Expert Systems with Applications 36 (2009) 4951–4957

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

An early warning system for global institutional investors at emerging stock markets based on machine learning forecasting

Il Suh Son [a], Kyong Joo Oh [a,*], Tae Yoon Kim [b], Dong Ha Kim [a]

[a] Yonsei University, Department of Information and Industrial Engineering, 134 Shinchon-Dong, Seodaemun-Gu, Seoul 120-749, Republic of Korea
[b] Department of Statistics, Keimyung University, Daegu 704-701, Republic of Korea

ARTICLE INFO

Keywords:
Early warning systems
Global institutional investors
Machine learning
Case-based reasoning

ABSTRACT

At local emerging stock markets such as Korea, Hong Kong, Singapore and Taiwan, global institutional investors (GII) comprised of global mutual funds, offshore funds, and hedge funds play a key role and more often than not cause severe turmoil via massive selling. Thus, for the concerned local governments or private and institutional investors, it is quite necessary to monitor the behavior of GII against a sudden pullout. The main aim of this article is to propose an early warning system (EWS) which purposes issuing warning signal against the possible massive selling of GII at the local market. For this, we introduce machine learning algorithm which forecasts the behavior of GII by predicting future conditions. Technically, this EWS is an advanced form of the EWS developed by Oh et al. [Oh, K. J., Kim, T. Y., & Kim, C. (2006). An early warning system for detection of financial crisis using financial market volatility. *Expert Systems, 23*, 83–98] which issues a warning based on classifying present conditions. This study is empirically done for the Korean stock market.

1. Introduction

Over the last 10 years, emerging stock markets such as Korea, Hong Kong, Singapore and Taiwan have been incorporated into the world financial market (Ghysels & Seon, 2005). Globalization and removal of the regulations at the local markets make global institutional investors (GII) to be major influences at the local markets. As a result, the movements of GII actually direct the local markets particularly when severe external or internal shocks hit the market. For example, GII, by May 2004, has occupied almost half the Korean stocks in terms of total market capitalization and in cases of external or internal shock such as Asian financial crisis in 1997, Russia moratorium in 1998, liquidity crises of Daewoo group in June–December 1999 and Hyundai group in June–December 1999, and 9/11 terrorists attack in 2001, the abnormal pullouts of GII has led the Korean stock market to the near collapse (Choe, Kho, & Stulz, 1999; Kim & Kwon, 2003; Kim & Wei, 1999). In order to prepare against such a devastating situation, a proper early warning system (EWS) that can detect or predict abnormal pullout of GII, especially global hedge funds, is strongly desired at the local markets.

In order to develop EWS for GII (EWSGII), we will follow a standard procedure developed by Oh, Kim, and Kim (2006) which employs machine learning algorithms for establishing EWS for financial markets. Technically, their procedure is based on classifying the current situation according to some standard measures such as market volatility. The core of the procedure is defining the gray zone as its main feature vector which may proceed to either stable condition or collapse and issuing a warning when the market enters the gray zone. In this article, we present an advanced EWS which issues warning by classifying and forecasting future conditions of a market. The core of the procedure is defining the oracle rule which determines the future market conditions in advance and then tracing (forecasting equivalently) the oracle rule with trained machine learning algorithm.

The rest of this study consists of as follows: Section 2 reviews the previous studies about behavior of GII and the existing EWS algorithm. A discussion of technical aspects of the EWS monitoring GII (EWSGII) is also given. Section 3 presents the detailed construction procedure of EWSGII. Section 4 is devoted to the empirical case study for building EWSGII for the Korean stock market. Concluding remarks are given in Section 5.

2. Background review for EWSGII

GII is becoming more influential at the Asian local markets. Stock holdings by GII have reached 18.8% in Japan, 20% in Taiwan, and 43.5% in Korea. Especially in the Korean stock market exchange (KSE), GII have increased their stock holdings from 11.9% in 1995 to 43.5% in 2004, which means KSE is extremely overweighed by GII. Therefore, there is always a calculated risk at the local market that GII may destabilize the market easily by dumping their stock holdings massively when severe shocks hit the market.

* Corresponding author. Tel.: +82 2 2123 5720; fax: +82 2 364 7807.
E-mail address: johanoh@yonsei.ac.kr (K.J. Oh).

0957-4174/$ - see front matter © 2008 Elsevier Ltd. All rights reserved.
doi:10.1016/j.eswa.2008.06.044

GII are well known for their herding behavior and positive feedback trading (Nofsinger & Sias, 1999). Several studies have explained this by way of examples during the 1997 financial crisis in Asia. Kim and Wei (1999) claim that the 1997 crisis should be attributed to emotional panic reaction and herding behavior of GII rather than economic vulnerability of those countries. Radelet and Sachs (1998), Kim and Wei (1999), Choe et al. (1999) noticed that the sudden pullout of GII accelerated the crisis by causing a financial chaos and policy failure during the 1997 financial crisis in the Korean financial market. As a recent study, Ghysels and Seon (2005) investigated the effect of derivative investment of GII such as index futures of the Korean stock market during the 1997 financial crisis. They found that net long (short) position of index futures by GII induces subsequent net buy (sell) of equities by domestic investor during the crisis. In other words, futures trading by global investor strengthened herding tendency and accelerated destabilizing the Korean stock market.

There are various EWS's which are designed to monitor economic or financial condition, e.g., Edison (2000), Eichengreen, Rose, and Wyplosz (1996), Frankel and Rose (1996), Goldstein (1996), Goldstein, Kaminsky, and Reinhart (2000), Kaminsky, Lizondo, and Reinhart (1998), Kaminsky and Reinhart (1999). Most of them are developed under the hypothesis that crisis is an eventual result of long-term deterioration of economic fundamentals (Eichengreen, Rose, & Wyplosz, 1995; Krugman, 1979; Obstfeld, 1986) and focus on the economic long-term variables. Recently, Kim, Oh, Sohn, and Hwang (2004c), Oh, Kim, Lee, and Lee (2005) and Oh et al. (2006) developed EWS under different hypothesis that crisis may result from short-term financial market instability without a significant deterioration of long-term economic fundamentals (Ozkan & Sutherland, 1995; Velasco, 1987). Thus, they focus on short-term variables such as daily movement of financial market indexes.

For formulating EWS as classification problem, Kim et al. (2004c) and Oh et al. (2006) introduce the gray zone which is characterized by abrupt reversals of market sentiments or sudden change of market volatility. What makes the gray zone uniquely important for their EWS is that it is regarded as a transition period (i.e., having gone through the gray zone, the financial market may proceed to either crisis or stable condition) and the EWS issues warning signal whenever financial market enters the gray zone. For establishing EWSGII, we will basically follow Kim et al. (2004c) and Oh et al. (2006) since EWSGII needs monitoring financial market daily. There are technical advances made for EWSGII in this paper, however, since warning signal is designed to be issued based on forecasting. Note that the main purpose of EWSGII is to prepare against possible massive pullout of GII in the near future.

3. EWSGII algorithm description

Building EWSGII based on forecasting consists of two phases. The first phase is to construct the oracle classifier and the second phase to construct the trained classifier (or EWSGII in a narrow sense). The oracle classifier is needed since trained classifier is designed to predict (or trace) the future market condition predetermined by the oracle classifier. For building the oracle classifier, we assume that GII react to external or internal stimulus over a period of time with some contingency plan, not instantly. At the appearance of Asian financial crisis in 1997, in fact, GII dumped stocks at Asian local markets over half a year. In the meantime building the trained classifier depends on training dataset and an assigned oracle classifier.

3.1. Phase 1: oracle classifier construction

We first obtain the oracle classifier *Of* classifying or defining stable period (SP), transition period (CP or the gray zone by Kim et al. (2004c)) and crisis period (CP) exclusively. For building *Of* we assume CP as the period during which GII sell enormous stocks with a contingency plan and TP as the period at which GII turns from net long position (or buying trend) to net short position (selling trend) to initiate the contingency plan. Assuming these, SP, TP and CP are defined in terms of the quarterly, monthly, weekly, and daily net sale of GII. Let the above four variables $Ox_1, Ox_2, \ldots Ox_4$ be the oracle predictor variables and Oy the oracle response assuming values 1, 2, 3 which corresponds to SP, TP, CP respectively. At the given time t oracle classifier

$$Of: \ OX \rightarrow OY \tag{1}$$

which maps $\mathbf{Ox}_t = (Ox_{1t}, Ox_{2t}, \ldots, Ox_{4t}) \in X$ to its classification label Oy_t.

Note that $Ox_1, Ox_2, \ldots, Ox_4$ are known to reflect selling trend with a contingency plan fairly well. In fact these quantities are being monitored by Financial Supervisory Service in Korea which has its own rule of thumb in using the quantities for detecting the abnormal massive selling of GII. One may refer to the expert's opinion in establishing valid oracle rule. Refer to Section 4 and Table 1.

3.2. Phase 2: lag l classifier (EWSGII) construction

Once the oracle classifier is assigned successfully, one may start to build the lag l classifier (or trained classifier for forecasting) as follows. Suppose that its selected predictor variables are $X_1, X_2, \ldots, X_p$ and lag l response is Y. For positive integer l, the training data set is given by

$$(X_{11}, \ldots, X_{p1}, Y_{1+l}), (X_{12}, \ldots, X_{p2}, Y_{2+l}), \ldots, (X_{1n}, \ldots, X_{pn}, Y_{n+l}) \tag{2}$$

Table 1
Oracle classification rules for CP, TP, and SP

	Classification rule
Of(1)	**If** Quarterly net sale more than 2.4 (or) monthly net sale more than 1.6 (or) weekly net sale more than 0.8 (or) daily net sale more than 0.4, **Then** $OY_t = 3$ **(CP)** **Else If** Quarterly net sale more than 1.2 (or) monthly net sale more than 0.8 (or) weekly net sale more than 0.4 (or) daily net sale more than 0.15, **Then** $OY_t = 2$ **(TP)** **Else** $OY_t = 1$ **(SP)**
Of(2)	**If** Quarterly net sale more than 3.0 (or) monthly net sale more than 2.0 (or) weekly net sale more than 1.0 (or) daily net sale more than 0.5, **Then** $OY_t = 3$ **(CP)** **Else If** Quarterly net sale more than 1.5 (or) monthly net sale more than 1.0 (or) weekly net sale more than 0.5 (or) daily net sale more than 0.2, **Then** $OY_t = 2$ **(TP)** **Else** $OY_t = 1$ **(SP)**
Of(3)	**If** Quarterly net sale more than 4.0 (or) monthly net sale more than 3.0 (or) weekly net sale more than 1.5 (or) daily net sale more than 0.7, **Then** $OY_t = 3$ **(CP)** **Else If** Quarterly net sale more than 2.0 (or) monthly net sale more than 1.5 (or) weekly net sale more than 0.8 (or) daily net sale more than 0.4, **Then** $OY_t = 2$ **(TP)** **Else** $OY_t = 1$ **(SP)**

Unit: 1 billion won.