Causality Analysis on Macroeconomic Variables: GDP and Four Key Factors

Hao Chen [1,a], Sanqing Hu [1,b], Jianhai Zhang [1], Wanzeng Kong [1]

[1]College of Computer Science, Hangzhou Dianzi University, Hangzhou, China

[b]sqhu@hdu.edu.cn, [a]1120103097@qq.com

Keywords: Granger Causality, New Causality, GDP, Consumption, Investment, Imports, Exports

Abstract. Causal relationships between different economic variables are of great significance. Granger causality (GC) is one of the most popular methods to explore causal influence in complex systems. It has been widely applied to economic variables. In 2011, Hu et. al pointed out shortcomings and/or limitations of GC by using a series of illustrative examples and showed that GC is only a causality definition in the sense of Granger and does not reflect real causality at all, and meanwhile proposed a new causality (NC) shown to be more reasonable and understandable than GC. It is a common belief that the factors related to one country's economic growth mainly include consumption, investment, imports and exports. In this paper, we select these five economic variables from five countries, America, France, Spain, Australia and China. We then apply GC and NC method to these data and find that i) the causal influence from consumption among all factors to GDP is the largest in all selected countries. ii) our results imply that NC method is more exact to reveal the causal influence between different economic variables than GC method. So, we believe that NC method will replaced GC method and will be widely applied to economics and many other fields.

1. Introduction

Since C.W.J Granger [1] has formalized the causality with a linear regression model of time series, Granger causality (GC) has been widely used to many different areas. The basic idea of GC can be briefly described as follows. If the historical information of time series X significantly improves the prediction accuracy of the future of time series Y in a multivariate autoregressive (MVAR) model, then Granger causality from time series X to Y is identified. People can use Granger causality to analyze direct interaction between multiple variables, thus make further investigation about internal relations between multiple variables. With the advancement of scientific research, many improved methods based on Granger causality appear, such as partial correlation Granger causality and nonlinear Granger causality in time domain and spectral Granger causality in frequency domain.

New causality (NC) method introduced by Hu [2] in 2011, who points out the GC value may not correctly reflect real causal influence between two variables, is shown to be better than GC method. NC is defined as a causality from any time series Y to any time series X in the linear regression model of multivariate time series, which describes the proportion that Y occupies among all contributions to X. NC method is a natural extension of GC and overcome GC's shortcomings and limitation.

Investigating the factors related to economic growth and the causality relationship among them has always been a hot topic studied by many people. However, there is few studies so far analyzing the strength of causal relationship. Maybe GDP reflects the total value of all final goods and services produced in a period of time by a country or region. It is the best index to measure the state of the economy. In China, we usually hear a term that consumption, investment, and net exports are three carriages driving economic growth in economic news. Therefore, it's essential to investigate their different contributions to economic growth. We select corresponding economic variables from developed countries, less developed countries and developing countries such as America, France, Australia, Spain and China, and finally analyze the causal influence to GDP from consumption, investment, imports and exports in these countries.

In this paper, firstly, we describe the definitions of Granger causality and New causality in time-invariant bivariate linear autoregressive models in detail. Secondly, we apply GC and NC to analyze the data sets from five different countries since 1960s. These data sets include the annual data of consumption, investment, imports and exports. We try to detect which factor has the largest causal influence on economic growth. The NC results show that consumption has a strong influence on economic growth, but the results by GC method fail to reveal this conclusion.

2. Causality Methods

Consider two stochastic time series which are assumed to be jointly stationary. Individually, under fairly general conditions, each time series admits an autoregressive representation

$$\begin{cases} X_{1,t} = \sum_{j=1}^{m} a_{11} X_{1,t-j} + \varepsilon_{1,t} \\ X_{2,t} = \sum_{j=1}^{m} a_{22} X_{2,t-j} + \varepsilon_{2,t} \end{cases} \tag{1}$$

and their joint representations are described as

$$\begin{cases} X_{1,t} = \sum_{j=1}^{m} a_{11,j} X_{1,t-j} + \sum_{j=1}^{m} a_{12,j} X_{2,t-j} + \eta_{1,t} \\ X_{2,t} = \sum_{j=1}^{m} a_{21,j} X_{1,t-j} + \sum_{j=1}^{m} a_{22,j} X_{2,t-j} + \eta_{2,t} \end{cases} \tag{2}$$

Where t=0,1, ...,N, the noise terms are uncorrelated over time, ε_i and η_i have zero means and variances of $\sigma^2(\varepsilon_i)$, and $\sigma^2(\eta_i)$, i=1,2. The covariance between η_1 and η_2 is defined by $\sigma_{\eta_1\eta_2}=\mathrm{cov}(\eta_1, \eta_2)$.

2.1.GC In Bivariate Autoregressive Model

Now consider the first equalities in (1) and (2), if $\sigma^2(\eta_1)$ is less than $\sigma^2(\varepsilon_1)$ in some suitable sense X_2 is said to have a causal influence on X_1. In this case, the first equality in (2) is more accurate than in (1) to estimate X_1. Otherwise, if $\sigma^2(\eta_1)=\sigma^2(\varepsilon_1)$, X_2 is said to have no causal influence on X_1. In this case, two equalities are same. Such kind of causal influence, called Granger causality(GC), is defined by

$$F_{X_2 \to X_1} = \ln \frac{\sigma_{\varepsilon_1}^2}{\sigma_{\eta_1}^2} \tag{3}$$

Obviously, $F_{X2\to X1}=0$ when there is no causal influence from X_2 to X_1 and $F_{X2\to X1}>0$ when there is. Similarly, the causal influence from X_1 to X_2 is defined by

$$F_{X_1 \to X_2} = \ln \frac{\sigma_{\varepsilon_2}^2}{\sigma_{\eta_2}^2} \tag{4}$$

2.2. NC In Bivariate Autoregressive Model

Based on (2), we can see contributions to $X_{1,t}$, which include $\sum_{j=1}^{m} a_{11,j} X_{1,t-j}$,$\sum_{j=1}^{m} a_{12,j} X_{2,t-j}$ and the noise term $\eta_{1,t}$ where the influence from $\sum_{j=1}^{m} a_{11,j} X_{1,t-j}$ is causality from X_1's own past values. Each contribution plays an important role in determining $X_{1,t}$. If $\sum_{j=1}^{m} a_{12,j} X_{2,t-j}$ occupies a larger portion among all those contributions, then X_2 has stronger causality on X_1, or vice versa. Thus, a good definition for causality from X_2 to X_1 in time domain should be able to describe what proportion X_2 occupies among all these contributions. So based on this general guideline new causality from X_2 to X_1 is defined as

$$n_{X_2 \to X_1} = \frac{\sum_{t=m}^{N} \left(\sum_{j=1}^{m} a_{12,j} X_{2,t-j}\right)^2}{\sum_{h=1}^{2} \sum_{t=m}^{N} \left(\sum_{j=1}^{m} a_{1h,j} X_{h,t-j}\right)^2 + \sum_{t=m}^{m} \eta_{1,t}^2} \tag{5}$$

Similarly, NC in time domain from X1 to X2 is defined by

$$n_{X_1 \to X_2} = \frac{\sum_{t=m}^{N} \left(\sum_{j=1}^{m} a_{21,j} X_{1,t-j}\right)^2}{\sum_{h=1}^{2} \sum_{t=m}^{N} \left(\sum_{j=1}^{m} a_{2h,j} X_{h,t-j}\right)^2 + \sum_{t=m}^{m} \eta_{2,t}^2} \tag{6}$$

The order of linear regression model is determined by AIC[3]~[5].

3. Experimental Methods

In this paper, we focus on four important factors: consumption (X_1), investment (X_2), imports (X_3) and exports (X_4) as well as GDP (Y) together. The five annual time series are represented by $X_{1,t}$, $X_{2,t}$,

$X_{3,t}$, $X_{4,t}$ and Y_t respectively. We choose four countries: America, France, Spain, Australia which are representative from developed countries and one country: China which is representative from developing countries. Chinese data sets are extracted from National Bureau of Statistics of China[6] and the data sets for other countries are from the database of Organization for Economic Co-operation and Development(OECD)[7]. To help to induce stationarity in the variance-covariance matrix, all data are preprocessed, that is, all variables are transformed into natural logs prior to analysis. Then each variable subtracts it's own mean. For the processed data, we then investigate the causal relationships from consumption, investment, exports, imports to GDP.

4. Simulation Results

For the American annual macroeconomic data sets from 1955 to 2013 which include GDP(Y), consumption volume (X_1), investment volume (X_2), exports volume (X_3) and imports volume (X_4), to calculate causality from X_i to Y, we need to estimate bivariate joint regressive model (2) for NC method as well as autoregressive model (1) for GC method according to the least square method and AIC criteria to determine the coefficients and the optimal order, i=1,2,3,4. After estimating the models we apply GC method to obtain , $F_{X1\to Y}$, $F_{X2\to Y}$, $F_{X3\to Y}$, $F_{X4\to Y}$ and apply NC method to obtain $n_{X1\to Y}$, $n_{X2\to Y}$, $n_{X3\to Y}$, $n_{X4\to Y}$. Similarly, we do so for the other remaining four countries where the data sets in French and Spanish are from 1970 to 2013, the data set in Australian is from 1960 to 2013, and the data set in China is from 1952 to 2013. The results are reported in Table 1 and Table 2.

Table 1.*Granger causality in bivariate model* ***Table 2.****New causality in bivariate model*

GC	Object1	Object2	Object3	Object4	Object5	NC	Object1	Object2	Object3	Object4	Object5
X1→Y	0.3828	0.0775	0.1445	0.0100	0.5410	X1→Y	0.0316	0.0892	0.07750	0.0448	0.00460
X2→Y	0.1752	0.4008	0.5548	-0.0011	0.8441	X2→Y	0.0165	0.0034	0.00025	0.0120	0.00008
X3→Y	0.9337	0.1375	0.1127	0.6867	0.2685	X3→Y	0.0003	0.0015	0.00037	0.0023	0.00002
X4→Y	0.0706	0.1763	0.0031	-0.0067	0.0577	X4→Y	0.0001	0.0005	0.00017	0.0013	0.00002

Where Object1~5 are American, France, Spain, Australia, and China respectively.

From Table 2 one can see that consumption has the largest causal influence on GDP among all five countries compared to the other three factors: investment, imports and exports. But from Table1 1 one can see that there is not any country whose consumption has the largest causal influence on GDP among four factors. So, the conclusions are totally different for both methods. Now a question is arising: which conclusion is true? To provide more evidence to show NC results are more true than GC results, next we will use move window technology to study causal influence from each of four factors to GDP to more clearly see the detailed changing process of causal influence as year increases.

We select 42 as the length of the window for America, Australia and China and select 33 as the length of the window for France and Spain. Take America as an example, the first window covers 1955-1996, then next window moves a year from left to right and so on. In each window we obtain one GC or NC value. The results of GC and NC methods are shown in Fig. 1 and Fig. 2 respectively.

Comparing Fig. 1 with Fig. 2, one can see that i) GC values from X_1, X_2, X_3, X_4 to Y are all much more fluctuant and less stable than NC values. Since two neighbor windows only have two points (two years) are different, true causality in the two neighbor windows should not change too much. So, the much fluctuant GC values should not be real; ii) NC results show that the causality from X_1 to Y in each country is greater than that from X_2, X_3, X_4 to Y in all moving windows, that is, X_1 has major causal influence on GDP for all five countries. This conclusion is consistent with NC results in the whole sample period; iii) Although X_1 has obvious major causal influence to Y in all five countries, in recent years the causal influence from X_1 to Y is greatly reduced in China because the investment in China is greatly increased and weakens the causal influence from consumption to GDP. This shows NC method is better than GC method to reveal the true causality.

By recalling the economic development process of these countries since 1960s, we know western countries realize industrialization very early and per capital GDP is generally high. So people living in these countries have very strong consumption awareness so that consumption has major causal influence on GDP. Therefore, the results obtained by NC method by using the whole available years data and moving window technology are all in line with the reality. However, GC results got by using the whole years data and moving window technology cannot show the phenomenon that consumption has major causal influence on GDP. So, NC method can better reveal the true causality than GC.

5. Conclusions

In bivariate regressive models, our NC results in the whole analyzed years demonstrated that the

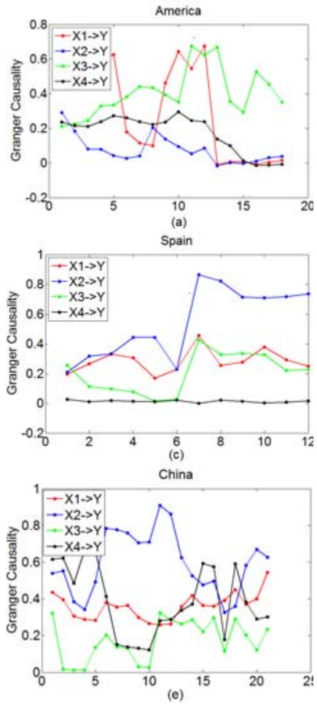

Figure 1. Granger causality from X_1 (consumption), X_2 (investment), X_3 (imports), X_4 (exports) to Y (GDP): (a) America. (b) France. (c) Spain. (d) Australia. (e) China.

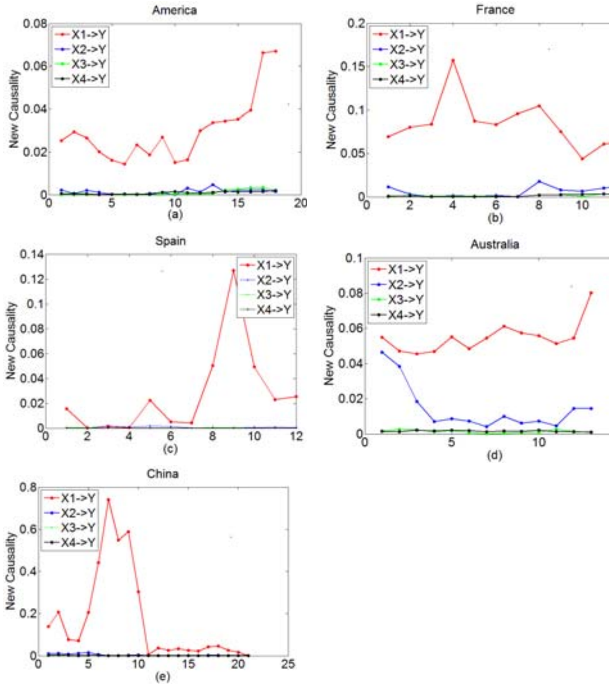

Figure 2. New causality from X_1 (consumption), X_2 (investment), X_3 (imports), X_4 (exports) to Y (GDP): (a) America. (b) France. (c) Spain. (d) Australia. (e) China.

causal influence from consumption among all factors to GDP is the largest and most obvious in all five analyzed countries. This may be the underlying common inherent economic law at least for these five countries. But based on those GC results there is not any country whose consumption has the largest causal influence on GDP among four factors. Our NC results by moving window technology further demonstrated that the causal influence from consumption to GDP among all factors is the largest in each moving window in all five analyzed countries. On the contrary, GC results by moving window technology are much fluctuant, less stable, and cannot draw such a conclusion at all.

Acknowledgements

This work was funded by National Natural Science Foundation of China under Grant No.61473110, Natural Science Foundation of Zhejiang Province, China, under Grant No. Z13F030002, International Science & Technology Cooperation Program of China, Grant No. 2014DFG12570, Key Lab of Complex Systems Modeling & Simulation , Ministry of Education, China.

References

[1] C.W.J.Ganger,“investigating causal relations by econometric models and cross-spectral ethods”,Econometrica,vol. 37, no. 3, pp.424-438 (1969).

[2] S. Hu, G. Dai, G. Worrell, Q. Dai, and H. Liang, “Causality analysis of neural connectivity: Critical examination of existing methods andadvances of new methods”, IEEE Trans Neural Networks, vol. 22, no.6, pp. 829–844 (2011).

[3] H. Akaike, “A new look at the statisticalmodel identification” ,IEEETransactions on Automatic Control, vol. 19, no. 6, pp.716–723 (1974).

[4] A. Seghouane and S. Amari,“The AIC criterion and symmetrizing theKullback Leibler divergence”, IEEE Trans. on Neural Networks, vol.18, no.1, pp.97–106, 2007.[11] K.Evan, C. Snigdhansu, and R.G. Auroop, “Exploring Granger causality between global average observed timeseries of carbondioxide and temperature”, Theor.Appl.Climatol, pp.325-335, (2011).

[5] A. Seghouane, “Model selection criteria for image restoration”, IEEETrans.on Neural Networks, vol.20, no.8,pp.1357–1363,2009.[13] J.B.Elasner, "Granger causality and Atlantic hurricanes”,Tellus-Series A:Dynamic Meteorology and Oceanography, pp.476-485, (2007).

[6] http://data.stats.gov.cn/.

[7] http://data.oecd.org/.