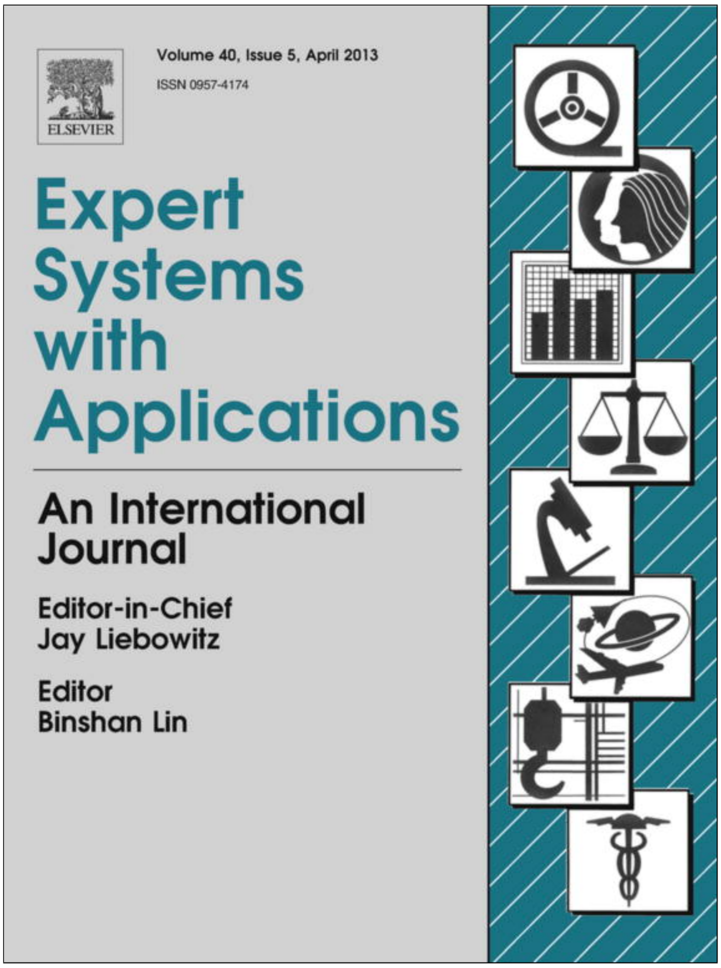
ELSEVIER
Volume 40, Issue 5, April 2013
ISSN 0957-4174
Expert Systems with Applications
An International Journal
Editor-in-Chief
Jay Liebowitz
Editor
Binshan Lin

Expert Systems with Applications 40 (2013) 1542–1554

Contents lists available at SciVerse ScienceDirect

Expert Systems with Applications

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

Data mining investigation of co-movements on the Taiwan and China stock markets for future investment portfolio

Shu-Hsien Liao *, Shan-Yuan Chou

Department of Management Sciences, Tamkang University, No. 151, Yingjuan Road, Danshuei Dist., New Taipei City, 251, Taiwan, ROC

ARTICLE INFO

Keywords:
Cross-national stock market
Stock market investment portfolio
Co-movements
Data mining
Association rules
Cluster analysis

ABSTRACT

On June 29, 2010, Taiwan signed an Economic Cooperation Framework Agreement (ECFA) with China as a major step to open markets between Taiwan and China. Thus, the ECFA will contribute by creating a closer relationship between China and Taiwan through economic and market interactions. Co-movements of the world's national financial market indexes are a popular research topic in the finance literature. Some studies examine the co-movements and the benefits of international financial market portfolio diversification/integration and economic performance. Thus, this study investigates the co-movement in the Taiwan and China (Hong Kong) stock markets under the ECFA using a data mining approach, including association rules and clustering analysis. Thirty categories of stock indexes are implemented as decision variables to observe the behavior of stock index associations during the periods of ECFA implementation. Patterns, rules, and clusters of data mining results are discussed for future stock market investment portfolio.

1. Introduction

With the increasing significance of international flows of goods, services and capital, co-movements of economic variables in different countries is becoming increasingly evident. The extent to which globalization causes domestic economies to move along with economies in the rest of the world or in their particular region, is a major concern for policy-makers. Thus, it is believed by many that regional trade integration and regional free trade agreements (RFTAs) are beneficial to a nation's economy (Edwards & Ginn, 2011; Kearney & Muckley, 2007). In the case of Taiwan and China, due to the nature of the political relationship across the Taiwan Straits, the R.O.C. Taiwan has been excluded from the rising trend of ASEAN economic integration, and thus is facing the risk of marginalization. However, since R.O.C. Taiwan President, Ma Ying-jeou took office in May 2008, cross-strait relationships have experienced their most rapid improvement in decades. On April 26, 2009, at the third round of cross-strait talks in Nanjing, Taipei and Beijing inked three new agreements. Such initiatives are crucial to normalize cross-strait economic ties. More importantly, Taiwan recognizes the need to establish more institutionalized cooperation platforms with its neighbors. On June 29, 2010, Taiwan signed an Economic Cooperation Framework Agreement (ECFA) with China as a major step in this direction. For Taiwan, the ECFA may promote greater market opportunities and the possibility of signing Free Trade Agreements (FATs) with other ASEAN countries. For China, the ECFA may contribute in eliminating economic imbalances caused by the rapid economic progress since China's reform and open-up policy in 1978. Thus, the ECFA will contribute by creating a closer relationship between China and Taiwan through economic and market interactions, thus establishing mutual understanding and trust, the basis of peace across the Straits.

July 29, 2011, the Ministry of Economic Affairs (MOEA) expressed satisfaction in the results of the early harvest program in the Economic Cooperation Framework Agreement (ECFA) with mainland China. Taiwan's gross export value to mainland China amounted to US$61.56 billion over the first six months this year, with tariff exemptions and reductions of US$53.71 million. On the other hand, Taiwan's import value from mainland China, during the same period, grossed US$21.92 billion, saving the Chinese about US$9.43 million on tariffs. It is also noteworthy that the American Chamber of Commerce (AmCham) in Taipei announced that nearly half of Taiwan's 2010 GDP growth came from its trade with China. Taiwan has become reliant on economic and trade relationships with China, which drove 47% of Taiwan's economic growth in 2010. Therefore, there is a co-movement relationship between the Straits not only on trade exchange, but also on industry and financial markets under the ECFA.

On the other hand, co-movements of the world's national financial market indexes are a popular research topic in the finance literature (Meric, Ratner, & Meric, 2008; Chow, Liu, and Niu, 2011; Graham, Kiviaho, & Nikkinen, 2012; Liao, Chu, & You, 2011). Some studies examine the co-movements and the benefits of interna-

* Corresponding author.
E-mail address: michael@mail.tku.edu.tw (S.-H. Liao).

0957-4174/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
http://dx.doi.org/10.1016/j.eswa.2012.08.075

tional financial market portfolio diversification/integration and economic performance (Liljeblom, Löflund, & Krokfors, 1997; Meric et al., 2001; Chu, 2002; Aslanidis, Osborn, & Sensier, 2010; Beine & Candelon, 2010; Madaleno & Pinho, 2012). In addition, the Asian financial crisis has stimulated a great deal of interest in how economic shocks are transmitted across different countries (Arestis, Caporale, Cipollini, & Spagnolo, 2005; Brown, Rhee, & Zhang, 2008; Chang, 2002; Jang & Sul, 2002). The international stock market is one of the most popular forms of investment due to the expectations of high profit. However, higher expected profit also implies higher risk. Thus, numerous studies have proposed different analysis methods to assist investors in analysis and decision-making. In addition, many individual investors, stockbrokers, and financial analysts attempt to predict stock market price activities and their potential development. This mass behavior runs counter to the counsel of the many academic studies, which contend that the prediction of international stock market development is ineffective. This point of view is codified in what is referred to as the efficient markets hypothesis (Fama, 1991; Haugen, 1997). In particular, Forbes and Rigobon (1999) and Rigobon (1999), in their studies of international stock market co-movements, discovered no significant changes in the international transmission mechanism of shocks during the financial crises, but found it puzzling why the degree of co-movement is so high at all times. Thus, the international financial capital market efficiency becomes an interesting issue for research on international financial market co-movements.

In addition, there are three degrees of international financial capital market efficiency. The first degree is the strong form of the efficient market hypothesis, which states that all information that is knowable is immediately factored into the market's price for security. If this is true, then all price predictors are definitely wasting their time, even if they have access to private information. The second degree is the semi-strong form of the efficient market hypothesis, in which all public information is considered to possess private information, which can be used for profit. The third degree is the weak form, which holds only that any information gained from examining a security's past trading history is reflected in price. Indeed, past trading history is public information, implying that the weak form is a specialization of the semi-strong form, which itself is a specialization of the strong form of the efficient market hypothesis. Thus, integration of co-movement and portfolio analysis in financial market, in terms of investment and risk management, has become a critical research issue (Alexakis, Niarchos, Patra, & Poshakwale, 2005; Bohl, Brzeszczyński, & Wilfling, 2009; Boyer & Zheng, 2009; Liao, Chu, & Teng, 2011; Oh & Parwada, 2007).

Due to the different degrees of international financial capital market efficiency, academic researchers investigate the efficient market hypothesis by exploring unknown and valuable knowledge from historical data, using techniques such as data mining. Enke and Thawornwong (2005) introduced an information gaining technique used in machine learning for data mining to evaluate the predictive relationships of numerous financial and economic variables. Neural network models for the estimation and classification of levels are then examined for their ability to provide an effective forecast of future values. Boginski, Butenko, and Pardalos (2006) proposed a network representation of stock market data referred to as a market graph. This graph is constructed by calculating cross correlations between pairs of stocks based on opening price data over a certain period of time. Chun and Park (2005) proposed a learning technique to extract new case vectors using Dynamic Adaptive Ensemble CBR (DAE CBR). The main idea of DAE CBR originates from finding combinations of parameters and updating and applying an optimal CBR model to an application or domain area. These concepts are investigated against the backdrop of a practical application involving the prediction of a stock market index. In addition, Rapach and Wohar (2006) analyzed in-sample and out-of-sample tests of stock return predictability to better understand the nature of the empirical evidence in return predictability. Their study found that certain financial variables display significant in-sample and out-of sample predictive ability with respect to stock returns. Overall, most studies consider stock market analysis as a time series problem, and there have been few studies using stock market efficiency to explore the possible cause-and-effect relationships among different stock categories or the influence of outside factors (Liao, Ho, & Lin, 2008).

Thus, this study investigates the co-movement in the Taiwan and China (Hong Kong) stock markets under the ECFA using a data mining approach, including association rules and cluster analysis. Specifically, this study investigates the following research issues: (1) the study of the relationships among Taiwan, China and Hong Kong stock market indexes by association rules to find a similar trend in transaction data, and also to identify any co-movement of market performance; (2) the use association rules to understand the co-movement between stock market indexes and their categorical stock indexes in Taiwan, China and Hong Kong stock market; (3) according to the findings, this study puts forward recommendations for investment portfolios and management as a follow-up reference. The rest of this study is organized as follows. In Section 2, we present the background of the Taiwan and China (Hong Kong) stock markets. Section 3 describes the methodology, including the research framework, data sources, and database design. Section 4 presents the data mining approach, association rules, Cluster analysis (K-means), and data mining tool – SPSS Clementine - and discusses research findings. Section 5 illustrates the data mining results. Finally, Section 6 presents a brief conclusion and discussion.

2. Taiwan and China (Hong Kong) stock market

2.1. Taiwan stock market

The TSEC, Taiwan Stock Exchange Corporation, maintains stock price indices to allow investors to conveniently grasp both overall market movement and the performances of different industrial sectors. These indices may be grouped into market value indices and price average indices. The former are similar to the Standard and Poor's Index, which is weighted by the number of outstanding shares, and the latter are similar to the Dow Jones Industrial Average and the Nikkei Stock Average. The Taiwan Stock Exchange Capitalization Weighted Stock Index ("TAIEX") is the most widely quoted of all TSEC indices. The base year value as of 1966 was set at 100. TAIEX is adjusted in the event of new listings, de-listings and new share offerings to offset the influence on TAIEX owing to non-trading activities. TAIEX covers all listed stocks, excluding preferred stocks, full-delivery stocks and newly listed stocks that have been listed for less than one calendar month. The other market value indices are calculated and adjusted similarly to that of the TAIEX, but with different groupings of stocks included for calculation. Out of the TAIEX Component Stocks, the non-Finance Sub-Index, Non-Electronics Sub-Index, and Non-Finance Non-Electronics Sub-Index include stocks not in the financial sector, not in the electronics sector, and not in either sector. Similarly, the Industrial Sub-Indices are calculated for different industrial sectors. In 1986, eight Industrial Sub-Indices were introduced, i.e., Cement/Glass/Ceramics, Textiles, Foods, Plastics/Chemicals/Rubber, Electric Machinery/Electric Appliance/Cable/Electronics, Paper/Pulp, Construction, and Finance. In 1995, the TSEC introduced an additional 14 Industrial Sub-Indices, i.e. Cement, Plastics, Electric machinery, Electric appliance/cable, Automobile, Chemicals, Glass/ceramics, Iron/steel, Rubber, Electronics, Transportation, Tourism, Retail and Others. This expansion was intended to give a broader perspective of industrial performance and a more comprehensive comparison with overall market trends. Total Return Indices add back cash dividends to the index calculations, and are published at the end of each trading day. This expansion can serve as a better indicator to measure the performance of funds.

In addition, the Industrial Price Average Index and the Composite Price Average Index contain 20 and 30 issues, respectively. The samples are chosen based on their representation in the market as a whole and are adjusted every year by considering profitability, operational efficiency and trading liquidity of the shares, so that the indices can mirror the market trends. All of the TSEC indices (excluding Total Return Indices) are constantly computed and broadcast every minute during the trading hours through the TSEC MIS system and information vendors' networks. This information can be easily accessed on the systems of local and international information vendors, such as Reuters, Bridge, Quick, Bloomberg, Primark, etc.

2.2. China stock market – Shanghai stock market

The economic reforms of China in the late 1980s and the emphasis on its socialist market economy resulted in the revival of the Chinese stock market in Shanghai. At present the Shanghai Stock Exchange (SSE) is the largest stock market in Mainland China. Currently administrated by the China Securities Regulatory Commission (CSRC), it is registered as a non-profit organization. Stocks, bonds and funds are the three main categories of securities traded in the Chinese stock market. Treasury bonds (T-bond), corporate bonds and convertible corporate bonds are the main bonds traded and the T-bond is the most active bond in China. Two types of shares are issued in the *Chinese stock market,* A-shares and B-shares. A-shares are priced in the local Renminbi (RMB) Yuan currency and the B-shares are quoted in U.S dollars. Until recently possession of A-shares was allowed only for domestic investors and B-shares were available to both domestic and foreign investors. With the implementation of more reforms in the stock market sector, foreign investors are now allowed to trade in A-shares as well under certain conditions specified in the Qualified Foreign Institutional Investor (QFII) system. Plans are also being formulated to merge the two types of shares. The Chinese stock market conducts business every week from Monday to Friday, and is closed on Saturday and Sunday and other Chinese public holidays.

The largest group of stock listings in the Chinese Stock Market is the A-shares. The Shanghai Stock Exchange index is the most commonly used index, and it reflects the market performance of the listed shares on any given day. The most commonly listed A-shares and B-shares normally determine the rise and fall of the stock market index. A-shares are priced in the local renminbi yuan currency, while B-shares are quoted in U.S. dollars. Initially, trading in A-shares was restricted to domestic investors only while B-shares were available to both domestic (since 2001) and foreign investors. However, after reforms were implemented in December 2002, foreign investors are now allowed (with limitations) to trade in A-shares under the Qualified Foreign Institutional Investor (QFII) program, which was officially launched in 2003.

The SSE Composite (also known as the Shanghai Composite) Index is the most commonly used indicator to reflect the SSE's market performance, and the SSE Composite Index is composed of all listed stocks (A shares and B shares) on the Shanghai Stock Exchange. The Base Day for the SSE Composite Index is December 19, 1990, and the Base Period is the total market capitalization of all stocks on that day. Other important indexes used in the Shanghai Stock Exchange include the SSE 50 Index and SSE 180 Index.

2.3. China stock market – Shenzhen stock market

However Mainland China has another much smaller stock exchange, called the Shenzhen Stock Exchange, which is located in the city of Shenzhen. With the merger of Hong Kong with Mainland China, the Hong Kong Stock Exchange, situated in the special administrative region of Hong Kong and with its own separate history, is the largest stock exchange in China presently. The Shenzhen Stock Exchange (SZSE) is one of the People's Republic of China's two stock exchanges in addition to the Shanghai Stock Exchange. The market capitalization of its listed companies was about US $1.3 trillion in 2010.

While the Hong Kong Stock Exchange handles larger companies, the Shenzhen stock market is composed mostly of small companies. State owned enterprises (SOEs) make up most of the company listings in the Shenzhen stock market, which means that the Chinese government has a controlling interest over these SOEs. The Shenzhen stock market makes use of numerous indexes to track the market. The blue-chip composite index was launched in 1995. Another popular index on the market is the Shenzhen Stock Exchange 100 Index.

2.4. Hong Kong stock market

The Hong Kong stock market began in the late 19th century with the first formal establishment in 1891, though informal securities exchanges had taken place since 1861. Since then, this stock exchange has been the main exchange for Hong Kong despite co-existing with other exchanges at different times. After a series of complex mergers and acquisitions, HKSE remains the core. The Hong Kong Stock Exchange (HKEX; SEHK: 0388) is a stock exchange located in Hong Kong.

The Hong Kong Stock Exchange (HKEX) is the primary stock market of Hong Kong. It is one of the largest stock markets in China, Japan and the outlying Asian economic ecosystem. In an increasingly global economic atmosphere, the fluctuating prices of the Hong Kong Stock Exchange can influence both eastern and western stock markets. Though the Hong Kong Stock Exchange operates under the basic principles that also govern stock markets such as the United State's NASDAQ, the Hong Kong stock market differs from western stock markets in several considerable ways. It is Asia's third largest stock exchange in terms of market capitalization, behind the Tokyo Stock Exchange and the Shanghai Stock Exchange, and is the fifth largest in the world. As of 31 Dec 2010, the Hong Kong Stock Exchange had 1413 listed companies with a combined market capitalization of $2.7 trillion. Hong Kong Exchanges and Clearing is the holding company for the exchange.

The Hong Kong Stock Market wields significant financial clout in the Asian world, mostly due to its large size and the multinational corporations that trade daily through its exchange. Thus, the health of the Hong Kong Stock Market is often used as a barometer of the health of Asian economies. Many western corporations also use the Hong Kong Stock Market as a significant source of revenue. The stability of Hong Kong's financial districts, as well as various international businesses, depends heavily on the health and stability of Hong Kong's stock exchange.

3. Methodology

3.1. Research framework

The research framework of this study is shown in Fig. 1. It involves collecting indices of 30 categories from TAIEX, SSE/SZSE (industrial indices), and HSI (industrial indices). According to the collected information, a database system is built to effectively store and inquire about a huge number of transactions, in accordance with the database system to further the use of association rule technology and to explore possible investment portfolios.

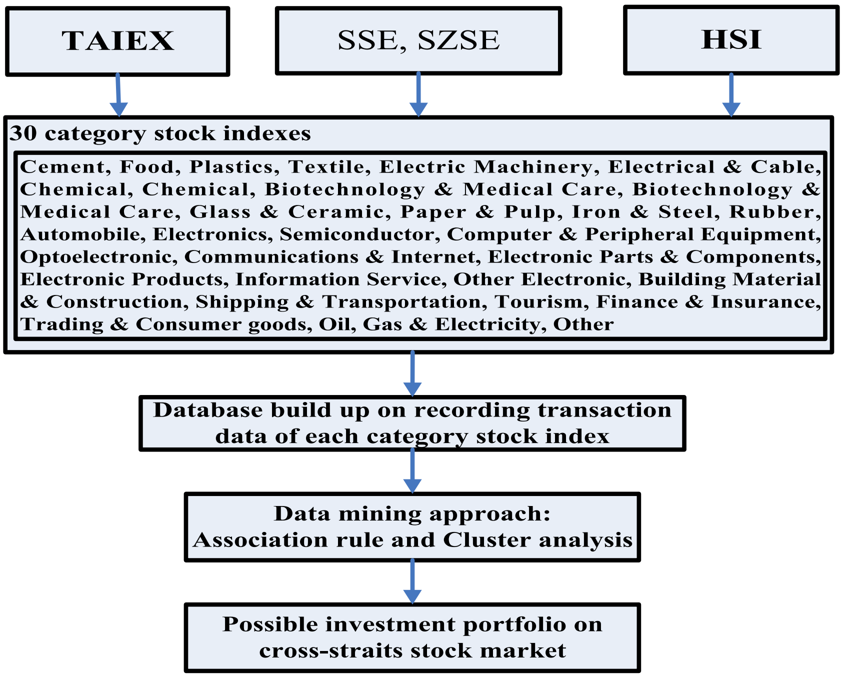

Fig. 1. Research framework.

According to the Taiwan Stock Exchange Web site, it relies on stock indexes database comprised of 30 categories. These indexes were collected from the Taiwan Economy Journal which provides transaction data on those four stock markets, for the period from June 2008 to March 2011, a total 795 transaction days.

In order to reduce the volatility of each trading day, this duration is divided into two periods to study the co-movement between stock market indexes and stock categorical indexes on Cross-traits stock markets (one period is called the Early Harvest Program, from June 2008 to Dec 2009, the other period is ECFA implementation, from Feb 2010 to March 2011. In this study, we use 0 and 1 to represent a rise or fall of the index, when the index over the previous period is lower, the value is set to "0", when the index over the previous period is higher, the value is set to "1", as shown in Table 1. For example, in case of Taiwan stock market, in the period from 2008/xx/x1-1 the TAIEX index is 1, the Cement index is 0, indicating that the TAIEX index rose and the Cement exchange rate index fell in that period. The Shanghai, Shenzhen, and Hong Kong stock indexes act the same way, forming multi-dimensional tables.

3.2. Database — the star schema

A star schema is a simple database design (particularly suited to ad hoc queries) in which dimensional data (describing how data are commonly aggregated) are separated from fact or event data (Devlin, 1997). A star schema consists of two types of tables: fact tables and dimension tables. Fact tables contain factual or quantitative data and dimension tables have descriptive data. Each dimension table has a one-to-many relationship to the central fact table. Each dimension table generally has a simple primary key, as well as several non-key attributes. In this study, a star schema is used to design the database. To build up the normalized relation is an important objective of database design. In general, with higher orders of normalization, more joint operations are needed to produce a specific output. Increasing the normalization will increase the tables, which thus increases the input/output actions of the hard drive and also slows the operation. Therefore, the database allows a certain degree of de-normalization in order to speed up the operation of data, but it does so at the cost of producing duplicate data. Because the data of stock price indices are non-duplicate, this study adopts the star schema for database design to speed up the data operation. In this study, the star schema contains four elements including a fact table, dimension table, attributes, and attribute hierarchies with eight dimensions (Fig. 2). Each element can be described as follows.

(1) Fact table: In general, facts are stored in a fact table, which is linked to n-dimension entities. The primary key of the fact table must be comprised of foreign keys, which link to the relative dimension table. In this study, the fact tables include four stock markets, time period, frequency, date, and data format.
(2) Dimension table: A dimension table can provide additional viewpoints for given facts. The data of the decision support system are always checked for their associations with other data. In this study, the dimensions include the indices of four stock markets, date and the format of data appearance.

Table 1
Example of data form.

Date	Taiwan			Hong Kong			Shanghai			Shenzhen		
	TAIEX	Cement	...	HSI	HSUTI	...	SSE	Commercial	...	SZSE	Manufacturing	...
2008/xx/x1	1	0	...	1	1	...	0	1	...	1	1	...
2008/xx/x2	1	0	...	1	0	...	1	1	...	1	0	...
2008/xx/x3	0	0	...	1	0	...	0	1	...	1	0	...
...	...	...	...	...	...	...	...	...	...	...	...	...

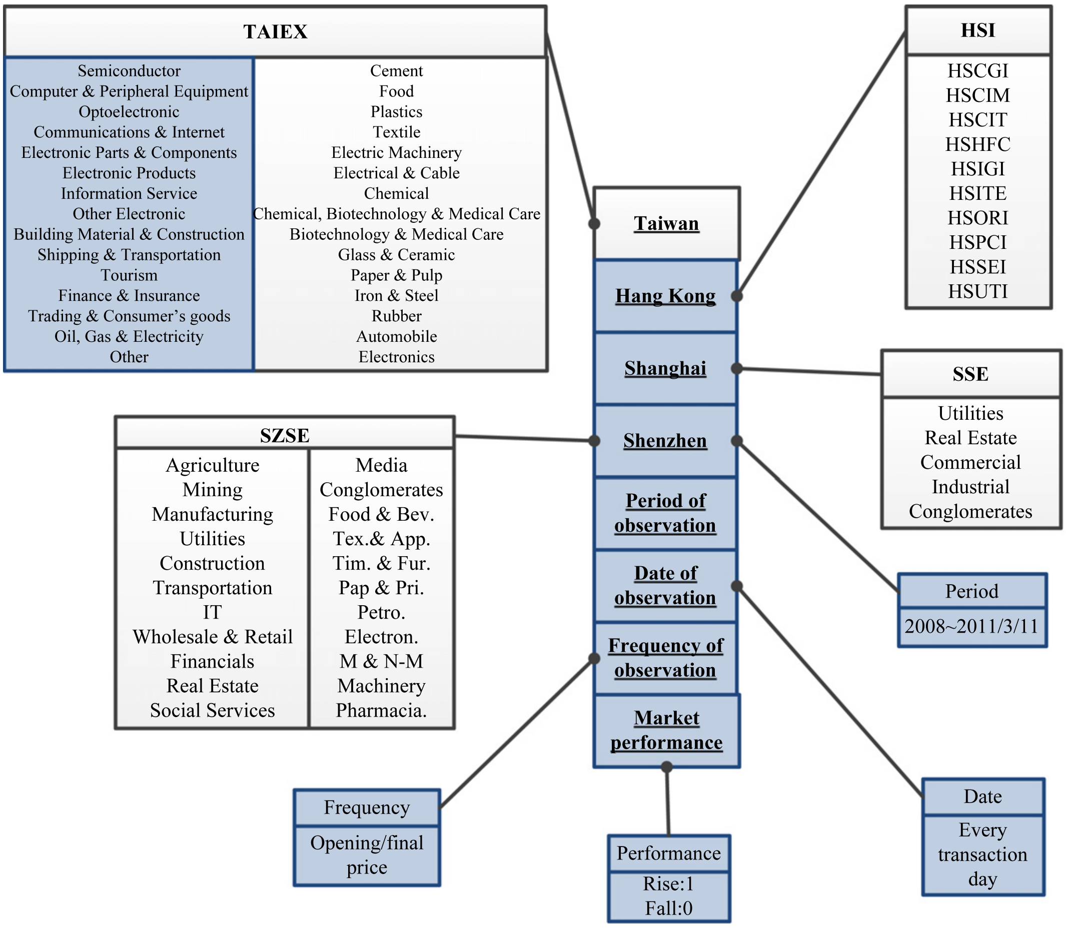

Fig. 2. Star schema map.

(3) Attribute: Attributes are used to search, filter or classify the facts, and attributes are always contained in each dimension. Dimensions acquire the descriptive characteristics of facts via attributes. This study uses 30 categories of stock indices from Taiwan and three China stock indices as attributes.
(4) Attribute hierarchies: Attributes in the dimension table can be ordered by using attribute hierarchies that have been defined specifically. Attribute hierarchies can be used to operate the drill down or roll up in data analysis.

The star schema is the simplest style of data warehouse schema. It consists of a fact tables for the reference numbers of the dimension tables. Although the primary key value must be unique in each row of a dimension table, that value can occur multiple times in the foreign key in the fact table in a many-to-one relationship. The star schema of this study is shown in Fig. 2.

4. Data mining

4.1. Association rules

As stated in Agrawal R and Swami A. (1993) discovering association rules is an important data mining issue, and there has been considerable research on using association rules in the field of data mining problems. The association rules algorithm is used mainly to determine the relationships between items or features that occur synchronously in databases. For instance, if people who buy item X also buy item Y, there is a relationship between item X and item Y, and this information is useful for decision makers. Therefore, the main purpose of implementing the association rules algorithm is to find synchronous relationships by analyzing random data and using these relationships as a reference during decision-making. The association rules are defined as follows (Wang, et al, 2004):

Make $I = \{i_1, i_2, \ldots, i_m\}$ as the item set, in which each item represents a specific literal. D stands for a set of transactions in a database in which each transaction T represents an item set such that $T \subseteq I$. That is, each item set T is a non-empty sub-item set of I. The *association rules* are an implication of the form $X \rightarrow Y$, where $X \subset I$, $Y \subset I$ and $X \cap Y = \Phi$. The rule $X \rightarrow Y$ holds in the transaction set D according to two measure standards - *support* and *confidence*. Support (denoted as $Sup(X, D)$) that represents the rate of transactions in D containing the item set X. *Support* is used to evaluate the statistical importance of D, and the higher its value, the more important the transaction set D is. Therefore, the rule $X \rightarrow Y$ has *support* $Sup(X \cup Y, D)$ that represents the rate of transactions in D containing $X \cup Y$. Each rule $X \rightarrow Y$ also has another measuring standard called Confidence (denoted as $Conf(X \rightarrow Y)$), representing the rate of transactions in D that contain X and also Y. That is, $Conf(X \rightarrow Y) = Sup(X \cap Y)/Sup(X, D)$.

In this case, $Conf(X \rightarrow Y)$ denotes that if the transaction includes X, the chance that the transaction also contains Y is relatively high. The measure confidence is then used to evaluate the level of confidence about the association rules $X \rightarrow Y$. Given a set of transactions D, the problem of mining association rules is to generate all transaction rules that have certain user-specified minimum support (called *Min* sup) and confidence (called *Minconf*) (Kouris, Makris, & Tsakalidis, 2005). According to Agrawal and Shafer (1996), the problem of mining association rules can be decomposed into

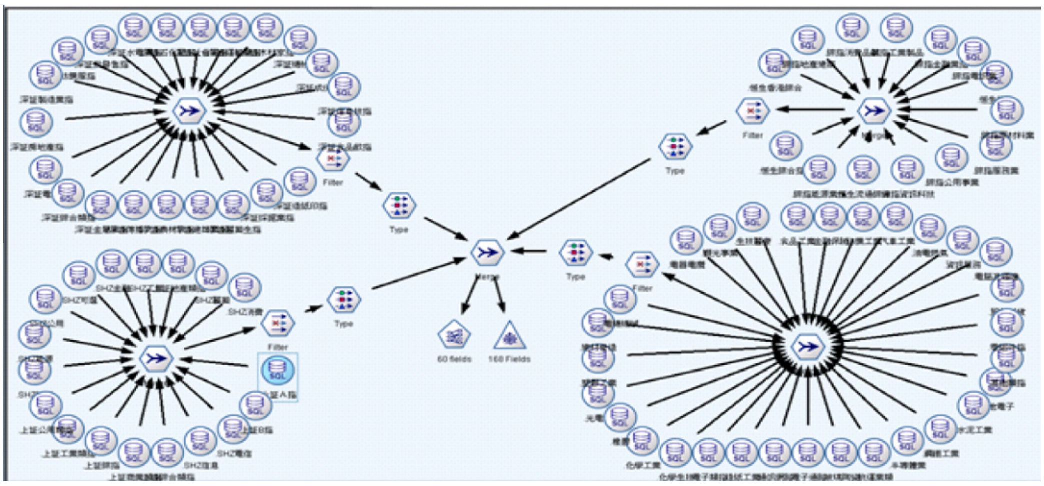

Fig. 3. Data stream of indexes of categories of four different stock markets.

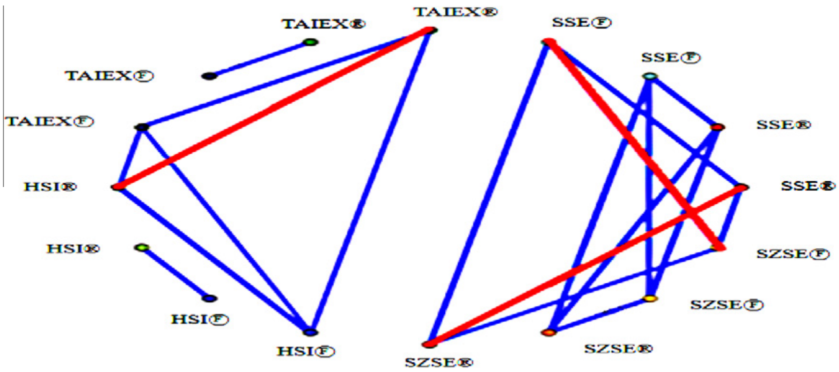

Fig. 4. Association map of four different stock markets.

two steps. The first step is to detect a large item set whose support is greater than *Min* sup and the second step is to generate association rules, using the large item set. Such rules must satisfy two conditions:

$$Sup(X \cup Y, D) \geq Min\,\mathrm{sup}$$
$$Conf(X \rightarrow Y) \geq Minconf$$

To explore the association rules, many researchers use the Apriori algorithm (Agrawal et al., 1993). In order to reduce the possible biases incurred when using these measure standards, the simplest way to judge the standard is to use the *lift* judgment. *Lift* is defined as: $Lift = Confidence(X \rightarrow Y)/Sup(Y)$ (Wang, Chuang, Hsu, & Keh, 2004).

4.2. Cluster analysis and K-means algorithm

The process of partitioning a large set of patterns into disjoint and homogeneous clusters is fundamental in knowledge acquisition. It is called *Clustering* in most studies and it is applied in various fields, including data mining, statistical data analysis, compression and vector quantization. The *k-means* is a very popular algorithm and is one of the best for implementing the clustering process. K-means clustering proceeds in the following order. Firstly, the K numbers of observations are randomly selected from all N number of observations, according to the number of clusters, and these become centers of the initial clusters. Secondly, for each of the remaining N–K observations, the nearest cluster found in terms of the Euclidean distance with respect to xi=(xi1,xi2,...;xip,...,xiP). After each observation is assigned to the nearest cluster, the center of the cluster is re-computed. Lastly, after the allocation of all observations, the Euclidean distance between each observation and the cluster's center point is calculated to confirm whether or not they have been allocated to the nearest cluster. In addition, several studies have discussed the implementation of the k-means algorithm for cluster analysis as a data mining approach (Ture, Kurt, Turhan, & Ozdamar, 2005; Vrahatis, Boutsinas, Alevizos, & Pavlides, 2002).

4.3. Data mining tool – SPSS Modeler

In this study, SPSS Modeler is employed as a data mining tool for analysis. The difference between Modeler and other software is that its data processing is carried out through the use of nodes, which are then connected together to form a stream frame. In addition, data visualization can be presented to users after the mining process is completed. All of the nodes can be divided into six categories: the source node, record options node, field options node, graphs node, modeling node, and output node (SPSS Inc., 2012).

5. Data mining results

The subject of this study is based on the 30 categories stock indexes on four different stock markets. The data source is the Taiwan Economy Journal website database. Data format conversion through streaming set analysis was the main basis for this chapter. As shown in Fig. 3, due to the process of Modeler data analysis, the

Table 2
Association rules of SZSE and Taiwan category stock indexes.

Rule	Lift	Sup.	Conf.	Consequent	Antecedent
R_{Q1}	1.613	5.42%	76.19%	SZSEⒻ	Information Service® Building Material & ConstructionⒻ Electrical & CableⒻ Computer & Peripheral EquipmentⒻ
R_{Q2}	1.252	21.68%	66.07%	SZSE®	Optoelectronic® Semiconductor® Electronics® Electronic Parts & Components® Electric Machinery®

primary stream nodes of this study set the nodes and links into the stream to complete the process, so the data through ODBC (Open Database Connectivity, an open data link) link the data into the Modeler system, completing the setup and analysis process.

5.1. Exploration of co-movement between Taiwan and China stock markets

In terms of different time periods of ECFA development, this study manipulates data by dividing research period into two observation periods: (1) 1/7/2008 to 31/12/2009 (Early Harvest Program), (2) 1/2/2010 to 31/3/2011 (ECFA implementation). By using association rules, this article intends to investigate possible co-movement on the Taiwan and China stock markets under the different implementation stages of ECFA.

An association map is a visual interactive data mining tool that illustrates association results among variables by describing the relationships between lines and nodes. Each line represents a strong/weak relationship between two nodes, with thicker lines for stronger associations among variables and thinner lines for weaker associations. After generating the association map, we can observe the association results for co-movements between the Taiwan and China stock markets.

In Fig. 4, stronger associations exist between strait stock market, Shanghai/Shenzhen and Taiwan/Hong Kong stock markets, showing that the Hong Kong stock market has higher co-movement on the Taiwan stock market. This implies that the Hong Kong stock market efficiency affects the Taiwan stock market efficiency for both rising and falling. In order to investigate further relationships between the Taiwan and China stock markets, this study explores the association of category indexes from four different stock markets. This allows us to consider more co-movement relationships between strait stock markets for generating possible investment alternatives.

5.2. Exploration of co-movement of stock category indexes between Taiwan and China stock markets

The analysis of the decision variables on Taiwan and China stock market (TAIEX, SSE, SZSE, HKEX) finds co-movement between the rise/fall of stock category indexes. In this study, the initial support and confidence are set to be 5.42% and 76.19%, respectively. In addition, the lift value should be greater than 1. After exploring the decision variables of Taiwan industrial indexes, the minimum support and confidence are 21.68% and 66.07% (Table 2). The mined association rules are shown in Table 2, in which two rules have been found on TAIEX and SZSE. Firstly, SZSE is set up as the consequent. R_{Q1} represents that when SZSE is falling, the associations with Taiwan industrial category indexes are: Information Service (rise), Building Material & Construction (fall), Electrical & Cable (fall), and Computer & Peripheral Equipment (fall). R_{Q2} represents that When SZSE is rising, the associations with Taiwan industrial category indexes are: Optoelectronic (rise), Semiconductor (rise), Electronics (rise), Electronic Parts & Components (rise), and Electric Machinery (rise).

Table 3
Association rules of SSE and Taiwan category stock indexes.

Rule	Lift	Sup.	Conf.	Consequent	Antecedent
R_{R1}	1.707	5.16%	80%	SSEⒻ	Information Service® Communications & InternetⒻ Electrical & CableⒻ OptoelectronicⒻ
R_{R2}	1.238	20.38%	65.82%	SSE®	Semiconductor® Textile® Oil, Gas & Electricity®

Secondly, SSE is set up as the consequent. The initial support and confidence are set to be 5.16% and 80%, respectively. In addition, the lift value should be greater than 1. After exploring the decision variables of Taiwan industrial indexes, the minimum support and confidence are 20.38% and 65.82% (Table 3). In Table 3, there are two rules that have been found on both TAIEX and SSE. Firstly, SSE is set up as the consequent. R_{R1} represents that When SSE is falling, the associations with Taiwan industrial category indexes are: Information Service (rise), Communications & Internet (fall), Electrical & Cable (fall), and Optoelectronic (fall). R_{R2} represents that When SSE is rising, the associations with Taiwan industrial category indexes are: Semiconductor (rise), Textile (fall), and Oil, Gas & Electricity (rise).

Thirdly, HSI is set up as the consequent. The initial support and confidence are set to be 10.9% and 65.43%, respectively. In addition, the lift value should be greater than 1. After exploring the decision variables of Taiwan industrial indexes, the minimum support and confidence are 14.27% and 70.5% (Table 4). In Table 4, there are two rules that have been found on TAIEX and HSI. Firstly, HSI is set up as the consequent. R_{S1} represents that When HSI is rising, the associations with Taiwan industrial category indexes are: Electrical & Cable (rise), Plastics (rise), and Tourism (fall). R_{S2} represents that When HSI is falling, the associations with Taiwan industrial category indexes are: Biotechnology & Medical Care (rise), and Plastics (fall).

In addition, this study investigates the association among Taiwan and China (Hong Kong) on category stock indexes, and 30 category indexes are set up as decision variables. Two observation periods are explored, including the period of Early Harvest Program (2008–2009), and the period of 2010 to 2011 (ECFA implementation). Antecedent is set up as China (Hong Kong) and consequent is set up as the Taiwan 30 category stock indexes. By doing so, during the Early Harvest Program, six rules been found, as illustrated in the Table 5 and Fig. 5, respectively. For example, R_{T1} has a greatest lift value as 1.14. Electronics (Taiwan) has a negative co-movement with Construction (Hong Kong) with a falling trend. On the other hand, five rules been explored for the period of ECFA. R_{U1} has a greatest lift value as 1.135. Semiconductor

Table 4
Association rules of HSI and Taiwan industrial category stock indexes.

Rule	Lift	Sup.	Conf.	Consequent	Antecedent
R_{S1}	1.405	10.90%	65.43%	HIS®	Electrical & Cable® Plastics® TourismⒻ
R_{S1}	1.324	14.27%	70.75%	HSⒻ	Biotechnology & Medical Care® PlasticsⒻ

Table 5
Association rules of Taiwan and China (Hong Kong) category stock indexes.

Time	Rule	Lift	Sup.	Conf.	Consequent	Antecedent
2008–2009	R_{T1}	1.140	57.06%	62.72%	ElectronicsⒻ	HSPCIⒻ
	R_{T2}	1.119	57.06%	62.01%	SemiconductorⒻ	HSPCIⒻ
	R_{T3}	1.062	55.62%	62.13%	Shipping & TransportationⒻ	HSCITⒻ
	R_{T4}	1.060	57.06%	62.01%	Shipping & TransportationⒻ	HSPCIⒻ
	R_{T5}	1.047	55.42%	61.25%	Shipping & TransportationⒻ	HSIGIⒻ
	R_{T6}	1.046	55.83%	61.17%	Shipping & TransportationⒻ	HSSEIⒻ
2010–2011	R_{U1}	1.135	50.24%	70.87%	SemiconductorⒻ	HSCIMⒻ
	R_{U2}	1.086	50.24%	69.90%	OptoelectronicⒻ	HSCIMⒻ
	R_{U3}	1.065	51.22%	68.57%	OptoelectronicⒻ	SHZ Ⓡ Food & Bev. Ⓡ Electron Ⓡ
	R_{U4}	1.045	53.66%	67.27%	OptoelectronicⒻ	SHZ Ⓡ
	R_{U5}	1.089	56.1%	66.96%	Building Material & ConstructionⒻ	HSCGI Ⓡ

(Taiwan) has a negative co-movement with Raw Material (Hong Kong), with a falling trend. More details as described on Table 5 and Fig. 6.

In Fig. 6, the trend of positive co-movement includes Food, Electronics, and category stock indexes. Compared with Fig. 5, more vivid market expansion on stock market shows that China shares a positive economy growth on consumed products. One of the reasons is China's early five-year policy of "China's Home Appliance Subsidy Program for Rural Areas". In the beginning of 2010, China expanded the items of this program from main electronic products to general consumer electronic products. Thus, positive co-movement indicates that Taiwan has a positive trend on these stock indexes for further co-operation and investment on China market.

5.3. Exploration of cluster of stock category indexes between the Taiwan and China stock markets

In order to explore possible clusters of stock category indexes between the Taiwan and China stock markets, this study implements a K-means approach by clustering the pattern of rise/fall trend with a percentage comparison (Table 6 and Fig. 7). By doing so, further investment considerations might be determined for the Taiwan and China stock markets.

In the period of global financial crisis, the Taiwan and China (Hong Kong) stock markets have had a falling trend. However, from Table 6, we might conclude that TAIEX generally has better market performance than SSE and SZSE in. Thus, besides the results of association rules, performance of an individual category stock index is an evaluation index for possible investment on Taiwan and China (Hong Kong) stock market.

6. Discussion and conclusion

6.1. Discussions

6.1.1. Better stock market structure, higher international market efficiency between Taiwan and China (Hong Kong)

The efficient international market hypothesis states that all information is knowable and thus is immediately factored into the market's price for security. China's stock market has experienced amazing growth since establishing its two exchanges in 1990, although the growth has been uneven and irregular, and the market remains in the early stages of its development (Wang, 2004). Gao (2002) discusses 14 special features of China's stock market, such as: abnormal performance, tremendous volatility, insulated market, substantial government ownership, irregular expansion, influence of IPOs, typical emerging markets, pyramid structure, unstable core, outperformance of micro stocks, incredible speculation, manufacturing orientation, disappointing earnings of companies, and low dividend yield. Although China's stock market has made great strides forward over the past ten years, market structure is still a main concern for market globalization. For example, whereas most developed markets are dominated by a limited number of large-cap stocks, China's market is cramped by a multitude of small-cap stocks. This feature allows for increased speculation and higher turnover for both investors and indexes, among other problems, such as the market liquidity risk factor on the Chinese stock market (Narayan and Zhang, 2010). In addition, a related matter is the reliance of China's market on external expansion, that is, expansion through the issuance of new shares rather than the appreciation in value of existing stocks. Since these shares generally do not experience sustained growth, often because of market manipulation, they contribute to the dominance of smaller size stocks in China's market (Chen, Firth, & Kim, 2004). Ultimately, the current structure is a major obstacle to the creation of viable index-related products in China. Finally, the government seems to have too much influence on the market. It keeps a tight control on the issuance of IPOs, and, as a result of widespread government holdings, many listed companies in China have very low free-float ratios (Gao, 2010). Meanwhile, due at least in part to the unusual market structure, market manipulation and speculation are common. The solution is simply a matter of strengthening controls in certain areas while relaxing them in others in order to foster an environment in which China's stock market can continue to thrive on the global market (Kang, Liu, & Ni, 2002).

In Table 6, both TAIEX and HSI seem have better market efficiency, not only for the period of financial crisis (2008–2009), but also during 2010–2011 in average. Therefore, adjustment of the current market structure is crucial in reforming the market.

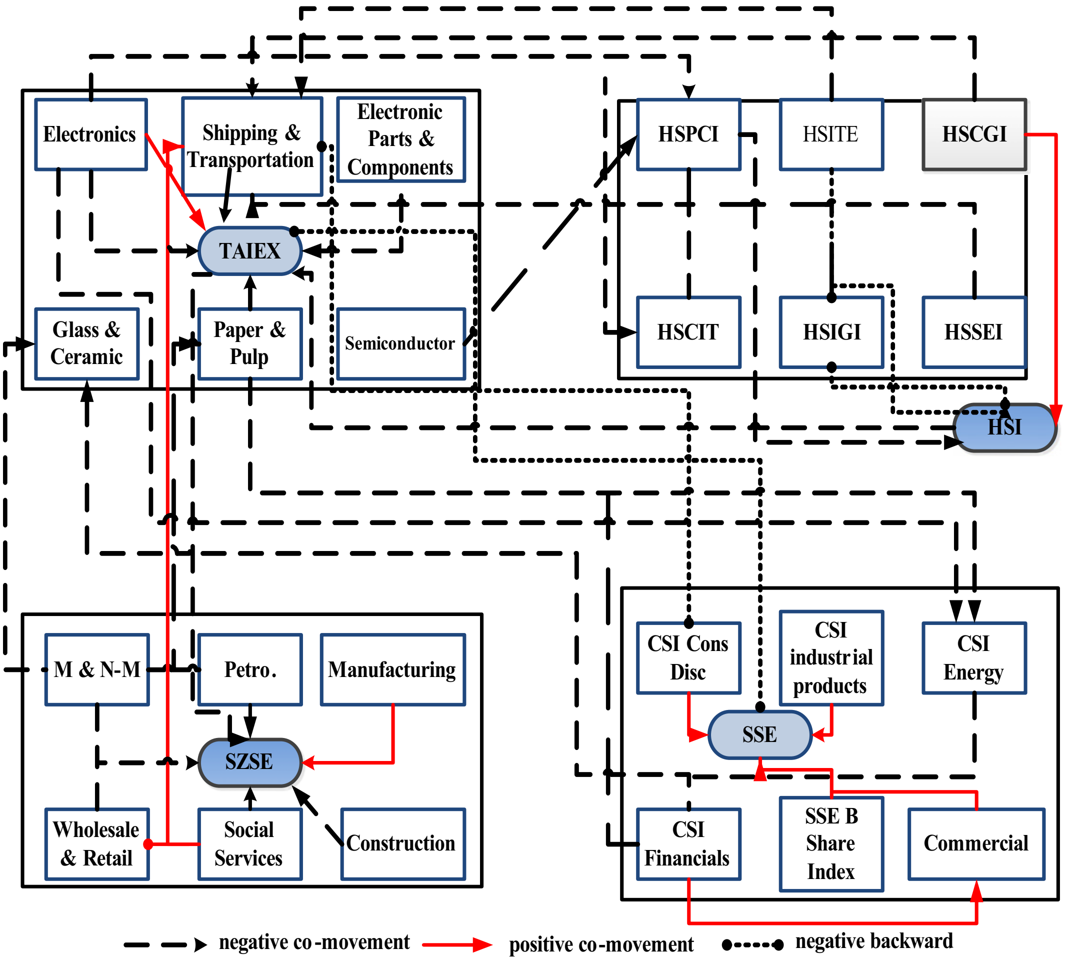

Fig. 5. Association maps of Taiwan and China (Hong Kong) category stock indexes for the period of Early Harvest Program.

What China's stock market needs most at present is to achieve long-term stable development, rather than simply getting out of the current difficulties. Given this, China should make it a priority to open its market to the world and include as soon as possible a batch of international industry leaders with real profitability, solidify the market foundation and optimize the market structure. Since we can see that co-movements do exist between the Taiwan and China (Hong Kong) stock market, we may infer that China has a better stock market structure, and there is higher international market efficiency between the Taiwan and China (Hong Kong) markets.

6.1.2. Better understanding for stock market co-movement and larger investment portfolio between Taiwan and China (Hong Kong)

According to Figs. 5 and 6, this study investigates the co-movement of category stock indexes between Taiwan and China (Hong Kong) for two periods, including the Early Harvest Program and ECFA. Results show that most frequent co-movement categories of stock indexes that show rising prices include Tourism, Energy Resources, Cultural, Biotechnology and Medicine, Agricultural, and Information technology. Thus we propose some possible investment portfolios for categories of stock between Taiwan and China (Hong Kong) as follows:

(1) Portfolio by selecting stocks in the categories of Information Service and Electronics:
In fact, stock in the category of Electronics has a high co-movement in both the Early Harvest Program and the ECFA observation periods for the cross-straits stock market. We find that the crowding-out effect exists not only in the industrial products market (Electronics and Machinery products), but also in the stock market between Taiwan and China. This infers that co-competition relationships can influence Cross-Straits industrial development for a specific manufacturing market. Industry upgrade is a possible solution for Taiwan to increase the value of its products in the supply chain by providing more advanced technology and a more competitive business model in order to release its labor-intensive market and create a high value-added industrial product market. Cloud computing, for example, is a thriving information service on the global marketplace. In terms of supply chain integration for Cross-Straits endeavors, Taiwan and China may implement cloud computing together on their industrial specialization chains in order to attract more investment, enhance complementary effects, and upgrade co-competition relationships by creating mutual market profit.

(2) Portfolio by selecting Electronics Components, Bicycles, and Electronic Motor Car product category stock:
Electronics components and bicycles are product items listed in the Early Harvest Program. Since 2009 China has been the largest consumer of automobiles, and so stocks of the automobile industry category have high expectation for profits in the future market. In addition, due to environmental concerns and rising oil prices, bicycles and electronic automobiles are not only products supported by governmental subsidies product item, but they are also the focus of new technology/product development. Thus, bicy-

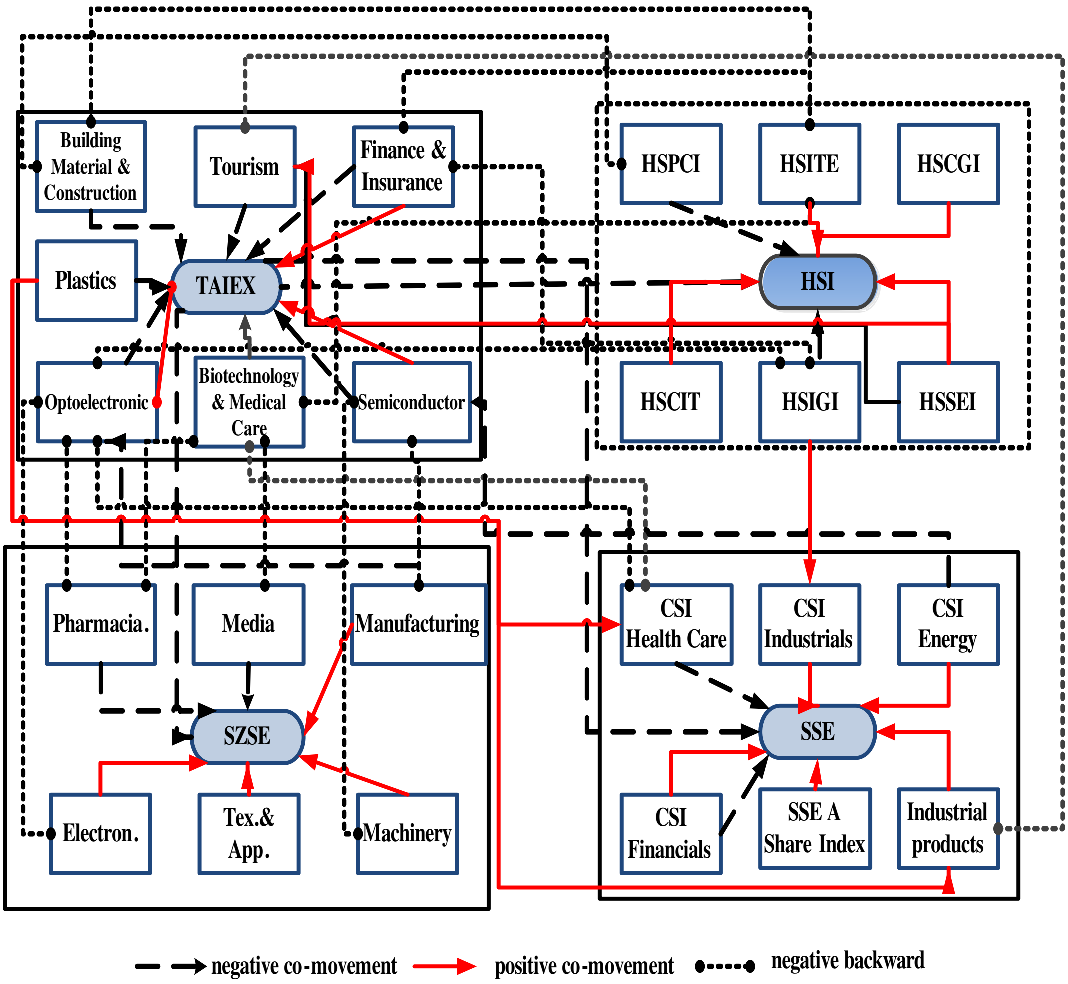

Fig. 6. Association map of Taiwan and China (Hong Kong) category stock indexes for the period of ECFA implementation.

cles, electronic automobiles, and their components, such as: lithium batteries, traction induction motors, and power electronic units, are future development products for Taiwan and China to co-operate on in the transportation industry.

(3) Portfolio composed of Tourism stocks:
Since 1981 Taiwanese businessmen and have visited China, while Taiwan opened the door on the Straits for business and humanitarian reasons. In 2008 Taiwan began to allow Chinese people to visit Taiwan for traveling and family visits using group tours and in 2011 it was opened to independent travelers. This Cross-Strait open policy provides mutual benefits by increasing mutual understanding for people, culture, and society. In this regard, the tourism market is becoming increasingly important for both Taiwan and China (Hong Kong). Thus, an investment portfolio might be assembled by selecting tourism-related stocks.

(4) Portfolio of Biotechnology and Medicine stocks:
Due to the increased population of older people and higher standards of living, the biotechnology and medicine industries are expected to provide high profits in the future, from both Western and Chinese medicine. In Jun 2011, the Cross-Strait Medicine and Health Agreement was published to support further cooperation on technology and product development. This agreement has broadened the co-operation scope of the biotechnology and medicine industry on Taiwan and China. Especially Chinese medicine, in particular, is a core competence for close cooperation with Chinese society on the global market. By implementing biotechnology into the medicine business, there is the possibility for great profit and opportunity on future investment.

In Fig. 8, this study proposes possible investment portfolios of for different categories of stock on the Taiwan and the China (Hong Kong) stock markets. For further alternatives about co-movements and other investment plans, please refer to the information in Figs. 5 and 6, and Table 6.

6.2. Conclusion

Data mining results show that the Taiwan stock market has strong co-movement on Electronics, Financial and Insurance, and Semi-conductor stock indexes with the TAIEX index. On the other hand, Hong Kong stock market has clear co-movement on Real Estate, Tele-communications, and Financial Services stock indexes with the HIS index. Manufacturing, Machinery, and Electronics stock indexes have co-movement with the SZSE for the Shenzhen stock market. In addition, Industrial Products, Energy, and the financial stock indexes have co-movement with the SSE for the Shanghai stock market. Each stock market reflects its industrial and business features on market development. For example, the financial services industry is the main feature of the Hong Kong stock market. Shenzhen has a strong foundation in manufacturing industries. Shanghai is the center for financial operations and has a strong market niche for developing the financial industry. Taiwan,

Table 6
Two clusters of performance comparison on Taiwan and China (Hong Kong) stock market.

Category stock indexes (Market)	Cluster-1 (405 Rec) (Higher possibility on rise)		Cluster-2 (338 Rec) (Higher possibility on fall)	
0 (fall)/ 1 (rise)	0	1	0	1
Chemical (TAIEX)	50.62%	49.38%	52.96%	47.04%
Chemical, Biotechnology & Medical Care (TAIEX)	52.84%	47.16%	52.66%	47.34%
Cement (TAIEX)	52.59%	47.41%	58.88%	41.12%
Semiconductor (TAIEX)	54.57%	45.43%	60.65%	39.35%
Biotechnology & Medical Care (TAIEX)	59.26%	40.74%	57.69%	42.31%
Optoelectronic (TAIEX)	56.79%	43.21%	60.95%	39.05%
Automobile (TAIEX)	54.57%	45.43%	57.99%	42.01%
Other Electronic (TAIEX)	54.57%	45.43%	60.95%	39.05%
Other Electronic (TAIEX)	53.33%	46.67%	57.99%	39.05%
Oil, Gas & Electricity (TAIEX)	44.94%	55.06%	52.37%	47.63%
Finance & Insurance (TAIEX)	53.58%	46.42%	56.21%	43.79%
Building Material & Construction (TAIEX)	53.09%	46.91%	55.92%	44.08%
Glass & Ceramic (TAIEX)	52.84%	47.16%	58.58%	41.42%
Food (TAIEX)	52.10%	47.90%	58.88%	41.12%
Textile (TAIEX)	50.12%	49.88%	54.44%	45.56%
Shipping & Transportation (TAIEX)	58.27%	41.73%	56.51%	43.49%
Communications & Internet (TAIEX)	51.36%	48.64%	51.78%	48.22%
Paper & Pulp (TAIEX)	53.83%	46.17%	60.36%	39.64%
Trading & Consumer's goods (TAIEX)	49.63%	50.37%	57.40%	42.60%
Plastics (TAIEX)	49.63%	50.86%	51.48%	48.52%
Information Service (TAIEX)	58.52%	41.48%	57.69%	42.31%
Electronic Parts & Components (TAIEX)	53.09%	46.91%	58.28%	41.72%
Electronics (TAIEX)	54.57%	45.43%	57.99%	42.01%
Computer & Peripheral Equipment (TAIEX)	53.83%	46.17%	59.76%	40.24%
Electrical & Cable (TAIEX)	53.58%	46.42%	56.51%	43.49%
Electric Machinery (TAIEX)	48.89%	51.11%	57.10%	42.90%
Other Electronic (TAIEX)	52.10%	47.90%	57.69%	42.31%
Rubber (TAIEX)	52.59%	47.41%	56.80%	43.20%
Iron & Steel (TAIEX)	55.80%	44.20%	55.62%	44.38%
Tourism (TAIEX)	57.78%	42.22%	57.40%	42.60%
TAIWEX (TAIEX)	52.35%	47.65%	57.10%	42.90%
HSCGI (HSI)	52.10%	47.90%	55.03%	44.97%
HSCIM (HSI)	51.11%	48.89%	49.70%	50.30%
HSCIT (HSI)	49.63%	50.37%	48.82%	51.18%
HSHFC (HSI)	53.58%	46.42%	55.33%	44.67%
HSIGI (HSI)	52.59%	47.41%	50%	50%
HSITE (HSI)	67.65%	32.35%	65.98%	34.02%
HSORI (HSI)	52.59%	47.41%	52.96%	47.04%
HSPCI (HSI)	51.36%	48.64%	50.30%	49.70%
HSSEI (HSI)	48.64%	51.36%	52.37%	47.63%
HSUTI (HSI)	52.84%	47.16%	53.55%	46.45%
HSIWEX (HSI)	55.56%	44.44%	53.25%	46.75%
SBI (SSE)	10.86%	89.14%	84.32%	15.68%
SAI (SSE)	10.86%	89.14%	90.83%	9.17%
SWEX (SSE)	10.86%	89.14%	90.83%	9.17%
Industrial (SSE)	9.88%	90.12%	92.31%	7.69%
Utilities (SSE)	11.60%	88.40%	92.01%	7.69%
Real Estate (SSE)	20.74%	79.26%	80.47%	19.53%
Commercial (SSE)	9.38%	90.62%	86.69%	13.31%
Conglomerates (SSE)	20%	80%	81.66%	18.34%
SZSEAI (SZSE)	10.62%	89.38%	91.42%	8.58%
SZSEBI (SZSE)	18.02%	81.98%	75.15%	24.85%
SZSEWEX (SZSE)	10.62%	89.38%	91.42%	8.58%
Mining (SZSE)	20.25%	79.75%	81.07%	18.93%
Agriculture (SZSE)	13.09%	86.91%	76.33%	23.67%
Manufacturing (SZSE)	2.72%	97.28%	89.05%	10.95%
Tim. & Fur. (SZSE)	13.83%	86.17%	72.49%	27.51%
Electron. (SZSE)	87.90%	12.10%	13.61%	86.39%
Petro. (SZSE)	7.41%	92.59%	84.91%	15.09%
Wholesale & Retail (SZSE)	14.81%	85.19%	83.73%	16.27%
Real Estate (SZSE)	20%	80%	79.59%	20.41%
Social Services (SZSE)	9.88%	90.12%	83.73%	16.27%
Financials (SZSE)	22.72%	77.28%	80.47%	19.53%
M & N-M (SZSE)	7.90%	92.10%	86.39%	13.61%
IT (SZSE)	10.12%	89.88%	83.14%	16.86%
Construction (SZSE)	10.86%	89.14%	79.29%	20.71%
Food & Bev. (SZSE)	14.07%	85.93%	82.25%	17.75%
Tex.& App. (SZSE)	5.68%	94.32%	83.73%	16.27%
Pap & Pri. (SZSE)	8.89%	91.11%	84.91%	15.09%
Media (SZSE)	17.04%	82.96%	73.37%	26.63%
Transportation (SZSE)	6.67%	93.33%	88.76%	11.24%
Electronics. (SZSE)	7.41%	92.59%	82.54%	17.46%
Conglomerates (SZSE)	6.91%	93.09%	83.43%	16.57%
Machinery (SZSE)	6.17%	93.83%	85.21%	14.79%
Pharmaceuticals (SZSE)	15.56%	84.44%	77.22%	22.78%

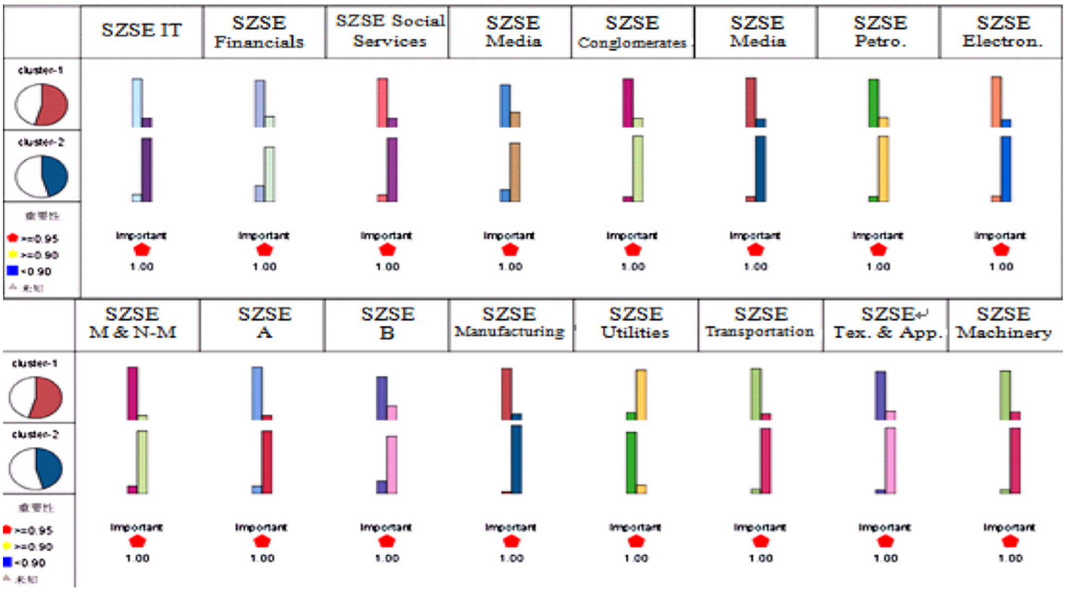

Fig. 7. Clusters of Taiwan and China (Hong Kong) stock market performance.

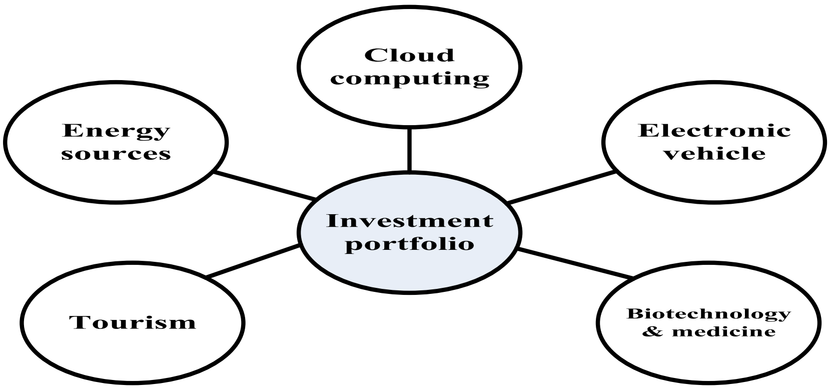

Fig. 8. Possible investment portfolios on the Cross-Strait stock market.

for its part, has great potential for the market in high-tech products.

In addition, during the Early Harvest Program, there has been co-movement on Transportation, and Electronics category stock between Taiwan and Hong Kong. Once Hong Kong Real Estate and Information Technology stock indexes begin to fall, Taiwan Electronics, Semi-conductor, and Shipping stock indexes also begin to fall. At the same time, Taiwan Shipping and Electronics stock indexes also tend to drop, which influences the Shanghai and Shenzhen Wholesale, and Manufacturing stock indexes with the same falling co-movement. In the ECFA period, Hong Kong, Shanghai, and Shenzhen Medical and Electronics stock indexes are influenced by the Crowding-Out Effect on Taiwan similar category stock indexes. In addition, the Hong Kong Industrial and Information stock indexes could affect the market performance of the Taiwan Financial stock index. This indicates that Hong Kong might tend to lead co-movement of Financial, and Service category stock indexes to Taiwan. Another interesting finding was that before ECFA implementation Taiwan Papermaking, and Glass stock indexes performance were influenced by the Shanghai and the Shenzhen stock markets. However, after ECFA implementation, new co-movements on the market performance began to affect Electronics and Biotechnology stock indexes.

This study investigates co-movements on the Taiwan and the China stock markets under ECFA using a data mining approach. Thirty categories of stock indexes are implemented as decision variables to observe the behavior of stock index associations during the periods of the Early Harvest Program and the ECFA implementation from the standpoint of Taiwan. This paper considers that a stock market has strong associations with both inside and outside factors. Some stock index categories of stock rise or fall together at the same time or are mutually influenced by domestic or foreign economic, social, and political situations. For individual or institutional investors, finding indications of trends in stock market association is an important ability. Thus, this case study of implementing data mining approaches and integrating them into international or regional stock market research on the Taiwan and China (Hong Kong) stock markets is an example for future cross-national stock research and implementation.

Acknowledgements

This research was funded by the National Science Council, Taiwan, Republic of China, under contract No. NSC 100-2410-H-032 -018 –MY3.

References

Agrawal, R., Imilienski, T., & Swami, A., 1993. Mining association rules between sets of items in large databases. In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data* (pp. 207–216).

Agrawal, R., & Shafer, J. (1996). Parallel mining of association rules. *IEEE Transactions on Knowledge and Data Engineering, 8*(6), 962–969.

Alexakis, C., Niarchos, N., Patra, T., & Poshakwale, S. (2005). The dynamics between stock returns and mutual fund flows: empirical evidence from the Greek market. *International Review of Financial Analysis, 14*(5), 559–569.

Arestis, P., Caporale, G. M., Cipollini, A., & Spagnolo, N. (2005). Testing for financial contagion between developed and emerging markets during the 1997 East Asian crisis. *International Journal of Finance and Economics, 10*(4), 359–367.

Aslanidis, N., Osborn, D. R., & Sensier, M. (2010). Co-movements between US and UK stock prices: the role of time-varying conditional correlations. *International Journal of Finance and Economics, 15*(4), 366–380.

Beine, M., & Candelon, B. (2010). Liberalisation and stock market co-movement between emerging economies. *Quantitative Finance, 99*(1), 1–14.

Boginski, V., Butenko, S., & Pardalos, P. M. (2006). Mining market data: A network approach. *Computers & Operations Research, 33*, 3171–3184.

Bohl, M. T., Brzeszczyński, J., & Wilfling, B. (2009). Institutional investors and stock returns volatility: Empirical evidence from a natural experiment. *Journal of Financial Stability, 5*(2), 170–182.

Boyer, B., & Zheng, L. (2009). Investor flows and stock market returns. *Journal of Empirical Finance, 16*(1), 87–100.

Brown, S., Rhee, S. G., & Zhang, L. (2008). The return to value in Asian stock markets. *Emerging Markets Review, 9*(3), 194–205.

Chang, Y. (2002). The pricing of foreign exchange risk around the Asian financial crisis: evidence from Taiwan's stock market. *Journal of Multinational Financial Management, 12*(3), 223–238.

Chen, G., Firth, M., & Kim, J. B. (2004). IPO underpricing in China's new stock markets. *Journal of Multinational Financial Management, 14*(3), 283–302.

Chow, G. C., Liu, C., & Niu, L. (2011). Co-movements of Shanghai and New York stock prices by time-varying regressions. *Journal of Comparative Economics, 39*(4), 577–583.

Chue, T. K. (2002). Time-varying risk preferences and emerging market co-movements. *Journal of International Money and Finance, 21*(7), 1053–1072.

Chun, S. H., & Park, Y. J. (2005). Dynamic adaptive ensemble case-based reasoning: Application to stock market prediction. *Expert Systems with Applications, 28*, 435–443.

Devlin, B. (1997). *Data warehouse: From architecture to implementation*. Reading, MA: Addison Wesley Longman.

Edwards, J. A., & Ginn, V. (2011). Evaluating growth volatility susceptibility within regional free trade agreements. *International Journal of Finance and Economics, 16*(1), 32–40.

Enke, D., & Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications, 29*, 927–940.

Fama, E. (1991). Efficient capital markets. *Journal of Finance XLVI*, 1575–1617.

Forbes, K., & Rigobon, R. (1999). No contagion, only interdependence. measuring stock market co-movements. NBER Working Paper No. 7267.

Gao, R. (2002). China stock market in a global perspective. Dow Jones Index, pp. 1–48.

Gao, Y. (2010). What comprises IPO initial returns: Evidence from the Chinese market. *Pacific-Basin Finance Journal, 18*(1), 77–89.

Graham, M., Kiviaho, J., & Nikkinen, J. (2012). Integration of 22 emerging stock markets: A three-dimensional analysis. *Global Finance Journal, 23*(1), 34–47.

Haugen, R. (1997). *Modern investment theory*. Upper Saddle River, NJ: Prentice-Hall.

Jang, H., & Sul, W. (2002). The Asian financial crisis and the co-movement of Asian stock markets. *Journal of Asian Economics, 13*(1), 94–104.

Kang, J., Liu, M. H., & Ni, S. X. (2002). Contrarian and momentum strategies in the China stock market: 1993–2000. *Pacific-Basin Finance Journal, 10*(3), 243–265.

Kearney, C., & Muckley, C. (2007). Is North and Southeast Asia becoming a yen block? *International Journal of Finance and Economics, 12*(3), 337–351.

Kouris, I. N., Makris, C. H., & Tsakalidis, A. K. (2005). Using Information Retrieval techniques for supporting data mining. *Data & Knowledge Engineering, 52*, 353–383.

Liao, S. H., Ho, H. H., & Lin, H. W. (2008). Mining stock category association and cluster on Taiwan stock market. *Expert Systems with Applications, 35*, 19–29.

Liao, S. H., Chu, P. H., & You, Y. L. (2011). Mining the co-movement between foreign exchange rates and category stock indexes in the Taiwan financial capital market. *Expert Systems with Applications, 38*, 4608–4617.

Liao, S. H., Chu, P. H., & Teng, T. K. (2011). Mining the co-movement in the Taiwan stock funds market. *Expert Systems with Applications, 38*, 5276–5288.

Liljeblom, E., Löflund, A., & Krokfors, S. (1997). The benefits from international diversification for Nordic investors. *Journal of Banking & Finance, 21*(4), 469–490.

Madaleno, M., & Pinho, C. (2012). International stock market indices co-movements: A new look. *International Journal of Finance and Economics, 17*, 89–102.

Meric, I., Ratner, M., & Meric, G. (2008). Co-movements of sector index returns in the world's major stock markets in bull and bear markets: portfolio diversification implications. *International Review of Financial Analysis, 17*(1), 156–177.

Narayan, P. K., & Zheng, X. (2010). Market liquidity risk factor and financial market anomalies: Evidence from the Chinese stock market. *Pacific-Basin Finance Journal, 18*(5), 509–520.

Oh, N. Y., & Parwada, J. T. (2007). Relations between mutual fund flows and stock market returns in Korea. *Journal of International Financial Markets, Institutions and Money, 17*(2), 140–151.

Rapach, D. E., & Wohar, M. E. (2006). In-sample vs. out-of-sample tests of stock return predictability in the context of data mining. *Journal of Empirical Finance, 13*, 231–247.

Rigobon, R. (1999). On the measurement of the international propagation of shocks. NBER Working Paper No. 7354.

SPSS Inc. (2012). PASW User's Guide. Chicago, SPSS Inc.

Ture, M., Kurt, I., Turhan, K. A., & Ozdamar, K. (2005). Comparing classification techniques for predicting essential hypertension. *Expert Systems with Applications, 16*(4), 379–384.

Vrahatis, M. N., Boutsinas, B., Alevizos, P., & Pavlides, G. (2002). The new *k*-windows algorithm for improving the *k* -means clustering algorithm. *Journal of Complexity, 18*(1), 375–391.

Wang, Y. F., Chuang, Y. L., Hsu, M. H., & Keh, H. C. (2004). A personalized recommender system for the cosmetic business. *Expert Systems with Applications, 26*, 427–434.

Wang, C. (2004). Relative strength strategies in China's stock market: 1994–2000. *Pacific-Basin Finance Journal, 12*(2), 159–177.