

文章编号:1672-058X(2013)04-0030-05

基于分类信息 GARCH 模型的高频数据波动率研究*

潘海峰

(安徽工程大学 数理学院,安徽 芜湖,241000)

摘要:提出了基于分类信息的 C-GARCH 模型和 S-GARCH 模型,并结合传统未考虑分类信息下的 GARCH 模型,以上证综指五分钟数据为样本,对波动率进行了实证分析;研究表明:分类信息 GARCH 模型优于未考虑分类信息的模型,最优模型为 C-GARCH 模型,其次为 S-GARCH 模型;好消息和坏消息对高频数据方差的影响程度相对较小,但却提高了描述精度;好消息与方差波动负相关,坏消息与方差波动正相关;坏消息对波动率的影响比好消息大,具有非对称性。

关键词:分类信息;高频数据;C-GARCH;S-GARCH

中图分类号:F224

文献标志码:A

金融市场中,无论是金融风险分析,还是金融产品定价及投资组合研究,波动率都扮演着重要的角色。Engle(1982)^[1]首先提出了 ARCH 模型对方差进行建模。Bollerslev(1986)^[2]在 ARCH 模型基础上引入自回归结构创建了 GARCH 模型,随后的几十年里,该模型已发展成为包含众多方法的模型类别,大量实证研究表明,GARCH 类模型特别适合于捕捉时间序列波动的持续性和丛集性现象。

关于价格波动的 GARCH 效应和持续性,Osborn(1959)^[3]研究表明交易量与价格波动存在显著地相关关系。Peker K. Clark(1973)^[4]首次提出了股价波动的混合分布假说(MDH),该假说认为,资产回报和交易量是由一个潜在的不可观测的信息流过程共同趋动的,信息流的冲击将同时产生交易量和价格波动。Tauchen 和 Pitts(1983)^[5]建立了量价关系的二元混合模型,说明了交易量序列可以作为产生价格持续性波动的因素,成为信息过程的代理指标。Lamoureux 和 Lastrapes(1990)^[6]把交易量作为信息流的替代指标,加入 GARCH 模型中,结果表明交易量的系数非常显著,对价格波动确实有很强的解释能力。Terben G. Andersen(1996)^[7]建立了修正的混合分布模型,该模型在噪声理性预期框架下,交易量分解为流动性需求交易和信息趋动性交易两部分。Graig A. Depken(1999)^[8]假设正的价格变化即为正的信息流的和,负的价格变化即为负的信息流的和,1 d 内价格的变化所带来的成交量就可以用来表示正的信息流和负的信息流,并将好消息和坏消息带来的成交量加入到 GARCH 模型的方差方程中。但该模型没有考虑高频数据对整个模型解释合理性的贡献,及非信息交易量和由于私有信息差异引起的信息交易量对方程回归效果的影响。

寻明辉,桂峰(2005)^[9]建立了一个股票价格冲击混合分布分类信息 GARCH 模型,解释了股价变动的自相关性和异方差特性。凌士勤等(2005)^[10]将非预期成交量作为信息交易量的替代指标加入 GARCH 模型,但在条件方差方程中没有引入非对称性波动结构。

收稿日期:2012-07-09;修回日期:2012-09-07.

* 基金项目:安徽工程大学青年基金重点项目(2008YQ026zd);教育部人文社会科学研究规划基金项目(12YJA790041);安徽省自然科学基金项目(1208085MG116).

作者简介:潘海峰(1982-),男,安徽宿州人,讲师,硕士研究生,从事金融数学与金融工程研究.

在此构建了基于分类信息的 C-GARCH 模型,并考虑时间和未来价格因素对成交量的影响,提出了 S-GARCH 模型。由于投资者的瞬时决策常会考虑买卖的时间和价格高低,且会有一定的犹豫,以上证综指 5 min 的高频数据为研究对象,应用 C-GARCH、S-GARCH 及传统的 GARCH、EGARCH 和 PARCH 模型,对波动率进行实证分析。

1 研究方法与研究模型

1.1 分类信息 C-GARCH 模型

(1) 假设。假设五分钟作为每日交易的基本单位;假设存在非信息交易和信息交易(好消息和坏消息)两种模式,若价格上涨,说明买方力量大于卖方力量,成交量视为非信息成交量和好消息带来的成交量;相反的,视为非信息成交量和坏消息带来的成交量。对日交易量进行回归,用估计值作为非信息交易量,非预期的交易量为信息交易量。

(2) 信息交易量的获取。处理步骤如下: V_t^+ 表示为当天 5 min 内股价上升时的非信息交易量加好消息带来的成交量,则:

$$V_t^+ = \sum_{i=1}^{24} I^+ V_{ti} \quad (1)$$

式(1)中: V_{ti} 表示第 t 天第 i 个 5 min 成交量, I^+ 为示性函数,当价格上升时为 1,否则为 0。类似地,可定义:

$V_t^- = \sum_{i=1}^{48} I^- V_{ti}$ 。对 V_t^+ 和 V_t^- 进行处理,采用 HP 滤波方法进行趋势过滤,平滑序列 V_t^{p+} 和 V_t^{p-} ,平稳的交易量序列为 $V_t^{h+} = V_t^+ - V_t^{s+}$ 和 $V_t^{h-} = V_t^- - V_t^{s-}$ 。对 V_t^{h+} 和 V_t^{h-} 进行回归,去除序列相关性;模型估计的残差分别作为去除了非信息交易量的好消息带来的成交量 V_t^{h+} 和坏消息带来的成交量 V_t^{h-} 。

(3) 最终模型。区分了好消息和坏消息的高频数据 GARCH(p, q) 模型如下:

$$\sigma^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 + \psi_1(V_t^{h+}) + \psi_2(V_t^{h-}) \quad (2)$$

式(2)中: $\alpha_0 > 0, \alpha_i \geq 0, \beta_j \geq 0, \sum_{i=1}^p \alpha_i + \sum_{j=1}^q \beta_j < 1$ 。

1.2 S-GARCH 模型

针对高频数据,进一步考虑股票价格随机性的影响。实际操作中,是否购买,时间和价格是投资者的重要参考因素。Bhar 和 Malliaris(1998)^[11]、Malliaris 和 Urrutia(1998)^[12] 提出成交量是未来均衡价格和时间的函数。

$$V = V(t, F) \quad (3)$$

式(3)中 V 表示成交量, F 表示未来价格, t 表示时间。假设价格 F 服从 Ito 过程,

$$dF = \mu dt + \sigma dZ \quad (4)$$

式(4)中 Z 表示标准的维纳过程。通过漂移项, Ito 过程更好的描述了随机游走的持续性。

由 Ito 公式有:

$$dV = \left[V_t + V_p \mu + \frac{1}{2} V_{pp} \sigma^2 \right] dt + V_p \sigma dZ \quad (5)$$

式(5)中 V_t, V_p, V_{pp} 表示偏导数。

式(3)-(5)描述了理论的成交量和随机游走性。如果未来价格遵循随机游走,成交量也表现为随机游走。进一步地,由式(5),有

$$E(dV) = V_t + V_p \mu + \frac{1}{2} V_{pp} \sigma^2 \quad (6)$$

式(6)表明,成交量的变化依赖于 V_t , 漂移率 μ , 波动率 σ^2 。实证研究中,也可以采用下述模型检验:

$$E(dV) = \alpha t + \beta\mu + \lambda\sigma^2 \quad (7)$$

该式价格变动与成交量之间的正相关关系。进一步,可得:

$$\text{var}(dV) = V_t^2 \sigma^2 \quad (8)$$

式(8)中成交量波动率是未来价格波动率的函数。可采用下式检验:

$$\text{var}(dV) = a + \delta\sigma^2 \quad (9)$$

实证研究中,式(8)、(9)可采取下式表示:

$$|\Delta V_t| = a + \delta |\Delta F_t| \quad (10)$$

通过式(10)分析对价格波动率是否对成交量波动产生显著影响。结合 GARCH 模型,建立 S-GARCH 模型:

$$\sigma^2 = \alpha_0 + \sum_{i=1}^p \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^q \beta_j \sigma_{t-j}^2 + \psi(V_t) \quad (11)$$

2 实证分析

2.1 样本数据及数据处理

选取上证指数 5 min 的收盘价及其成交量,跨度为 2011 年 5 月至 2011 年 10 月 23 日,样本总数 5 616 个,数据来源于“飞狐交易软件”。收益率采用对数收益率:

$$R_t = \ln(F_t) - \ln(F_{t-1}) \quad (12)$$

从上证指数对数收益率序列的时序图中(图 1),可以观察到对数收益率波动的丛集性现象,即波动在有的时间段内较小,在有的时间段内较大。

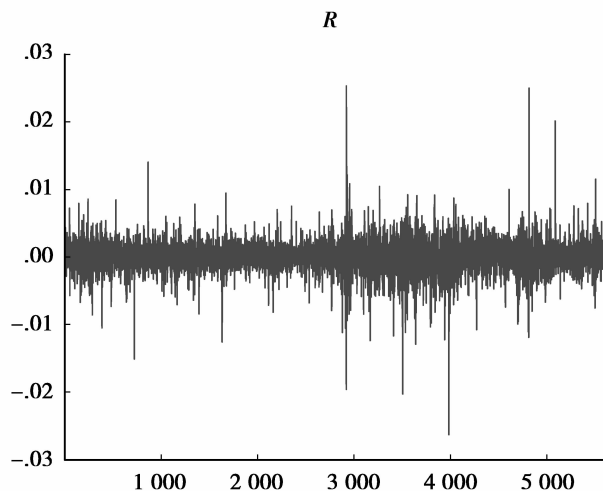


图 1 对数收益率时序图

通过对收益率描述性统计分析,可得偏度为 -0.249355 , 小于 0, 说明序列分布有长的左拖尾, 峰度为 12.75484 , 高于正态分布的峰度值 3, 说明收益率序列具有尖峰和厚尾性。Jarque-Bera 统计量为 22320.94 , P 值为 0, 拒绝该对数收益率序列服从正态分布的假设。应用软件进行 ADF 检验, 结果表明 t 统计量值为 -74.60945 , 对应 P 值为 0.0001 , 表明时间序列平稳。

对数收益率的自相关图表明, 序列的自相关和偏自相关系数均落在两倍的估计置信度内, 且 Q -统计量对应的 p 值均大于置信度 0.05 , 故序列在 5% 的显著性水平下不存在显著的相关性。

进一步进行 ARCH 效应检验,对 R 去均值化,并计算残差的平方相关图,结果表明序列存在自相关,所以有 ARCH 效应。对收益率进行条件异方差的 ARCH-LM 检验,滞后结束位 3 时,检验结果: F 值为 102.588 0,对应 P 值为 0,拒绝原假设,说明残差序列存在 ARCH 效应。

成交量采用对数成交量,即对成交量取自然对数。对上证指数的对数成交量进行处理,估计出去除了非信息交易量的好消息带来的成交量和坏消息带来的成交量。

2.2 实证结果

表 1 GARCH 类模型实证结果

	α_1	β_1	ψ_1	ψ_2	γ	δ	AIC	SC
GARCH(1,1)	0.067 29	0.904 21					-9.269 77	-9.266 22
EGARCH(1,1)	0.069 07	0.986 46			-0.060 43		-9.299 59	-9.294 87
PARCH(1,1)	0.028 59	0.960 70			0.995 57	1.214 82	-9.299 19	-9.293 28
C-GARCH(1,1)	6.66E-18	0.971 50	-4.15E-20	4.32E-20			-82.553 7	-82.547 8
S-GARCH(1,1)	2.66E-17	0.971 50	-3.75E-20				-79.499 2	-79.493 3

(1) 高频数据收益率的波动具有较强的丛集性,时间序列表现出尖峰厚尾特征;

(2) 从上述结果(表 1)可以看出,结合 AIC 准则,在未考虑分类信息情况下,EGARCH 模型结果最优, α_1 的估计值为 0.069 07,非对称项 γ 的估计值为 -0.060 43,则当 $u_{t-1} > 0$,该信息冲击对条件方差的对数有一个 0.008 64 倍的冲击,当 $u_{t-1} < 0$,它给条件方差的对数带来的是 0.129 5 倍的冲击。

(3) 波动率的计算对于 GARCH 类模型的种类不太敏感,但对于是否考虑分类信息比较敏感。考虑了分类信息的 GARCH 模型优于未考虑分类信息的模型。在考虑分类信息的模型中,最优的为 C-GARCH 模型,其次为 S-GARCH 模型。从 C-GARCH 模型看出,好消息和坏消息对方差的影响相对较小,但却提高了波动率的描述精度。从具体数值看, $\psi_1 = -4.15E-20$,为负数,说明好消息对方差的波动成负相关; $\psi_2 = 4.32E-20$,为正数,说明坏消息与方波动正相关,与实际相符合。另外, $|\psi_1| < |\psi_2|$,说明坏消息对波动率的影响比好消息大,体现了高频数据波动的非对称性。

3 结束语

综合研究表明,考虑分类信息的 GARCH 类模型提高了高频数据波动性刻画的精度,动态描述金融时间序列的多类特征。在对上证指数的实证分析中,可以发现 C-GARCH 模型较其他模型更符合实际情况,能够准确的衡量分类信息对波动的影响。

参考文献:

- [1] ENGLE R. Autoregressive conditional heteroskedasticity with estimates of the variance of U. K. inflation [J]. *Econometrica*, 1982,50(4):987-1008
- [2] BOLLERSLEV T. Generalized autoregressive conditional heteroskedasticity [J]. *Journal of Econometrics*, 1986,31(3):307-327
- [3] OSBORNE M. Brownian motion in the stock market [J]. *Operation Research*, 1959(7):145-173
- [4] CLARK P. A subordinated stochastic process model with finite variance for speculative prices [J]. *Econometrica*, 1973,41(1):135-156
- [5] TAUCHEN G, PITTS M. The Price variability-volume relationship on speculative markets [J]. *Econometrica*, 1983,51(2):485-506

- [6] LAMOURENX C, WILLIAM D. Heteroskedasticity in stock return data: volume versus garch effects[J]. Journal of Finance, 1990(45):221-229
- [7] TORBEN G. Return volatility and trading volume: an information flow interpretation of stochastic volatility [J]. Journal of Finance, 1996, 7(1):169-204
- [8] CRAIG A. Good news, bad news and GARCH effects in stock return data [J]. Journal of Applied Economics, 1999, 4(2): 313-327
- [9] 寻明辉, 桂峰. 股票价格冲击混合分布分类信息 GARCH 模型及其应用[J]. 上海交通大学学报, 2007, 41(9):1105-1109
- [10] 凌士勤, 杨波, 袁开洪. 分类信息对股市波动的影响研究[J]. 中国管理科学, 2005, 13(3):20-25
- [11] BHAR R, MALLIARIS A. Volume and volatility in foreign currency futures markets [J]. Review of Quantitative Finance and Accounting, 1998(10):285-302
- [12] MALLIARIS A, URRUTIA J. Volume and price relationships: hypotheses and testing for agricultural futures [J]. Journal of Futures Markets, 1998(18):53-72

Research on High Frequency Data Volatility Rate Based on Classification Information GARCH Model

PAN Hai-feng

(School of Mathematics and Physics, Anhui Polytechnic University, Anhui Wuhu 241000, China)

Abstract: C-GARCH Model and S-GARCH Model based on classification information are proposed by combining traditional GARCH Model without considering classification information. Taking Shanghai Composite Index in five minutes as an example, empirical analysis is conducted on its volatility rate. Research results show that classification information GARCH Model is better than the Model without considering classification information, that C-GARCH Model is the optimal model and S-GARCH Model is the second, that the influence degree of good news and bad news on high frequency data variance is small but improves its descriptive accuracy, that good news is negatively related to variance volatility but bad news is positively related to variance volatility and that the influence of bad news on volatility rate is bigger than that of good news and is asymmetric.

Key words: classification information; high frequency data; C-GARCH; S-GARCH

责任编辑:代小红