Testing Granger Causality under Dynamic Covariance

Manabu ASAI

Abstract
The paper considers the Granger causality tests based on the heteroskedasticity-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 satisfactorily. 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 from 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 \( Y_t = (y_{1t}, \ldots, y_{mt})' \) be an \( m \)-dimensional vectors. Consider a VAR(\( p \)) model with dynamic covariance, as follows:

\[
Y_t = \mu + \Phi_1 Y_{t-1} + \ldots + \Phi_p Y_{t-p} + H_{t}^{1/2} \varepsilon_t,
\]

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

\[
Y = XB + U,
\]

where

\[
y_Y = (y_1 \ldots y_T), \quad X = (Y_{-1} \ldots Y_{-p}), \quad B = (\mu \Phi_1 \ldots \Phi_p)',
\]

\[
Y_{t \times m}, \quad X_{t \times h}, \quad B_{h \times m}.
\]
\[ \begin{align*}
U & = (u_1, \ldots, u_T)'_x, \quad t = (1, \ldots, 1)', \\
Y_{-h} & = (y_{1-h}, \ldots, y_{T-h})', \\
\end{align*} \]

with \( k = 1 + mp \). We may have another form of (3) as
\[ y = (X \otimes I_m) \beta + u \]

where
\[ y = \text{vec}(Y'), \quad \beta = \text{vec}(B'), \quad u = \text{vec}(U'), \]

with \( E(u|H_1, ..., H_T) = 0 \) and
\[ E(uu' | H_1, ..., H_T) = \Sigma = \begin{pmatrix} H_1 & 0 \\ 0 & H_T \end{pmatrix}. \]

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

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
\[ (X'X)^{-1} \otimes \Sigma \]
with \( \Sigma = T^{-1} \bar{U}' \bar{U} \) and \( \bar{U} = Y - \bar{X} \bar{B} \), is inappropriate. Generally, conventional \( t \) tests and \( F \) tests over-reject the null hypothesis.

The paper employs the Heteroskedasticity-Consistent Covariance matrix (HCC) estimator of White (1980), in order to have
\[ \begin{align*}
\hat{V}(\hat{\beta}) & = [(X'X)^{-1}X' \otimes I_m] \\
& = \sum_{t=1}^{T} [(X'X)^{-1}x_t' \otimes I_m] (\bar{u}_t \bar{u}_t') [x_t(X'X)^{-1} \otimes I_m] \end{align*} \]

where \( x_t(1 \times k) \) is the row vector of \( X \) and \( \bar{u}_t(m \times 1) \) is the column vector of \( \bar{U}' \). 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
\[ W = (R \hat{\beta} - q) \left[ R \hat{V}(\hat{\beta}) R' \right]^{-1} (R \hat{\beta} - q) \]
has the asymptotic \( \chi^2 \) 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 \( y_i \) affects \( y_t \), 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 \( h = 1, 2, \ldots, p \), where \( \phi_{ij,h} \) is the \((i,j)\)-th element of \( \Phi_h \).

For convenience, let us denote \( y_j \overset{G}{\rightarrow} y_i \) for the case that \( y_j \) Granger-causes \( y_i \). For the Granger causality test, the test statistic (5) has the asymptotic \( \chi^2 \) 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

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

\[
H_t = \begin{pmatrix}
\exp(\alpha_{1t}/2) & 0 \\
0 & \exp(\alpha_{2t}/2)
\end{pmatrix}
\begin{pmatrix}
P_1 & 0 \\
0 & P_2
\end{pmatrix}
\begin{pmatrix}
\exp(\alpha_{1t}/2) & 0 \\
0 & \exp(\alpha_{2t}/2)
\end{pmatrix},
\]

\[
\alpha_t = \begin{pmatrix}
0 \\
0
\end{pmatrix} + \begin{pmatrix}
0.95 & 0 \\
0 & 0.98
\end{pmatrix} \alpha_{t-1} + \begin{pmatrix}
0.2 & 0 \\
0 & 0.1
\end{pmatrix} \eta_t, \quad \eta_t \sim N(0,P_t),
\]

\[
\mu = \begin{pmatrix}
0 \\
0
\end{pmatrix}, \quad B_1 = \begin{pmatrix}
0 & 0 \\
0 & 0
\end{pmatrix}, \quad B_2 = \begin{pmatrix}
0.5 & a \\
b & 0.5
\end{pmatrix}, \quad P_1 = \begin{pmatrix}
1 & c \\
c & 1
\end{pmatrix}, \quad P_2 = \begin{pmatrix}
1 & 0.4 \\
0.4 & 1
\end{pmatrix},
\]

where \( H_t^{1/2} \) 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 DGP's 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 empirical 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 \( T=100 \)

<table>
<thead>
<tr>
<th>DGP</th>
<th>Granger-Causality Test</th>
<th>w/o Correction</th>
<th>HCCM Estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Y2 to Y1</td>
<td>Y1 to Y2</td>
</tr>
<tr>
<td>DGP1</td>
<td>0.127</td>
<td>0.127</td>
<td>0.093</td>
</tr>
<tr>
<td>DGP2</td>
<td>0.126</td>
<td>--</td>
<td>0.099</td>
</tr>
<tr>
<td>DGP3</td>
<td>0.137</td>
<td>0.129</td>
<td>0.103</td>
</tr>
<tr>
<td>DGP4</td>
<td>--</td>
<td>0.131</td>
<td>--</td>
</tr>
<tr>
<td>DGP5</td>
<td>0.133</td>
<td>--</td>
<td>0.095</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>DGP</th>
<th>Granger-Causality Test</th>
<th>w/o Correction</th>
<th>HCCM Estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Y2 to Y1</td>
<td>Y1 to Y2</td>
</tr>
<tr>
<td>DGP1</td>
<td>0.077</td>
<td>0.078</td>
<td>0.060</td>
</tr>
<tr>
<td>DGP2</td>
<td>0.078</td>
<td>--</td>
<td>0.057</td>
</tr>
<tr>
<td>DGP3</td>
<td>0.075</td>
<td>--</td>
<td>0.055</td>
</tr>
<tr>
<td>DGP4</td>
<td>--</td>
<td>0.072</td>
<td>--</td>
</tr>
<tr>
<td>DGP5</td>
<td>0.080</td>
<td>--</td>
<td>0.057</td>
</tr>
</tbody>
</table>

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

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

<table>
<thead>
<tr>
<th>DGP</th>
<th>Granger-Causality Test with HCCM Estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Y1 to Y1</td>
</tr>
<tr>
<td>DGP1</td>
<td>0.958</td>
</tr>
<tr>
<td>DGP2</td>
<td>0.983</td>
</tr>
<tr>
<td>DGP3</td>
<td>0.950</td>
</tr>
<tr>
<td>DGP4</td>
<td>0.937</td>
</tr>
<tr>
<td>DGP5</td>
<td>0.976</td>
</tr>
<tr>
<td>DGP6</td>
<td>0.931</td>
</tr>
</tbody>
</table>

Note: 'HCCM estimator' stands for the heteroskedasticity-consistent covariance matrix estimator of White (1980).
Table 4: Rejection Frequencies at 5% Nominal Level for 
\(T=500\)

<table>
<thead>
<tr>
<th>DGP</th>
<th>Granger-Causality Test with HCCM Estimator</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Y1 to Y1</td>
</tr>
<tr>
<td>DGP1</td>
<td>1.000</td>
</tr>
<tr>
<td>DGP2</td>
<td>1.000</td>
</tr>
<tr>
<td>DGP3</td>
<td>1.000</td>
</tr>
<tr>
<td>DGP4</td>
<td>1.000</td>
</tr>
<tr>
<td>DGP5</td>
<td>1.000</td>
</tr>
<tr>
<td>DGP6</td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Y2 to Y1</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Y1 to Y2</td>
</tr>
<tr>
<td></td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>Y2 to Y2</td>
</tr>
<tr>
<td></td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>--</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
<tr>
<td></td>
<td>1.000</td>
</tr>
</tbody>
</table>

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, \(\alpha_t^r\), \(\alpha_t^d\), \(\alpha_t^{r1}\) and \(\alpha_t^{r2}\).

With the unobservable processes, they may be specified as

\[
y_t = \sigma_u u_t^r \exp(0.5 \alpha_t^r),
\]
\[
\ln d_t = c_d + \alpha_t^d + u_t^d,
\]
\[
v_t = c_v \exp(\alpha_t^{r1}) + \sigma_u u_t^r \exp(0.5 \alpha_t^{r2}),
\]

where \(\sigma_u\), \(\sigma_v\), \(c_d\) and \(c_v\) are unknown parameters, and \((u_t^r, u_t^d, u_t^v)\) 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 $\alpha_t^2 = \alpha_t^2$ 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 $\alpha_t^2 = \alpha_t^2$, based on the framework of the continuous-time stochastic volatility models. Andersen (1996) and Liesenfeld (1998) specified the relation as $\alpha_t^2 = \alpha_t^2$, but found a significant difference, compared to the univariate case regarding equation (9).

As the true structure of $\alpha_t = (\alpha_t^1, \alpha_t^2, \alpha_t^3, \alpha_t^2)^T$ 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 $(\ln y_t, \ln d_t, \ln v_t)$, in order to derive implications regarding the structure. The former can be considered as a proxy of $(y_t, \alpha_t^1, \exp(\alpha_t^2))$, while the latter as $(\alpha_t^1, \alpha_t^2, \alpha_t^3)$.

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 $r_t = 100 \times (\ln P_t - \ln P_{t-1})$, 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 $y \rightarrow (\ln d, v)$, $\ln d \rightarrow \ln d$ and $v \rightarrow (y, v)$, where $x_i \rightarrow \rightarrow x_j$ 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, $\alpha_t^1$ depends on the past values of $y_t$ and itself, while $\alpha_t^2$ also depends on the past values of $y_t$ and itself. An interesting result is that $y_t$ depends on the past values of $v_t$ ($\alpha_t^3$).

| Table 5 : Results for Granger Causality : $(y_t, \ln d_t, v_h)$ |
|------------------------|------------------------|------------------------|
|                         | 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.
Table 6: Test Results for Granger Causality:
\((\ln y^2, \ln d^t, \ln v^t)\)

<table>
<thead>
<tr>
<th>From</th>
<th>(\ln y^2)</th>
<th>(\ln d^t)</th>
<th>(\ln v^t)</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\ln y^2)</td>
<td>15.644*</td>
<td>11.811*</td>
<td>4.627</td>
</tr>
<tr>
<td>(\ln d^t)</td>
<td>18.337*</td>
<td>53.075*</td>
<td>4.097</td>
</tr>
<tr>
<td>(\ln v^t)</td>
<td>4.449</td>
<td>8.452</td>
<td>35.499*</td>
</tr>
</tbody>
</table>

Note: "*" denotes significant at five percent level.

Table 6 and Figure 2 give the results for the Granger causality under the dynamic covariance with respect to \((\ln y^2, \ln d^t, \ln v^t)\), indicating that \(\ln y^2 \rightarrow (ln y^2, \ln d)\), \(\ln d \rightarrow (ln y^2, \ln d)\) and \(\ln v^2 \rightarrow \ln v^2\). The causal relationships between \(ln y^2\) and \(ln d\) seems to be brought by a common factor between \((a^x, a^y)\).

Summing up the empirical results, the implications are as follows; (R1) \(a^{x1}\) is not equal
to $\alpha^2$; (R2) $\alpha^\delta$ and $\alpha^2$ should be specified separately; (R3) $\alpha^\delta$ and $\alpha^2$ may be specified by the bivariate VAR model; (R4) $\alpha^\delta$ also depends on the past values of $y_t$; (R5) $\alpha^2$ depends on the past values of itself; (R6) The conditional mean of $y_t$ depends on the past values of $\alpha^\delta$.

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


