中国股市价格—交易量的线性及非线性因果关系研究①

王承炜, 吴冲锋
(上海交通大学管理学院, 上海 200030)

摘要: 实证研究了沪市和深市价格和交易量之间的线性和非线性因果关系, 研究结果表明, 两个市场之间存在着收益对交易量的线性 Granger 因果关系和双向的非线性 Granger 因果关系. 在经过周末效应和 GARCH 模型调整之后, 沪深两市量价之间的非线性因果关系消失了.

关键词: 非线性; Granger 因果关系; GARCH 模型
中图分类号: F831 文献标识码: A 文章编号: 1007-9807(2002)04-0007-06

0 引 言

股票市场中价格与交易量的关系已引起了相当多的理论研究者的注意. 虽然在金融市场微观结构理论的框架中, 不同的模型会导致量价关系的不同结论[1,2]. 但正如文[3]的两阶段模型所指出, 交易者可以通过同时考察交易量和价格来推测该股票的信息. 实证方面, Karpoff 总结性地指出, 价格绝对变化量与交易量正相关, 且这种关系是不对称的, 他把这归因于卖空的限制[4]. Harris 用混合分布理论(MDH)解释了量价之间的正相关关系, 并给出了实证的检验[5,6]. 与此同时, 国内的研究者也发现了量价之间的这种非对称的关系[7]. 另一方面, 股票价格与交易量之间的非线性关系也被 Gallant、Rossi 和 Tauchen[8], 以及 Campell、Grossman 和 Wang[9] 的研究所记载. 因此, 有必要研究价格和交易量之间的因果关系到底是线性还是非线性的, 如果是非线性的, 那么这种非线性关系的来源是什么.

Jain 和 Joh、Smirlock 和 Starks 用传统的 Granger 方法证实了量价之间的线性因果关系[10,11]. 文[12]则用非参数方法, 证实了 NYSE 的市场收益和交易量之间的非线性双向因果关系. 也许正如 Granger(1989) 指出, 这个世界几乎肯定是由非线性关系构成的. 在我们的实证研究中发现: 中国股市——沪市和深市的交易量与价格之间也是双向的非线性因果关系. 但在经过周末效应和 GARCH 模型的调整后, 非线性的因果关系消失了, 这就说明了量价之间的非线性因果关系体现在它们方差的记忆过程中, 这一结论与 Hiemstra 和 Jones 对 NYSE 作的研究结论不同. 对此 MDH 也许是一个解释, Clark 指出, 量价是由共同因素——信息到达驱动的. 信息之间的相关性造成了 GARCH 类现象, 因此, 在滤掉这一因素后, 量价之间的相关性将消失. Hiemstra 和 Jones 的研究中, 滤过后序列因果关系的显著性明显减小了, 但在 5% 的单边水平上仍然存在. 国内也有研究者检验了沪市一定区间周数据原始时间序列的因果关系[13], 但本文与他们研究的区别是明显的.

本文的结论对市场监管者和实际投资者都有着重要的意义. 市场监管者可以从价格和交易量的关系中看出一个市场波动的性质, 信息在市场中传递的方向及效率. 实践方面, 华尔街有句老话: "It takes volume to make prices move." 国内任何一本关于股市投资的技术分析书籍中, 都可

① 收稿日期: 2001-05-18; 修订日期: 2002-04-16.
基金项目: 国家杰出青年科学基金资助项目(70025303); 教育部跨世纪优秀人才资助项目.
作者简介: 王承炜(1974-), 男, 四川自贡人, 博士生.

以看到量价的结合在做出投资决策方面的重要性,因此本文的结论对实际投资者的重要性是不言而喻的

1 数据和方法分析

本文的数据来源是香港理工大学和深圳国泰君安信息公司提供的市场交易数据库(CSMAR). 采用的样本是 1993 年 1 月 1 日至 2000 年 12 月 30 日的数据 研究对象是整个市场,因此价格采用实践中应用广泛的上证综合指数和深证综合指数 交易量也采用整个市场总的交易量(A 股交易量加B 股交易量). 沪深两市样本数分别为 1974 和 1893 由于 Granger 因果关系对时间序列的平稳性要求,对两个序列采用了对数差分. 以后除非特别声明,收益和交易量序列 R_t 和 V_t 指: $R_t = 100^*(\ln P_t - \ln P_{t-1})$, $V_t = 100^*(\ln V_t - \ln V_{t-1})$. 对 R_t 和 V_t 的ADF 检验表明,两个时间序列都是平稳的

Granger 线性因果关系指出:

设 $\{x_t\}$ 和 $\{y_t\}$ 是两个稳定的时间序列, $F(x_t | \Phi_{t-1})$ 是在给定信息集 Φ_{t-1} 下 x_t 的条件概率分布,其中, $\Phi_{t-1} = \{x_{t-lx}^{lx}, y_{t-ly}^{ly}\}$, $x_{t-lx}^{lx} = \{x_{t-lx}, x_{t-lx-1}, \cdots, x_{t-1}\}$, $y_{t-ly}^{ly} = \{y_{t-ly}, y_{t-ly-1}, \cdots, y_{t-1}\}$, lx 和 ly 是 x 和 y 的滞后阶数 如果有 $F(x_t | \Phi_{t-1}) = F(x_t | \Phi_{t-1} - y_{t-ly}^{ly})$, 则称 y 不是 x 的(严格的)Granger 原因①. 具体的统计方法可采用如下的VAR 进行:

$$X_t = A(L)X_t + B(L)Y_t + U_{X,t} \tag{1}$$

$$Y_t = C(L)X_t + D(L)Y_t + U_{Y,t} \tag{2}$$

上式中, $A(L)$、$B(L)$、$C(L)$ 和 $D(L)$ 是差分算子,并假设 $U_{X,t}$ 和 $U_{Y,t}$ 独立同分布 考虑 y 对 x 的Granger 因果关系检验用 H_0: $B(L)$ 的所有滞后同时为零 x 对 y 的 Granger 因果关系检验用 H_0: $C(L)$ 的所有滞后同时为零 本文中最优滞后阶数采用Akaike 信息准则(AIC) 或 Schwarz 贝叶斯信息准则(SBC).

非线性因果关系的检验由Baek 和Brock 提出: 设 x_t 和 y_t 是两个稳定的时间序列 $x_{t-lx}^{lx} = (x_{t-lx}, x_{t-lx-1}, \cdots, x_{t-1})$, $y_{t-ly}^{ly} = \{y_{t-ly}, y_{t-ly-1}, \cdots, y_{t-1}\}$ 是 x 和 y 的滞后向量 $x_t^m = (x_t, x_{t+1}, \cdots, x_{t+m+1})$, $y_t^m = (y_t, y_{t+1}, \cdots, y_{t+m+1})$ 是 x 和 y 的超前向量 对给定的 $m, lx, ly \geq 1$ 和 $e > 0$, 若

$$\Pr(\| x_t^m - x_s^m \| < e \,|\, \| x_{t-lx}^{lx} - x_{s-lx}^{lx} \| < e, \| y_{t-ly}^{ly} - y_{s-ly}^{ly} \| < e) = \Pr(\| x_t^m - x_s^m \| < e \,|\, \| x_{t-lx}^{lx} - x_{s-lx}^{lx} \| < e) \tag{3}$$

其中, $\|\cdot\|$ 是向量元素的最大模 则说 y 不是 x 的(严格的)Granger 非线性原因

对沪市和深市的非线性因果关系的检验是建立在方程(1)、(2) 的标准化残差基础上的 如Hiemstra 指出,方程(1)、(2) 的残差很难满足i i d 的要求,故要采用修正的Baek 和Brock 方法[14]. 方程(3) 的修正的非参数估计如下,令

$$C1(m + lx, ly, e, n) = \frac{2}{n(n-1)} \cdot \sum\sum I(x_{t-lx}^{m+lx}, x_{s-lx}^{m+lx}, e) \cdot I(y_{t-ly}^{ly}, y_{s-ly}^{ly}, e) \tag{4}$$

$$C2(lx, ly, e, n) \equiv \frac{2}{n(n-1)} \cdot \sum\sum I(x_{t-lx}^{lx}, x_{s-lx}^{lx}, e) \cdot I(y_{t-ly}^{ly}, y_{s-ly}^{ly}, e) \tag{5}$$

$$C3(m + lx, e, n) \equiv \frac{2}{n(n-1)} \cdot \sum\sum I(x_{t-lx}^{m+lx}, x_{s-lx}^{m+lx}, e) \tag{6}$$

$$C4(lx, e, n) \equiv \frac{2}{n(n-1)} \sum\sum I(x_{t-lx}^{lx}, x_{s-lx}^{lx}, e) \tag{7}$$

则式(3) 可以表达为

$$\frac{C1(m + lx, ly, e)}{C2(lx, ly, e)} = \frac{C3(m + lx, e)}{C4(lx, e)} \tag{8}$$

检验采用下式:

$$\sqrt{n}\left(\frac{C1(m + lx, ly, e, n)}{C2(lx, ly, e, n)} - \frac{C3(m + lx, e, n)}{C4(lx, e, n)}\right)^a \sim N(0, \sigma^2(m, lx, ly, e)) \tag{9}$$

式中, $t < s$, $t, s = \max(lx, ly) + 1, \cdots, T - m + 1$, $n = T + 1 - m - \max(lx, ly)$. $I(Z_1, Z_2, e)$ 及 $\sigma^2(m, lx, ly, e)$ 的含义及估计方法见文[12]. 在Hiemstra 和Jones 的蒙特卡洛模拟基础上,也同文[12],取 $lx = ly$, $m = 1$, 滞后值从1至8, $\sigma = 1$, $e = 1.5\sigma$ 值得一提的是,所有这些结果是在大

① 同文[12],本文不检验即时的 Granger 线性因果关系

样本的情况下成立的, 这对金融数据的样本长度有一定的要求 本文样本长度达 8 年, 故有一定的代表性 但是将整体样本拆成子集以检验结果的鲁棒性是不合适的

Clark 和 Harris 提出了 MDH, 每日的股价变化和交易量是信息到达数量的条件正态分布, Anderson[15] 进一步指出, 在放宽 Clark 的每日到达信息数量是 i i d 的条件下, MDH 理论暗示了股价序列的 ARCH 族现象 因此, 采用交易量和价格的原时间序列, 有可能识别到的是由共同因素: 信息到达导致的因果关系 对此, 本文将两个时间序列用 GARCH 模型调整, 并且, 与文[12] 不同, 调整 GARCH 模型的同时, 考虑了这些金融数据的周末效应 这样原始的时间序列的结果可以得出一个量价关系的粗略认识, 具体的调整方程如下:

$$r_t = \varphi_0 + \varphi_1 D_{1t} + \varphi_2 D_{2t} + \varphi_4 D_{4t} + \varphi_5 D_{5t} + \epsilon_t \tag{10}$$

$$\epsilon_t = \sqrt{h_t e_t} \tag{11}$$

$$h_t = k + \alpha\epsilon_{t-1}^2 + \beta h_{t-1} + \lambda_1 D_{1t} + \lambda_2 D_{2t} + \lambda_4 D_{4t} + \lambda_5 D_{5t} \tag{12}$$

$$\hat{r}_t = \hat{\epsilon}_t / \hat{h}_t^{0.5} \tag{13}$$

其中, $\hat{\epsilon}_t$、$\hat{h}_t$ 分别是回归残差和条件方差拟合值 $\hat{r}_t$ 是调整后的时间序列, D_1、D_2、D_4、D_5 分别代表周一、二、四、五的虚变量

2　计算结果

首先对未作调整的时间序列进行 Granger 线性和非线性因果关系检验, 检验结果如表 1 和表 2 所示:

表 1　未作调整的量价 Granger 线性因果关系检验

	H_0: V_t 不是 R_t 的 Granger 原因				H_0: R_t 不是 V_t 的 Granger 原因			
	lr	lv	F	p 值	lv	lr	F	p 值
沪市	12	20[a]	1.30	0.17	9	20[a]	12.68	0.00
深市	11	20[a]	0.97	0.49	9	20[a]	9.31	0.00

a: 此处采用 AIC 时会产生很大的滞后阶数, 因此直接选用 20, 延长滞后阶数对本文的定性结论不会产生影响

表 2　未作调整的量价 Granger 非线性因果关系检验*

	H_0: V_t 不是 R_t 的 Granger 原因		H_0: R_t 不是 V_t 的 Granger 原因	
$lx = ly$	CS	TVAL	CS	TVAL
沪市价格交易量关系				
1	0.003 44	1.82	0.004 84	2.31
2	0.005 247	2.05	0.009 23	2.81
3	0.008 56	2.05	0.009 92	2.78
4	0.008 15	2.34	0.009 32	2.49
5	0.007 95	2.03	0.008 48	2.25
6	0.008 698	2.11	0.009 27	2.58
7	0.009 53	2.18	0.009 82	2.70
8	0.008 64	1.80	0.010	2.68
深市价格交易量关系				
1	0.003 37	1.89	0.005 83	2.72
2	0.004 16	1.67	0.008 19	2.80
3	0.006 51	2.09	0.010 9	3.14
4	0.007 00	1.92	0.009 61	2.51
5	0.007 46	1.80	0.007 68	2.05
6	0.008 25	1.84	0.006 89	2.26
7	0.009 97	2.10	0.007 19	1.76
8	0.020 9	2.19	0.007 16	1.72

* CS 和 TVAL 分别表示式(8) 中方程两边的差和式(9) 中标准化后的检验值

从表 1 可以看出, 可以在 1% 的水平上拒绝收益不是交易量的线性严格 Granger 原因的假设. 但在两个市场上, 无法显著地检验到交易量是收益的线性严格 Granger 原因. 对交易量和收益经线性方程滤过后的残差进行的非参数检验表明, 交易量和收益之间存在双向的非线性因果关系, 所有的滞后中, 至少在 0.1 的水平上拒绝了原假设. 以上这些结论与美国 NYSE 市场的结论相同.

对交易量和股价的时间序列调整的统计结果如表 3.

表 3　量价周末效应和 GARCH 现象

名　　称	φ_1	φ_2	φ_4	φ_5
沪市收益率	-0.207 6 (-1.581)	-0.497 2* (-3.741)	-0.24*** (-1.753)	-0.005 3 (-0.042)
沪市交易量	-1.781 9 (-0.645)	-5.958 1** (-1.971)	3.643 35 (1.281)	4.666*** (1.634)
深市收益率	-0.045 5 (-0.298)	-0.280** (-2.084)	0.010 1 (0.061)	0.119 1 (0.832)
深市交易量	-0.626 (-0.219)	-7.953* (-2.671)	5.283** (1.81)	5.132*** (1.725)

表 3(续)　量价周末效应和 GARCH 现象

名　　称	k	α	β	λ_1	λ_2	λ_4	λ_5
沪市收益率	1.05E-8* (11.475)	0.125 8* (19.960)	0.889 6* (216.6)	0.022 11 (0.256)	0.00 (0.00)	0.241 04* (2.919)	0.00 (0.00)
沪市交易量	1 223.8* (30.36)	0.119* (4.729)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)
深市收益率	1.05E-8* (459.5)	0.239 8* (13.13)	0.655* (28.26)	1.646* (6.1)	0.213 7 (1.191)	2.188 0* (18.344)	0.00 (0.00)
深市交易量	1 215.28* (1 215.2)	0.079 6* (3.401)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)	0.00 (0.00)

表 3 的结果可以给关于深市和沪市价格和交易量序列的周末效应和 GARCH 现象的一个初步认识. 两个市场上的价格和交易量都表现出了“周末现象”, 但是各个时间序列的表现方式不同, 沪深两式收益在周二都是显著的, 交易量在周二和周五是显著的. 沪深两市收益表现出了强烈的方差持续现象: β 值在 1% 的水平上显著. 且在周四两个市场的方差都会明显变大. 两个市场交易量的方差持续现象不长: β 值显著为零. 且四个虚变量对方差没有影响, 也许 ARCH 模型对交易量时间序列更加合适. 但为了普遍性, 本文仍采用 GARCH (1, 1) 这个最有代表性的模型来调整时间序列.

调整后的量价之间的 Granger 线性和非线性因果关系如表 4 和表 5 所示:

表 4　调整后的量价 Granger 线性因果关系检验

	H_0: V_t 不是 R_t 的 Granger 原因				H_0: R_t 不是 V_t 的 Granger 原因			
	lr	lv	F	p 值	lv	lr	F	p 值
沪市	6	20[a]	1.26	0.19	10	19	13.9	0.00
深市	6	20[a]	0.73	0.79	9	20[a]	11.08	0.00

a: 此处采用 AIC 时会产生很大的滞后阶数, 因此直接选用 20, 延长滞后阶数对本文的定性结论不会产生影响.

表 5　调整后的量价 Granger 非线性因果关系检验*

H_0: V_t 不是 R_t 的 Granger 原因			H_0: R_t 不是 V_t 的 Granger 原因	
lx = *ly*	CS	TVAL	CS	TVAL
沪市价格交易量关系				
1	- 2.31E-05	- 0.010 8	0.001 9	0.591
2	7.64E-04	0.202	0.002 18	0.611
3	0.002 2	0.429	0.002 25	0.459
4	1.13E-03	0.160	0.001 95	0.315
5	9.85E-04	0.110	- 3.561E-4	- 0.046
6	0.001 35	0.127	0.001 34	0.157
7	0.003 00	0.253	0.002 58	0.279
8	0.001 01	0.071	0.003 10	0.301
深市价格交易量关系				
1	0.006 40	0.862	0.003 78	1.77
2	0.004 03	1.25	0.004 09	1.26
3	0.005 19	1.17	0.006 93	1.73
4	0.006 32	1.14	0.006 19	1.24
5	0.008 45	1.32	0.006 43	1.11
6	0.008 51	1.20	0.007 51	1.20
7	0.010 4	1.36	0.008 79	1.31
8	0.012 4	1.54	0.008 44	1.15

* 深市收益对交易量的非线性因果关系在滞后阶数取 1 和 3 时是 0.1 显著的, 但延长滞后阶数时无法拒绝原假设

调整后的时间序列有着完全相反的结论, 虽然收益对于交易量仍然有线性的 Granger 因果关系, 但是没有发现非线性的双向因果关系. 这说明两个时间序列的非线性因果关系表现在它们的方差持续之中, 一旦用 GARCH 模型将这种效应过滤, 非线性因果关系将消失.

3　结束语

本文检验了沪深两市价格和交易量之间的线性和非线性 Granger 因果关系, 结果表明, 沪深两市存在着双向的量价之间的非线性因果关系, 但将价格和交易量的时间序列经周末效应和 GARCH 现象调整后, 量价之间不存在任何方向的非线性因果关系, 这一结论与美国 NYSE 市场不同, 对此, MDH 也许是一种解释. 其暗含的经济含义是, 每日的信息到达会同时地影响到交易量和价格的变化, 而交易量和价格变化的关系是通过信息到达这一共同因素联系起来的. 对国内市场而言, 信息的这种作用是极其显著且呈主导作用, 这就说明了目前而言我国市场量价的波动对信息流的到达是相当依赖的.

参 考 文 献:

[1]　Kyle A S. Informed speculations with imperfect competition[J]. Review of Economics Studies, 1985, 56: 317-355

[2]　Wang J. A model of competitive stock trading volume[J]. Journal of Political Economy, 1994, 102(1): 127-168

[3]　Blume I E. Easley D, O'Hara. Market statistics and technical analysis: the role of volume[J]. Journal of Finance, 1994, 49(1): 153-182

[4]　Karpoff J M. The relation between price changes and trading volume: a survey[J]. Journal of Financial Quantitative Analysis, 1987 22(1): 109-126

[5]　Harris L. Cross-security tests of the mixture of distribution hypothesis[J]. Journal of Financial and Quantitative

Analysis, 1986, 21(1): 39- 46

[6] Harris L. Transaction data tests of the mixture distribution hypothesis[J]. Journal of Financial and Quantitative Analysis, 1987, 22(2): 127- 141

[7] 陈怡玲, 宋逢明. 中国股市价格变动与交易量关系的实证研究[J]. 管理科学学报, 2000, 3(2): 62- 68

[8] Gallant A R. Rossi P E. Tauchen G. Stock price and volume[J]. The Review of Financial Studies, 1992, 3(2): 142- 199

[9] Campell J, Crossman S, Wang J. Trading volume and serial correlation in stock returns[J]. Quarterly Journal of Econimics, 1993, 108: 905- 939

[10] Jain P, Joh G. The dependence between hourly prices and trading volume[J]. Journal of Financial and Quantitative Analysis, 1988, 23(2): 269- 283

[11] Smirlock M, Starks L. An empirical analysis of the stock price-volume relationship[J]. Journal of Banking and Finance, 1988, 12(1): 31- 42

[12] Hiemstra C, Jones J D. Testing for linear and nonlinear Granger causality in the stock price-volume relation[J]. Journal of Finance, 1994, 54(5): 1639- 1664

[13] 张 维, 闫冀楠. 关于上海股市量价因果关系的实证探索[J]. 系统工程理论与实践, 1998, 6: 111- 114

[14] Baek E, Brock W. A nonparametric test for independence of a multivariate time series[J]. Statistica Sinica, 1992, 2: 137- 156

[15] Andersen T. Return volatility and trading volume: An information flow interpretation of stochastic volatility[J]. Journal of Finance, 1996, 51(1): 169- 204

Linear and nonlinear Granger causality test of stock price-volume relation: Evidences from Chinese markets

WANG Cheng-wei, WU Chong-feng

School of Management, Shanghai Jiaotong University, Shanghai 200030, China

Abstract: This paper empirically test the linear and nonlinear Granger causality relation between price and volume of Shanghai and Shenzhen stock market. It is showed that there exits linear Granger causality from stock return to trading volume and bi-directional nonlinear Granger causality between these time series. But after filtering weekend and GARCH effects, the nonlinear Granger causality relation disappears.

Key words: nonlinear; Granger causality; GARCH