

文章编号:1007-6735(2011)05-0499-09

# 股票价格运动的跳跃和杠杆效应研究

赵久伟, 肖庆宪

(上海理工大学 管理学院, 上海 200093)

**摘要:** 基于已实现变差理论和双幂次变差理论,在股票价格为半鞅过程的前提下将已实现波动率分解为连续波动部分和离散跳跃部分,研究了上证综指收益率和波动率的杠杆效应以及连续波动部分和离散跳跃部分分别在杠杆效应中所起的作用,并且分析了跳跃部分和连续波动部分在价格运动过程中的区别.由独立方程模型的估计结果发现各个模型扰动项之间存在着一种类似杠杆效应的非线性关系,然后再通过联立方程组模型进一步验证扰动项之间的内在依赖性.实证分析结果表明:收益率和波动率的杠杆效应主要是通过连续波动部分起作用.

**关键词:** 双幂次变差; 内在依赖性; 杠杆效应; 联立方程组

**中图分类号:** F 830      **文献标志码:** A

## Empirical analysis on jumps and leverage effect of equity price movements based on high-frequency data

ZHAO Jiu-wei, XIAO Qing-xian

(Business School, University of Shanghai for Science and Technology, Shanghai 200093, China)

**Abstract:** The leverage effect between returns and volatility of Shanghai Composite Index(SCI), was investigated and it was studied whether the observed so-called leverage effect is caused by a negative correlation between the lagged returns and the current continuous variance component and/or current jumps by using a nonparametric decomposition of the total price variation into two separate components based on the theories of realized variation and realized bipower variation. A discrete-time stochastic volatility model was introduced to distinguish the jump component and continuous volatility component of price movements. Then, a nonlinear relationship among the residuals was presented by use of univariate estimation results of the separate equations, which is similar to the commonly assumed lagged leverage effect. The modeling of the joint equation system allows to further assess the inter-dependencies among the residuals, and the estimation results of the joint model suggest that the leverage effect works primarily through the continuous volatility component.

**Key words:** bipower variation; inter-dependency; leverage effect; joint model

收稿日期: 2011-01-12

基金项目: 上海市重点学科建设资助项目(S30501)

作者简介: 赵久伟(1987-),男,硕士研究生.研究方向:数量经济学. E-mail: zhaojiuwei28@126.com.

肖庆宪(联系人),男,教授.研究方向:金融工程、数量经济学. E-mail: qxxiao@163.com

股票价格运动在金融市场中是一种复杂现象,尤其是股票价格波动规律引起了学者们的兴趣.上世纪末以来,金融市场波动率的测量和建模在实证金融领域和时间序列计量经济学领域一直是一个热门专题.在国外,首先利用参数估计法对连续时间模型的研究工作可以追溯到 Merton<sup>[1]</sup>,此后,Andersen 等<sup>[2]</sup>,Chernov 等<sup>[3]</sup>和 Eraker<sup>[4]</sup>研究了带跳跃的随机波动率模型.但是基于日数据或者更高频率数据的实证研究并没有严格区分模型的扩散部分和跳跃部分,所以,跳-扩散模型的参数估计方法仍然没有很好解决.近期一些文献采用相对次高频的日内数据测量波动率,如 Andersen 等<sup>[5]</sup>利用道琼斯 30 只股票 5 分钟收益率研究了个股波动率之间的相互依赖性. Barndorff-Nielsen 等<sup>[6]</sup>引入非参数方法,发现高频收益率的平方和收敛于价格过程的二次变差(quadratic variation),该测度包含了连续样本路径过程和跳跃两部分的相关测度. Barndorff-Nielsen 等<sup>[7-8]</sup>将二次变差方法扩展到幂变差和双幂次变差的非参数方法,认为当价格过程服从随机波动模型时,价格过程的二次变差依概率收敛于连续波动部分与跳跃部分之和,从而能挖掘出其潜在的跳跃过程. Andersen 等<sup>[9]</sup>利用非参数分解方法分析外汇现货市场、证券期货市场和利率期货市场,并指出跳跃部分对价格方差有显著贡献.

国内学者针对已实现波动率理论做了大量研究.李胜歌等<sup>[10]</sup>利用深成指高频数据对已实现双幂次变差和已实现多幂次变差的有效性进行了研究,得出双幂次变差在一般条件下比已实现波动率更有效.闵素芹等<sup>[11]</sup>和李胜歌等<sup>[12]</sup>研究了国内金融高频数据最优抽样方法理论.唐勇等<sup>[13]</sup>证明了在一定条件下已实现极差波动率优于已实现波动率,同时还研究了市场微观结构噪音对波动率测量的干扰程度.苗晓宇<sup>[14]</sup>基于超高频数据介绍了 5 种不同的风险度量方法.汪春峰等<sup>[15]</sup>基于双幂次变差的测量理论,利用价格过程的非参数分解方法研究了我国上证综指的已实现波动率和跳跃,结果发现已实现波动率的可测性均来自连续样本路径过程.综观上述研究工作,尚无文献从已实现变差理论和双幂次变差理论的角度阐述连续波动部分和跳跃部分在杠杆效应中所起的作用.

本文以上证指数为样本,基于二次变差理论和幂变差理论将跳跃部分从价格过程中分离出来,严格区分连续样本路径方差和跳跃方差,进而检验两种不同成分的统计特征,并在此基础上建

立收益率、连续波动部分和跳跃部分的离散时间模型,通过联立方程组分析收益率与波动率之间的非对称关系,以及不同成分对波动率的估计和杠杆效应的影响.

## 1 研究方法

### 1.1 两个重要概念

在无套利的前提下,假设对数价格  $p(t)$  为半鞅过程,则连续时间跳-扩散过程可表示为

$$p_t = \int_0^t \mu(s)ds + \int_0^t \sigma(s)dW(s) + \sum_{j=1}^{N(t)} \kappa(s_j) \quad (1)$$

均值过程  $\mu(t)$  是连续的局部有界变差过程,  $\sigma(t) > 0$  是右连左极(càdlàg)随机波动率过程,  $W(t)$  是标准布朗运动,  $N(t)$  为跳跃次数,  $\kappa(s)$  表示跳跃幅度. 则价格过程式(1)的二次变差可以表示为

$$[p]_t = p \lim \sum_j^n (p_{\tau_{j+1}} - p_{\tau_j})^2 \quad (2)$$

假设  $\tau_0 = 0 \leq \tau_1 \leq \tau_2 \leq \dots \leq \tau_n = t$  表示时间区间  $[0, t]$  的时间间隔,则当  $n \rightarrow \infty$  时有

$$\sup_j \{\tau_{j+1} - \tau_j\} \rightarrow 0$$

因而

$$[p]_t = \int_0^t \sigma^2(s)ds + \sum_{j=1}^{N(t)} \kappa^2(s_j) \quad (3)$$

式(3)表示可积波动率与跳跃平方的和,其中,当跳跃的次数  $N(t) = 0$  时,二次变差  $[p]_t$  等于可积波动率  $\int_0^t \sigma^2(s)ds$ .

为了简单起见,记第  $t$  天的收益率为

$$r_t = p_t - p_{t-1}, \quad t = 1, 2, \dots \quad (4)$$

第  $t$  天的第  $j$  个日内收益率表示为

$$r_{t,j} = p_{t-1+\frac{j}{M}} - p_{t-1+\frac{j-1}{M}}, \quad j = 1, 2, \dots, M \quad (5)$$

式中,  $M$  表示日内收益率的抽样频率,该收益率序列的平方和为

$$RV_t = \sum_{j=1}^M r_{t,j}^2 \quad (6)$$

$RV_t$  为已实现二次变差的自然估计量(或称为已实现方差,已实现波动率).特别的,当频率  $M$  趋近于无穷大时,已实现波动率依概率一致收敛于二次变差

$$RV_t \xrightarrow{M \rightarrow \infty} \int_{t-1}^t \sigma^2(s)ds + \sum_{j=N(t-1)+1}^{N(t)} \kappa^2(s_j) \quad (7)$$

为了区别连续部分与跳跃部分,Barndorff-Nielsen 等<sup>[7]</sup>第一次引入了双幂次变差测度

$$BV_t = \frac{\pi}{2} \sum_{j=2}^M |r_{t,j}| |r_{t,j-1}| \quad (8)$$

当抽样频率  $M$  趋近于无穷大时,双幂次变差  $BV_t$  一致收敛于可积波动率

$$BV_t \xrightarrow{M \rightarrow \infty} \int_{t-1}^t \sigma^2(s) ds \quad (9)$$

因此,已实现波动率和双幂次变差的这种不同特点为整个价格变差过程中的跳跃部分提供了一种非参数估计方法. Huang 等<sup>[16]</sup>提出了一种更强的 (robust) 跳跃测度

$$J_t = \log RV_t - \log BV_t \quad (10)$$

在实证研究中,本文依赖  $BV_t$  和  $J_t$  的联立模型解释这两种成分分别在日价格变差过程中的贡献. 跳跃测度  $J_t$  在理论上应该是严格非负的,然而,在实际操作过程中,由于抽样频率  $M$  有限,所以  $BV_t$  可能会出现大于  $RV_t$  的情况以至于造成  $J_t$  小于零.

## 1.2 基本模型

Corsi<sup>[17]</sup>提出异质自回归模型 (heterogenous autoregressive), 即 HAR-RV 模型. Andersen 等在文献[9]和文献[18]中应用该方法描述了双幂次变差  $BV_t$  的动态依赖性. Bollerslev 等<sup>[19]</sup>提出了 HAR-BV-GARCH 模型作为 HAR-RV 模型的一种扩展形式,本文通过波动率的波动率模型 (volatility-of-volatility, 即 HAR-BV-GARCH 模型) 描述  $BV_t$  模型中残差序列的异方差性. 考虑到对数双幂次变差序列呈现出的长记忆性,定义多时期 (multiperiod) 双幂次变差测度

$$(\log BV_t)_{t+1-k:t} = \frac{1}{k} \sum_{j=1}^k \log BV_{t-j} \quad (11)$$

当  $k=5$  和  $k=22$  时分别表示一周和一个月,则双幂次变差模型可表示为

$$\begin{aligned} \log BV_t = & \alpha_0 + \alpha_d \log BV_{t-1} + \alpha_w (\log BV)_{t-5:t-1} + \\ & \alpha_m (\log BV)_{t-22:t-1} + \theta_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \theta_2 I[r_{t-1} < 0] + \\ & \theta_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I[r_{t-1} < 0] + \sqrt{h_t} u_t \end{aligned} \quad (12)$$

$$h_t = \omega + \alpha_1 (\log BV_{t-1} - \chi'_{BV} \beta_{BV})^2 + \beta_1 h_{t-1} \quad (13)$$

在此假设  $u_t \sim N(0,1)$ . 式(13)中,  $\chi_{BV}$  表示对数双幂次变差模型的回归元,  $\beta_{BV}$  表示相应的系数. 式(12)右端的第二项、第三项和第四项分别表示滞后一天、滞后一周和滞后一个月的双幂次变差测度,而其余几项则通过加入虚拟变量  $I$  来描述连续波动部分的杠杆效应. 除此之外, Barndorff-Nielsen 等<sup>[8]</sup>通

过研究发现已实现波动率的波动对已实现波动率大幅度的变化反应程度不同,这也是本文采用 HAR-BV-GARCH 的主要原因.

其次,在跳跃模型中加入双幂次变差模型中的杠杆项,形式为

$$\begin{aligned} \log \left( \frac{RV_t}{BV_t} \right) = & \delta_0 + \phi_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \phi_2 I[r_{t-1} < 0] + \\ & \phi_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I[r_{t-1} < 0] + \\ & \sum_{j=1}^n \delta_j \log \left( \frac{RV_{t-j}}{BV_{t-j}} \right) + v_t \end{aligned} \quad (14)$$

在此假设扰动项  $v_t$  服从均值为 0, 方差为  $\sigma^2$  的正态分布. 众所周知,股票收益率与波动率之间存在非对称的负相关性,通过该模型使得我们可以侦查这种不对称性是由于连续波动部分的作用还是跳跃部分的作用.

根据混合分布假说<sup>[20]</sup> (mixture-of-distributions hypothesis), 为了描述整个价格波动过程,考虑收益率模型

$$r_t = \gamma_0 + \sum_{j=1}^d \gamma_j r_{t-j} + \sqrt{RV_t} \epsilon_t \quad (15)$$

其中,日收益率依赖于已实现波动率测度  $RV_t$ ,  $\epsilon_t$  服从标准正态分布.

## 2 样本描述和单方程参数估计

双幂次变差测度  $BV_t$  和跳跃测度  $J_t$  的一致性取决于抽样频率  $M$  的大小,然而在实际中,抽样频率会受到实际报价和交易频率等限制,以及高频分时收益率可能受到市场微观结构摩擦的影响,比如离散的价格采样过程和买卖报价差 (bid-ask spreads) 等. 此外,关于高频数据最优抽样频率的选择本身是一个热点问题,不同市场的抽样频率选择会有一些差异,一般来说 5 分钟的抽样频率包含了较大信息量. 本文选取 2000~2009 年上证综指 5 分钟数据作为高频采样数据进行研究,数据来自于万得 (WIND) 金融数据库 (www.wind.com.cn).

上证综指日收益率时间序列、对数已实现波动率序列、对数双幂次变差序列以及跳跃序列如图 1 (见下页) 所示,图中所有序列均呈现出了明显的波动率聚类效应 (clustering effect), 往往“大”的值的出现会造成这种异常值成群地出现. 在跳跃时间序列图中,可以观察到大量的正观察值,并且它们的值都很小. 同时,这些观察值中不乏还包含

了负值,原因可以归咎于价格过程的离散化和有 限的抽样频率.

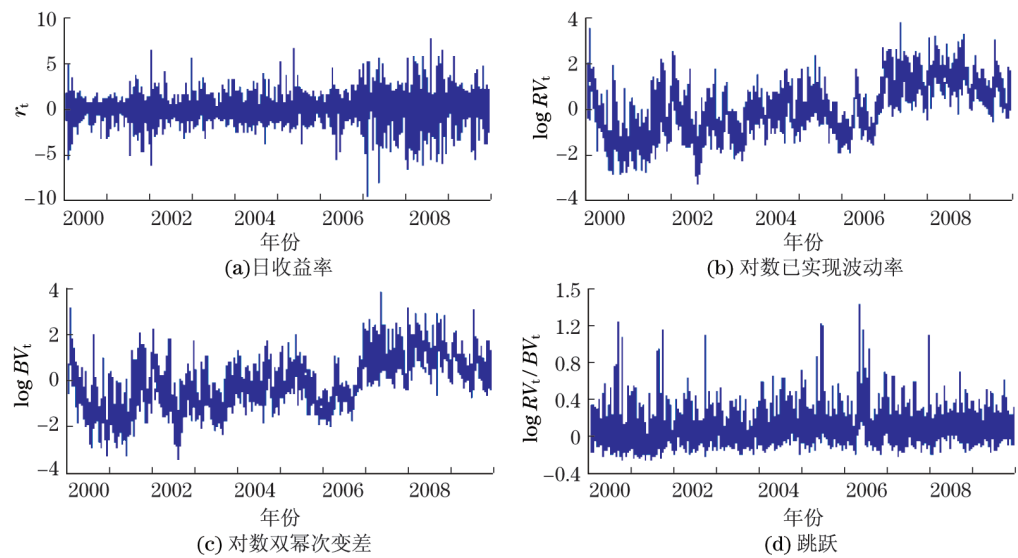


图 1 日收益率、对数已实现波动率、对数双幂次变差和跳跃的时间序列图

Fig.1 Time series of daily returns,logarithmic realized variance,logarithmic Bipower variation and jumps

表 1 呈现的是本文相关序列的描述性统计分析指标. 由表可见, 已实现波动率序列  $\sqrt{RV_t}$  的均值和标准差均大于双幂次变差的平方根  $\sqrt{BV_t}$  的均值和标准差, 并且这两种波动率序列的非条件分布均呈现出明显的右偏和高峰特征. 虽然这两组序列存在明显的右偏和高峰现象, 但是, 取对数后, 偏度系数和峰度系数明显下降, 近似地服从正态分布. 表中 Q(10) 表示滞后 10 阶的 Ljung-Box Q 统计量. 类似的, 从表 1 和图 2 可以观察到跳跃测度  $J_t$  的描述性统计量以及相应的核密度分布图清晰地显示出正的

偏度和尖峰态分布; 日收益率序列的非条件分布也呈现出了预期的超额峰度和负的偏度.

表 1 最后一列显示的为相关序列的 LB 统计量中, 所有波动率序列均呈现出显著的自相关, 这同样由图 3 的自相关函数图和偏相关函数图得到证实. 相比之下, 图 3 中收益率序列和跳跃测度没有表现出如此强烈的自相关性, 日收益率序列表现出 1 阶滞后, 而跳跃测度表现出 5 阶滞后, 也就是说当天的跳跃依赖于过去 5 天的跳跃.

虽然, 已经得到了每个序列各自的序列相关性,

表 1 已实现波动率、双幂次变差、跳跃和日收益率序列的统计量  
Tab.1 Descriptive statistics of logarithmic realized variance,logarithmic Bipower variation,jumps and daily returns

| 序列                       | 均值       | 标准差     | 中位数      | 偏度系数     | 峰度系数    | Q(10)   |
|--------------------------|----------|---------|----------|----------|---------|---------|
| $\sqrt{RV_t}$            | 1.165 8  | 0.726 5 | 0.976 6  | 1.651 8  | 7.345 8 | 9 572.3 |
| $\log RV_t$              | -0.037 7 | 1.173 7 | -0.047 4 | 0.069 1  | 2.502 6 | 12 612  |
| $\sqrt{BV_t}$            | 1.110 9  | 0.687 7 | 0.929 1  | 1.751 1  | 8.229 1 | 9 323.6 |
| $\log BV_t$              | -0.121 0 | 1.148 3 | -0.144 0 | 0.100 3  | 2.606 6 | 12 199  |
| $\log \frac{RV_t}{BV_t}$ | 0.082 9  | 0.182 3 | 0.054 6  | 1.558 4  | 8.487 4 | 381.71  |
| $r_t$                    | 0.050 3  | 1.605 8 | 0.078 9  | -0.210 4 | 5.818 7 | 28.83   |
| $r_t/\sqrt{RV_t}$        | 0.047 6  | 1.284 5 | 0.085 3  | -0.140 6 | 3.001 1 | 36.27   |

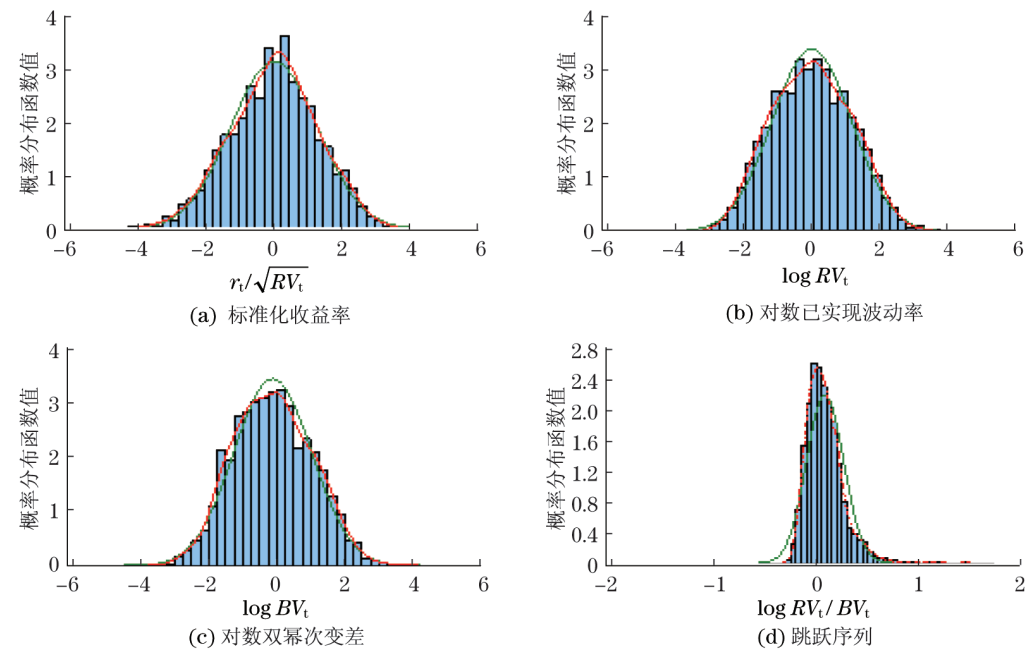


图 2 标准化收益率、对数已实现波动率、对数双幂次变差和跳跃序列的无条件分布图(实线)和核密度估计图(虚线)  
Fig.2 Unconditional distributions of standardized returns, logarithmic realized variance, logarithmic Bipower variation and jumps

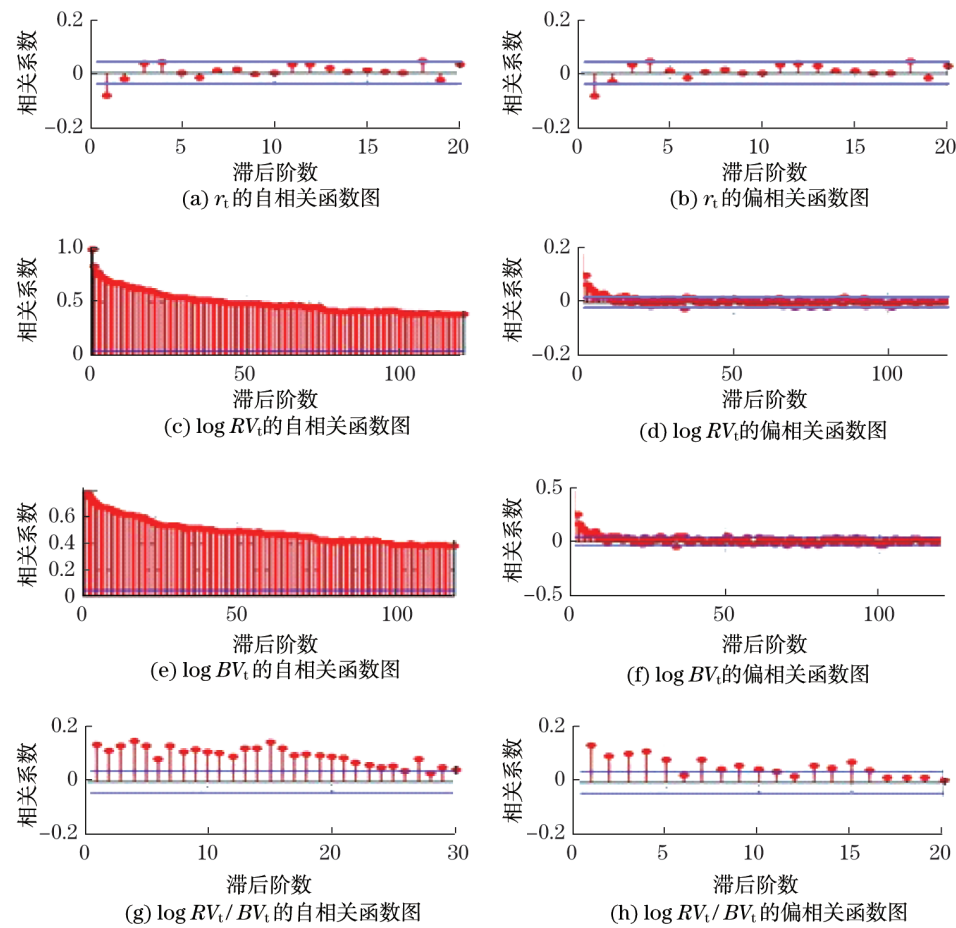


图 3 日收益率、对数已实现波动率、对数双幂次变差和跳跃序列的自相关函数图和偏相关函数图  
Fig.3 Sample autocorrelations and partial autocorrelations of daily returns, logarithmic realized variance, logarithmic Bipower variation and jumps

但是,在联立方程模型的建模中,序列之间存在的相关性以及依赖性则显得更为重要.一些文献提出了非对称性的观点,即“利空”消息相比“利好”消息会导致更大的波动率,过去的收益率扰动和现在的波动率之间存在负相关性,如 Engle 等<sup>[21]</sup>就提出了信息冲击曲线(news impact curve)的观点. Bollerslev 等<sup>[19]</sup>认为波动率的非对称性主要是双幂次变差序

列  $BV_t$  的非对称性造成的.

如果模型之间的扰动项相互独立,则每个模型的参数可以用最大似然方法单独估计.但是,这种假设可能存在一定的问题<sup>[9]</sup>,在此我们将独立地估计每个模型以便检验这种假设的有效性.表 2 分别给出了双幂次变差模型、跳跃模型和收益率模型通过 Eviews 计算出来的估计结果.

表 2 独立方程估计结果  
Tab.2 Single equation estimation results

| 双幂次变差方程    |         |         |         | 跳跃模型       |         |         |         | 收益率模型      |          |         |         |
|------------|---------|---------|---------|------------|---------|---------|---------|------------|----------|---------|---------|
| 参数         | 估计值     | 标准差     | p 值     | 参数         | 估计值     | 标准差     | p 值     | 参数         | 估计值      | 标准差     | p 值     |
| $\alpha_0$ | -0.154  | 0.025 2 | 0       | $\delta_0$ | 0.061 8 | 0.009 6 | 0       | $\gamma_0$ | 0.038 5  | 0.010 2 | 0.000 2 |
| $\alpha_d$ | 0.429 4 | 0.022 9 | 0       | $\phi_1$   | -0.013  | 0.007 1 | 0.06    | $\gamma_1$ | -0.034 2 | 0.013 8 | 0.013 5 |
| $\alpha_w$ | 0.322 4 | 0.036 3 | 0       | $\phi_2$   | -0.018  | 0.012 4 | 0.145 3 |            |          |         |         |
| $\alpha_m$ | 0.214 8 | 0.028 7 | 0       | $\phi_3$   | 0.015   | 0.009 5 | 0.113 8 |            |          |         |         |
| $\theta_1$ | 0.076 4 | 0.020 2 | 0.000 2 | $\delta_1$ | 0.093 8 | 0.019 1 | 0       |            |          |         |         |
| $\theta_2$ | 0.006 7 | 0.036 6 | 0.865 3 | $\delta_2$ | 0.064 5 | 0.019 2 | 0.000 8 |            |          |         |         |
| $\theta_3$ | 0.151 1 | 0.028 7 | 0       | $\delta_3$ | 0.087 9 | 0.020 2 | 0       |            |          |         |         |
| $\omega$   | 0.016 5 | 0.003 7 | 0       | $\delta_4$ | 0.106 2 | 0.017 3 | 0       |            |          |         |         |
| $\alpha_1$ | 0.066 8 | 0.009 5 | 0       | $\delta_5$ | 0.085 3 | 0.017 4 | 0       |            |          |         |         |
| $\beta_1$  | 0.882 8 | 0.017 8 | 0       | $\sigma^2$ | 0.031 1 | 0.000 6 | 0       |            |          |         |         |

首先,从表 2 给出的双幂次变差模型的参数估计结果中可以看到,  $\alpha_d$ 、 $\alpha_w$  和  $\alpha_m$  的估计值分别为 0.429 4, 0.322 4 和 0.214 8, 均为高度显著的. GARCH(1,1)模型充分地刻画了条件方差的变动过程,因为该模型的系数  $\alpha_1$  和  $\beta_1$  的估计结果同样显著异于零,并且满足稳定性条件  $\alpha_1 + \beta_1 < 1$ .  $\theta_1$  和  $\theta_3$  则直接说明了连续波动部分的杠杆效应的确存在,与预期结论一致,相对于同样大小的正的扰动,负的扰动会导致更大的波动率变化.相反,不显著的  $\theta_2$  说明了负面新闻不能对波动率造成有效的影响.此外,残差项的自相关函数和偏相关函数均没有显示出序列相关性.其次,跳跃模型的参数估计结果显示,滞后 1 至 5 阶的自相关参数估计量  $\delta_j (j = 1, 2, 3, 4, 5)$  均显著异于零,并且它们的值都很小,这与图 3 的情况一致.同时,在表 2 中发现一个非常值得注意的情况,那就是解释杠杆效应的参数  $\phi_1$ 、 $\phi_2$  和  $\phi_3$  的估计值均不显著,也就是说跳跃没有受到滞后 1 阶的收益率扰动的非对称影响,这与连续波动部分关于杠杆项的估计结果形成了鲜明的对比.最后,收益率模型中的参数估计值均显著,滞后 1 阶的自相关参数值也很小,  $\gamma_1 = -0.034 2$ .

3 残差项的内在依赖性和联立方程组

基于残差项之间相互独立的假设,3 个模型的参数估计已经在上文中给出,但是这种假设的正确与否值得关注.我们知道,收益率方程和波动率方程中的扰动项存在相关性主要是因为杠杆效应或者波动率反馈效应. Bollerslev 等<sup>[22]</sup>讨论了高频收益率的相关性,认为大样本的高频收益率能够更精确地评价波动率的非对称性.如果扰动项之间存在内在依赖性就会导致不一致的估计结果,但是,联立方程组能够解决这种依赖性并给出有效的估计结果<sup>[19]</sup>.通过计算双幂次变差方程、跳跃方程和收益率方程残差序列的相关系数得到相关系数矩阵  $\hat{\mathbf{A}}$

$$\hat{\mathbf{A}} = \begin{bmatrix} 1 & -0.170\ 4 & -0.086\ 6 \\ -0.170\ 4 & 1 & 0.028\ 8 \\ -0.086\ 6 & 0.028\ 8 & 1 \end{bmatrix}$$

双幂次变差扰动项与跳跃扰动项和收益率扰动项均呈现出负相关系数,跳跃扰动项和收益率扰动项的相关系数较小.由于非对称性的影响,扰动项除了存在线性关系还可能不存在非线性关系,我们将通过扰

动项  $u_t$ 、 $\varepsilon_t$  和  $v_t$  的散点图和非线性关系拟合以及核估计观察非线性关系,如图 4. 首先,  $u_t$  和  $\varepsilon_t$  之间存在非对称关系,这种关系类似于杠杆效应. 其次,  $v_t$  和  $\varepsilon_t$  之间的非对称关系较弱. 最后,  $v_t$  和  $u_t$  同样显示出了非对称且非线性关系. 通过图形分析可以确定 3 组残差序列的非对称依赖关系的确存在,为了进一步检验残差序列非对称关系的显著性和修正独立方程估计的误差,在此建立联立方程组模型并将非对称关系包含于模型内,所以,考虑联立方程组模型

$$r_t = \gamma_0 + \sum_{j=1}^d \gamma_j r_{t-j} + \sqrt{RV_t} \varepsilon_t$$

$$\log BV_t = \alpha_0 + \alpha_d \log BV_{t-1} + \alpha_w (\log BV)_{t-5:t-1} +$$

$$\alpha_m (\log BV)_{t-22:t-1} + \theta_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \theta_2 I[r_{t-1} < 0] +$$

$$\theta_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I[r_{t-1} < 0] + \sqrt{h_t} (u_t + g(\varepsilon_t))$$

$$h_t = \omega + \alpha_1 (\log BV_{t-1} - \chi'_{BV} \beta_{BV})^2 + \beta_1 h_{t-1}$$

$$\log \left( \frac{RV_t}{BV_t} \right) = \delta_0 + \psi_1 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} + \psi_2 I[r_{t-1} < 0] +$$

$$\psi_3 \frac{|r_{t-1}|}{\sqrt{RV_{t-1}}} I[r_{t-1} < 0] + \sum_{j=1}^n \delta_j \log \left( \frac{RV_{t-j}}{BV_{t-j}} \right) +$$

$$(v_t + m(u_t) + k(\varepsilon_t)) \quad (16)$$

其中,  $\begin{cases} g(\varepsilon_t) = g_1 \varepsilon_t + g_2 \varepsilon_t^2 \\ k(\varepsilon_t) = k_1 \varepsilon_t + k_2 \varepsilon_t^2 \\ m(u_t) = m_1 u_t + m_2 u_t^2 \end{cases}$

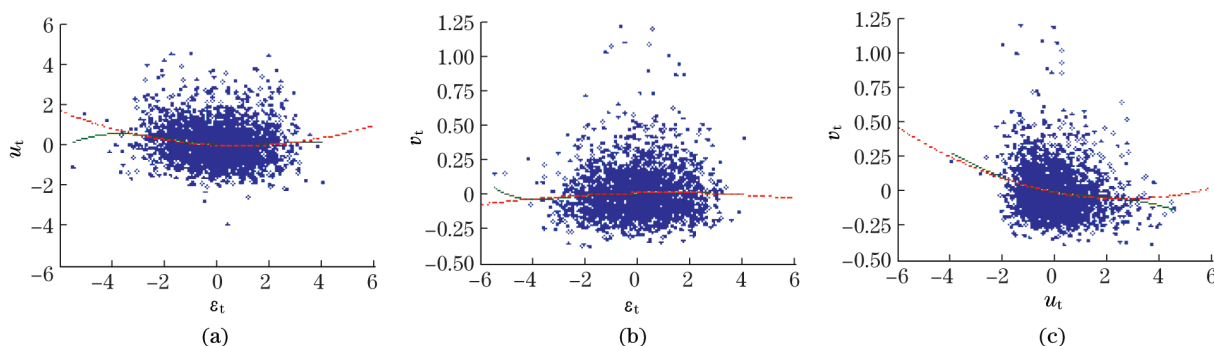


图 4 双幂次变差模型、跳跃模型和收益率模型残差序列之间的依赖性分析(虚线表示多元线性估计,实线表示核估计)

Fig.4 Dependency analysis of the residuals among the return equation, Bipower variation equation and jump equation

与独立模型相比,联立方程组通过函数  $g(\varepsilon_t)$ 、 $k(\varepsilon_t)$  和  $m(u_t)$  使得连续波动率模型扰动项依赖于  $\varepsilon_t$ , 跳跃模型扰动项依赖于  $\varepsilon_t$  和  $u_t$ . 鉴于此,考虑递归结构

$$f_y(y_t | z_{t-1}; \theta) = \frac{1}{\sqrt{h_t} \sqrt{RV_t}} f_\varepsilon \left( \frac{r_t - x'_r \beta_r}{\sqrt{RV_t}}; \theta_\varepsilon \right) \cdot$$

$$f_u \left( \frac{\log BV_t - x'_{BV} \beta_{BV}}{\sqrt{RV_t}} - g \left( \frac{r_t - x'_r \beta_r}{\sqrt{RV_t}} \right); \theta_u \right) \cdot$$

$$f_v \left( \log \left( \frac{RV_t}{BV_t} \right) - x'_{RV} \beta_{RV} - m(u_t) - k(\varepsilon_t); \theta_v \right)$$

其中,  $y_t = (\log BV_t, \log(RV_t/BV_t), r_t)'$ ,  $z_{t-1}$  为  $t-1$  时刻的信息集,  $x_{BV}$ 、 $x_r$  和  $x_{RV}$  分别表示对应模型的解释变量,  $f_\varepsilon$ 、 $f_u$  和  $f_v$  为残差项的密度函数, 而  $\theta_\varepsilon$ 、 $\theta_u$  和  $\theta_v$  表示密度函数的参数,  $\varepsilon_t$ 、 $u_t$  和  $v_t$  分别服从  $N(0,1)$ 、 $N(0,1)$  和  $N(0, \sigma^2)$ . 联立方程组的估计结果如表 3(见下页)所示,与独立方程估计结果相比,大部分参数值没有产生较大的变化. 从表 3 中可以看到,杠杆效应依然存在并且与第 2 节中

的推断相一致. 除此之外,高度显著的参数  $g_1$  和  $g_2$  进一步说明了连续波动率方程残差项和收益率方程残差项之间的非线性关系. 相反,对于跳跃方程和收益率方程残差项的函数关系式  $k(\varepsilon_t)$ ,  $k_1$  和  $k_2$  均为不显著参数. 最后,参数  $m_1$  和  $m_2$  的显著性进一步支持了连续波动率方程和跳跃方程残差项的非线性关系.

从表 3 还可以看到,估计结果中还包括了其它几个非显著参数,限制非显著参数的值为零,即  $\theta_2 = \psi_1 = \psi_2 = \psi_3 = k_1 = k_2 = 0$ , 然后重新给出模型估计结果,如表 4 所示. 通过对比表 3 和表 4 的估计结果发现,调整后的模型的估计结果没有因非显著参数限制为零而受到较大的影响,表 4 中的结果只表现出细微的变化. 最后,还有必要诊断残差序列  $u_t$ 、 $v_t$  和  $\varepsilon_t$  之间的相关性. 通过计算残差序列的相关系数,得到相关系数矩阵

$$\hat{B} = \begin{bmatrix} 1 & -0.0087 & -0.0034 \\ -0.0087 & 1 & 0.0209 \\ -0.0034 & 0.0209 & 1 \end{bmatrix}$$

和相关系数矩阵  $\hat{A}$  相比,矩阵  $\hat{B}$  中的数值更接近于 0,换一个角度说,二次多项式成功地描述了残差序列之间的非线性依赖性及其相互独立的假设并不可靠.从估计结果可以看出,上证指数的杠杆效应主要通过连续波动部分起作用,通过联立方程组模型可以抵消残差序列的相互依赖性.

表 3 联立方程模型的参数估计结果  
Tab.3 Estimation results of the joint model

| 双幂次变差方程    |          |         |         | 跳跃模型       |          |         |         | 收益率模型      |          |         |         |
|------------|----------|---------|---------|------------|----------|---------|---------|------------|----------|---------|---------|
| 参数         | 估计值      | 标准差     | p 值     | 参数         | 估计值      | 标准差     | p 值     | 参数         | 估计值      | 标准差     | p 值     |
| $\alpha_0$ | -0.201 0 | 0.026 8 | 0       | $\delta_0$ | 0.058 7  | 0.010 8 | 0       | $\gamma_0$ | 0.028 1  | 0.010 4 | 0.006 9 |
| $\alpha_d$ | 0.444 7  | 0.022 8 | 0       | $\delta_1$ | 0.102 0  | 0.019 2 | 0       | $\gamma_1$ | -0.039 0 | 0.013 7 | 0.004 8 |
| $\alpha_w$ | 0.315 9  | 0.036 2 | 0       | $\delta_2$ | 0.064 3  | 0.018 7 | 0.000 6 |            |          |         |         |
| $\alpha_m$ | 0.229 0  | 0.028 6 | 0       | $\delta_3$ | 0.086 4  | 0.020 2 | 0       |            |          |         |         |
| $\theta_1$ | 0.094 0  | 0.020 2 | 0       | $\delta_4$ | 0.102 5  | 0.017 3 | 0       |            |          |         |         |
| $\theta_2$ | 0.006 7  | 0.036 5 | 0.854 6 | $\delta_5$ | 0.082 6  | 0.017 2 | 0       |            |          |         |         |
| $\theta_3$ | 0.146 2  | 0.028 7 | 0       | $\psi_1$   | -0.012 0 | 0.007 3 | 0.091 0 |            |          |         |         |
| $\omega$   | 0.021 5  | 0.004 8 | 0       | $\psi_2$   | -0.016 0 | 0.012 6 | 0.197 3 |            |          |         |         |
| $\alpha_1$ | 0.067 2  | 0.009 9 | 0       | $\psi_3$   | 0.013 9  | 0.009 6 | 0.147 1 |            |          |         |         |
| $\beta_1$  | 0.864 7  | 0.021 9 | 0       | $\sigma^2$ | 0.030 2  | 0.000 6 | 0       |            |          |         |         |
| $g_1$      | -0.066 0 | 0.015 2 | 0       | $m_1$      | -0.056 0 | 0.006 3 | 0       |            |          |         |         |
| $g_2$      | 0.037 2  | 0.007 4 | 0       | $m_2$      | 0.015 9  | 0.005 9 | 0.007 5 |            |          |         |         |
|            |          |         |         | $k_1$      | 0.002 9  | 0.002 9 | 0.306 7 |            |          |         |         |
|            |          |         |         | $k_2$      | -0.002 0 | 0.001 6 | 0.203 5 |            |          |         |         |

表 4 剔除不显著参数后的联立方程模型的估计结果  
Tab.4 Restricted estimation results of the joint model

| 双幂次变差方程    |          |         |     | 跳跃模型       |          |         |         | 收益率模型      |          |         |         |
|------------|----------|---------|-----|------------|----------|---------|---------|------------|----------|---------|---------|
| 参数         | 估计值      | 标准差     | p 值 | 参数         | 估计值      | 标准差     | p 值     | 参数         | 估计值      | 标准差     | p 值     |
| $\alpha_0$ | -0.193 0 | 0.026 8 | 0   | $\delta_0$ | 0.041 3  | 0.005 6 | 0       | $\gamma_0$ | 0.026 5  | 0.010 4 | 0.010 5 |
| $\alpha_d$ | 0.444 6  | 0.022 8 | 0   | $\delta_1$ | 0.101 6  | 0.019 1 | 0       | $\gamma_1$ | -0.038 0 | 0.013 6 | 0.005 4 |
| $\alpha_w$ | 0.316 6  | 0.036 2 | 0   | $\delta_2$ | 0.067 4  | 0.018 5 | 0.000 3 |            |          |         |         |
| $\alpha_m$ | 0.228 3  | 0.028 5 | 0   | $\delta_3$ | 0.087 2  | 0.020 0 | 0       |            |          |         |         |
| $\theta_1$ | 0.088 1  | 0.016 7 | 0   | $\delta_4$ | 0.104 0  | 0.017 2 | 0       |            |          |         |         |
| $\theta_2$ |          |         |     | $\delta_5$ | 0.084 8  | 0.017 2 | 0       |            |          |         |         |
| $\theta_3$ | 0.152 4  | 0.017 4 | 0   | $\psi_1$   |          |         |         |            |          |         |         |
| $\omega$   | 0.021 2  | 0.004 7 | 0   | $\psi_2$   |          |         |         |            |          |         |         |
| $\alpha_1$ | 0.067 2  | 0.009 9 | 0   | $\psi_3$   |          |         |         |            |          |         |         |
| $\beta_1$  | 0.865 6  | 0.021 7 | 0   | $\sigma^2$ | 0.030 3  | 0.000 6 | 0       |            |          |         |         |
| $g_1$      | -0.063 0 | 0.015 0 | 0   | $m_1$      | -0.057 0 | 0.006 3 | 0       |            |          |         |         |
| $g_2$      | 0.030 2  | 0.007 4 | 0   | $m_2$      | 0.015 8  | 0.005 9 | 0.007 4 |            |          |         |         |
|            |          |         |     | $k_1$      |          |         |         |            |          |         |         |
|            |          |         |     | $k_2$      |          |         |         |            |          |         |         |

4 结 论

近几年,波动率一直是实证金融研究和时间序列计量经济学研究领域的一个热点,因为金融资产收益波动率的估计和预测对资产定价、投资组合的

选择和金融衍生产品的设计起着至关重要的作用.采用“已实现”波动率测度和双幂次变差测度作为市场波动率的度量方法,避免了较低频率金融数据中复杂的参数估计,给其它参数波动率估计模型提供了一种间接的估计方法.

本文利用非参数分解方法将已实现波动率分解

为连续波动部分和离散跳跃部分,基于已实现变差理论和双幂次变差理论,在价格过程符合随机波动模型的前提下,对上证综指高频数据进行分析得出以下结论:

a. 通过联立方程组建模,发现上证指数的杠杆效应主要通过连续波动部分起作用,而不是跳跃部分。

b. 模型有效地检验了残差序列之间的内在依赖性,同时指出残差序列之间存在类似的非对称现象。

c. 基于高频数据的“已实现”波动率测度和双幂次变差测度衍生出价格变动的跳跃过程,通过建立关于波动率、跳跃和收益率的方程组模型,清晰地刻画了连续波动部分和跳跃部分的动态特征并发现跳跃部分的持续性较差,跳跃部分的可测性弱于连续波动部分,因此,有效地改善了预测波动率的预测能力,这对波动率的区间预测和风险管理决策的研究意义重大。

#### 参考文献:

- [1] MERTON R C. On estimating the expected return on the market: An exploratory investigation[J]. *Journal of Financial Economics*, 1980, 8(4): 323 - 361.
- [2] ANDERSEN T G, BENZONI L, LUND J. An empirical investigation of continuous-time equity return models[J]. *Journal of Finance*, 2002, 57(3): 1239 - 1284.
- [3] CHERNOV M, GALLANT A R, GHYSELS E, et al. Alternative models for stock price dynamics[J]. *Journal of Econometrics*, 2003, 116(1/2): 225 - 257.
- [4] ERAKER B. Do stock prices and volatility jump? Reconciling evidence from spot and option prices[J]. *Journal of Finance*, 2004, 59(3): 1367 - 1404.
- [5] 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.
- [6] BARNORFF-NIELSEN O E, SHEPHARD N. Estimating quadratic variation using realized variance[J]. *Journal of Applied Econometrics*, 2002, 17(5): 457 - 477.
- [7] 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.
- [8] BARNDORFF-NIELSEN O E, SHEPHARD N. How accurate is the asymptotic approximation to the distribution of realised variance[Z]. Oxford: University of Oxford, 2001.
- [9] ANDERSEN T G, BOLLERSLEV T, DIEBOLD F X. Roughing it up: Including jump components in the measurement, modeling and forecasting of return volatility[J]. *The Review of Economics and Statistics*, 2007, 89(4): 701 - 720.
- [10] 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.
- [11] 闵素芹, 柳会珍. “已实现”波动率中最优抽样频率的选择[J]. *统计与决策*, 2009, 13(13): 13 - 15.
- [12] 李胜歌, 张世英. 基于金融高频数据波动率计算方法的比较研究[J]. *中国地质大学学报(社会科学版)*, 2008, 8(1): 80 - 83.
- [13] 唐勇, 张世英. 已实现波动和已实现极差波动的比较研究[J]. *系统工程学报*, 2007, 22(4): 437 - 442.
- [14] 苗晓宇. (超)高频数据视角下金融风险度量研究进展[J]. *经济论坛*, 2010, (8): 202 - 207.
- [15] 汪春峰, 姚宁, 房振明, 等. 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.
- [16] HUANG X, TAUCHEN G. The relative contribution of jumps to total price variance[J]. *Journal of Financial Econometrics*, 2005, 3(4): 456 - 499.
- [17] CORSI F. A simple long memory model of realized volatility[D]. Lugano: University of Lugano and Swiss Finance Institute, 2004.
- [18] ANDERSEN T G, BOLLERSLEV T, HUANG X. A reduced form framework for modeling and forecasting jumps and volatility in speculative prices[J]. *Journal of Econometrics*, 2011, 160(1): 176 - 189.
- [19] BOLLERSLEV T, KRETSCHMER U, PIGORSCH C, et al. A discrete-time model for daily S & P500 returns and realized variations: Jumps and leverage effects[J]. *Journal of Econometrics*, 2009, 150(2): 151 - 166.
- [20] FORSBERG L, BOLLERSLEV T. Bridging the gap between the distribution of realized (ECU) volatility and ARCH modeling (of the euro): The GARCH - NIG model[J]. *Journal of Applied Econometrics*, 2002, 17(5): 535 - 548.
- [21] ENGLE R F, NG V K. Measuring and testing the impact of news on volatility[J]. *Journal of Finance*, 1993, 48(5): 1749 - 1778.
- [22] BOLLERSLEV T, LITVINOVA J, TAUCHEN G. Leverage and volatility feedback effects in high-frequency data[J]. *Journal of Financial Econometrics*, 2006, 4(3): 353 - 384.