Statistics and Application 统计学与应用, **2020, 9(2), 172-181**
Published Online April 2020 in Hans. http://www.hanspub.org/journal/sa
https://doi.org/10.12677/sa.2020.92019

Research and Application of New GARCH Model

Dekang Zhu, Shushan Li

Shandong University of Science and Technology, Qingdao Shandong
Email: zdk12345679@163.com, 1767931666@qq.com

Received: Mar. 6th, 2020; accepted: Mar. 20th, 2020; published: Mar. 27th, 2020

Abstract

Combined with GARCH model and GARCH model with exogenous variables, a generalized GARCH model with exogenous variables is established. In this paper, 1642 sets of data of Shanghai Composite Index and NASDAQ composite index corresponding to their time are selected to study the impact of NASDAQ Composite Index on Shanghai Composite Index, and the change law between the two composite indexes is similar, that is, the two composite index models have correlation.

Keywords

GARCH Model, Volatility, Exogenous Variables

新型GARCH模型的研究与应用

朱得康，李述山

山东科技大学，山东 青岛
Email: zdk12345679@163.com, 1767931666@qq.com

收稿日期：2020年3月6日；录用日期：2020年3月20日；发布日期：2020年3月27日

摘 要

本文结合GARCH模型与加入外生变量的GARCH模型，建立了广义带有外生变量的GARCH模型，本文选取了上证综指和与其时间相对应的纳斯达克综合指数各1642组数据来进行研究纳斯达克综合指数对上证综指的影响，得到两个综指之间的变化规律是相似的，即两个综指模型有相关关系。

关键词

GARCH模型，波动性，外生变量

文章引用: 朱得康, 李述山. 新型 GARCH 模型的研究与应用[J]. 统计学与应用, 2020, 9(2): 172-181.
DOI: 10.12677/sa.2020.92019

1. 引言

一直以来，对于金融方向的研究中，一个重要的课题便是对于金融时间序列波动性的研究，因为往往是波动带来风险，风险带来收益。而对于股票市场收益率波动性的研究便是对于研究金融时间序列的波动的重要手段，因为股票的波动象征着股票价格的变化，而股票价格的变化便与整个金融市场息息相关。

对于随机时间序列预测，美国学者 George Box 和英国统计学家 Gwilym Jenkins 共同建立了 ARMA 模型，由于利用时间序列建立的 ARMA 模型要求序列是平稳的，但是实际上时间序列并不可能全是表现出平稳性，就像单位根过程便是时间序列数据中的非平稳过程；于是建立了一种特殊的描述单位根过程的模型——随机游动模型，Engle 于 1982 年提出了自回归条件异方差模型即 ARCH 模型，但是对于高阶 ARCH 效应的时候参数往往不能通过显著性检验，于是就有了广义 ARCH 模型，即 GARCH 模型[1]，而当 GARCH 模型不能很好地描述序列的变化的时候，为了提高精度，ARCH 模型还有很多的拓展形式(EGARCH，ARCH-M，TARCH，幂 ARCH，成分 ARCH，非对称成分 ARCH)；为了研究虚拟变量对于回归模型的影响，Hamilton 和 Susmel 建立了 SWARCH 模型；以此来研究加入外生变量对于股票股指的影响[2]。

本文旨在建立一个广义的加入外生变量的 GARCH 模型，以此来进行对于股指变化的研究；

2. 新型广义具有外生变量的 GARCH 模型的建立

2.1. 多元线性回归模型——ARMA 模型

ARMA 模型包含两部分即 AR(p)、MA(q)模型；

(1) AR(p)模型是仅用时间序列$\{Y_t\}$不同滞后项作为解释变量的模型：

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + e_t \tag{1.1}$$

如上式，$\phi_1, \cdots, \phi_p$ 是自回归系数；p 为自回归阶数；式(1.1)便是 p 阶自回归模型；

(2) MA(q)模型仅用误差$\{e_t\}$的不同滞后项作为解释变量的模型：

$$Y_t = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \tag{1.2}$$

如上式，$\theta_1, \theta_2, \cdots, \theta_q$ 是移动平均系数；q 为移动平均阶数；式(1.2)便是 q 阶移动平均模型[3]；

由(1) (2)可知

$$Y_t - \phi_1 Y_{t-1} - \phi_2 Y_{t-2} - \cdots - \phi_p Y_{t-p} = e_t - \theta_1 e_{t-1} - \theta_2 e_{t-2} - \cdots - \theta_q e_{t-q} \tag{1.3}$$

令 $\phi(B) = 1 - \phi_1 B - \phi_2 B^2 - \cdots - \phi_p B^p$，$\theta(B) = 1 - \theta_1 B - \theta_2 B^2 - \cdots - \theta_q B^q$，于是，式(1.3)便可以写成

$$\phi(B) Y_t = \theta(B) e_t \tag{1.4}$$

式(1.4)便是 ARMA (p,q)模型[1]。

2.2. 自回归条件异方差模型——GARCH 模型

对于常用的回归模型 $y_t = x_t'\beta + \varepsilon_t$，如果随机扰动项的平方服从 AR (q)过程，即 $\varepsilon_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \alpha_2 \varepsilon_{t-2}^2 + \cdots + \alpha_q \varepsilon_{t-q}^2 + \eta_t$，其中 η_t 是独立同分布的，则称上述模型为自回归条件异方差模型即 ARCH 模型。

广义 ARCH 模型——GARCH 模型

GARCH 模型被称为广义 ARCH 模型，是在 ARCH 模型基础上提出的，可以应用的范围更广，GARCH 模型通常应用于对回归和自回归模型的随机扰动进行建模，GARCH (p,q)模型的表现形式如下：

$$\begin{cases} Y_t = \mu_t + a_t, \mu_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_m Y_{t-m} \\ a_t = \sigma_t \varepsilon_t, \varepsilon_t \sim iid(0,1) \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 \end{cases}$$

2.3. 新型广义具有外生变量的 GARCH 模型

式(2.1)是 ARMA-GARCH (p,q)模型[4]；第一个式子是 ARMA 均值模型，现在我们要做的是在 ARMA 均值模型上加入一个线性方程 δ_t，这个线性模型用来表示外生变量，δ_t 的公式是：$\delta_t = \delta_1 X_{t-1} + \delta_2 X_{t-2} + \cdots + \delta_k X_{t-k}$，然后 σ_t^2 便变成了 $\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{\gamma=1}^{l} \gamma_l \sigma_{t-l}^2$，所以新型广义含有外生变量的 GARCH 模型的表达式如下：

$$\begin{cases} Y_t = \mu_t + a_t + \delta_t, \mu_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_m Y_{t-m}, \delta_t = \delta_1 X_{t-1} + \delta_2 X_{t-2} + \cdots + \delta_k X_{t-k} \\ a_t = \sigma_t \varepsilon_t, \varepsilon_t \sim iid(0,1) \\ \sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i a_{t-i}^2 + \sum_{j=1}^{q} \beta_j \sigma_{t-j}^2 + \sum_{\gamma=1}^{l} \gamma_l \sigma_{t-l}^2 \end{cases} \quad (3.1)$$

3. 数据来源与实证分析

3.1. 数据来源

本文选取的是从 2013 年开始到 2019 年的上证综指和纳斯达克综合指数的日收盘价，日收益率 SC_t 的计算方法是 $SC_t = \log(p_t) - \log(p_{t-1})$，其中 p_t 为日收盘价，p_{t-1} 为滞后一日的收盘价。

3.2. 数据描述

图 1 和图 2 分别给出纳斯达克综合指数的日收盘价和日收益率的曲线图：

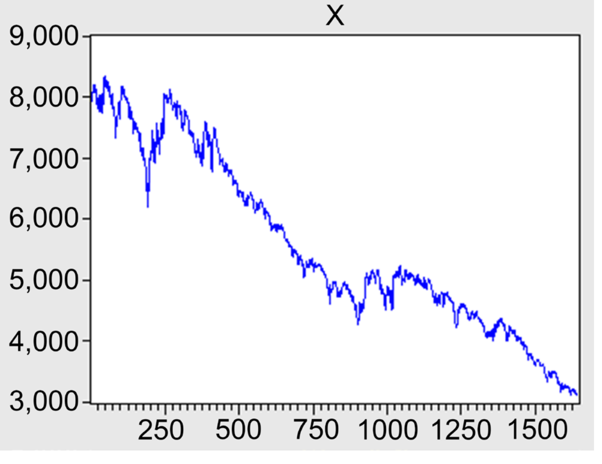

Figure 1. Daily closing price of NASDAQ Composite Index
图 1. 纳斯达克综合指数日收盘价

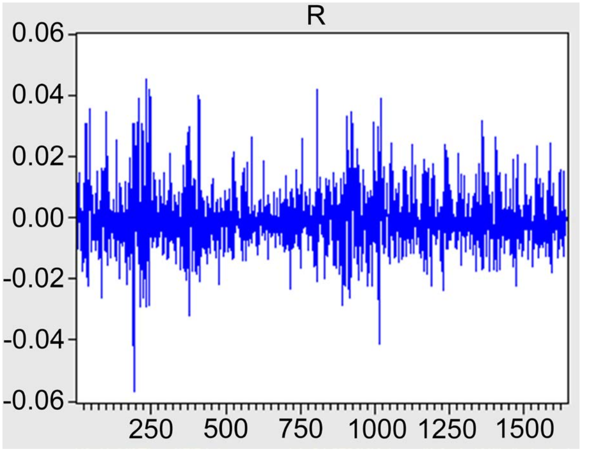

Figure 2. Daily yield of NASDAQ Composite Index
图 2. 纳斯达克综合指数日收益率

图 3 和图 4 分别给出的是上证综指的日收盘价和日收益率的曲线图：

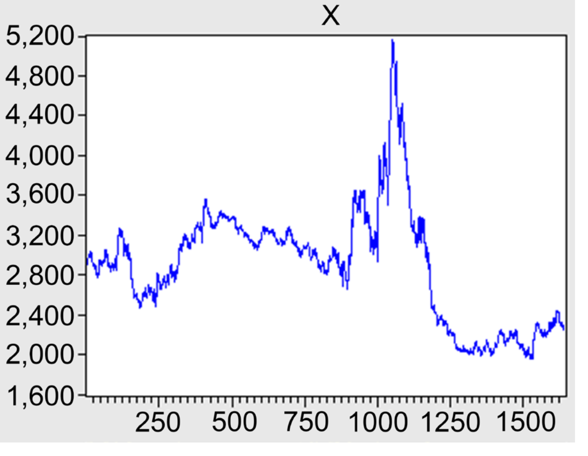

Figure 3. Daily closing price of Shanghai Composite Index
图 3. 上证综合指数日收盘价

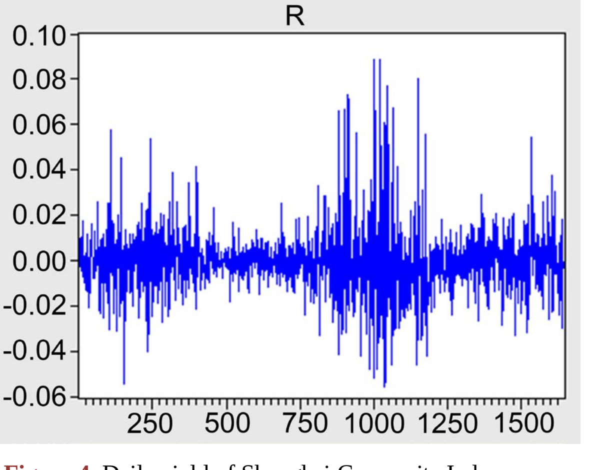

Figure 4. Daily yield of Shanghai Composite Index
图 4. 上证综合指数日收益率

从图 1、图 2 和图 3、图 4 中我们可以发现，纳斯达克综合指数和上证综指的日收盘价是非平稳的，而经过处理后所选用的对数收益率趋于平稳，这样，我们便可以将对以收盘价的研究改为对以日收益率的研究。

4. 数据的统计检验

4.1. 收益率的正态性检验

我们以纳斯达克综合指数作为变量来研究纳斯达克综合指数和上证综指之间的关联，那么我们用 EVIEWS 软件测出上证综指的日收益率序列的正态性检验：

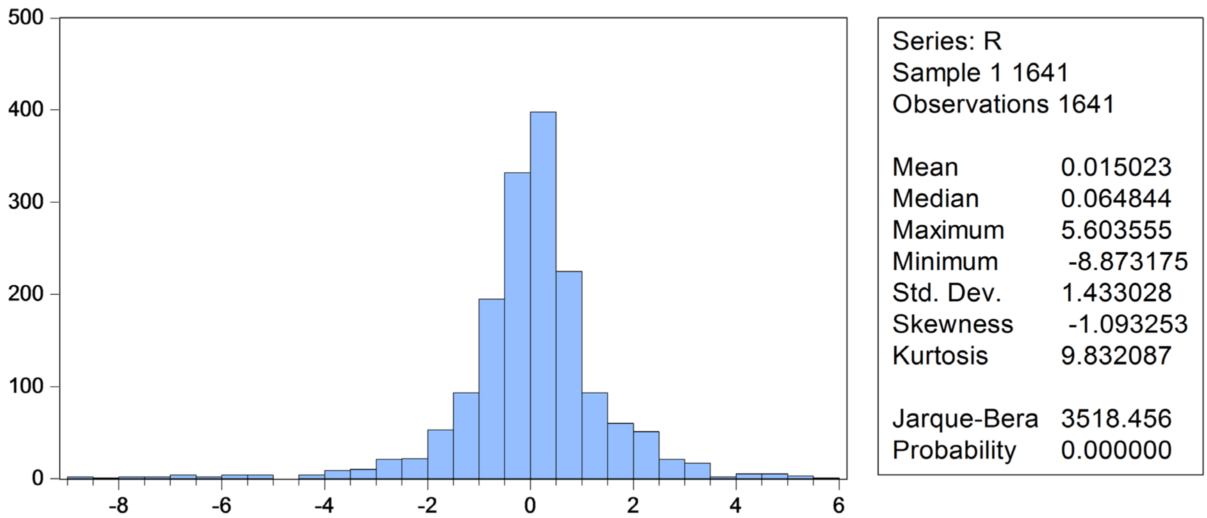

Figure 5. Normality test of daily yield series of Shanghai Composite Index
图 5. 上证综指的日收益率序列的正态性检验

从图 5 的检验结果来看，JB 值为 3518.456，p 值为 0，检验序列的分布于正态分布无显著差异的原假设，即序列不服从于正态分布。

4.2. 平稳性检验

由图 6 ADF 检验结果显示，序列为平稳序列。

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		-38.50348	0.0000
Test critical values:	1% level	-2.566373	
	5% level	-1.941017	
	10% level	-1.616569	

*MacKinnon (1996) one-sided p-values.

Figure 6. ADF test
图 6. ADF 检验

4.3. 序列自相关检验

图 7 是序列的自相关图，从图中可以看出，序列存在自相关性。

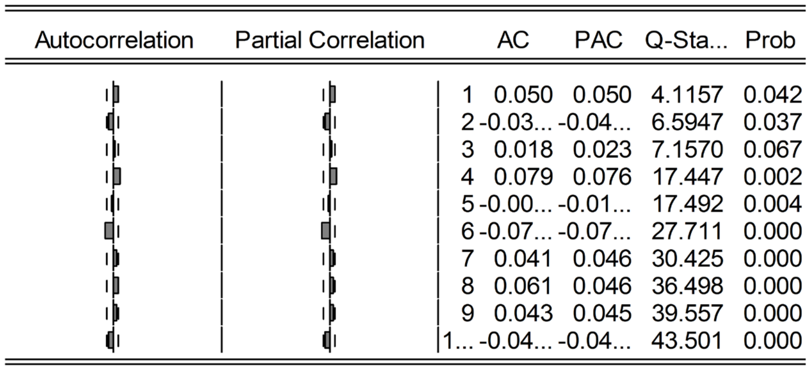

Figure 7. Autocorrelation test results of Shanghai Composite Index Series
图 7. 上证综指序列的自相关检验结果

5. GARCH 模型的建立

在滞后 10 阶的检验中，自相关和偏相关在 1、4、6 阶截尾，
构建 ARMA (3,3)模型，如下：

Table 1. ARMA (3,3) model
表 1. ARMA (3,3)模型

Variable	Coefficien...	Std.Error	t-Statistic	Prob.
C	0.015984	0.026661	0.599531	0.5489
AR(1)	0.139715	0.031922	4.376760	0.0000
AR(4)	0.861118	0.024721	34.83371	0.0000
AR(6)	−0.006711	0.021404	−0.313540	0.7539
MA(1)	−0.111201	0.038130	−2.916342	0.0036
MA(4)	−0.821925	0.207979	−3.951966	0.0001
MA(6)	−0.066874	0.021323	−3.136214	0.0017
SIGMASQ	1.992629	0.044360	44.91946	0.0000
R-squared	0.029084	Mean dependent var		0.015023
Adjusted R-squared	0.024922	S.D.dependent var		1.433028
S.E.of regression	1.415059	Akaike info criterion		3.537763
Sum squared resid	3269.905	Schwarz criterion		3.564103
Log likelihood	−2894.734	Hannan-Quinn criter		3.547531
F-statistic	6.988063	Durbin-Watson stat		1.955300
Prob(F-static)	0.000000			
Inverted AR Roots	1.00 −0.09	0.09 −0.93	0.03 − 0.96i	0.03 + 0.96i
Inverted MA Roots	1.00 −0.00 + 0.29i	−0.03 − 0.93i −0.95	0.03 + 0.93i	−0.00 − 0.29i

从表 1 的回归结果来看，变量在 AR (6)不显著，因此剔除，构建 ARMA (1,1)模型。构建的 ARMA (2,3)结果如下：

Table 2. ARMA (3,3) model
表 2. ARMA (3,3)模型

Variable	Coefficien...	Std.Error	t-Statistic	Prob.
C	0.015446	0.032642	0.473180	0.6361
AR(1)	0.132176	0.025761	5.130849	0.0000
AR(4)	0.864408	0.025451	33.96309	0.0000
MA(1)	−0.102980	0.026670	−3.861254	0.0001
MA(4)	−0.825392	0.214554	−3.847018	0.0001
MA(6)	−0.071629	0.014447	−4.958038	0.0000
SIGMASQ	1.993658	0.042715	46.67307	0.0000
R-squared	0.028583	Mean dependent var		0.015023
Adjusted R-squared	0.025016	S.D.dependent var		1.433028
S.E.of regression	1.414991	Akaike info criterion		3.536882
Sum squared resid	3271.592	Schwarz criterion		3.559930
Log likelihood	−2895.012	Hannan-Quinn criter		3.545430
F-statistic	8.013053	Durbin-Watson stat		1.955693
Prob(F-static)	0.000000			
Inverted AR Roots	1.00	0.03 − 0.96i	0.03 − 0.96i	−0.93
Inverted MA Roots	1.00	0.03 − 0.93i	0.03 + 0.93i	−0.00 − 0.30i
	−0.00 + 0.30i	−0.95		

表 2 即为 ARMA (2,3)模型，图 8 为模型的残差序列自相关图，即 ARMA (2,3)模型不存在自相关性。

Autocorrelation	Partial Correlation		AC	PAC	Q-Sta...	Prob
		1	0.022	0.022	0.8049	
		2	-0.02...	-0.02...	1.9153	
		3	0.025	0.026	2.9259	
		4	0.016	0.014	3.3456	
		5	-0.04...	-0.04...	6.3962	
		6	0.000	0.002	6.3963	0.011
		7	0.061	0.058	12.500	0.002
		8	0.001	0.000	12.502	0.006
		9	0.008	0.012	12.612	0.013
		1...	0.022	0.017	13.380	0.020

Figure 8. Auto correlation test of residual sequence of ARMA (2,3) model
图 8. ARMA (2,3)模型的残差序列自相关检验

图 9 为 ARCH_LM 检验结果，说明构建的自回归模型存在 ARCH 效应。

Heteroskedasticity Test: ARCH

F-statistic	77.22173	Prob. F(1,1638)	0.0000
Obs*R-squared	73.83514	Prob. Chi-Square(1)	0.0000

Figure 9. ARCH_LM test
图 9. ARCH_LM 检验

如下，表 3 是构建的 GARCH (1,1)模型结果。模型的 AR 根均在单位圆内，表明构建的模型是稳健的。

Table 3. GARCH (1,1) model
表 3. GARCH (1,1)模型

GARCH = C(7) + C(8) * RESID(−1)^2 + C(9) * GARCH(−1)				
Variable	Coefficien...	Std.Error	z-Statistic	Prob.
C	0.022924	0.023941	0.957506	0.3383
AR(1)	−0.405982	0.010532	−38.545660	0.0000
AR(4)	−0.711142	0.010143	−70.11235	0.0000
MA(1)	0.447706	0.021530	20.79411	0.0000
MA(4)	0.716766	0.009450	75.84843	0.0000
MA(6)	−0.051315	0.017636	−2.909747	0.0036
Variance Equation				
C	0.007187	0.001847	3.890853	0.0001
RESID(−1)^2	0.065276	0.005791	11.27270	0.0000
GARCH(−1)	0.933587	0.005106	182.8359	0.0000
R-squared	0.008351	Mean dependent var		0.015023
Adjusted R-squared	0.005318	S.D.dependent var		1.433028
S.E.of regression	1.429212	Akaike info criterion		3.131509
Sum squared resid	3339.729	Schwarz criterion		3.161142
Log likelihood	−2560.403	Hannan-Quinn criter		3.142499
Durbin-Watson stat	1.974143			
Inverted AR Roots	0.56−0.64i	0.56 + 0.64i	−0.76 + 0.63i	−0.76 − 0.63i
Inverted MA Roots	0.54 − 0.66i −0.76 + 0.64i	0.54 + 0.66i −0.76 − 0.64i	0.27	−0.27

6. 加入 X 变量的 GARCH 模型构建

(1) X 变量自相关检验

由图 10 可知：X 在 4 阶和 8 阶存在自相关。

Autocorrelation	Partial Correlation		AC	PAC	Q-Sta...	Prob
		1	-0.01...	-0.01...	0.2655	0.606
		2	-0.02...	-0.02...	1.4103	0.494
		3	0.008	0.007	1.5079	0.680
		4	-0.05...	-0.05...	6.1280	0.190
		5	-0.02...	-0.02...	6.8235	0.234
		6	-0.00...	-0.00...	6.8785	0.332
		7	0.031	0.030	8.4193	0.297
		8	-0.06...	-0.07...	16.052	0.042
		9	-0.01...	-0.02...	16.571	0.056
		1...	-0.00...	-0.01...	16.651	0.082

Figure 10. Auto correlation test of NASDAQ Composite Index series
图 10. 纳斯达克综合指数序列的自相关检验

(2) 新的 GARCH 模型构建

Table 4. New GARCH model
表 4. New GARCH 模型

GARCH = C(9) + C(10) * RESID(−1)^2 + C(11) * GARCH(−1)				
Variable	Coefficien...	Std.Error	z-Statistic	Prob.
C	0.016098	0.024645	0.653198	0.5136
X(−4)	−2.979228	2.419876	−1.231149	0.2183
X(−8)	7.843263	2.739426	2.863104	0.0042
AR(1)	−0.694831	0.284043	−2.446214	0.0144
AR(4)	0.003986	0.306994	−0.012985	0.9896
MA(1)	0.722414	0.277123	2.606834	0.0091
MA(4)	0.035556	0.304910	0.116612	0.9072
MA(6)	0.042787	0.024397	−1.753757	0.0795
Variance Equation				
C	0.006743	0.001702	3.960632	0.0001
RESID(−1)^2	0.063751	0.005774	11.04165	0.0000
GARCH(−1)	0.935288	0.004991	187.3812	0.0000
R-squared	0.017034	Mean dependent var		0.014228
Adjusted R-squared	0.012800	S.D.dependent var		1.433636
S.E.of regression	1.424431	Akaike info criterion		3.131218
Sum squared resid	3297.133	Schwarz criterion		3.167581
Log likelihood	−2545.640	Hannan-Quinn criter		3.144707
Durbin-Watson stat	1.939084			
Inverted AR Roots	0.09 + 0.14i	0.09 − 0.14i	−0.20	−0.68
Inverted MA Roots	0.49 −0.44 − 0.42i	0.23 − 0.49i −0.79	0.23 + 0.49i	−0.44 + 0.42i

(3) LB 检验

Table 5. Coefficient equality test of X lag variable
表 5. X 滞后变量的系数相等检验

Test Statistic	Value	df	Probability
t-statistic	−2.832574	1622	0.0047
F-statistic	8.023473	(1.1622)	0.0047
Chi-square	8.023473	1	0.0046
Null Hypothesis: C(2) = C(3)			
Null Hypothesis Summary:			

拒绝原假设，即系数不相等。

X 滞后变量的系数相等。

Table 6. Coefficient equality test of new GARCH model
表 6. new-GARCH 模型系数相等检验

Test Statistic	Value	df	Probability
F-statistic	5.852527	1	0.0001
Chi-square	23.41011	4	0.0001
Null Hypothesis: C(4) = C(5) = C(6) = C(7) = C(8)			
Null Hypothesis Summary:			

拒绝原假设。

r 的 GARCH 模型系数相等。

由表 4~6，由于对 r 的 GARCH 模型系数相等的检验中，我们发现实证拒绝原假设，所以得出$\{X_t\}$是$\{Y_t\}$的原因，即纳斯达克综合指数的日收益率与上证综指日收益率有着密切的联系。

7. 结论与建议

本文主要运用了加入外生变量的 GARCH 模型来对上证综指和纳斯达克综合指数之间的日收益率的波动来研究两种综合指数是否具有一定的联系，实证研究表明，两种综合指数之间具有很高的契合度，所以说两种综指之间具有联系。GARCH 模型中 $Y_t = \mu_t + a_t, \mu_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_m Y_{t-m}$ 中的 μ_t 是自回归模型，新型 GARCH 模型的建立是在 μ_t 的基础上加上一个符合自回归模型的变量，以此来组建一个新模型，根据新模型中系数是否相等来判断，变量和原量之间是否有关。

本文的模型建立可以在加入一个符合自回归模型的变量的基础上改为加入含有二阶项的变量，由于本文选取的数据不符合二阶模型，所以在此提出一个想法，可以建立更好更广义的 GARCH 模型。

参考文献

[1] 王沁. 时间序列分析及其应用[M]. 西安: 西安交通大学出版社, 2018.

[2] 李金凤, 张德生, 井霞霞. 基于考虑外生变量的 SWARCH 模型对中国股市波动性的实证研究[J]. 陕西科技大学学报(自然科学版), 2012, 30(1): 82-86.

[3] 林国华. 时间序列分析法在移动通信数据分析中的研究与应用[D]: [硕士学位论文]. 广州: 广东工业大学, 2013.

[4] 郭娜. 基于 skst-Copula-GARCH 模型的 VaR 计算研究[D]: [硕士学位论文]. 重庆: 重庆大学, 2010.