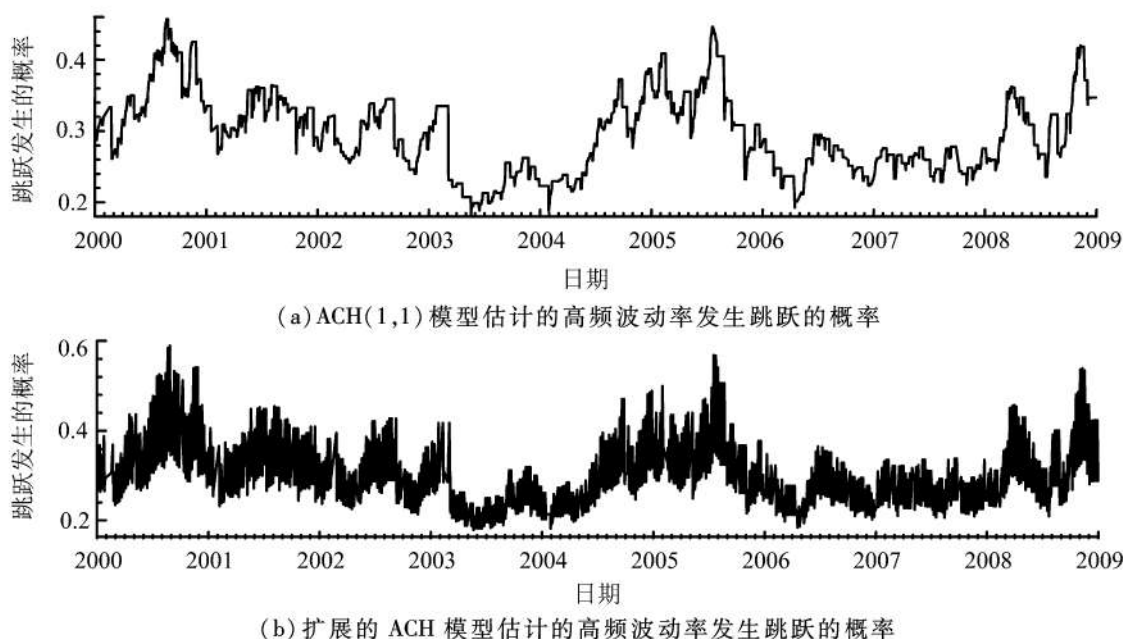


图5 ACH模型和扩展的ACH模型估计的高频波动率发生跳跃的概率的时序图  
Figure 5 Conditional Hazard Rates from ACH Model and Augmented ACH Model

跳跃的持续期的特征。实证结果表明,①基于上证综指的高频数据,用  $C_{TZ}$  统计量甄别日间波动率发生显著跳跃的比例比  $Z$  统计量的高,远高于对日或者更低频的数据应用参数跳跃扩散模型估计的跳跃强度;②在不同的时间区间内,无论是基于  $Z$  统计量还是  $C_{TZ}$  统计量,高频波动率跳跃的幅度都具有时变性,当跳跃发生时,跳跃对高频波动率的贡献占相当大的比例,并且在不同的时间区间内这个比例的均值具有相对稳定性,在整个考察期间,高频波动率跳跃强度和跳跃幅度的分布也同样具有时变性;③在整个考察期间,中国股市高频波动率跳跃表现出较强的正相关性(即聚集的特征),在牛市行情中高频波动率跳跃的自相关性要比熊市的强;④中国股市高频波动率跳跃的持续期在整个考察期间存在较强的长记忆性,牛市高频波动率跳跃持续期的长记忆性稍微比熊市强,并且在整个考察期间,高频波动率跳跃的持续期具有周日历效应,高频波动率发生跳跃的概率表现出一个较清晰的时序模式。

中国股票市场投机气氛比较严重,投资者较少关注当前收益,而注重资本利得,在市场上容易追涨杀跌,带来股价和股市波动率的频繁大幅度变动,因此应加强对投资者的理性投资教育,改善市场投资主体的构成,发挥投资基金的理性投资主导作用。

#### 参考文献:

- [1] Chan W H, Maheu M J. Conditional jump dynamics in stock market returns [J]. *Journal of Business and Economic Statistics*, 2002, 20(3): 377-389.
- [2] Eraker B, Johannes M S, Polson N. The impact of jumps in volatility and returns [J]. *Journal of Finance*, 2003, 58(3): 1269-1300.
- [3] Maheu J M, McCurdy T H. News arrival, jump dynamics and volatility components for individual stock returns [J]. *Journal of Finance*, 2004, 59(2): 755-793.
- [4] Barndor-Nielsen O E, Shephard N. Realized power variation and stochastic volatility [J]. *Bernoulli*, 2003, 9(2): 243-265.
- [5] Barndor-Nielsen O E, Shephard N. Econometric analysis of Realised covariation: High frequency covariance, regression and correlation in financial economics [J]. *Econometrica*, 2004, 72(3): 885-925.
- [6] Barndor-Nielsen O E, Shephard N. Impact of jumps on returns and realized variances: Econometric analysis of time-deformed levy process [J]. *Journal of Econometrics*, 2006, 131(2): 217-252.
- [7] Bollerslev T, Kretschmer U, Pigorsch C, Tauchen G. A discrete-time model for daily S&P 500 returns and realized variations: Jumps and leverage effects [J]. *Journal of Econometrics*, 2009, 150(2): 151-166.
- [8] Thomakos D D, Vortelinos D I. Realized volatility and jumps in the Athens stock exchange [R]. Greece: University of Peloponnese Working Paper, 2006: 1-27.
- [9] 潘祺. 跳跃扩散模型下上证指数估值研究 [J]. *金融经济*, 2006(6): 130-131.  
Pan Qi. Valuing the Shanghai index using jump diffusion model [J]. *Finance & Economy*, 2006(6): 130-131. (in Chinese)
- [10] Merton R C. Optimum consumption and portfolio rules in a continuous-time model [J]. *Journal of Economic Theory*, 1971, 3(4): 373-413.

- [11] 童汉飞,刘宏伟. 中国股市收益率与波动率跳跃性特征的实证分析[J]. 南方经济,2006(5): 61-72.  
Tong Hanfei, Liu Hongwei. Empirical investigation on rate and volatility with jump-GARCH model in Chinese stock market [J]. South China Journal of Economics, 2006(5):61-72. (in Chinese)
- [12] 杜军,刘建江. 上证指数混合跃变 GARCH 效应的实证分析[J]. 统计与决策,2006(13):31-33.  
Du Jun, Liu Jianjiang. Empirical investigation on the mixing jump-GARCH effects for SSE [J]. Statistics and Decision, 2006(13):31-33. (in Chinese)
- [13] 王春峰,姚宁,房振明,李晔. 中国股市已实现波动率的跳跃行为研究[J]. 系统工程,2008, 26(2):1-6.  
Wang Chunfeng, Yao Ning, Fang Zhenming, Li Ye. An empirical research on jump behavior of realized volatility in Chinese stock markets [J]. Systems Engineering, 2008, 26(2):1-6. (in Chinese)
- [14] Huang X, Tauchen G. The relative contribution of jumps to total price variance [J]. Journal of Financial Econometrics, 2005, 3(4):456-499.
- [15] Corsi F, Pirino D, Reno R. Threshold bipower variation and the impact of jumps on volatility forecasting [J]. Journal of Econometrics, 2010, 159(2):276-288.
- [16] 李胜歌. 基于高频数据的金融波动率研究 [D]. 天津:天津大学,2008:1-21.  
Li Shengge. The study of financial volatility based on high-frequency data [D]. Tianjin: Tianjin University, 2008:1-21. (in Chinese)
- [17] Hansen P R, Lunde A. Realized variance and market microstructure noise [J]. Journal of Business and Economic Statistics, 2006, 24(2):127-161.
- [18] Andersen T G, Bollerslev T, Diebold F X, Labays P. Modeling and forecasting realized volatility [J]. Econometrica, 2003, 71(2):579-625.
- [19] Fan J, Yao Q W. Nonlinear time series [M]. New York: Springer, 2003:193-212.
- [20] Engle R F, Russell J R. Autoregressive conditional duration: A new model for irregularly spaced transaction data [J]. Econometrica, 1998, 66(5):1127-1162.
- [21] Hamilton J D, Jorda O. A model of the federal funds rate target [J]. Journal of Political Economy, 2002, 110(5):1135-1167.
- [22] Tauchen G, Zhou H. Realized jumps on financial markets and predicting credit spreads [J]. Journal of Econometrics, 2011, 160(1):102-118.

## Empirical Research on Jump Behavior of High-Frequency Volatility Based on C\_TMPV in Chinese Stock Markets

Yang Ke, Chen Langnan

Lingnan College, Sun Yat-sen University, Guangzhou 510275, China

**Abstract:** Based on the every 5 minutes high-frequency financial data of SSE from 2000 to 2008. We utilized kernel-based estimator to estimate the high-frequency volatility, and corrected realized threshold multipower variation (C\_TMPV) to investigate the jump dynamics of volatility and the ACD model and the ACH model to analyze the features of jump duration. The results that both show the volatility and jump have agglomeration effect, the proportion of the trade day with the volatility jump is very high, and the size, the intensity and distribution of jumps in Chinese stock market are time-varying, but the jump's contribution to high-frequency volatility is relatively stable. In addition, during our analytical term of sample, the jumps of high-frequency volatility exhibit strong positive correlations, and the jump durations have the long term memory and weekday calendar effects.

**Keywords:** high-frequency volatility; jumps; C\_TMPV; C\_TZ test

**Received Date:** April 15<sup>th</sup>, 2010    **Accepted Date:** August 29<sup>th</sup>, 2010

**Funded Project:** Supported by the National Natural Science Foundation of China (70673116), the Humanities and Social Sciences Research base of Ministry of Education (05JJD790075), the National Social Science Foundation of China (07BJY167), the industry and regional development research innovation base of 985 Project of Sun Yat-sen University, the Humanities and Social Science research base of Guangdong Provincial college and the Natural Science Foundation of Guangdong Province (9151027501000032)

**Biography:** Yang Ke, a Hunan Yueyang native (1983 - ), is a Ph. D. candidate in the Lingnan College at Sun Yat-sen University. His research interests include financial economics and monetary economics, etc. E-mail: yangkeda@foxmail.com    □