

The relation between news events and stock price jump: an analysis based on neural network

Wanbin (Walter) Wang^a, Kin-Yip Ho^a, Wai-Man (Raymond) Liu^a, Kun (Tracy) Wang^b

^a*Research School of Finance, Actuarial Studies and Applied Statistics
The Australian National University
Canberra, Australia 0200*

^b*Research School of Accounting and Business Information Systems
The Australian National University
Canberra, Australia 0200
Email: walter.wang@anu.edu.au*

Abstract: The efficient-market hypothesis states that price movements are extremely efficient in reflecting information flows. Some studies have shown that stock prices are related to news events such as earnings announcements, political events and corporate takeovers, while others have failed to find convincing evidence to relate price changes to news. The aim of this study is to explore the relation between news and abnormal financial market volatility.

We first investigate the Granger causality between news and stock returns. Our results show that stock price change is the Granger cause of news volume and news sentiment; news volume is not the Granger cause of stock price change while news sentiment is the Granger cause of stock price change.

Moreover, we utilize an artificial neural network model to predict stock market collapses by using different volatility parameters. The findings from this study will further our understanding of stock price movements and the reasons for stock market collapses.

Keywords: *Stock price jump, News sentiment, Granger causality, Neural network*

1. INTRODUCTION

The efficient-market hypothesis (Fama, 1970) states that securities markets are extremely efficient in reflecting information about individual stocks and the stock market as a whole. Some studies have shown that stock prices jumps are related to unexpected extreme news (Asgharian *et al.*, 2011), prescheduled earnings announcements (Francis *et al.*, 1992, Skinner, 1994), political events (Kim and Mei, 2001), or corporate takeovers (Pound and Zeckhauser, 1990). However others have failed to find convincing evidence to relate price changes to news (Joulin *et al.*, 2008).

Stock market collapses can cause significant loss of investor wealth, and deeply affect listed corporations and the whole economy. Some researchers argue that the reason of stock market collapse is a significant portion of investors revising their expectations on future cash flows (Chen *et al.*, 2013). Zouaoui *et al.* (2011) examine the influence of investor sentiment on the probability of stock market crises. They find that investor sentiment increases the probability of occurrence of stock market crises within a one-year horizon. It seems natural accordingly to conclude that the information of investors such as news events affect their sentiment.

The objectives of this study are to investigate the ability of news events to predict stock price jumps and the occurrence of stock market collapses. To achieve these objectives, we first use causality tests to find factors affecting stock price movements, and then we build an artificial neural network to predict stock price jumps using these factors.

The remainder of this paper proceeds as follows. In the second section, we discuss the Granger causality test model and report the results of causality tests for news and stock returns. Causality tests can provide useful information on whether knowledge of company news improves short-run forecasts of current and future movements in stock prices, and vice versa. In the third section, we discuss our artificial neural networks model and report the prediction ability of our model. The last section concludes this study and discusses potential future research directions.

2. THE GRANGER GAUSALITY TEST FOR NEWS AND STOCK RETURNS

2.1. Granger Causality

Granger (1969) defines the causality for two scalar-valued, stationary, and ergodic time series $\{X_t\}$ and $\{Y_t\}$ using a simple model:

$$X_t = \sum_{j=1}^m a_j X_{t-j} + \sum_{j=1}^m b_j Y_{t-j} + \varepsilon_t$$

$$Y_t = \sum_{j=1}^m c_j X_{t-j} + \sum_{j=1}^m d_j Y_{t-j} + \eta_t$$

Where ε_t, η_t are two uncorrelated white-noise series.

If some b_j is not zero, the knowledge of past Y values helps to predict current and future X values, and Y is said to Granger cause X. Similarly, X causes Y if some c_j is not zero. Linear least squares predictors are used in implementing this test.

The Granger causality test has been used to find factors causing stock price changes. Hiemstra and Jones (1994) show unidirectional Granger causality from stock returns and percentage volume changes. Ray (2012) report bi-directional causality exist between stock price and foreign exchange reserve, stock price and money supply, stock price and crude oil price, and stock price and whole price index.

2.2. Data and Methodology

We use daily closing prices (2004-2012) of the Dow Jones Price Index. We compute stock returns as 100 times the first difference of the natural logarithm of the daily stock price, i.e. $100 * \ln(P_t/P_{t-1})$ and obtain stock price return series $\{Stock_return_t\}$. The news data used in this study are provided by RavenPack Inc., a leading provider of news analytic data. The data in RavenPack database are derived from all news articles and press releases that appeared in the Dow Jones newswire. The database contains a unique observation for every article and includes the date and time each news article was released, a unique firm identifier, and several variables that quantify the content and form of the article. For example, the "relevance" score ranging from 0 to 100 indicates how strongly an entity is related to the underlying news story, and a score of 100

indicates the article is highly relevant. For a news story with a relevance score of 100, the "ENS-Event Novelty Score" represents how novel this news story is. The first story reporting a categorized event receives a novelty score of 100. The novelty scores of subsequent stories about the same event following a decay function (i.e. 100 75 56 42 32 24 18 13 10 8 6 4 3 2 2 1 1 1 1 0 ...). The "ESS – Event Sentiment Score" represents the news sentiment for a given entity, ranging from 0 to 100, where 0 indicates extremely negative news, 50 indicates neutral news, and 100 indicates extremely positive news.

During 2004-2012, there are 20,354,107 news articles in RavenPack database. Among them, there are 856,071 (4.21%) news articles whose relevance score is 100, and 342,098 (1.68%) news articles whose event novelty score is 100. Among news articles with novelty score of 100, the number of positive news articles, negative news articles and neutral news articles are 151,309 (44.23%), 150,458 (43.98%) and 40,331(11.79%) respectively.

In this study, we explore the relation between stock return with news volume and the relation between stock returns and news sentiment. We calculate the number of news in a trading day and obtain time series $\{News_number_t\}$. The news sentiment series $\{News_ESS_t\}$ are calculated as the $\sum(ESS - 50)$ in a trading day, i.e. the sum of the event sentiment score minus 50. We examine whether there exists any causal linkage between stock prices and news by conducting the Granger causality test.

2.3. Empirical Results

We first conduct Dickey-Fuller tests (Dickey and Fuller, 1979) to examine whether there a unit root present in the three time series $\{Stock_return_t\}$, $\{News_number_t\}$ and $\{News_ESS_t\}$. The results indicate that the null hypothesis of a unit root is rejected for all three series.

Table 1 reports the results of our Granger causality tests for stock price returns, news volume, and news sentiments. It shows that at 5% significance level, stock price change is the Granger cause of news volume and news sentiment. News volume is not the Granger causality of stock price change while news sentiment is the Granger cause of stock price change.

Table 1. The results of Granger causality tests

Null hypothesis	Significance level	Results
Stock return does not Granger cause news volume	0.05	Reject
	0.01	Accept
Stock return does not Granger cause news sentiment	0.05	Reject
	0.01	Accept
News volume does not Granger cause stock return	0.05	Accept
	0.01	Accept
News sentiment does not Granger cause stock return	0.05	Reject
	0.01	Reject

3. THE ANN MODEL TO PREDICT STOCK PRICE JUMPS

3.1. Artificial Neural Network

The Artificial Neural networks (ANN) mimic the human brain and nervous system. Neural networks have been used to solve complicated practical problems in various fields, such as character recognition, machine faults diagnosis, etc. A neural network consists of a set of fundamental processing elements (called neurons) and processes information using a connectionist approach to computation. Most neural networks contain three types of layers: input, hidden, and output (as shown in Figure 1). Each neuron in a hidden layer receives the input data attributes x_m from each of the neurons in an input layer and the attributes are added through applied weights w_m and converted to an output value by an activation function (u). Then, the output is passed to neurons in the next layer, providing a feed forward path to the output layer (z).

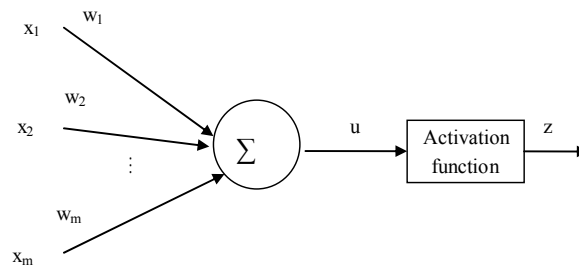


Figure 1. The structure of a neuron with its summation node.

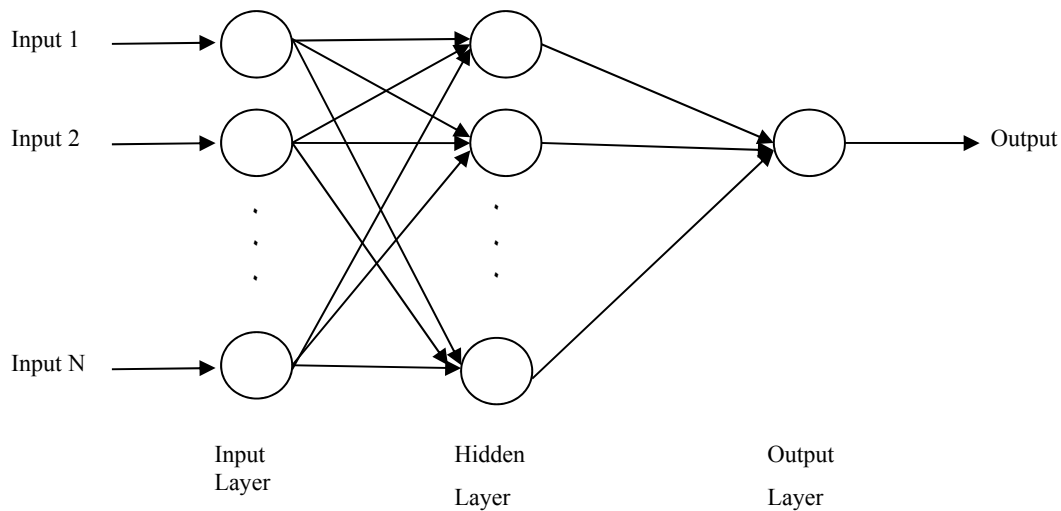


Figure 2. The structure of a neuron network.

In general, the activation function is a nonlinear function. Activation functions that are commonly used include the threshold function, the piecewise-linear function and the sigmoid function. A neural network is an adaptive system that changes its weights based on external or internal information that flows through the network during the learning phase. These learning rules include supervised learning, unsupervised learning, reinforcement learning. In supervised learning the network is trained by being provided with input and matching output patterns.

In this study, the PNN (probabilistic neural network) is implemented using the MATLAB Neural Network Toolbox, where the network structures were set to the default. A probabilistic neural network is based on the classical Bayesian classifier, whose goal is to statistically minimize the risk of misclassifications. PNN creates a two-layer network. The first layer has radial basis network neurons, and calculates its weighted inputs by vector distance between its weight vector and the input vector, multiplied by the bias. The second layer has competitive transfer function neurons, and calculates its weighted input with dot product weight function and its network inputs with the sum of network inputs.

3.2. Data and Methodology

In this study we define stock price "jumps" as stock price change bigger than 2% in one trading day, i.e., $|(P_t - P_{t-1})/P_{t-1}| > 2\%$. We use whether stock prices "jump" or not as our training patterns. Figure 3 shows our neural network classification framework for the prediction of stock price jumps. The neural network model is trained on the training data and subsequently tested to measure its performance on the testing data. In this study, we employ 80% of the data from the data set as the training set; the remaining 20% as the testing set. Given these preparations, this study uses stock price returns in the last three trading days (i.e., $Stock_return_{t-3}, Stock_return_{t-2}, Stock_return_{t-1}$) and news sentiments in the last trading days

(i.e., $News_ESS_{t-1}$) as input features in the probabilistic neural network. Totally this approach comes to 4 indices. So there will be 4 input nodes and one output node. We do not normalize the data because neural networks are able to recognize the high-level feature.

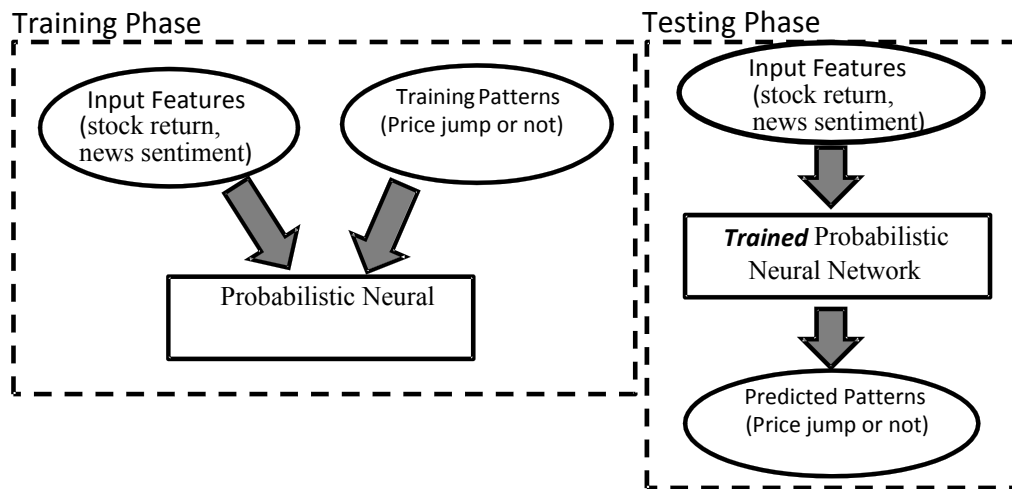


Figure 3. The probabilistic neural network classification framework for the prediction of stock price jumps

3.3. Empirical Results

For the application of binary classification, sensitivity and specificity are used to measure the performance. In this study, we define the pattern as: "price jumps" and "price not jumps". Then we calculate

$$Sensitivity = \frac{Number\ of\ true\ "price\ jumps"}{Number\ of\ true\ "price\ jumps" + Number\ of\ false\ price\ jumps}$$

$$Specificity = \frac{Number\ of\ true\ "price\ not\ jumps"}{Number\ of\ true\ "price\ not\ jumps" + Number\ of\ false\ "price\ not\ jumps"}$$

$$Prediction\ rate = \frac{Number\ of\ true\ "price\ not\ jumps" + Number\ of\ true\ "price\ not\ jumps"}{Number\ of\ prediction\ days}$$

Table 2 shows the performance characteristics of our PNN prediction model. The specificity is good but the sensitivity is not good enough to predict stock price jump.

Table 2. Performance Characteristics of the PNN prediction model

	Sensitivity	Specificity	Prediction rate
PNN	37.04%	89.25%	84.66%

4. CONCLUSION AND FUTURE WORKS

This study applies Granger causality tests and the PNN forecast model to investigate the predictability of stock price jumps. The Granger causality tests show that news sentiment is the Granger cause of stock price change, and stock price change is the Granger cause of news volume and news sentiment. The results of our PNN forecast model is promising.

Estimating stock market volatility has received considerable attention by both academics and practitioners. There is much work to improve the prediction accuracy of our model, especially the sensitivity of our predictions. One possible direction is the use of high frequency data of stock price and news in the forecast model.

REFERENCES

- Asgharian, H., Holmfeldt, M. & Larson, M. 2011. An event study of price movements following realized jumps. *Quantitative Finance*, 11, 933-946.
- Chen, L., Da, Z. & Zhao, X. 2013. What Drives Stock Price Movements? *Review of Financial Studies*, 26, 841-876.
- Dickey, D. A. & Fuller, W. A. 1979. Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American statistical association*, 74, 427-431.
- Fama, E. F. 1970. Efficient capital markets: A review of theory and empirical work. *Journal of Finance*, 25, 383-417.
- Francis, J., Pagach, D. & Stephan, J. 1992. The stock market response to earnings announcements released during trading versus nontrading periods. *Journal of Accounting Research*, 30, 165-184.
- Granger, C. W. 1969. Investigating causal relations by econometric models and cross-spectral methods. *Econometrica: Journal of the Econometric Society*, 424-438.
- Hiemstra, C. & Jones, J. D. 1994. Testing for Linear and Nonlinear Granger Causality in the Stock Price - Volume Relation. *The Journal of Finance*, 49, 1639-1664.
- Joulin, A., Lefevre, A., Grunberg, D. & Bouchaud, J. P. 2008. Stock price jumps: news and volume play a minor role. *Arxiv Preprint Arxiv:0803.1769*.
- Kim, H. & Mei, J. 2001. What makes the stock market jump? An analysis of political risk on Hong Kong stock returns. *Journal of International Money and Finance*, 20, 1003-1016.
- Pound, J. & Zeckhauser, R. 1990. Clearly heard on the street: The effect of takeover rumors on stock prices. *Journal of Business*, 63, 291-308.
- Ray, S. 2012. Testing Granger Causal Relationship between Macroeconomic Variables and Stock Price Behaviour: Evidence from India. *Advances in Applied Economics and Finance*, 3, 470-481.
- Skinner, D. 1994. Why firms voluntarily disclose bad news. *Journal of Accounting Research*, 32, 38-60.
- Zouaoui, M., Nouyrigat, G. & Beer, F. 2011. How Does Investor Sentiment Affect Stock Market Crises? Evidence from Panel Data. *Financial Review*, 46, 723-747.