

· 证券与投资 ·

基于高频数据的金融市场波动溢出分析

张瑞锋,汪同三

(中国社会科学院 数量经济与技术经济研究所,北京 100732)*

摘要:在讨论“已实现”波动率、“已实现”协方差基础上,针对金融市场的高频数据,引入“已实现”波动变结构,分阶段计算“已实现”波动率的相关系数,检验“已实现”波动率相关系数,判断在变结构点前后是否发生显著变化,从而分析金融市场之间的波动溢出效应,并进行实证分析。

关键词:高频数据;金融市场;波动溢出

中图分类号:F224

文献标识码:A

文章编号:1003-7217(2013)01-0021-05

一、引言

波动溢出指的是波动从一个金融市场传递到另一个金融市场,普遍存在于不同金融市场之间。因此,波动溢出效应可能存在于不同区域的市场之间,也可能存在于不同类型的金融市场之间。

在金融市场中,由于信息是连续地影响证券市场价格的变动,如果使用离散的数据就会造成信息不同程度的缺失,若采集频率高,则信息丢失就越少;反之,信息丢失越多。高频时间序列指以每小时、每分钟甚至每秒为频率所采集的数据。由于使用高频时间序列比低频时间序列包含更多的信息,因此,有必要对更高频金融数据加以研究,以深入分析金融市场的波动特征。其次,随着信息化进程的推进,高频的金融数据收集更方便。另外,金融市场决策者在进行短期决策时,不仅需要市场的长期信息,而且还需要即时的信息,因此,使用高频数据研究金融市场之间波动溢出,能够为金融决策者提供更及时的决策信息。

本文在讨论高频时间序列的“已实现”波动率及协方差的基础上,通过判断在变结构点前后不同金融市场之间相关系数是否显著发生变化来分析金融市场之间的波动溢出问题。最后,以股票市场的价格为例进行了实证分析。

二、高频时间序列的“已实现”波动率

(一)高频数据的收益率

高频时间序列是在等时间间隔(如每1分钟,每5分钟,每30分钟等)上采集的时间序列,股票市场收盘价就是交易期间最后一笔交易的交易价格,如果在最后时间间隔内没有发生交易,就使用上一时间间隔的收盘价。

由于收益率的统计性质比价格的统计性质好,因此,在研究金融市场时,使用收益率(价格的差分)建模。即:

$$R_{t,n} = \log(P_{t+n/N}) - \log(P_{t+(n-1)/N}) \quad (1)$$

其中, $R_{t,n}$ 是收益率, $P_{t+n/N}$ 是股票市场在 $t+n/N$ 时刻的价格, N 是抽样频率下的一天中总的样本个数, n 是日内时段序列号。

(二)“已实现”波动率

金融市场中假设不存在风险套利,则金融资产的对数收益率是一个特殊半鞅过程。设 $(p(t))_{t \in [0,T]}$ 是定义在完备空间上的金融资产的对数价格过程, $p(t)$ 是 t 时刻包含在空间信息集合内。则投资于该金融资产 Δ 时段上的对数收益率为:

$$r(t, \Delta) = p(t) - p(t - \Delta) \quad (2)$$

其中, $\Delta > 0$ 表示时间间隔,特殊地 $r(t, t) \equiv r(t) = p(t) - p(0)$ 。

根据 Andersen 和 Bollerslev 等学者对西方发达国家发达金融市场的高频金融时间序列的研究,“已实

* 收稿日期: 2012-05-27

基金项目: 博士后基金项目(20100481209)、国家社科基金项目(10BGL056)、国家自然科学基金项目(71171056)

作者简介: 张瑞锋(1972—),男,河北廊坊人,中国社会科学院数量经济与技术经济研究所博士后,河北经贸大学经济管理实验中心副教授,研究方向:金融计量、数量经济、金融市场波动溢出;汪同三(1948—),男,江苏南京人,博士,中国社会科学院数量经济与技术经济研究所研究员,研究方向:经济模型理论、方法论及经济预测和政策分析。

现”波动率通常具有下列性质:^[1-4]

(1) “已实现”方差与“已实现”标准差的无条件分布均为极端右偏,而且峰度极高;

(2) 取对数后的“已实现”标准差的无条件分布是近似正态分布;

(3) $\frac{\text{日间收益率}}{\text{“已实现”标准差}}$ 的条件分布是正态分布;

(4) “已实现”波动率的自相关系数是按双曲线的趋势下降;

(三) 多维高频时间序列的“已实现”协方差

在 Barndorff-Nielsen 和 Shepard 的研究中^[5-7]提出了“已实现”协方差阵(Realized Covariance Matrix, 简称 RCM)的概念,徐正国^[8]在他的硕士论文中讨论了“已实现”协方差阵的性质。

若金融资产的价格过程是 $P(t) = (P_1(t), \dots, P_N(t))'$, 其中, $P_i(t)$ 是第 i 项资产在 t 时刻的价格。对价格向量取对数, 得:

$$p(t) = (\log [P_1(t)], \dots, \log [P_N(t)])' \quad (3)$$

根据(3)式定义收益率向量:

$$r(t+\Delta, \Delta) = p(t+\Delta) - p(t) \quad (4)$$

“已实现”协方差阵就是把所有 t 时刻到 $t+1$ 时刻的时段等分成 n 个时间段, 每个小时时间段上的收益率为:^[5-7]

$$r(t+j/n, 1/n) = p(t+j/n) - p(t) \quad (5)$$

因此, 从 t 时刻到 $t+1$ 时刻上的收益率矩阵为:

$$R_{t,t+1} = (r(t+1/n, 1/n), \dots, r(t+1, 1/n)) \quad (6)$$

“已实现”协方差阵定义为:

$$RCM_{t+1} = R_{t,t+1} R'_{t,t+1} \quad (7)$$

N 维“已实现”协方差阵是 $N \times N$ 的对称矩阵, 当 $N=1$ 时, 协方差阵就退化成“已实现”波动率了; 当 $N>1$ 时, 对角线上的元素 $RCM_t(i, i)$ 是变量 i 的“已实现”波动率 $RV_{t(i,i)}$; 非对角线元素 $RCM_t(i, j)$ 则是变量 i 与变量 j ($i \neq j$) 的“已实现”协方差 $RC_{t(i,j)}$ 。“已实现”协方差 $RC_{t(i,j)}$ 主要描述的是变量 i 和变量 j 之间的相关性。

$$RCM_t = \begin{bmatrix} RV_t(1,1) & \dots & RC_t(1,N) \\ \dots & \dots & \dots \\ RC_t(N,1) & \dots & RV_t(N,N) \end{bmatrix} \quad (8)$$

由式(8)可知: “已实现”协方差阵既包括对角线元素——单个变量的“已实现”波动率, 也包括了非对角线元素——不同变量之间的“已实现”协方差。

三、基于高频数据的金融市场间波动溢出分析

由于外部环境的变化可能使变量间的相关结构发生改变, 因此, 要分析不同金融市场之间是否存在波动溢出, 可以先找出各金融市场波动的变结构点; 根据“已实现”协方差矩阵的结果, 计算不同变量间的“已实现”相关系数, 通过检验“已实现”相关系数在波动变结构点前后是否发生显著变化, 依此分析判断不同金融市场之间是否存在波动溢出。

(一) “已实现”波动变结构点的诊断

关于波动变结构的检测, 国内外学者进行了大量的讨论与尝试, 黄违洪、张世英提出了一种针对线性模型变结构的 GBV (General Bayesian method with vague prior information) 检测法, GBV 法具有计算简单、过程直观、通用性强等优点, 但 GBV 法无法判定结构变化点的数目^[9]; Kim^[10]提出一种与 GBV 法类似的方法, 他构造了检验线性模型平稳性的 Bayes 显著检验量, 以检测到单个参数的非平稳性; Kim 与 Kon^[11]将这种方法直接用于检测方差平稳性, 这种程序不仅在计算上更加有效, 而且能估计出未知的变结构点数目。本文采用 Kim 与 Kon 提出的检测方法诊断“已实现”波动率变结构点。

将“已实现”协方差矩阵 RCM_t 中“已实现”波动率 $RV_{t(i,i)}$ 形成新的矩阵:

$$\begin{bmatrix} RV_1(1,1) & RV_2(1,1) & \dots & RV_T(1,1) \\ RV_1(2,2) & RV_2(2,2) & \dots & RV_T(2,2) \\ \dots & \dots & \dots & \dots \\ RV_1(N,N) & RV_2(N,N) & \dots & RV_T(N,N) \end{bmatrix} \quad (9)$$

其中 T 是样本容量。对 $\{RV_t(i,i)\}_{t=1}^T, i=1, 2, \dots, N$ (N 表示金融市场的个数) 进行变结构点诊断。得到波动变结构点:

$$\begin{bmatrix} \tau_{11} & \tau_{12} & \dots & \tau_{1k_1} \\ \tau_{21} & \tau_{22} & \dots & \tau_{2k_2} \\ \dots & \dots & \dots & \dots \\ \tau_{N1} & \tau_{N1} & \dots & \tau_{Nk_N} \end{bmatrix} \quad (10)$$

其中, τ_{ik_i} 表示第 i 个金融市场“已实现”波动的第 k_i 个变结构点, 共 k_i 个变结构点。

(二) 金融市场间相关系数计算

借鉴 Barndorff-Nielsen 和 Shepard 等人的研究^[12], 得到基于“已实现”波动率的金融市场之间相关系数的计算公式:

$$\rho_{ij} = \frac{\sum_{t=1}^T RC_t(i, j)}{\sqrt{\sum_{t=1}^T RV_t(i, i) \sum_{t=1}^T RV_t(j, j)}} \quad (11)$$

其中, ρ_{ij} 表示了第 i 个金融市场与第 j 个金融市场的“已实现”相关系数。

为了便于分析判断第 j 个金融市场“已实现”波动对第 i 个金融市场“已实现”波动是否存在溢出, 要根据变结构点 $\tau_{j1}, \dots, \tau_{jk_j}$ 分阶段计算出相应的“已实现”相关系数:

$$\begin{aligned} \rho_{ij,1} &= \frac{\sum_t RC_t(i, j)}{\sqrt{\sum_t RV_t(i, i) \sum_t RV_t(j, j)}} \quad \dots \\ t &= 1, \dots, \tau_{j1} \\ \rho_{ij,2} &= \frac{\sum_t RC_t(i, j)}{\sqrt{\sum_t RV_t(i, i) \sum_t RV_t(j, j)}} \quad \dots \\ t &= \tau_{j1}, \dots, \tau_{j2} \\ \dots & \quad \dots \quad \dots \\ \rho_{ij,k_j} &= \frac{\sum_t RC_t(i, j)}{\sqrt{\sum_t RV_t(i, i) \sum_t RV_t(j, j)}} \quad \dots \\ t &= \tau_{jk_j}, \dots, T \end{aligned} \quad (12)$$

(三) 金融市场间波动溢出判断分析

若要分析判断第 j 个金融市场“已实现”波动对第 i 个金融市场“已实现”波动是否存在溢出, 可以依次判断“已实现”相关系数 ρ_{ij} 在变结构点 $\tau_{j1}, \dots, \tau_{jk_j}$ 处是否显著发生变化。

对变结构点 $\rho_{ij,1}, \rho_{ij,2}$ 前后的“已实现”相关系数 $\rho_{ij,1}, \rho_{ij,2}$, 假设:

$$H_0: \rho_{ij,1} = \rho_{ij,2} \quad H_1: \rho_{ij,1} \neq \rho_{ij,2}$$

为了研究的精确性, 最大程度的减小抽样分布与正态分布的偏差, 这里, 对“已实现”相关系数 ρ_{ij} 进行 Fisher 转换^[13]:

$$\bar{\rho}_{ij,l} = \frac{1}{2} \ln \left(\frac{1 + \rho_{ij,l}}{1 - \rho_{ij,l}} \right), l = 1, \dots, k_j \quad (13)$$

根据 Pedhazur 的研究^[13], 这里构造 Z 检验的统计量为:

$$Z = \frac{\bar{\rho}_{ij,1} - \bar{\rho}_{ij,2}}{\sqrt{\frac{1}{\tau_{j1} - 3} + \frac{1}{(\tau_{j2} - \tau_{j1}) - 3}}} \quad (14)$$

统计量 Z 近似地服从 $(0, 1)$ 正态分布^[13]。给定置信水平 α , 若 $|Z| \geq z_{\alpha/2}$, 则拒绝 $H_0: \rho_{ij,1} = \rho_{ij,2}$, 否则接受 $H_0: \rho_{ij,1} = \rho_{ij,2}$ 。

若接受原假设, 则说明在变结构点 τ_{j1} 前后的“已实现”相关系数没有发生显著变化, 即在时间段

$1 \sim \tau_{j2}$ 内, 第 j 个金融市场“已实现”波动没有传递到第 i 个金融市场, 说明两个金融市场彼此之间虽然存在相互影响, 但影响不显著, 即没有形成波动溢出; 若拒绝原假设, 说明在变结构点 τ_{j2} 前后的“已实现”相关系数发生显著变化, 即在时间段 $1 \sim \tau_{j2}$ 内, 第 j 个金融市场“已实现”波动传递到第 i 个金融市场, 说明第 j 个金融市场与第 i 个金融市场存在波动溢出。

对变结构点 $\tau_{j1}, \dots, \tau_{jk_j}$ 前后的“已实现”相关系数, 依次重复上述步骤并进行 Z 检验, 根据检验结果分析判断第 j 个金融市场与第 i 个金融市场在不同变结构点区间段内是否存在波动溢出。

重复上述步骤, 根据检验结果分析判断其它金融市场之间是否存在波动溢出。

四、高频数据的股票市场间波动溢出实证分析

(一) 数据描述

由于国外股票市场的高频数据获得比较困难, 这里以上证指数和深圳成份指数的高频时间序列为例, 数据来源于天相数据分析系统, 收集的是每分钟采集频率的收盘价, 时间区间: 2004-01-02 ~ 2006-04-14, 共涉及到 551 个交易日。上海与深圳证券交易所每天的交易时间为上午 9:30 ~ 11:30、下午 13:00 ~ 15:00, 根据徐正国^[13] 对上证指数和深圳成份指数的实证结果, 10 分钟的抽样频率其微观结构误差最小, 且这样的抽样频率也足够高。所以, 这里选择 10 分钟的抽样频率, 共得到 $551 \times 48 = 26448$ 个数据。

首先, 采用式(4) 分别对每个指数数据计算每 10 分钟的高频收益率。其次, 采用式(8) 计算“已实现”协方差阵。

$$RCM_t = \begin{bmatrix} RV_t(SH) & RC_t(SH, SZ) \\ RC_t(SH, SZ) & RV_t(SZ) \end{bmatrix}$$

其中, $RV_t(SH)$ 是上证指数(SH) 的“已实现”方差, $RV_t(SZ)$ 是深圳成份指数(SZ) 的“已实现”方差, $RC_t(SH, SZ)$ 是上证指数和深圳成份指数之间的“已实现”协方差。

(二) “已实现”波动变结构点的诊断及相关系数计算、检验

为了研究分析上海股票市场与深圳股票市场之间的波动溢出问题, 首先, 采用 Bayes 诊断程序, 对样本内的深圳成份指数、上证指数收益率的“已实现”波动, 分别诊断变结构点; 其次, 根据深圳成份收益率的“已实现”波动变结构点, 分阶段计算深圳股

票市场与上海股票市场之间的相关系数 ρ_{SZ-SH} , 根据上证指数收益率的“已实现”波动变结构点, 分阶段计算上海股票市场与深圳股票市场之间的相关系数 ρ_{SH-SZ} ; 最后, 根据 Z 检验, 检验在变结构点前后, 相应的相关系数是否显著发生的变化。表 1 列出了深圳成份指数收益率的“已实现”波动变结构点的诊断结果, 以及“已实现”相关系数计算结果和 Z 检验结果。表 2 列出了上证指数收益率的“已实现”波动变结构点的诊断结果, 以及“已实现”相关系数计算结果和 Z 检验结果。

表 1 深圳成份指数收益率“已实现”波动变结构点诊断及 Z 检验结果

k	τ	日期	ρ_{SZ-SH}	Z
1	19	2004-02-09	0.98752	-1.2011
2	44	2004-03-15	0.99431	-1.4704
3	77	2004-04-29	0.99729	9.7061*
4	104	2004-06-14	0.80466	-4.2887*
5	152	2004-08-19	0.97174	1.3158
6	233	2004-12-17	0.95573	-5.3592*
7	267	2005-02-04	0.99072	4.974*
8	290	2005-03-18	0.916	-2.1141*
9	321	2005-05-09	0.97167	-3.6412*
10	359	2005-07-01	0.99445	-3.7573*
11	377	2005-07-27	0.99892	3.8179*
12	430	2005-10-17	0.99049	-5.4264*
13	487	2006-01-06	0.99851	15.96*
14	551	2006-04-14	0.75273	

表 2 上证指数收益率“已实现”波动变结构点诊断及 Z 检验结果

k	τ	日期	ρ_{SZ-SH}	Z
1	15	2004-02-03	0.98932	-0.75932
2	44	2004-03-15	0.9937	-1.9132
3	75	2004-04-27	0.99745	8.4707*
4	133	2004-07-23	0.90298	-8.2541*
5	173	2004-09-17	0.9938	11.374*
6	185	2004-10-12	0.38797	-4.4812*
7	233	2004-12-17	0.96519	-2.0444*
8	271	2005-02-21	0.98352	4.3333*
9	287	2005-03-15	0.89451	-1.9552
10	321	2005-05-09	0.96783	-4.5917*
11	385	2005-08-08	0.99512	2.4858*
12	430	2005-10-17	0.98918	5.1749*
13	551	2006-04-14	0.93684	

注: 表中的 τ 为“已实现”波动的变结构点; $\tau_0=0, \tau_K=T$, 其中 K 代表相应的变结构点总数, T 代表样本容量, $T=551, \rho$ 为 τ_{k-1} 至 τ_k 之间深圳与上海股市之间的相关参数; Z 是根据式(14)计算的检验值, 带*表示在给定置信水平 0.05 下 ($Z_{0.025}^*=1.96$), 拒绝原假设。

(三) 股票市场之间波动溢出分析

根据表 1、表 2 所列估计结果, 分析如下:

(1) 从波动变结构点诊断结果分析, 样本内, 由于我国经济处于调整阶段, 因此上海、深圳股市的波动变化比较频繁, 而变结构点一部分处在节假日附近, 或重要日期 3 月 15 日“消费者权益日”, 或有重大的政策出台引起的。

(2) 根据相关系数及 Z 检验结果可知, 深圳与上海股票市场之间相关系数变化比较大, 无论是深圳对上海股票市场, 还是上海对深圳股票市场都存在着长期的波动溢出, 其原因比较明确, 在深圳与上海股票市场上市的公司均为我国的企业, 受相同政策变化的影响, 而且不同企业之间的经济关联比较密切。

(3) 通过诊断深圳与上海股票市场高频数据收益率“已实现”波动变结构点、分阶段计算深圳与上海股票市场“已实现”相关系数、检验“已实现”相关系数在变结构点前后是否显著变化来分析判断股票市场之间是否存在波动溢出, 其结果符合现实经济的发展规律。由于根据高频数据所计算的“已实现”波动率及协方差, 损失的信息比低频数据的要小的多, 能够包含更多的信息, 所以, 此方法所刻画的特征更符合的金融市场变化的规律。由于诊断出了样本内波动变结构点, 因此不仅能判断分析样本内不同股票市场波动溢出是否存在, 而且还可以分析出产生波动溢出的时间段。

五、结 论

高频数据的分析是近年来兴起的一个领域, 相对低频数据而言, 高频数据所包含的信息更多, 所以, 研究基于高频数据的金融市场波动溢出是非常必要的。另外, 金融市场决策者在进行短期决策时, 不仅需要长期的市场信息, 而且还需要即时的信息, 因此, 使用高频数据研究金融市场之间波动溢出, 能够为金融决策者提供更及时的决策信息。本文在讨论“已实现”波动率及“已实现”协方差矩阵理论基础上, 将波动变结构引入高频时间序列, 通过诊断“已实现”波动率的变结构点、计算不同金融市场之间“已实现”相关系数, 检验“已实现”相关系数在变结构点前后是否显著发生变化来分析金融市场之间的波动溢出问题。最后选取上证指数、深圳成份指数每分钟数据为原始数据, 选择了 10 分钟的抽样频率, 验证分析了上证指数、深圳成份指数之间波动溢出问题, 实证结果不仅验证了方法的可行性, 由于高

频数据损失信息小,而且“已实现”波动率考虑了金融市场数据日内的波动情况,所以,所计算的“已实现”相关系数更能够反映金融市场之间的相关关系,以此为基础进行金融市场之间波动溢出的分析就更符合金融市场发展规律。

参考文献:

- [1] Andersen, T. G., T. Bollerslev, Diebold, F. X., Labys, P. Modeling and forecasting realized volatility[J]. *Econometrica*, 2003, 71(2): 579—625
- [2] Andersen, T. G., T. Bollerslev, Diebold, F. X. Some like it smooth, and some like it rough: untangling continuous and jump components in measuring, modeling, and forecasting asset return volatility[J]. *PIER Working Paper 03—025*, University of Pennsylvania, 2003.
- [3] Andersen, T. G., T. Bollerslev, Diebold, F. X., Labys, P. Exchange rate returns standardized by realized volatility are (Nearly) gaussian[J]. *Multinational Finance Journal*, 2000, 4(3&4): 159—179.
- [4] Andersen T. G., Tim Bollerslev, et. al., Analytic evaluation of volatility forecasts[D]. Working Paper, 2002.
- [5] Barndorff-Nielsen, O. E. and N. Shephard. Power and bipower variation with stochastic volatility and jumps[J]. *Journal of Financial Econometrics*, 2004, 2(1): 1—37.
- [6] Barndorff-Nielsen, O. E. and N. Shephard (b). Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics[J]. *Econometrica*, 2004, 72(3): 885—925.
- [7] Barndorff-Nielsen, O. E. and N. Shephard. Econometrics of testing for jumps in financial economics using bipower variation[J]. *Journal of Financial Econometrics*, 2006 forthcoming.
- [8] 徐正国. 金融市场高频/超高频时间序列的分析、建模与应用[D]. 天津: 天津大学, 2004.
- [9] 黄违洪, 张世英. 模型结构变化点检测法[J]. *应用数学学报*, 1987, 10(3): 267—275.
- [10] Kim D. A Bayesian significance test on the stationarity of regression parameters[J]. *Biometrika*, 1991, 78: 667—675.
- [11] Kim D, Kon S J. Structural change and time dependence in models of stock returns[J]. *Journal of Empirical Finance*, 1999, (9): 283—308.
- [12] Barndorff-Nielsen, O. E. and N. Shephard (b). Econometric analysis of realized covariation: high frequency based covariance, regression, and correlation in financial economics[J]. *Econometrica*, 2004, 72(3): 885—925.
- [13] Pedhazur E J. Multiple regression in behavioral research[M]. Orlando: Harcourt Brace College Publishers, 1997.

(责任编辑:王铁军)

Volatility Spillover Analysis on the Financial Market Based on the High Frequency Data

ZHANG Rui-feng, WANG Tong-san

(Institute of Quantitative & Technical Economics, Chinese Academy of Social Sciences 100732, China)

Abstract:iming at the high frequency data in the financial markets, based on the study of Realized Volatility and Realized covariance, the Structural change of Realized volatility has been introduced, and the correlation coefficient of "realized" volatility for the grading data has been calculated, then volatility spillover in the financial markets has been studied through testing whether the correlation coefficient of Realized volatility has significant changes around the structure point, and conducts the empirical analysis.

Key words: High frequency data; Financial markets; Volatility spillover