Testing Granger Causality under Dynamic Covariance

Manabu ASAI

Abstract

The paper considers the Granger causality tests based on the hetersokedasticity-consistent covariance matrix of White (1980). The Monte Carlo experiments show that the conventional test over-rejects the null hypothesis of non-causality, while the proposed test works satisfactory. The proposed test also has enough power. Using the data for the Standard and Poor's 500 Composite Index, the paper examines the causalities among the return, volatility and trading volume.

Keywords : Granger Causality ; Heteroskedasticity.

1. Introduction

In general, it may be stated that there is the causality from X to Y, if the change in X brings the change in Y directly. At least, it requires three conditions for such a direct connection ; (i) X appears before Y does, (ii) there is a strong correlation between X and Y, and (iii) the correlation between X and Y is not caused by other variables. We may test theoretical connection such as a hypothesis derived by economic theories, using regression models based on data of X and Y. The concept of causality proposed by Granger (1969) is widely spread, after Sims (1972) proposed an approach to test the Granger causality, in the framework of Vector AutoRegression (VAR) models. The idea of the Granger causality is that a cause cannot come after the effect. The test for Granger causality examines whether X helps forecast Y.

The VAR models considered in the paper accommodate the dynamic covariance, as in multivariate GARCH and stochastic volatility (SV) models. For the survey of multivariate GARCH and SV models, see the papers including McAleer (2005), Asai, McAleer and Yu (2006), and Chib, Omori and Asai (2009). Since high frequency time series, such as weekly, daily or intraday data exhibits heteroskedasticity, and since some monthly data shows ARCH effects, it is necessary to deal with dynamic heteroskedasticity with unknown form.

We may take two approaches ; the first one is to construct Wald statistic based on the heteroskedasticity-consistent covariance matrix estimator of White (1980), while the second one is work with the robust Lagrange multiplier (RLM) tests proposed by Wooldridge (1990). Regarding the latter approach for an univariate process, the conditional mean test of Wooldridge (1990) is the regression-based one, and it is robust to departures form the distributional assumptions which is not tested. Furthermore, it is valid in the presence of the unknown form of the conditional or stochastic variance. For example, Wooldridge (1991) proposed the RLM test for autocorrelation in the presence of ARCH effects, employing the results of Wooldridge (1990). The test is valid in the presence of the GARCH or stochastic volatility process, which is also examined by the Monte Carlo simulations conducted by Silvapulle and Evans (1998) and Asai (2000). Unlike constructing the RLM tests for VAR models, it is straightforward to develop Wald statistic based on the White's estimator. Hence, the paper works with the Wald tests.

The paper is organized as follows. Section 2 introduces our VAR model with heteroskedasticity, and constructs the heteroskedasticity-corrected Granger causality tests. Section 3 conducts Monte Carlo simulations for examining the finite sample properties of the heteroskedasticity corrected tests. Section 4 investigates the causal relationships of the return, range, and trading volume for the daily data of Standard and Poor's 500 Composite Index. Section 5 gives some concluding remarks.

2. Granger Causality Test

Let $Y_t = (y_{1t}, ..., y_{mt})'$ be an m-dimensional vectors. Consider a VAR(p) model with dynamic covariance, as follows :

$$Y_t = \mu + \Phi_1 Y_{t-1} + ... + \Phi_p Y_{t-p} + H_t^{1/2} \varepsilon_t, \tag{1}$$

where $\varepsilon_t \sim i.i.d.(0, I_m)$, H_t is the process of m-dimensional positive definite matrix, μ is the m-vector of parameters, and $\Phi_h(h=1,, ..., p)$ are $m \times m$ matrices of parameters. Assume that the structure of H_t is unknown, and let the forth moment of ε_t exist.

Let $u_t = H_t^{1/2} \varepsilon_t$ for convenience. Then, $E(u_t|H_t) = 0$ and $E(u_t u_t'|H_t) = H_t$. Now, the matrix form of the VAR model is given by

$$Y = XB + U, \tag{2}$$

where

$$\underset{T \times m}{Y} = (y_1 ... y_T)', \quad \underset{T \times k}{X} = (\iota\, Y_{-1} ... Y_{-p}), \quad \underset{k \times m}{B} = (\mu\, \Phi_1 ... \Phi_p)',$$

$$\underset{T\times m}{U}=(u_1...u_T)', \quad \underset{T\times 1}{\iota}=(1, ..., 1)', \quad \underset{T\times m}{Y_{-h}}=(y_{1-h}...y_{T-h})',$$

with $k=1+mp$. We may have another form of (3) as

$$y=(X\otimes I_m)\beta+u$$

where

$$\underset{Tm\times 1}{y}=\text{vec}(Y'), \quad \underset{km\times 1}{\beta}=\text{vec}(B'), \quad \underset{Tm\times 1}{u}=\text{vec}(U'),$$

with $E(u|H_1, ...H_T)=0$ and

$$E(uu'|H_1, ...H_T)=\Xi=\begin{pmatrix} H_1 & & O \\ & \ddots & \\ O & & H_T \end{pmatrix}.$$

The OLS estimator of β is given by $\hat{\beta}=[(X'X)^{-1}X'\otimes I_m]y$. Hence the covariance matrix of $\hat{\beta}$ is given by $V(\hat{\beta})=[(X'X)^{-1}X'\otimes I_m]\Xi[X(X'X)^{-1}\otimes I_m]$.

For the case of homoscedasticity, it is straightforward to show $E(uu'|\Sigma_1, ...\Sigma_T)=I_T\otimes\Sigma$, and hence we have $V(\hat{\beta})=(X'X)^{-1}\otimes\Sigma$ for constant covariance model. However, in the presence of heteroskedasticity, the conventional estimator for $V(\hat{\beta})$, that is

$$(X'X)^{-1}\otimes\hat{\Sigma} \tag{3}$$

with $\hat{\Sigma}=T^{-1}\hat{U}'\hat{U}$ and $\hat{U}=Y-X\hat{B}$, is inappropriate. Generally, conventional t tests and F tests over-reject the null hypothesis.

The paper employs the Heteroskedasicity-Consistent Covariance matrix (HCC) estimator of White (1980), in order to have

$$\widehat{V(\hat{\beta})}=[(X'X)^{-1}X'\otimes I_m]\begin{pmatrix} \hat{u}_1\hat{u}_1' & & O \\ & \ddots & \\ O & & \hat{u}_T\hat{u}_T' \end{pmatrix}[X(X'X)^{-1}\otimes I_m] \tag{4}$$

$$=\sum_{t=1}^{T}[(X'X)^{-1}x_t'\otimes I_m](\hat{u}_t\hat{u}_t')[x_t(X'X)^{-1}\otimes I_m]$$

where $x_t(1\times k)$ is the row vector of X and $\hat{u}_t(m\times 1)$ is the column vector of $\hat{U}'$. With the covariance matrix estimator (4), we can construct test statistics for linear restrictions, including the test for Granger causality. Under the null hypothesis $H_0:R\beta=q$, the test statistic

$$W=(R\hat{\beta}-q)'\left[R\widehat{V(\hat{\beta})}R'\right]^{-1}(R\hat{\beta}-q) \tag{5}$$

has the asymptotic χ^2 distribution with the degree of freedom which is equal to the number of restrictions.

The idea of causality proposed by Granger (1969) is that a cause cannot come after the effect. Thus, if y_j affects y_i, the former should help improving the predictions of the latter variable. In the context of the VAR model, the concept is interpreted as follows.

Granger Non-causality

y_i is not Granger-caused by y_j if and only if $\phi_{ij,h}=0$ for $h=1, 2, ..., p$, where $\phi_{ij,h}$ is the (i,j)-th element of Φ_h.

For convenience, let us denote $y_j \overset{G}{\longrightarrow} y_i$ for the case that y_j Granger-causes y_j. For the Granger causality test, the test statistic (5) has the asymptotic χ^2 distribution with the degree of freedom p, under the null hypothesis of non-causality.

3. Monte Carlo Experiments

This section conducts Monte Carlo experiments to the accuracy of the size and power of the Granger causality test based on the HCC estimator of White (1980).

The data generating process (DGP) is the bivariate VAR(2) model with multivariate stochastic volatility, which is given by

$$Y_t = \mu + B_1 Y_{t-1} + B_2 Y_{t-2} + H_t^{1/2} \varepsilon_t, \quad \varepsilon_t \sim N(0, I_2),$$

$$H_t = \begin{pmatrix} \exp(\alpha_{1t}/2) & 0 \\ 0 & \exp(\alpha_{2t}/2) \end{pmatrix} P_\varepsilon \begin{pmatrix} \exp(\alpha_{1t}/2) & 0 \\ 0 & \exp(\alpha_{2t}/2) \end{pmatrix},$$

$$\alpha_t = \begin{pmatrix} 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0.95 & 0 \\ 0 & 0.98 \end{pmatrix} \alpha_{t-1} + \begin{pmatrix} 0.2 & 0 \\ 0 & 0.1 \end{pmatrix} \eta_t, \quad \eta_t \sim N(0, P_\eta), \tag{6}$$

$$\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, B_1 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, B_2 = \begin{pmatrix} 0.5 & a \\ b & 0.5 \end{pmatrix}, P_\varepsilon = \begin{pmatrix} 1 & c \\ c & 1 \end{pmatrix}, P_\eta = \begin{pmatrix} 1 & 0.4 \\ 0.4 & 1 \end{pmatrix},$$

where $H^{1/2}$ is based on the Choleski decomposition, and (a, b, c) are specified as follows

$$(a, b, c) = \begin{cases} (0.0, 0.0, 0.0) & \text{DGP1} \\ (0.4, 0.0, 0.0) & \text{DGP2} \\ (0.0, 0.4, 0.0) & \text{DGP3} \\ (0.0, 0.0, 0.3) & \text{DGP4} \\ (0.4, 0.0, 0.3) & \text{DGP5} \\ (0.0, 0.4, 0.3) & \text{DGP6} \end{cases}$$

The sample size considered here is $T = 100$ and 500. Monte Carlo simulations are based on 10,000 replications. For the benchmark, we work with the conventional tests based on (3).

Table 1 shows the empirical sizes of the tests at 5% nominal level, with 6 DGPs for T =100. All empirical sizes of the conventional test are larger than the corresponding tests based on the HCC estimator of White (1980). The empricall size of the corrected test for T =100 is about 10% in the experiment. Table 2 presents the empirical size for T=500, which shows the improvement of the empirical sizes for the tests with HCC, which are close to 0.05.

Table 1 : Empirical Sizes at 5% Nominal Level for T=100

DGP	Granger-Causality Test			
	w/o Correction		HCCM Estimator	
	Y2 to Y1	Y1 to Y2	Y2 to Y1	Y1 to Y2
DGP1	0.127	0.127	0.093	0.103
DGP2	--	0.126	--	0.099
DGP3	0.126	--	0.097	--
DGP4	0.137	0.129	0.103	0.100
DGP5	--	0.131	--	0.103
DGP6	0.133	--	0.095	--

Note : 'HCCM estimator' stands for the heteroskedasticity-consistent covariance matrix estimator of White (1980).

Table 2 : Empirical Sizes at 5% Nominal Level for T=500

DGP	Granger-Causality Test			
	w/o Correction		HCCM Estimator	
	Y2 to Y1	Y1 to Y2	Y2 to Y1	Y1 to Y2
DGP1	0.077	0.078	0.060	0.061
DGP2	--	0.078	--	0.057
DGP3	0.078	--	0.055	--
DGP4	0.075	0.074	0.057	0.059
DGP5	--	0.072	--	0.053
DGP6	0.080	--	0.057	--

Note : 'HCCM estimator' stands for the heteroskedasticity-consistent covariance matrix estimator of White (1980).

Table 3 : Rejection Frequencies at 5% Nominal Level for T=100

DGP	Granger-Causality Test with HCCM Estimator			
	Y1 to Y1	Y2 to Y1	Y1 to Y2	Y2 to Y2
DGP1	0.958	--	--	0.966
DGP2	0.983	0.899	--	0.970
DGP3	0.950	--	0.922	0.991
DGP4	0.937	--	--	0.961
DGP5	0.976	0.855	--	0.946
DGP6	0.931	--	0.888	0.984

Note : 'HCCM estimator' stands for the heteroskedasticity-consistent covariance matrix estimator of White (1980).

Table 4: Rejection Frequencies at 5% Nominal Level for
$T=500$

| DGP | Granger-Causality Test with HCCM Estimator | | | |
	Y1 to Y1	Y2 to Y1	Y1 to Y2	Y2 to Y2
DGP1	1.000	--	--	1.000
DGP2	1.000	1.000	--	1.000
DGP3	1.000	--	1.000	1.000
DGP4	1.000	--	--	1.000
DGP5	1.000	1.000	--	1.000
DGP6	1.000	--	1.000	1.000

Note: 'HCCM estimator' stands for the heteroskedasticity-consistent
covariance matrix estimator of White (1980).

With respect to the Granger causality test with HCC, Table 3 and 4 give the rejection frequencies under the alternative hypothesis at the nominal size of 5% for $T=100$ and $T=500$, respectively. Those tables indicate that tests have enough power, even for the small sample size of $T=100$.

4. Empirical Example

The section examines causalities among the stock returns, range and trading volume, based on the SVAR-causality, the partially instantaneous correlation, and the Granger causality for the daily data of S&P 500 index. The sample period for the S&P 500 index is 8/4/2005 to 8/2/2007, giving $T=500$ observations.

Denote the stock return, range, and trading volume as y_t, d_t and v_t, respectively. The range is defined as the difference between highest and lowest returns on day t, while the trading volume is the number of shares traded at day t. The log-range is sometimes used as a proxy of the return volatility. Recent development including Liesenfeld (2001), Alizadeh, Brandt and Diebold (2002) and Fleming, Kirby and Ostdiek (2006) show that these time series are supposed to contain unobservable stochastic process, say, α_t^y, α_t^d, α_t^{v1} and α_t^{v2}.

With the unobservable processes, they may be specified as

$$y_t = \sigma_y u_t^y \exp(0.5\alpha_t^y), \tag{7}$$

$$\ln d_t = c_d + \alpha_t^d + u_t^d, \tag{8}$$

$$v_t = c_v \exp(\alpha_t^{v1}) + \sigma_v u_t^v \exp(0.5\alpha_t^{v2}), \tag{9}$$

where σ_y, σ_v, c_d and c_v are unknown parameters, and (u_t^y, u_t^d, u_t^v) are disturbances. The equation (7) is the conventional specification of the stochastic volatility model. See Shephard (2005) and Chernov et al. (2003) for the recent development in the fields. The equation (8) is

obtained by Alizadeh, Brandt and Diebold (2002), implying that the log-range is the sum of the log-volatility and noise. The equation (9) with the restriction of $a_t^{v1} = a_t^{v2}$ corresponds to the bivariate mixture model of Tauchen and Pitts (1983), and is applied to the analysis of Andersen (1996), Liesenfeld (1998) and so on.

Alizadeh, Brandt and Diebold (2002) assume $a_t^y = a_t^d$, based on the framework of the continuous-time stochastic volatility models. Andersen (1996) and Liseenfeld (1998) specified the relation as $a_t^y = a_t^v$, but found a significant difference, compared to the univariate case regarding equation (9).

As the true structure of $a_t = (a_t^y, a_t^d, a_t^{v1}, a_t^{v2})'$ is unknown, the current paper will examine the relationships by Granger Causality tests on two kinds of trivariate processes, i.e., $(y_t, \ln d_t, v_t)$ and $(\ln y_t^2, \ln d_t, \ln v_t^2)$, in order to derive implications regarding the structure. The former can be considered as a proxy of $(y_t, a_t^d, \exp(a_t^{v1}))$, while the latter as (a_t^y, a_t^d, a_t^{v2}).

The sample period for the S&P 500 index is 8/4/2005 to 8/2/2007, giving $T = 500$ observations. Let P_t is the price of stock on day t. Especially, the closing price is denoted as P_t^C. Then, the return is defined by $r_t = 100 \times \{\ln P_t - \ln P_{t-1}^C\}$, while the return for the closing price is given by $y_t = 100 \times \{\ln P_t^C - \ln P_{t-1}^C\}$. The range is defined by $d_t = \max\{r_t\} - \min\{r_t\}$. The trading volume is the number of shares traded at day t. It was divided by 10^9, and the trend effect was removed.

Table 5 presents the results for the tests for the Granger causality under the dynamic covariance, showing that $y \xrightarrow{G} (\ln d, v)$, $\ln d \xrightarrow{G} \ln d$ and $v \xrightarrow{G} (y, v)$, where $x_i \xrightarrow{G} x_j$ means that x_i causes x_j in Granger's sense. Figure 1 shows the causality graph for return, trading volume and log-range. For the latent variables, a_t^d depends on the past values of y_t and itself, while a_t^v also depends on the past values of y_t and itself. An interesting result is that y_t depends on the past values of v_t (a_t^{v1}).

Table 5 : Results for Granger Causality :
$(y_t, \ln d_t, v_h)$

From	Granger Causality		
	to y_t	to $\ln d_t$	to v_h
y_t	3.232	67.751*	12.376*
$\ln d_t$	3.747	35.871*	2.985
v_h	13.071*	3.341	31.708*

Note : '*' denotes significant at five percent level.
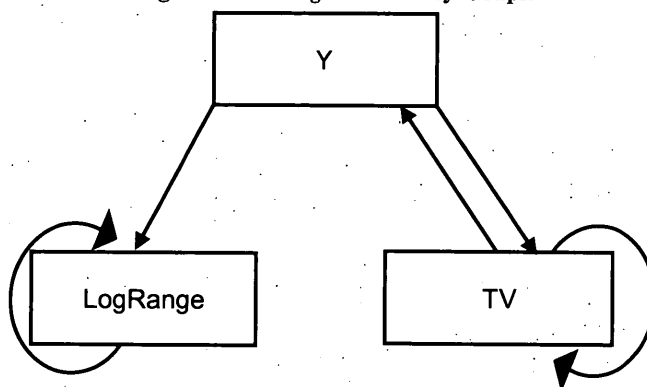
Figure 1 : Granger Causality Graph

Table 6 : Test Results for Granger Causality :
$(\ln y_t^2, \ln d_t, \ln v_t^2)$

From	Granger Causality		
	to $\ln y_t^2$	to $\ln d_t$	to $\ln v_t^2$
$\ln y_t^2$	15.644*	11.811*	4.627
$\ln d_t$	18.337*	53.075*	4.097
$\ln v_t^2$	4.449	8.452	35.499*

Note : '*' denotes significant at five percent level.
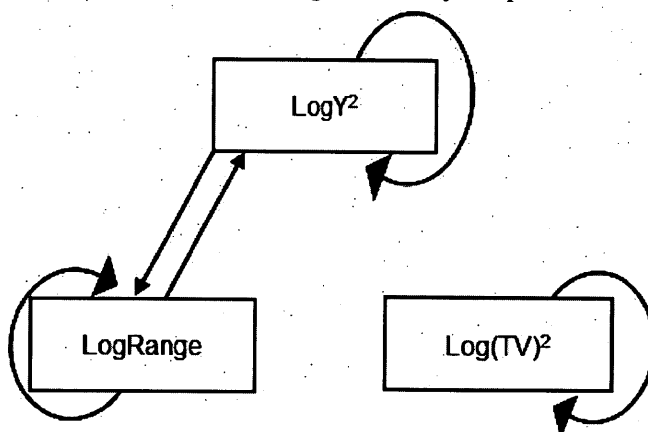
Figure 2 : Granger Causality Graph

Table 6 and Figure 2 give the results for the Granger causality under the dynamic covariance with respect to $(\ln y_t^2, \ln d_t, \ln v_t^2)$, indicating that $\ln y^2 \xrightarrow{G} (\ln y^2, \ln d)$, $\ln d \xrightarrow{G} (\ln y^2, \ln d)$ and $\ln v^2 \xrightarrow{G} \ln v^2$. The causal relationships between $\ln y^2$ and $\ln d$ seems to be brought by a common factor between (a_t^y, a_t^d).

Summing up the empirical results, the implications are as follows ; (R1) a_t^{v1} is not equal

to a_t^{v2} ; (R2) a_t^d and a_t^{v2} should be specified separately ; (R3) a_t^d and a_t^{v1} may be specified by the bivariate VAR model ; (R4) a_t^d also depends on the past values of y_t ; (R5) a_t^{v2} depends on the past values of itself ; (R6) The conditional mean of y_t depends on the past values of a_t^v.

5. Conclusions

Under the dynamic covariance, the paper considers the heteroskedasticity-corrected Granger causality tests, based on White's (1980) estimator. Monte Carlo experiments show that the heteroskedasticity-corrected test works satisfactory, and that the conventional tests over-reject the null hypothesis of non-causality. The paper investigated the causalities among the S&P 500 return, trading volume and volatility, by using the new approach. Several new results are obtained.

References

Alizadeh, S., M.W. Brandt and F.X. Diebold (2002), "Range-Based Estimation of Stochastic Volatility Models", *Journal of Finance*, **57**, 1047-1091.

Andersen, T.G. (1996), "Return Volatility and Trading Volume: An Information Flow Interpretation of Stochastic Volatility", *Journal of Finance*, **51**, 169-204.

Asai (2000), "Testing for Serial Correlation in the Presence of Stochastic Volatility", *Asia-Pacific Financial Markets*, 2000, **7**, 321-337.

Asai, M., M. McAleer and J. Yu (2006), "Multivariate Stochastic Volatility: A Review", *Econometric Reviews*, **25**, 145-175.

Asai, M. and A. Unite (2008), "The Relationship between Stock Return Volatility and Trading Volume: The case of The Philippines", forthcoming in *Applied Financial Economics.*, **18**, 1333-1348.

Chernov, M., A.R. Gallant, E. Ghysels, and G. Tauchen (2003), "Alternative Models for Stock Price Dynamics", *Journal of Econometrics*, **116**, 225-257.

Chib, S., Y. Omori and M. Asai (2009), "Multivariate Stochastic Volatility", in: T.G. Andersen, R.A. Davis, J.-P. Kreiss and T. Mikosch (eds.), *Handbook of Financial Time Series*, Springer-Verlag, New York, 365 -400.

Fleming, J., C. Kirby, and B. Ostdiek (2006) "Stochastic Volatility, Trading Volume, and the Daily Flow of Information", *Journal of Business*, **79**, 1551-1590.

Garman, M.B. and M.J. Klass (1980), "On th estimation of Security Price Volatilities from Historical Data", *Journal of Business*, **53**, 67-78.

Granger, C.W.J. (1969), "Investigating Causal Relations by Econometric Models and Cross-Spectral Methods", *Econometrica*, **37**, 424-438.

Liesenfeld, R. (1998), "Dynamic Bivariate Mixture Models: Modeling the Behavior of Prices and Trading Volume", *Journal of Business & Economic Statistics*, **16**, 101-109.

Liesenfeld, R. (2001), "A generalized bivariate mixture model for stock price volatility and trading volume", *Journal of Econometrics*, **104**, 141-178.

McAleer, M. (2005), "Automated Inference and Learning in Modeling Financial Volatility", *Econometric Theory*, **21**, 232-261.

Shephard, N. (2005), *Stochastic Volatility*, Oxford University Press, Oxford.

Silvapulle, P. and M. Evans (1998), "Testing for Serial Correlation in the Presence of Dynamic Heteros-

cedasticity", *Econometric Reviews* **17**, 31-55.

Sims, C.A. (1972), "Money, Income, and Causality", *American Economic Review*, **62**, 540-552.

Sims, C.A. (1980), "Macroeconomics and Reality," *Econometrica*, **48**, 1-48.

Tauchen, G.E. and M. Pitts (1983), "The Price Variability-Volume Relationship on Speculative Markets", *Econometrica*, **51**, 485-505.

White, H. (1980), "A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity," *Econometrica*, **50**, 1-26.