ISSN 1916-971X

INTERNATIONAL JOURNAL OF ECONOMICS AND FINANCE

Vol. 1, No.1
February 2009

CCSE

Canadian Center of Science and Education

International Journal of Economics and Finance
Vol. 1, No. 1
February 2009
www.ccsenet.org/journal.html

Contents

The Interaction between Exchange Rates and Stock Prices: An Australian Context 3
Noel Dilrukshan Richards, John Simpson & John Evans
The Relationship between Earnings and Stock Returns: Empirical Evidence from the Greek Capital Market 40
Panagiotis E. Dimitropoulos & Dimitrios Asteriou
Order Aggressiveness on the ASX Market 51
Ying Xu
Cognitive Pattern Analysis Employing Neural Networks: Evidence from the Australian Capital Markets 76
Edward Sek Khin Wong
Study on the Transnational Venture Capital Investments: A Theoretical Framework 81
Yijun Zeng
Globalization of the Malaysian Manufacturing Sector: An Agenda for Social Protection Enhancement 86
Evelyn S. Devadason & Chan Wai Meng
Study on the Attributes and Correlations of Production Factors 98
Zuohua Yue
The Relationship between Human Resource Management and Firm Performance in Malaysia 103
May-Chiun Lo, Abang Azlan Mohamad & Maw King La
An Overview Study on Dornbusch Overshooting Hypothesis 110
Wenwen Tu & Junwen Feng
AFTA and the Intra-Trade Patterns among ASEAN-5 Economies: Trade-Enhancing or Trade-Inhibiting? 117
Kim-Lan Siah, Chee-Keong Choong & Zulkornain Yusop
A Literature Review on the Relationship between Foreign Trade and Economic Growth 127
Huan Chen
Import Inflows of Bangladesh: the Gravity Model Approach 131
Md. Mahmudul Alam, Md. Gazi Salah Uddin & Khan Md. Raziuddin Taufique
An Analysis on Influencing Factors of Stockholders Interests Redistribution-from a Consideration View of Non-Tradable Shares Reform in Medicine Industry 141
Yuhe Wang
Board Structure and Corporate Performance in Malaysia 150
Zubaidah Zainal Abidin, Nurmala Mustaffa Kamal & Kamaruzaman Jusoff
Audit Independence and Expectation Gap: Empirical Evidences from Iran 165
Mahdi Salehi, Ali Mansoury & Zhila Azary
Theory Survey of Stock Yield Prediction Models 175
Wei Shen

International Journal of Economics and Finance

www.ccsenet.org/journal.html

Vol. 1, No. 1

February 2009

Contents

Estimating Software Cost with Security Risk Potential 183

Nur Atiqah Sia Abdullah, Rusli Abdullah, Hasan Selamat, Azmi Jaafar & Kamaruzaman Jusoff

Strengthening the Construction of the Constitutionalism and Promoting the Healthy Development of the Socialism Market Economy 193

Bijun Xiong

Customers Satisfaction in Malaysian Islamic Banking 197

'Ismah Osman, Husniyati Ali, Anizah Zainuddin, Wan Edura Wan Rashid & Kamaruzaman Jusoff

Asymmetric Information and Dividend Policy in Emerging Markets: Empirical Evidences from Iran 203

Hashem Valipor, Vahab Rostami & Mahdi Salehi

Empirical Forecasting Practices of a British University 212

Rozainun Abdul Aziz, D.F. Percy & Kamaruzaman Jusoff

Tacit Knowledge Sharing and SMEs' Organizational Performance 216

Rohana Ngah & Kamaruzaman Jusoff

Analysis on the Service Trade between China and ASEAN 221

Chunmei Yang

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

The Interaction between Exchange Rates and Stock Prices: An Australian Context

Noel Dilrukshan Richards & John Simpson
School of Economics and Finance
Curtin University of Technology, Perth
Kent St. Bentley, GPO Box U1987
Perth, 6845 Western Australia
E-mail: simpsonj@cbs.curtin.edu.au

John Evans
Curtin University of Technology
Sarawak, Malaysia, CDT 250, 98009 MIRI

Abstract

The aim of this paper is to examine the interaction between stock prices and exchange rates in Australia. During the period of the study, the value of the stock market increased by two-thirds and the Australian dollar exchange rate appreciated by almost one-third. The empirical analysis employed provides evidence of a positive co-integrating relationship between these variables, with Granger causality found to run from stock prices to the exchange rate during the sample period. Although commodity prices have not been included, the significance of the results lends support to the notion that these two key financial variables interacted in a manner consistent with the portfolio balance model, that is, stock price movements cause changes in the exchange rate. This challenges the traditional view of the Australian economy as export-dependent, and also suggests that the Australian stock market has the depth and liquidity to adequately compete for both domestic and international capital against other larger markets.

Keywords: Exchange rates, Stock prices, Causality, Flow orientated, Portfolio balance

1. Introduction

The objective of this study is to ascertain the significance of the strength and direction of the influence of Australian stock price movements on the Australian dollar exchange rate between 2 January 2003 and 30 June 2006. This period was characterised by a high degree of co-movement between the two variables (Note 1). Indeed, there has never been a period in which these two key macroeconomic variables have moved so strongly and in the same direction since the float of the Australian dollar in 1983. The initial analysis investigates the broad relationship between stock prices and exchange rates in Australia and is then expanded to investigate the changes in these key economic variables and the relationship between those changes.

The interaction between equity and currency markets has been the subject of much academic debate and empirical analysis over the past 25 years. This is understandable given the crucial role that equity and currency markets play in facilitating economic activity. Classical economic theory hypothesises that stock prices and exchange rates can interact by way of the ‘flow oriented’ and ‘portfolio balance’ models. Flow oriented models, first discussed by Dornbusch and Fisher 1980, postulate that exchange rate movements cause movements in stock prices. This approach is built on the macroeconomic view that because stock prices represent the discounted present value of a firm’s expected future cash flows, then any phenomenon that affects a firm’s cash flow will be reflected in that firm’s stock price if the market is efficient as the Efficient Market Hypothesis suggests. Movements in the exchange rate are one such phenomenon.

Portfolio balance approaches, or ‘stock oriented’ models developed by Branson et. al. 1977 postulates the opposite to flow models, that is, that movements in stock prices can cause changes in exchange rates via capital account transactions. The buying and selling of domestic securities in foreign currency (either by foreign investors or domestic residents moving funds from offshore into domestic equities) in response to domestic stock market movements has a flow through effect into the currency market. Although the literature on this subject has examined the relationship between stock prices and exchange rates in various economies, the results have been mixed in terms of the evidence as to which of the above models is most applicable to, or prevalent within an economy.

Ramasamy and Yeung (2005) suggest that the reason for these divergent results is that the nature of the interaction between stock and currency markets is sensitive to the stage of the business cycle and wider economic factors, such as developments or changes in market structures within an economy. So the period of time in which the interaction between stock and currency markets is observed is critical to the end result. This observation is a key platform on which the current study of the interaction between stock prices and exchange rates in Australia is developed given the high degree of co-movement between Australian stock prices and the Australian dollar exchange rate during the period of the study.

This positive relationship is intriguing given the traditional importance of export earnings to the growth profile of the Australian economy. Indeed, this view of the economy lends itself to the flow oriented model, whereby exchange rate appreciation would be expected to cause stock prices to fall. This is also consistent with the conclusions of Mao and Ka (1990), who found that an appreciation in the currency of export-dominant economies tends to negatively influence the domestic stock markets of those economies. Reinforcing this view is the fact that the Australian stock market lacks the depth and liquidity of other larger markets in Asia, Europe and North America. Hence, rises in stock prices here would not normally be expected to result in an appreciation in the value of the Australian dollar as the portfolio balance model postulates, and as is observed by the trends in these variables during the said period.

The results of this study has value for policy makers and market practitioners in that it sheds light on the nature of the strong co-movement between stock prices and the Australian dollar. Indeed, any evidence that stock price movements are found to Granger cause movements in the Australian dollar exchange rate would certainly challenge the traditional view that Australian financial markets reflect the economy's traditional commodity base. Section 2 examines the economic theory surrounding stock and currency market interactions, and also reviews the literature on the interaction between stock prices and exchange rates. Section 3 reviews the data used in the analysis and describes the hypotheses which underpin the study. Section 4 details the methodology employed in the study, and section 5 describes the results of the analysis. Section 6 provides concluding comments.

2. Theory and Literature Review

Classical economic theory hypothesises that stock prices and exchange rates can interact. The first approach is encompassed in 'flow oriented' models (Dornbusch and Fisher 1980), which postulate that exchange rate movements cause stock price movements. In the language of Granger-Sim causality, this is termed as 'uni-directional' causality running from exchange rates to stock prices, or that exchange rates 'Granger-cause' stock prices. This model is built on the macro view that as stock prices represent the discounted present value of a firm's expected future cash flows, then any phenomenon that effects a firm's cash flow will be reflected in that firm's stock price if the market is efficient as the Efficient Market Hypothesis suggests.

One of the earliest distinctions of how exchange rates affected stock prices was according to whether the firm was multinational or domestic in nature (Franck and Young 1972). In the case of a multinational entity, changes in the value of the exchange rate alter the value of the multinational's foreign operations, showing up as a profit or loss on its books which would then affect its share price. Flow oriented models postulate a causal relationship between exchange rates and stock prices. Clearly, the manner in which currency movements influence a firm's earnings (and hence its stock price) depends on the characteristics of that firm. Indeed, today most firms tend to be touched in some way by exchange rate movements, although the growing use of derivatives, such as forward contracts and currency options, might work to reduce the manner in which currency movements affect a firm's earnings.

In contrast to flow oriented models, 'stock oriented' or 'portfolio balance approaches' (Branson et. al 1977) postulate that stock prices can have an effect on exchange rates. In contrast to the flow oriented model - which postulate that currency movements influence a firm's earnings and hence causes change in stock prices - stock oriented models suggest that movements in stock prices Granger-cause movements in the exchange rate via capital account transactions. The degree to which stock oriented models actually explain real world stock and currency market reactions is critically dependent upon issues such as stock market liquidity and segmentation. For example, illiquid markets make it difficult and/or less timely for investors to buy and sell stock, while segmented markets entail imperfections, such as government constraints on investment, high transactions costs and large foreign currency risks, each of which may discourage or hinder foreign investment (Eiteman et. al. 2004).

It is clear from this theoretical review that there are various ways by which stock and currency markets can interact. This makes empirical analysis of the degree and direction of causality between stock prices and exchange rates particularly interesting and has provided the motivation for several studies in examining the interaction between stock prices and exchange rates. Although theory such as the flow and portfolio models and the money demand equation hypothesise that a relationship should exist between exchange rates and stock prices, the evidence provided by the literature on this subject matter has been mixed.

A study by Aggarwal (1981) provided some evidence in support of the flow model. This study examined the relationship between exchange rates and stock prices by looking at the correlation between changes in the US trade-weighted exchange rate and changes in US stock market indices each month for the period 1974 to 1978. The study found that the trade-weighted exchange rate and the US stock market indices were positively correlated during this period, leading Aggarwal (1981) to conclude that the two variables interacted in a manner consistent with the flow model. That is, movements in the exchange rate could directly affect the stock prices of multinational firms by influencing the value of its overseas operations, and indirectly effect domestic firms through influencing the prices of its exports and/or its imported inputs. Soenen and Hennigar (1988) found a significant negative correlation between the effective value of the US dollar and changes in US stock prices using monthly data between the period from 1980 to 1986. While this finding is in contrast to Aggarwal (1981), who found a positive correlation, it still provides evidence in support of the flow model.

While the above studies focussed exclusively on the United States, a later study by Ma and Kao (1990) examined the relationship between exchange rates and stock prices in six industrialised economies, including the UK, Canada, France, West Germany, Italy and Japan. Using monthly data between January 1973 and December 1983, the authors tested the degree of stock price reaction to exchange rate changes in each of the above jurisdictions. Their findings were consistent with the flow model, leading the authors to conclude that the relationship between exchange rates and stock prices hinged on the extent to which an economy depended on exports and imports. These early studies were useful in establishing a foundation for further studies on the interaction between exchange rates and stock prices, but they were limited in that they only applied simple regression analysis to establish a correlation between the variables, or only tested the ‘reaction’ of one variable to changes in the other.

Bahmani-Oskooee and Sohrabian (1992) were one of the first to utilise tests of causality in examining the relationship between stock prices and exchange rates in the US context. They also used a much longer time period (15 years) and also utilised tests of co-integration. Co-integration techniques allow one to establish if the variables share a long-run relationship, as the interactions uncovered by the Granger (1969), Sim (1972) method are intrinsically short-run in nature. Using monthly data of the US S&P 500 index and the effective exchange rate of the US dollar, the authors employed an autoregressive framework, finding that US stocks and the exchange rate shared a dual or bi-causal relationship (i.e. changes in the exchange rate effected stock prices and vice versa) in the sample period, 1973 to 1988. These results would seem to affirm both the portfolio and flow models. Meanwhile, the co-integration test (carried out using the methodology outlined by Engle and Granger 1987) found little evidence that the variables shared any relationship in the long-run.

A study by Ajayi et al. (1998) examined the relationship between exchange rates and stock prices among developing and developed nations. Like Bahmani-Oskooee and Sohrabian (1992) and Yu Qiao (1997), Ajayi et al. (1998) used Granger-Sim causality to examine the relationship between movements in the stock price indexes and movements in the exchange rates. Importantly, the findings of Ajayi et al. (1998) appeared to have uncovered a consistency in the relationships between stock prices and exchange rates among developed economies, which were in accordance with the portfolio model. On the contrary, the patterns of causality among the emerging Asian economies examined were mixed.

No significant causal relationships were detected in Hong Kong, Singapore, Thailand or Malaysia. Notably, this result is again in contrast with those of Yu Qiao (1997), which found uni-directional causality from exchange rates to stock returns in Hong Kong, although the findings of Ajayi et al. (1998) are consistent with those of Yu Qiao (1997) in that neither study found a relation between stock prices and exchange rates for Singapore. Ajayi et al. (1998) attributed the difference in their findings between developed and emerging economies to structural differences between the currency and stock markets of each. Specifically, the authors suggest that markets are likely to be more integrated and deep in advanced economies, and that emerging markets tend to be much smaller, less accessible to foreign investors and more concentrated. The authors also made note of wider risks such as political stability and the legislative environments which might make investment in emerging markets less attractive. Hence, the study concluded that activity in emerging stock markets tends to portray wider macroeconomic factors less strongly than in developed markets and as a result, these markets tend to have weaker linkages to the currency market.

While most literature in this context had previously focussed on developed markets or on comparisons between developed and emerging markets, the Asian financial crisis of the late 1990s sparked interest in the interaction between currency and stock markets solely in developing markets. Indeed, the Asian crisis was characterised by plunging currency and stock markets within South East Asia. Granger et al. (2000) was one such study which focussed on this region. It examined the interaction between stock and currency markets in Hong Kong, Indonesia, Japan, South Korea, Malaysia, the Philippines, Singapore, Thailand and Taiwan, all of which were effected by the crisis. The empirical results showed that, with the exception of Singapore (where exchange rate changes led stock prices as per the flow model), all countries displayed little evidence of interaction between currency and stock markets during the first period. In the second period, the exchange rate in Singapore again led its stock market, while the reverse (as per the portfolio model) was evident in the cases of Taiwan and Hong Kong.

The contrasting results across the body of literature regarding this issue suggest that there is no underlying or intrinsic causal relationship between exchange rates and stock markets across jurisdictions. Rather, the differing causal relationships uncovered through empirical analysis implies that the interaction between currency and stock markets are influenced by the business cycle and different economic structures present within individual countries, meaning causality between the two financial variables is sensitive to the time period in which the analysis is undertaken. This view is confirmed by Ramasamy and Yeung (2005), who suggest that causality is unique within jurisdictions, within specific time periods and is even sensitive to the frequency of data utilised. In their study, the authors examined the degree of exchange rate and stock price causality in the same nine Asian economies studied in Granger et al. (2000), but during the period 1 January, 1997 to 31 December, 2000 – the entire period of the Asian currency crisis. The empirical results of Ramasamy and Yeung (2005) differ from those of Granger et al. (2000). While Granger et al. (2000) found a bi-causality for Malaysia, Singapore, Thailand and Taiwan, Ramasamy and Yeung (2005) found that stock prices lead exchange rates for these countries. On the other hand, Granger et al. (2000) found that stock prices lead exchange rates for Hong Kong, but a bi-causality was detected by Ramasamy and Yeung (2005).

The current study on the interaction between exchange rates and stock prices in the Australian context differs from previous work in a number of ways. Firstly, it employs a current data set. Secondly, it does not seek to postulate the existence of some underlying causal relation between stock prices and exchange rates as early studies on thus subject have sought to. Rather, recognising the robust and changing dynamics between these variables, this study examines how these variables interacted during the sample period. This is done specifically with a view to challenging the traditional export-dependent view of the Australian economy which lends itself to the flow oriented model of stock price and exchange rate interaction. Hence, the focus is on ascertaining the significance of the strength and direction of the influence of Australian stock price movements on the Australian dollar exchange rate in the study period. Given the importance of both equity and currency markets to the functioning of an economy, the empirical results provide useful information to market practitioners and policy makers on the interaction between stock prices and exchange rates.

3. Data and Hypotheses

This study examines the interaction between Australian stock prices and the Australian-USD exchange rate from 2 January 2003 to 30 June 2006. Daily observations of Australian stock prices and the Australian-US dollar exchange rate was gathered and analysed using the *EViews 4* statistical package.

Stock prices are measured using the daily (five days a week) closing prices of the All Ordinaries stock price index. The All Ordinaries index is chosen as it is considered to be Australia's leading share market indicator, representing the 500 largest companies listed on the Australian Stock Exchange. Level stock price series is expressed by the symbol 'SP' and first difference data for SP (denoted SP1) is equal to Log (SP_t/SP_{t-1}).

Similarly for the Australian-US dollar exchange rate, five day-a-week daily, nominal observations at the close of market are gathered from the Reserve Bank of Australia. The exchange rate is expressed in terms of the number of Australian dollars per unit of US currency. Using this form of quotation is consistent with previous empirical studies (Granger et. al. 2000, Ajayi et. al. 1998 and Nieh and Lee 2001). The level exchange rate series is expressed by the symbol 'EX' and first difference data for EX (denoted EX1) is equal to Log (EX_t/EX_{t-1}). Although both sets of data are at close of trade in Australian markets, some date synchronisation was required to ensure that the trading days of both time-series matched. In total, there are 877 observations in the sample data series.

Three hypotheses are explored in this study in examining the interaction between stock prices and exchange rates in Australia during the period in question. Each of the ensuing hypotheses is stated in the null format. Both the flow and portfolio models postulate that a relationship exists between stock prices and exchange rates. Hence, the first step in the empirical analysis of this study is to investigate the broad relationship between stock prices and exchange rates using OLS regression analysis. Because the exchange rate series in this study is expressed in terms of Australian dollars per unit of US currency (i.e. direct quotation), a negative correlation between stock prices and exchange rates would be indicative of a positive co-movement between the variables. Hence, the first hypothesis is as follows:

H_o 1a: There is a significant positive relationship between level series of Australian stock prices and the Australian dollar exchange rate.

H_o 1b: There is a significant positive relationship between first differences in Australian stock prices and first differences in the Australian dollar exchange rate.

According to Brooks (2002), if one financial variable significantly and consistently influences another, the two variables should be co-integrated. In the context of this study, a co-integrating relationship will provide evidence that

Australian stock price movements significantly explain expected movements in the Australian dollar exchange rate over the long term. The second hypothesis follows:

H_o 2a: There are no co-integrating relationships between level series of Australian stock prices and the Australian dollar exchange rate.

H_o 2b: There are no co-integrating relationships between changes (first differences) in Australian stock prices and changes (first differences) in the Australian dollar exchange rate.

According to Granger (1981, 1988), if a pair of variable series are co-integrated, the bi-variate co-integrating system must possess a causal order in at least one direction. If the evidence is such that exchange rate variability is linked to stock price movements, it can also be shown that the change in the exchange rate either lags or leads movements in stock prices. Based on this theory, the third and most important hypothesis of the study is:

H_o 3a: There is no directional causality between the level series of Australian stock prices and the level series of the Australian dollar exchange rate.

H_o 3b: There is no directional causality between the changes in Australian stock prices and changes in the Australian exchange rate.

4. Methodology and Results

The level series are tested first and an unrestricted vector autoregression (VAR) model is applied. A VAR model is required to investigate causality as standard regression models are limited to examining the degree of correlation between two variables and can not establish a causal connection between the variables. Standard regression analysis assumes that the relationship between the dependent variable and the explanatory variable is contemporaneous, that is, that the variables interact at the same point in time (Brooks 2002). Hence, the standard regression framework is inadequate to test the causal relationship between variables. The regression of the Australian dollar exchange rate against Australian stock prices is analysed in order to examine the relationship between the two variables. The variables have been previously defined. The regression undertaken is as follows:

$$\mathrm{Log}\ (Ex_t) = a + \beta_1 \mathrm{Log}\ (SP_t) + e_t \qquad (1)$$

According to Brooks 2002, the key premise of causal analysis lies in the assumption that the variables are non-contemporaneous in that the value of a variable in the current time period is influenced by its value in some prior time period. This difference is known as the lag. This is essentially the foundation of autoregressive models. The standard auto-regression process is based on the standard regression process, except that the value of the dependent variables in the system depends only on the lagged values of the dependent variable plus an error term. Extending this model one step further gives the vector autoregressive model which applies when the dependent variable in the system not only depends on its own lags, but also on the lags of another explanatory variable.

4.1 Ordinary Least Squares Regression

(a) Level series

When OLS regression analysis is run on the level series data as in equation 1, the adjusted R-square value is found to be 0.4818, with an F-value at 815.7186 (highly significant at $p = 0.0000$). With regard to the coefficients (the significance level is in parenthesis), the intercept t-statistic is 32.2487 ($p = 0.0000$) and the stock price t-statistic is -28.5607 ($p = 0.0000$). See Table A in Appendix 1 for more details on this regression.

The correlation between stock prices and exchange rates is highly negatively correlated (-0.6946) during the sample period. Note that the exchange rate series is expressed in terms of direct quotation (Australian dollars per unit of US currency), and therefore a decline in the exchange rate in direct quotation terms is indicative of an appreciation of the Australian dollar. Therefore, a negative correlation between the two variables is indicative of a positive co-movement between them.

The DW statistic is equal to 0.0152, which is far less than the adjusted R-square value (0.4818). This would indicate that if the level series of stock prices is integrated, the regression may be spurious. In addition, the DW statistic is very close to zero and substantially less than two. This indicates a high degree of positive serial correlation in the series which supports the rejection of the DW test null hypothesis of zero autocorrelation, indicating that there may be a high degree of time dependence in the series.

The relatively high adjusted R-square value and the significance of the coefficients in the above regression provide some support for accepting the null hypotheses 1a, which states there is a significant positive relationship between the level series of Australian stock prices and the Australian dollar exchange rate. But this evidence needs to be treated with caution in light of the spurious nature of the regression.

(b)First Differences

As reported in the section on *Data and Hypothesis*, first difference data for SP is denoted 'SP1' and is equal to Log (SP_t/SP_{t-1}). Similarly, first difference data for EX is denoted EX1 and is equal to Log (EX_t/EX_{t-1}).

When OLS regression analysis is run on the first difference data (in the same form as in equation 1 except with SP1 and EX1), it is found that the adjusted R-square value falls to just 0.00428, with an F-statistic of 4.7687 which is significant at the 5 per cent level (with p = 0.02924). See Table B in Appendix 1 for detailed results on this regression.

It also evident from the first differences regression that exchange rate changes are negatively related to changes in stock prices, with a t-statistic of -2.1837 (where p = 0.0292) which is significant at the 5 per cent level. As reported earlier, the exchange rate series is expressed in terms of direct quotation (Australian dollars per unit of US currency) and therefore, a decline in the exchange rate in direct quotation terms is indicative of an appreciation of the Australian dollar. Therefore, a negative relation between the two variables is indicative of a positive co-movement between them.

Notably, the DW statistic at 1.9987, which is greater than the adjusted R-square value and sufficiently higher than zero and close to two to conclude that the regression may be relied upon. That is, unlike the level series regression the relationship uncovered by the regression of the first differences series is unlikely to be spurious. Nevertheless, it is apparent that substantial information has been lost in the first differencing process, given the very low adjusted R-square value.

Therefore, Null Hypothesis 1b which states that there is a significant positive relationship between first differences in Australian stock prices and the Australian dollar exchange rate, cannot be rejected. Again, as in the case of the level series, this result needs to be treated with caution due to the low explanatory power of the model.

4.2 Testing for Unit Roots

(a) Level series

Each of the level series was tested for a unit root using the ADF test. The results indicate that the level series of stock prices and exchange rates are non-stationary processes at the 1 per cent ADF critical level. See Tables A, B and C in Appendix 2 for more details on these ADF test results.

In the case of stock prices, the ADF statistic was 0.0712 which compares against the 1 per cent, 5 per cent and 10 per cent critical values of -3.5000, -2.8918 and -2.5827 respectively. As this ADF test statistic is greater than the 1 per cent, 5 per cent and 10 per cent critical values, the ADF test-null hypotheses of a unit root is accepted.

In the case of the exchange rate series, the ADF statistic was -2.9038, which is greater than the 1 per cent critical value, but less than the 5 per cent and 10 per cent critical values. Hence, the ADF test of a unit root is accepted at the 1 per cent critical level.

When the residual of the stock price regression was tested for a unit root, it was found that the ADF test statistic was less than the 1 per cent, 5 per cent and 10 per cent critical values at -12.7339, meaning that the residuals are stationary. Therefore, some evidence is provided to suggest that there are stationary processes in the level series regression even if the variables themselves are non-stationary.

(b)First Differences

Each of the first differenced series was tested for a unit root using the ADF test. The results indicate that the first differenced series of stock prices and exchange rates are stationary processes at the 1 per cent, 5 per cent and 10 per cent ADF critical levels. See Tables D and E in Appendix 2 (page 13) for detailed results on the ADF tests.

In the case of stock prices, the ADF statistic was -12.9268 which compares against the 1 per cent, 5 per cent and 10 per cent critical values of -3.5000, -2.8918 and -2.5827 respectively. As this ADF test statistic is lower than the 1 per cent, 5 per cent and 10 per cent critical values, the ADF test null hypotheses of a unit root is rejected.

In the case of the exchange rate series, the ADF statistic was -12.8839, which is also less than the 1 per cent, 5 per cent and 10 per cent critical values indicating rejection of a unit root.

When the ADF test was applied to the error terms of the first difference regression of stock prices, the test statistic was found to be -13.1280 which is also less than the 1 per cent, 5 per cent and 10 per cent ADF critical values, meaning that the residuals are also stationary.

As evidence is provided that the first difference data are non-integrated, non-stationary processes, checks of co-integration are not required. Null Hypothesis 2b, which states that there are no co-integrating relationships between the first difference series, therefore cannot be rejected.

4.3 Heteroskedasticity

(a) Level series

Before estimating any ARCH type models, the Engle (1982) test for ARCH effects is first carried out to ensure this class of models is appropriate for the data. The ARCH LM test is undertaken on the level series regression of exchange rates against stock prices to test the null hypothesis that there is no ARCH up to order five in the residuals.

The test results show that both the F-statistic (4539.163) and the LM statistic (839.9502) are very significant (both with p-values of 0.0000), suggesting the presence of ARCH in the level series data.

An ARCH model was then applied to the regression of stock prices against exchange rates. The ML-ARCH model was applied to the data of 877 observations with convergence achieved after 230 iterations. The variance equation coefficients for ARCH 1 and GARCH 1 respectively were 1.0033 and -0.0293. The sum of the coefficients is close to unity (approximately 0.99), meaning that shocks to the conditional variance are persistent in the data. This confirms autoregressive conditional heteroskedasticity is present in the level series data.

With the OLS regression re-specified as an ARCH-ML model, the adjusted R-square value falls to 0.4765 with an F-statistic of 200.4115, which is highly significant (p = 0.0000). The z-statistic for the stock price is -76.9117, which is highly significant (p = 0.0000). However, at 0.0153 the DW test statistic remains near zero and less than two, indicating that the regression results remain spurious.

The use of the ARCH model again provides evidence to support the acceptance of the null hypotheses 1a, which states that there is a significant positive relationship between the level series of Australian stock prices and the Australian dollar exchange rate. However, this evidence again needs to be treated with caution in light of the spurious nature of the regression. See Tables A and B in Appendix 3 (page 17) for more information on the results of the ARCH-LM test and ARCH ML model.

(b)First Differences

The ARCH LM test is undertaken on the first differences regression of exchange rates against stock prices to test the null hypothesis that there is no ARCH up to order five in the residuals. The test shows that both the F-statistic (0.9760) and the LM statistic (4.8866) are not significant, with p-values of 0.4313 and 0.4298 respectively (Table C, Appendix 3). This suggests there is no presence of ARCH in the first differenced data series.

4.4 Co-Integration

With the level series established as being integrated, non-stationary processes, the study then proceeded to check if the level series are co-integrated. An unrestricted VAR model was applied to the level series data, with lag intervals of between 1 and 6. The VAR is expressed as follows:

$$\mathbf{LogSP_t = \alpha_0 + \sum_{i=1}^{P}\alpha_i\, LogSP_{t-i} + \sum_{i=1}^{P}\beta_i\, LogEX_{t-i} + \mu_t} \qquad \mathbf{(2)}$$

$$\mathbf{LogEX_t = \varphi_0 + \sum_{i=1}^{P}\varphi_i\, LogEX_{t-i} + \sum_{i=1}^{P}\gamma_i\, LogSP_{t-i} + v_t} \qquad \mathbf{(3)}$$

A critical issue in using VAR models is the choice of lag length. Prior research (notably Granger et al. 2000, Ajayi et al. 1998, and Ramasamy and Yeung 2005) intuitively employed a one-day lag length in their models sighting the fact that the highly integrated nature of financial markets is likely to mean that the flow of information to investors is very efficient, allowing them to react quickly to developments in either of the markets.

An unrestricted VAR model for the first differenced data is specified in order to undertake Granger-Sim causality. The VAR model, which is specified with lag intervals of between 1 and 6, is expressed as follows:

$$\mathbf{SP1_t = \alpha_0 + \sum_{i=1}^{P}\alpha_i\, SP1_{t-i} + \sum_{i=1}^{P}\beta_i\, EX1_{t-i} + \mu_t} \qquad \mathbf{(4)}$$

$$\mathbf{EX1_t = \varphi_0 + \sum_{i=1}^{P}\varphi_i\, EX1_{t-i} + \sum_{i=1}^{P}\gamma_i\, SP1_{t-i} + v_t} \qquad \mathbf{(5)}$$

Note again that SP1 and EX1 denote the first difference data, with SP1 equal to

Log (SP_t/SP_{t-1}), while EX1 is equal to Log (EX_t/EX_{t-1}).

This study employs maximum likelihood tests to establish the optimum lag length. Under this approach, the optimum length is the one in which the value of most information criteria are minimised. Lag length criteria tests were undertaken for lengths of between 1 and 8 for the sample period, with most criteria minimised at 1 lag length for the level series. Table A in Appendix 4 shows the results of this maximum likelihood test.

When lag order selection criteria are applied to the first difference data, it is found that Akaike's information criteria is at its minimum at 0 lags, with a value of -14.4907. Other information criteria, such as the Schwarz information criterion and the Hannan-Quinn information criterion are also at their minimum values at zero lags.

The sequential modified LR test statistic is minimised at four lag lengths, and using the lag exclusion Wald test, all lags except lag four are also rejected, where the joint

Chi-Square value of lag four is 8.6088. This value is significant at the 10 per cent level, with p = 0.0716. See Tables B and C in Appendix 4 for detailed results of maximum likelihood and lag exclusions tests. Pair-wise Granger causality tests were then run on the first differenced series for lags one to six.

At one lag length, uni-directional causality is found at the 5 per cent level, with causality running from stock prices to the exchange rate with an F-statistic equal to 5.3983 and a significance level of 0.0203. Meanwhile, the F-statistic for the test of causality running from exchange rates to stock prices at one lag is 0.4527, with a significance level of 0.5012.

At two lags, uni-directional causality from stock prices to exchange rates is evident again, although at a lower significance level. Here, causality is again seen to run from stock prices to the exchange rate, but only at a significance level of 10 per cent, as p = 0.08157. Meanwhile, the significance of the test of causality running from exchange rates to stock prices at two lags is 0.1410.

Beyond two lags, the direction of causality appears to switch from stock prices to exchange rates, to exchange rates influencing stock prices. Indeed, for three lags the significance of the test for stock prices Granger-causing exchange rates rises to 0.2167, while the test of significance for exchange rates Granger-causing stock prices falls to 0.0663 – significant at the 10 per cent level only. However, at the optimal 4 lags uni-directional causality from exchange rates to stock prices is evident at the 5 per cent significance level. In this case, the F-statistic is equal to 2.4506 and the p-value is equal to 0.0446. Meanwhile, the F-statistic for the test of causality running from stock prices to exchange rates at four lags is 1.416, with a significance level of 0.2265. At five lags, exchange rates are again seen to Granger cause stock prices, but only at the 10 per cent level of significance (p = 0.09588), while little evidence of causality is demonstrated from stock prices to exchange rates (p = 0.26012). At six lags, there does not appear to be any significant causal relationship in the first difference data for the model of stock prices and exchange rates.

The lag structure/AR Roots Test was also applied as a test of the VAR's stability condition. *EViews 4* undertakes the test and reports the roots of the characteristic autoregressive polynomial. The VAR is considered stable or stationary if all roots have a modulus less than one and lie inside the unit circle. The results of this test (Table A, Appendix 5) show that the unrestricted VAR satisfies the stability condition, as all polynomial roots have a value of less than one and lie within the unit circle. The lag structure/AR Roots Test was again applied as a test of the VAR's stability condition. The results of this test (Table B, Appendix 5) show that the unrestricted VAR satisfies the stability condition, as all polynomial roots have a value of less than one and lie within the unit circle.

When the Johansen co-integration test was applied (assuming an intercept and a linear deterministic trend in the data), it was found that the test null hypothesis of zero co-integration could be rejected. For the test of zero co-integrating relations (Appendix 6), the trace statistic (32.2484) and maximum eigenvalue statistic (25.9395) were each greater than the 5 per cent and 1 per cent critical values. In contrast, the trace and maximum eigenvalue statistics for the test of at least one co-integrating relation were both less than the 5 per cent and 1 per cent critical values. Therefore, there is evidence to support the rejection of the null hypotheses 2a of no co-integrating relationship between the level series data. It is therefore evident that even though the level series are integrated (i.e. contain one unit root or I (1)), a linear combination of these I (1) variables becomes I (0) when the variables are co-integrated. This indicates that the level series of Australian stock prices and exchange rates share a long-run relationship.

4.5 Causality

Pair-wise Granger causality tests were run on the level series at the optimal one lag length. It is found that the level series of the variables are independent during the sample period at the adopted 5 per cent level. The F-statistic for the test of causality running from stock prices to the exchange rate at one lag length is 1.0951, with a significance level of 0.2956. Meanwhile, the F-statistic for the test of causality running from exchange rates to stock prices is 0.0478, with a significance level of 0.8268.

However, uni-directional causality is found at the 5 per cent level at a lag length of two, with causality running from stock prices to exchange rates. The F-statistic of this test is equal to 3.3122, with a significance of 0.0368. At three, four, five and six lag lengths (the other lengths in the VAR), the variables again appear independent, with no significant causal relationship. Notably, at five lengths there is some evidence of causality running from exchange rates to stock prices, but only at the 10 per cent level of significance. Detailed results of Granger-Sim causality on the level series at various lag lengths are provided in Appendix 7. It is apparent therefore that causality is one-way, running from stock prices to exchange rates at a two-day lag, although one-day is the optimal lag. This would suggest a relationship in line with the portfolio model whereby stock price movements influence exchange rates via capital account transactions.

Support is therefore provided for the rejection of the Null Hypothesis 3a, that there is no directional causality between the level series of Australian stock prices and the level series of the Australian dollar exchange rate.

In summary, it is apparent from the first difference analysis, that exchange rate changes cause stock price changes with an optimal four-day lag. However, for lags less than three days, the opposite is true. At one and two lag lengths, causality runs from stock prices to exchange rates with the significance of causality at the one lag interval the strongest of any other causality test in this study, at p = 0.0203. Evidence is therefore provided for the rejection of Null Hypothesis 3b, that there is no directional causality between changes in Australian stock prices and changes in the Australian exchange rate.

5. Conclusion

This paper sought to investigate the relationship between stock prices and exchange rates in Australia and then expanded the analysis to investigate the changes in these key economic variables and the relationships between those changes through the process of the Granger-Sim causality test. The motivation for this paper was to test the degree of interaction between stock prices and exchange rate movements in Australia in the period 2 January 2003 to 30 June 2006. This period is somewhat intriguing considering that the value of the Australian stock market increased by two-thirds during the sample period, while the Australian dollar exchange rate appreciated by as much as 32 per cent relative to the US dollar. This would imply a strong positive relationship existed between the two variables during the period in question, although it is not known if the two markets interacted or caused movements in the other during this period as postulated by economic theory.

In terms of the broad relationship between the two variables, the level series regression results support the above observation of a short-term positive relationship between the Australian dollar exchange rate and stock prices during the sample period. However, these results should be treated with caution in light of the spurious nature of the OLS and ARCH regressions estimated. When first differences are examined, evidence of a positive relationship between these variables remains, although these results again need to be treated with caution in light of the low explanatory power of the first differences OLS regression. Evidence is also provided for co-integration of the subject level variables, implying that the variables not only appear to be related in the short-run of the sample period, but that longer-term expectations also play some part in activity in stock and currency markets in Australia. With the broader short and longer term relationship between stock prices and exchange rates established, the pair-wise Granger-Sim causality tests provide a deeper insight into the degree of interaction between the two variables during the sample period.

Literature on the study of the relationship refers to three methodologies that relate first to a flow oriented model, and then a portfolio balance approach and finally a cointegration and causality approach. Examples of flow oriented models in developed country studies are those of Franck and Young (1972), Dornbuscher and Fisher (1980) Aggarwal (1981) and Ma and Kao (1990) who generally found exchange rates driving stock prices. In the portfolio balance approach Branson et al (1977) and Eiteman et al (2004) had similar findings and in the latter case the direction of influence of the exchange rate was largely dependent on factors such as stock market liquidity and segmentation. Cointegration and causality studies by Bahmani, Oskooee and Sohrabian (1992), Yu Qiao (1997) Ajayi et al (1998), Granger et al (2000) and Ramasamy and Yeung (2005) have found a mixture of evidence of causality running from the exchange rate to stock prices and vice versa with the latter finding dual causality.

From this analysis in this paper, it is evident past evidence of causality running from the exchange rate to the stock price is not supported. The study finds that a significant (at the 5 per cent level) uni-directional causal relationship exists between the variables, with stock price changes found to Granger cause changes in the Australian dollar exchange rate during the sample period. This demonstrates a relationship more consistent with the portfolio balance approach. Although causality was evident in the opposite direction (i.e. from exchange rates to stock prices), the degree of causality from stock prices to exchange rates was the most statistically significant in the analysis of both the level series and first differences series.

In terms of policy implications of this research it may be that commodity prices are the ‘missing link’ between these two variables, with high commodity prices often bolstering both the Australian stock market and the domestic currency given the economy’s strong commodity base. Although this study has shown that the broad relation between stock prices and the exchange rate during the sample period to not be spurious (on first differences), the regression model utilised lacks explanatory power as noted above. Again, in terms of policy implications there are wider issues, which suggest that commodity prices might not have had a significant effect on these variables during the sample period. Importantly, the relationship between commodity prices and the Australian dollar broke down at the turn of the decade (Note 2) and this ‘marriage’ has not yet resumed at the time of writing this paper. Furthermore, commodity prices only began rising sharply in early 2005 while the exchange rate and stock prices trended higher together before this period. Even as commodity prices have accelerated sharply since 2005, the dollar has continued to trade between a narrow band, while stock prices have continued to increase.

In addition, the findings of this study are also consistent with broader macroeconomic trends given that the sample period studied was characterised by relatively low levels of economic growth in the United States, Japan and Europe. In comparison to the Australian economy experienced broad-based growth on the back of strong consumer spending and high investment into dwellings and fixed capital - as illustrated by the positive gains on the domestic stock market during this period. This would have made the Australian equity market a more attractive proposition for domestic and foreign investors, which lends itself to stock and currency markets interacting in a manner postulated by the portfolio model, as uncovered in this study. The result of this study has implications for both policy-makers and market practitioners alike, as it suggests that Australian stock prices and the exchange rate can interact in a manner in accordance with the portfolio model, whereby stock price movements influence exchange rates via capital account transactions. This is a shift from the traditional view of the Australian economy as an export-dependent economy. This notion lends itself more to the flow oriented model, which implies that exchange rate movements should cause movements in stock prices, or that a sharp appreciation in the Australian dollar (as is the case in this sample period) should negatively influence the domestic stock market. Therefore, the results of this study imply that the Australian stock market has the depth and liquidity to adequately compete for domestic and foreign capital against other larger markets in Asia, Europe and North America.

Future research will more closely focus on dividing a long sample period into periods of stability and volatility and re examine the connection between the variables in differing periods of economic activity according to the business cycles uncovered and in addition control for commodity prices will be introduced.. In light of the present global financial crisis it is evident that stock prices in the US are driving exchange rates in Australia as US dollar investments are repatriated back to the USA by risk averse international investors. The redemption of Australian dollar investments has created a downward movement in the Australian dollar. A longer period of data will need to be studied to provide evidence for this Australia dollar pressure and whether or not the stock market in another major developed country is the driver.

References

Aggarwal, R. (1981). “Exchange Rates and Stock Prices: A Study of the US Capital Markets under Floating Exchange Rates”, *Akron Business and Economic Review*, vol. 12, pp. 7–12.

Ajayi, R.A., Friedman, J. and Mehdian, S.M. (1998). “On the Relationship between Stock Returns and Exchange Rates: Tests of Granger Causality”, *Global Finance Journal*, vol. 9(2), pp. 241–51.

Bahmani-Oskooee, M. and Sohrabian, A. (1992). “Stock Prices and the Effective Exchange Rate of the Dollar”, *Applied Economics*, vol. 24, pp. 459–464.

Branson, W., Halttunen, H., and Masson, P. (1977). “Exchange rate in the short run: the dollar Deutsche mark rate”, *European Economic Review*, 10, pp. 303–324.

Brooks. C, 2002, *“Introductory Econometrics for Finance”,* Cambridge University Press. Cambridge, United Kingdom.

Dornbusch, R. and Fischer, S. (1980). “Exchange Rates and the Current Account”, *American Economic Review*, Vol. 70, No. 5 (Dec., 1980), pp. 960-971

Granger, C.W., Huang, B. and Yang, C. (2000). “A Bivariate Causality between Stock Prices and Exchange Rates: Evidence from Recent Asian Flu”, *Quarterly Review of Economics and Finance*, vol. 40, pp. 337–354.

Granger, C.W.J. (1969). “Investigating causal relations by econometric models and cross-spectral methods”. *Econometrica*, 37, pp. 428-438.

Krugman, P. and Obstfeld, M., (1997). *“International Economics Theory and Policy”,* 4th Edition, Addison-Wesley, New York, United States.

Ma, C.K. and Kao, G.W. (1990). “On Exchange Rate Changes and Stock Price Reactions”, *Journal of Business Finance and Accounting,* vol. 17(2), pp. 441–449.

McTaggart, D., Findlay, C. and Parkin, M. (1996). *“Economics”,* second edition, Addison-Wesley, Sydney.

Muhammad, N. and Rasheed, A. c. (2001). *“Stock Prices and Exchange Rates: Are they Related? Evidence from South Asian Countries”,* Department Economics & Finance - Institute of Business Administration, Karachi University Campus.

Ramasamy, B. and Yeung, M.C.H. (2005). “The Causality between Stock Returns and Exchange Rates: Revisited”. *Australian Economic Papers*, Vol. 44, No. 2, pp. 162-169, June 2005.

Sim, C. (1972). “Money, income and causality”, *American Economic Review*, 62, pp. 540-552.

Soenen, L.A. and Hennigar, E.S. (1988). “An Analysis of Exchange Rates and Stock Prices: The US Experience Between 1980 and 1986”, *Akron Business and Economic Review*, vol. 19(4), pp. 71–76.

Yu, Qiao. (1997). “Stock Prices and Exchange Rates: Experience in Leading East Asian Financial Centres: Tokyo, Hong Kong and Singapore,” *Singapore Economic Review*, 41, pp. 47-56.

Notes

Note 1. The value of the Australian stock market increased by two-thirds during this period, while the Australian dollar exchange rate appreciated by as much as 32 per cent relative to the US, implying a strong positive relationship existed between the two variables.

Note 2. Reserve Bank of Australia, Statement on Monetary Policy, February 2004. Page 17.

APPENDIX

Appendix 1 – OLS Regression Results

Appendix 2 – Testing for Unit Roots

Appendix 3 – Heteroskedasticity Tests

Appendix 4 – Maximum Likelihood Tests and Lag Exclusions Tests

Appendix 5 – VAR Stability Condition Test – AR Roots Table

Appendix 6 – Johansen Co-Integration Test

Appendix 7 – Pair-wise Granger Causality Tests

Appendix 1 – OLS Regression Results

A. Level Series Regression

Dependent Variable: LNEX
Method: Least Squares
Date: 10/07/06 Time: 15:11
Sample: 877
Included observations: 877

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	2.886034	0.089493	32.24877	0.0000
LNSP	-0.310427	0.010869	-28.56079	0.0000
R-squared	0.482469	Mean dependent var		0.330641
Adjusted R-squared	0.481877	S.D. dependent var		0.079328
S.E. of regression	0.057101	Akaike info criterion		-2.885704
Sum squared resid	2.852982	Schwarz criterion		-2.874811
Log likelihood	1267.381	F-statistic		815.7186
Durbin-Watson stat	0.015280	Prob(F-statistic)		0.000000

B. First Difference Regression

Dependent Variable: EX1
Method: Least Squares
Date: 10/15/06 Time: 08:53
Sample(adjusted): 2 877
Included observations: 876 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.000266	0.000235	-1.130305	0.2587
SP1	-0.085092	0.038966	-2.183748	0.0292
R-squared	0.005427	Mean dependent var		-0.000316
Adjusted R-squared	0.004289	S.D. dependent var		0.006945
S.E. of regression	0.006930	Akaike info criterion		-7.103645
Sum squared resid	0.041973	Schwarz criterion		-7.092743
Log likelihood	3113.397	F-statistic		4.768755
Durbin-Watson stat	1.998759	Prob(F-statistic)		0.029246

Appendix 2 – Testing for Unit Roots

A. Level Series Regression – Log of SP

ADF Test Statistic	0.071217	1% Critical Value*	-3.4405
		5% Critical Value	-2.8653
		10% Critical Value	-2.5688

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LNSP)
Method: Least Squares
Date: 10/08/06 Time: 09:25
Sample(adjusted): 6 877
Included observations: 872 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNSP(-1)	8.20E-05	0.001151	0.071217	0.9432
D(LNSP(-1))	-0.042473	0.034023	-1.248369	0.2122
D(LNSP(-2))	0.007044	0.034087	0.206660	0.8363
D(LNSP(-3))	0.029949	0.034121	0.877738	0.3803
D(LNSP(-4))	0.077086	0.034067	2.262808	0.0239
C	-0.000139	0.009479	-0.014624	0.9883

R-squared	0.008285	Mean dependent var	0.000577
Adjusted R-squared	0.002559	S.D. dependent var	0.006015
S.E. of regression	0.006007	Akaike info criterion	-7.384920
Sum squared resid	0.031249	Schwarz criterion	-7.352093
Log likelihood	3225.825	F-statistic	1.446994
Durbin-Watson stat	1.983016	Prob(F-statistic)	0.204969

B. Level Series Regression – Log of EX

ADF Test Statistic	-2.903822	1% Critical Value*	-3.4405
		5% Critical Value	-2.8653
		10% Critical Value	-2.5688

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(LNEX)
Method: Least Squares
Date: 10/08/06 Time: 09:28
Sample(adjusted): 6 877
Included observations: 872 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
LNEX(-1)	-0.008737	0.003009	-2.903822	0.0038
D(LNEX(-1))	0.003401	0.033966	0.100128	0.9203
D(LNEX(-2))	0.002522	0.033890	0.074412	0.9407
D(LNEX(-3))	-0.059829	0.033842	-1.767861	0.0774
D(LNEX(-4))	0.029432	0.033907	0.868037	0.3856
C	0.002575	0.001019	2.528076	0.0116

R-squared	0.013949	Mean dependent var	-0.000297
Adjusted R-squared	0.008256	S.D. dependent var	0.006944
S.E. of regression	0.006915	Akaike info criterion	-7.103357
Sum squared resid	0.041411	Schwarz criterion	-7.070531
Log likelihood	3103.064	F-statistic	2.450141
Durbin-Watson stat	1.990278	Prob(F-statistic)	0.032327

C. Level Series Regression –Residuals

ADF Test Statistic	-12.73394	1% Critical Value*	-3.4405
		5% Critical Value	-2.8653
		10% Critical Value	-2.5688

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESID1)
Method: Least Squares
Date: 10/08/06 Time: 09:40
Sample(adjusted): 7 877
Included observations: 871 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID1(-1)	-0.982424	0.077150	-12.73394	0.0000
D(RESID1(-1))	-0.023253	0.069024	-0.336875	0.7363
D(RESID1(-2))	-0.016439	0.058961	-0.278813	0.7805
D(RESID1(-3))	-0.071511	0.048191	-1.483886	0.1382
D(RESID1(-4))	-0.033381	0.034047	-0.980426	0.3272
C	2.51E-05	0.000329	0.076201	0.9393

R-squared	0.504195	Mean dependent var	-2.28E-05
Adjusted R-squared	0.501329	S.D. dependent var	0.013743
S.E. of regression	0.009705	Akaike info criterion	-6.425471
Sum squared resid	0.081473	Schwarz criterion	-6.392615
Log likelihood	2804.293	F-statistic	175.9275
Durbin-Watson stat	1.986549	Prob(F-statistic)	0.000000

ADF Test Statistic	-12.92684	1% Critical Value*	-3.4405
		5% Critical Value	-2.8653
		10% Critical Value	-2.5688

*MacKinnon critical values for rejection of hypothesis of a unit root.

D. First Difference Series Regression – SP1

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(SP1)
Method: Least Squares
Date: 10/15/06 Time: 09:29
Sample(adjusted): 7 877
Included observations: 871 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
SP1(-1)	-0.984940	0.076193	-12.92684	0.0000
D(SP1(-1))	-0.052718	0.069343	-0.760257	0.4473
D(SP1(-2))	-0.043965	0.060396	-0.727952	0.4668
D(SP1(-3))	-0.012591	0.049166	-0.256089	0.7979
D(SP1(-4))	0.063164	0.034099	1.852405	0.0643
C	0.000573	0.000208	2.760225	0.0059

R-squared	0.523836	Mean dependent var	2.01E-05
Adjusted R-squared	0.521084	S.D. dependent var	0.008666
S.E. of regression	0.005997	Akaike info criterion	-7.388156
Sum squared resid	0.031112	Schwarz criterion	-7.355300
Log likelihood	3223.542	F-statistic	190.3203
Durbin-Watson stat	1.990013	Prob(F-statistic)	0.000000

E. First Difference Series Regression – EX1

ADF Test Statistic	-12.88392	1% Critical Value*	-3.4405
		5% Critical Value	-2.8653
		10% Critical Value	-2.5688

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(EX1)
Method: Least Squares
Date: 10/15/06 Time: 09:32
Sample(adjusted): 7 877
Included observations: 871 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
EX1(-1)	-0.990629	0.076889	-12.88392	0.0000
D(EX1(-1))	-0.005654	0.068665	-0.082338	0.9344
D(EX1(-2))	-0.001441	0.058598	-0.024596	0.9804
D(EX1(-3))	-0.061527	0.047963	-1.282782	0.1999
D(EX1(-4))	-0.029880	0.034074	-0.876914	0.3808
C	-0.000287	0.000236	-1.213314	0.2253

R-squared	0.499292	Mean dependent var	-1.61E-05
Adjusted R-squared	0.496398	S.D. dependent var	0.009788
S.E. of regression	0.006946	Akaike info criterion	-7.094455
Sum squared resid	0.041733	Schwarz criterion	-7.061599
Log likelihood	3095.635	F-statistic	172.5110
Durbin-Watson stat	1.987376	Prob(F-statistic)	0.000000

F. First Difference Series Regression – Residuals

ADF Test Statistic	-13.12809	1% Critical Value*	-3.4406
		5% Critical Value	-2.8653
		10% Critical Value	-2.5688

*MacKinnon critical values for rejection of hypothesis of a unit root.

Augmented Dickey-Fuller Test Equation
Dependent Variable: D(RESID2)
Method: Least Squares
Date: 10/15/06 Time: 09:39
Sample(adjusted): 12 877
Included observations: 866 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
RESID2(-1)	-1.002452	0.076359	-13.12809	0.0000
D(RESID2(-1))	0.004988	0.068346	0.072980	0.9418
D(RESID2(-2))	0.002608	0.059123	0.044116	0.9648
D(RESID2(-3))	0.000411	0.048269	0.008519	0.9932
D(RESID2(-4))	0.003944	0.034199	0.115327	0.9082
C	1.06E-05	0.000330	0.032062	0.9744

R-squared	0.496457	Mean dependent var	-3.43E-05
Adjusted R-squared	0.493529	S.D. dependent var	0.013643
S.E. of regression	0.009709	Akaike info criterion	-6.424545
Sum squared resid	0.081074	Schwarz criterion	-6.391539
Log likelihood	2787.828	F-statistic	169.5793
Durbin-Watson stat	1.990574	Prob(F-statistic)	0.000000

Appendix 3 – Heteroskedasticity

A. ARCH LM Test – OLS regression, Log of EX on Log of SP

ARCH Test:

F-statistic	4539.163	Probability	0.000000
Obs*R-squared	839.9502	Probability	0.000000

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 10/14/06 Time: 10:58
Sample(adjusted): 6 877
Included observations: 872 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	8.52E-05	3.33E-05	2.559063	0.0107
RESID^2(-1)	0.966773	0.033977	28.45370	0.0000
RESID^2(-2)	-0.007510	0.047188	-0.159142	0.8736
RESID^2(-3)	-0.005353	0.047211	-0.113385	0.9098
RESID^2(-4)	0.076496	0.046728	1.637055	0.1020
RESID^2(-5)	-0.065224	0.032869	-1.984398	0.0475

R-squared	0.963246	Mean dependent var	0.003112
Adjusted R-squared	0.963033	S.D. dependent var	0.004074
S.E. of regression	0.000783	Akaike info criterion	-11.45917
Sum squared resid	0.000531	Schwarz criterion	-11.42634
Log likelihood	5002.197	F-statistic	4539.163
Durbin-Watson stat	1.989313	Prob(F-statistic)	0.000000

B. ML-ARCH Model, Level Series Regression – Log of EX on Log of SP

Dependent Variable: LNEX
Method: ML - ARCH (Marquardt)
Date: 10/14/06 Time: 11:01
Sample: 1 877
Included observations: 877
Convergence achieved after 230 iterations
Variance backcast: ON

	Coefficient	Std. Error	z-Statistic	Prob.
C	3.102778	0.035976	86.24471	0.0000
LNSP	-0.336704	0.004378	-76.91175	0.0000
	Variance Equation			
C	5.24E-05	1.12E-05	4.676868	0.0000
ARCH(1)	1.003319	0.229230	4.376914	0.0000
GARCH(1)	-0.029396	0.086830	-0.338548	0.7350

R-squared	0.478982	Mean dependent var	0.330641
Adjusted R-squared	0.476592	S.D. dependent var	0.079328
S.E. of regression	0.057392	Akaike info criterion	-3.858982
Sum squared resid	2.872203	Schwarz criterion	-3.831750
Log likelihood	1697.163	F-statistic	200.4115
Durbin-Watson stat	0.015315	Prob(F-statistic)	0.000000

C. ARCH LM Test – OLS regression, EX1 on SP1

ARCH Test:

F-statistic	0.976083	Probability	0.431309
Obs*R-squared	4.886698	Probability	0.429864

Test Equation:
Dependent Variable: RESID^2
Method: Least Squares
Date: 10/15/06 Time: 18:31
Sample(adjusted): 7 877
Included observations: 871 after adjusting endpoints

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	4.11E-05	4.46E-06	9.221683	0.0000
RESID^2(-1)	0.046003	0.034252	1.343077	0.1796
RESID^2(-2)	0.011021	0.034255	0.321724	0.7477
RESID^2(-3)	0.034953	0.034237	1.020913	0.3076
RESID^2(-4)	0.039425	0.034247	1.151210	0.2500
RESID^2(-5)	0.010909	0.034237	0.318626	0.7501

R-squared	0.005610	Mean dependent var	4.79E-05
Adjusted R-squared	-0.000137	S.D. dependent var	8.56E-05
S.E. of regression	8.56E-05	Akaike info criterion	-15.88653
Sum squared resid	6.34E-06	Schwarz criterion	-15.85367
Log likelihood	6924.583	F-statistic	0.976083
Durbin-Watson stat	1.986700	Prob(F-statistic)	0.431309

Appendix 4 – Maximum Likelihood Tests and Lag Exclusions Tests

A. Level Series Maximum Likelihood Test

VAR Lag Order Selection Criteria
Endogenous variables: LNEX LNSP
Exogenous variables: C
Date: 10/14/06 Time: 11:48
Sample: 1 877
Included observations: 869

Lag	LogL	LR	FPE	AIC	SC	HQ
0	1556.947	NA	9.57E-05	-3.578704	-3.567732	-3.574505
1	6303.336	9460.006	1.74E-09*	-14.49329*	-14.46038*	-14.48070*
2	6307.043	7.372459	1.74E-09	-14.49262	-14.43776	-14.47163
3	6308.806	3.497519	1.75E-09	-14.48747	-14.41067	-14.45808
4	6311.832	5.989174	1.75E-09	-14.48523	-14.38648	-14.44744
5	6316.818	9.845987*	1.75E-09	-14.48750	-14.36681	-14.44132
6	6319.086	4.466871	1.76E-09	-14.48351	-14.34087	-14.42893
7	6320.214	2.217680	1.77E-09	-14.47690	-14.31232	-14.41393
8	6324.472	8.349784	1.77E-09	-14.47750	-14.29097	-14.40612

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

B. First Difference Maximum Likelihood Test

VAR Lag Order Selection Criteria
Endogenous variables: EX1 SP1
Exogenous variables: C
Date: 10/15/06 Time: 16:11
Sample: 1 877
Included observations: 868

Lag	LogL	LR	FPE	AIC	SC	HQ
0	6291.003	NA	1.74E-09*	-14.49079*	-14.47981*	-14.48659*
1	6294.609	7.185471	1.75E-09	-14.48988	-14.45694	-14.47727
2	6296.342	3.446375	1.76E-09	-14.48466	-14.42975	-14.46365
3	6299.424	6.114548	1.76E-09	-14.48254	-14.40567	-14.45313
4	6304.316	9.682076*	1.76E-09	-14.48460	-14.38576	-14.44678
5	6306.589	4.488278	1.76E-09	-14.48062	-14.35982	-14.43439
6	6307.727	2.241863	1.77E-09	-14.47402	-14.33126	-14.41939
7	6311.939	8.279165	1.77E-09	-14.47451	-14.30978	-14.41148
8	6313.716	3.484910	1.78E-09	-14.46939	-14.28270	-14.39795

* indicates lag order selected by the criterion
LR: sequential modified LR test statistic (each test at 5% level)
FPE: Final prediction error
AIC: Akaike information criterion
SC: Schwarz information criterion
HQ: Hannan-Quinn information criterion

C. First Difference Wald Lag Exclusion Test

VAR Lag Exclusion Wald Tests
Date: 10/15/06 Time: 16:22
Sample: 1 877
Included observations: 870

Chi-squared test statistics for lag exclusion:
Numbers in [ ] are p-values

	EX1	SP1	Joint
Lag 1	4.990367	1.962579	7.359374
	[0.082481]	[0.374827]	[0.118073]
Lag 2	0.060186	3.686060	3.782670
	[0.970355]	[0.158337]	[0.436216]
Lag 3	2.907081	4.656516	7.180172
	[0.233741]	[0.097465]	[0.126668]
Lag 4	2.011021	6.672676	8.608873
	[0.365858]	[0.035567]	[0.071655]
Lag 5	1.597753	2.953329	4.359496
	[0.449834]	[0.228398]	[0.359534]
Lag 6	0.515443	1.832439	2.247771
	[0.772810]	[0.400029]	[0.690294]
df	2	2	4

Appendix 5 – VAR Stability Condition Test – AR Roots Table

A. Level Series

Roots of Characteristic Polynomial
Endogenous variables: LNEX LNSP
Exogenous variables: C
Lag specification: 1 6
Date: 10/14/06 Time: 13:02

Root	Modulus
0.999582	0.999582
0.988322	0.988322
-0.621231	0.621231
-0.149082 + 0.598596i	0.616881
-0.149082 - 0.598596i	0.616881
0.136542 - 0.536471i	0.553574
0.136542 + 0.536471i	0.553574
0.533136	0.533136
0.445279 + 0.238351i	0.505059
0.445279 - 0.238351i	0.505059
-0.410513 - 0.153454i	0.438256
-0.410513 + 0.153454i	0.438256

No root lies outside the unit circle.
VAR satisfies the stability condition.

B. First Difference Series

Roots of Characteristic Polynomial
Endogenous variables: EX1 SP1
Exogenous variables: C
Lag specification: 1 6
Date: 10/15/06 Time: 18:37

Root	Modulus
-0.665431	0.665431
-0.218424 + 0.579948i	0.619716
-0.218424 - 0.579948i	0.619716
0.077891 + 0.582732i	0.587914
0.077891 - 0.582732i	0.587914
0.557964	0.557964
-0.493742 + 0.218709i	0.540014
-0.493742 - 0.218709i	0.540014
0.448817 - 0.226242i	0.502615
0.448817 + 0.226242i	0.502615
0.218600 + 0.369721i	0.429511
0.218600 - 0.369721i	0.429511

No root lies outside the unit circle.
VAR satisfies the stability condition.

Appendix 6 – Johansen Co-Integration Test

Level Series

Date: 10/14/06 Time: 13:19
Sample(adjusted): 10 877
Included observations: 868 after adjusting endpoints
Trend assumption: Linear deterministic trend (restricted)
Series: LNEX LNSP
Lags interval (in first differences): 1 to 8
Unrestricted Co-integration Rank Test

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	5 Percent Critical Value	1 Percent Critical Value
None **	0.029442	32.24849	25.32	30.45
At most 1	0.007242	6.308928	12.25	16.26

*(**) denotes rejection of the hypothesis at the 5%(1%) level
Trace test indicates 1 co-integrating equation(s) at both 5% and 1% levels

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	5 Percent Critical Value	1 Percent Critical Value
None **	0.029442	25.93957	18.96	23.65
At most 1	0.007242	6.308928	12.25	16.26

*(**) denotes rejection of the hypothesis at the 5%(1%) level
Max-eigenvalue test indicates 1 co-integrating equation(s) at both 5% and 1% levels

Appendix 7 – Pair-wise Granger Causality Tests

A. Level Series – 1 Lag

Pairwise Granger Causality Tests
Date: 10/14/06 Time: 16:00
Sample: 1 877
Lags: 1

Null Hypothesis:	Obs	F-Statistic	Probability
LNSP does not Granger Cause LNEX	876	1.09516	0.29562
LNEX does not Granger Cause LNSP		0.04788	0.82684

B. Level Series – 2 Lags

Pairwise Granger Causality Tests
Date: 10/14/06 Time: 16:02
Sample: 1 877
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Probability
LNSP does not Granger Cause LNEX	875	3.31227	0.03689
LNEX does not Granger Cause LNSP		0.29125	0.74740

C. Level Series – 3 Lags

Pairwise Granger Causality Tests
Date: 10/14/06 Time: 16:02
Sample: 1 877
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Probability
LNSP does not Granger Cause LNEX	874	2.07498	0.10199
LNEX does not Granger Cause LNSP		1.38339	0.24647

D. Level Series – 4 Lags

Pairwise Granger Causality Tests
Date: 10/14/06 Time: 16:04
Sample: 1 877
Lags: 4

Null Hypothesis:	Obs	F-Statistic	Probability
LNSP does not Granger Cause LNEX	873	1.36064	0.24579
LNEX does not Granger Cause LNSP		1.86248	0.11499

E. Level Series – 5 Lags

Pairwise Granger Causality Tests
Date: 10/14/06 Time: 16:11
Sample: 1 877
Lags: 5

Null Hypothesis:	Obs	F-Statistic	Probability
LNSP does not Granger Cause LNEX	872	1.36982	0.23329
LNEX does not Granger Cause LNSP		2.02371	0.07308

F. Level Series – 6 Lags

Pairwise Granger Causality Tests
Date: 10/14/06 Time: 16:13
Sample: 1 877
Lags: 6

Null Hypothesis:	Obs	F-Statistic	Probability
LNSP does not Granger Cause LNEX	871	1.28196	0.26287
LNEX does not Granger Cause LNSP		1.60327	0.14309

G. First Difference Series – 1 Lag

Pairwise Granger Causality Tests
Date: 10/15/06 Time: 16:34
Sample: 1 877
Lags: 1

Null Hypothesis:	Obs	F-Statistic	Probability
SP1 does not Granger Cause EX1	875	5.39831	0.02039
EX1 does not Granger Cause SP1		0.45271	0.50123

H. First Difference Series – 2 Lags

Pairwise Granger Causality Tests
Date: 10/15/06 Time: 17:36
Sample: 1 877
Lags: 2

Null Hypothesis:	Obs	F-Statistic	Probability
SP1 does not Granger Cause EX1	874	2.51358	0.08157
EX1 does not Granger Cause SP1		1.96296	0.14106

I. First Difference Series – 3 Lags

Pairwise Granger Causality Tests
Date: 10/15/06 Time: 17:37
Sample: 1 877
Lags: 3

Null Hypothesis:	Obs	F-Statistic	Probability
SP1 does not Granger Cause EX1	873	1.48649	0.21674
EX1 does not Granger Cause SP1		2.40248	0.06633

J. First Difference Series – 4 Lags

Pairwise Granger Causality Tests
Date: 10/15/06 Time: 17:37
Sample: 1 877
Lags: 4

Null Hypothesis:	Obs	F-Statistic	Probability
SP1 does not Granger Cause EX1	872	1.41649	0.22652
EX1 does not Granger Cause SP1		2.45069	0.04469

K. First Difference Series – 5 Lags

Pairwise Granger Causality Tests
Date: 10/15/06 Time: 17:38
Sample: 1 877
Lags: 5

Null Hypothesis:	Obs	F-Statistic	Probability
SP1 does not Granger Cause EX1	871	1.30367	0.26012
EX1 does not Granger Cause SP1		1.87697	0.09588

L. First Difference Series – 6 Lags

Pairwise Granger Causality Tests
Date: 10/15/06 Time: 17:38
Sample: 1 877
Lags: 6

Null Hypothesis:	Obs	F-Statistic	Probability
SP1 does not Granger Cause EX1	870	1.06384	0.38260
EX1 does not Granger Cause SP1		1.75285	0.10593

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

The Relationship between Earnings and Stock Returns: Empirical Evidence from the Greek Capital Market

Panagiotis E. Dimitropoulos
Teaching & Research Associate
University of Peloponnese
Department of Sport Management
3-5 Lysandrou Str, P.C.23100, Sparta, Greece
Tel: 30-273-102-4096, 30- 693-689-8306
E-mail: dimitrop@uop.gr

Dimitrios Asteriou
Associate Professor
Hellenic Open University
City University, London
School of Social Sciences
Riga Feraiou 162 & Tsamadou
Patras, Greece
Tel: 44-694-710-0540
E-mail: D.A.Asteriou@city.ac.uk

Abstract

The relationship between earnings figures and stock returns has been a topic of international research since decades. The purpose of this paper is to investigate the above relationship in the context of the Greek capital market. Previous studies resulted in controversial results regarding the usefulness of models which were using earnings levels or earnings changes as the explanatory variable. In an introductory context, this study examines the earnings-return relation applying four models, proposed by Kothari and Zimmerman (Journal of Accounting and Economics, 20, 155-192, 1995), on individual Greek stocks as well as portfolios between 1994-2004. The overall results, demonstrated a significant value relevancy of accounting earnings prepared under the Greek GAAP. Specifically in the Greek stock market the price model produces less biased ERC's than the return model but suffers from various econometric problems. Also, the use of cross-sectional and time-series aggregated data results in a large increase in the explanatory power of earnings for returns (for the return and differenced model) yielding more significant Earnings Response Coefficients.

Keywords: Earnings response coefficients, Stock returns, Price and returns models, Earnings relevance, Greece

JEL classification: M41, G10

1. Introduction

The relation between stock returns and accounting earnings has been an internationally popular subject of research for more than thirty years. Since Ball and Brown (1968) first documented the aforementioned relation, an extensive number of accounting literature has been published, all trying to best accommodate the incentives of the returns-earnings association, as well as the ability to forecast future earnings from stock movements and vice-versa.

Of course the interest for this subject is more than intense since accounting numbers are published in order to assist investors to take rationale investing decisions. Beaver (1989) mentioned that earnings per share is the only figure in the financial statements that receives the greatest attention by the investors. Moreover, earning numbers are supposed to facilitate analysts and investors to forecast future cash flows and deal with relative investments risks. For the above reasons there have been many studies in US, Europe and Far East trying to evaluate the possibility of achieving this objective. If we manage to enlighten the causes that drive the return-earnings relation then we will have accomplished a great step into creating a less risky decision-making environment.

The aim of this paper is to assess the association between published accounting earnings and stock returns, in the context of the Hellenic Capital Market. The disclosure legislation, as well as the accounting practises in Greece are different compared to US, UK and other Western countries but they have some similarities to other countries such as France, Italy, the Baltic region, Cyprus and Czech Republic. Despite the fact that the aforementioned issue has received great attention from researchers, the majority of studies on this field have focused on well developed capital markets creating a gap on the literature of less developed and organized stock markets. The Greek accounting setting can be characterized by moderate tax and financial accounting conformity, moderate use of accruals, low importance of the capital market and weak corporate governance. Therefore, the present study adds to the existing literature by examining the issue of earnings relevance within the context of an emerging capital market like Greece.

Under this framework we examined different return-earnings model specifications (Price, Return, Differenced and Deflated models) following Kothari and Zimmerman (1995). Annual data from companies that have their titles listed on the Athens Stock Exchange over an 11 year period (1994-2004), were used. Results indicate a significant value relevancy of accounting numbers in explaining stock returns under the Price and Return model specifications. The differenced and deflated models did not yield any significant results supporting the results from the two previous specifications.

Additionally, the cross-sectional and time-series aggregated procedures (see Dumontier and Labelle, 1998) resulted in a large increase in the explanatory power of earnings for returns (for the return and differenced model) yielding more significant earnings response coefficients (Hereafter ERC’s).

The paper is organized as follows. The next section presents the review of relative literature that examined the existence of the return-earnings relation in the past, alongside with the empirical results that proved the above fact. Section 3 describes the data selection procedure and provides details on the methodological framework of the paper. Section 4 illustrates the empirical results from annual, cross-sectional and aggregated data and finally the last section is dedicated to concluding remarks, limitations and fruitful ways for future research.

2. Literature review

Several papers examined the ability of price and return models (along with some alternative forms) to accommodate the return-earnings relation. Gonedes and Dopuch (1974) declared that price models are conceptually inferior to return models under the presence of under-developed theories of valuation. Additionally Christie (1987) argues that price models present more econometric problems than return models. On the contrary there is a majority of studies that declare the superiority of the price model on explaining the return-earnings relation such as Bowen (1981), Olsen (1985), Landsman (1986), Barth et.al (1990, 1992) Barth (1991) and Harris et.al (1994). Furthermore, Kothari and Zimmerman (1995) documented that the earnings response coefficients are less biased and more economically sensible in the price model compared to the return and differenced models. On the other hand, price and differenced models suffer from heteroscedasticity and other specification problems relative to return model.

Consistent to Kothari and Zimmerman (1995) were the results by Martikainen et.al (1997) and Dumontier and Labelle (1998) in Finland and France respectively. The first used two different measures of the earnings variable, the published earnings numbers and the adjusted earnings numbers, estimated according to the recommendations by the Finnish Committee for Corporate Analysis (COC). They concluded that the return model specification (using the published accounting data) yields higher and more significant ERC’s compared to differenced model specification under the adjusted accounting data. Dumontier and Labelle (1998) reached in the same conclusion regarding the ability of the return model to evaluate the return-earnings relation by providing highly significant ERC’s. Also their tests indicated that a cross-sectional data cumulating procedure could yield a large increase in the explanatory power of the return model and the significance of the earnings response coefficient.

On the contrary, Vafeas et.al (1998) found conflicting evidence for the evaluating ability of the return model in the Cyprus stock market. They concluded that both specifications are important in explaining stock returns but the differenced model yield higher ERC’s compared to the return model. This finding agrees with studies by Martikainen et.al (1991), Board and Day (1989), Alford et.al (1993) and Harris et.al (1994).

Finally, Pritchard (2002) examined the return-earnings relation in the countries of the Baltic region (Estonia, Latvia, and Lithuania). She documented a substantial difference in the association of the above variables. Lithuania showed the weakest and Estonia the highest value relevance followed by Latvia. In total the return model proved to perform better compared to differenced model specification.

As a conclusion we can say that we have conflicting results regarding the valuation ability of different model specifications. There is a group of studies that favours for the price model’s ability to explain the return-earnings relation and another group which argues that return model performs better. Nevertheless all model specifications present serious econometric problems and they must be used with caution. As Kothari and Zimmerman (1995) argue, the use of all model specifications has the potential to provide more convincing and reliable evidence instead of considering each model separately.

3. Data selection and methodology

3.1 Sample selection

Our sample consists of 105 companies, all listed in the Athens Stock Exchange, with full annual data of reported earnings and stock prices during the period 1994-2004. Data were collected from the Athens Stock Exchange database. The initial sample contained 112 companies with full data for the period under investigation but we restricted it to companies with December fiscal year-end limiting the final number to 105. The inclusion of the same 105 firms in the sample was imposed by the time-aggregation procedure which requires the variables for each firm to be available for each year under investigation. According to Collins et.al (1994) this fact is likely to bias the results, but since the influence of this bias on the final results is hard to be determined a priori, it is supposed to be insignificant.

We use per share values of earnings and returns in order to reduce heteroscedasticity on the error terms (see also Barth et.al, 1992; Kothari and Zimmerman, 1995). The earnings variable is annual earnings per share (EPS_t) for the price model, annual earnings per share deflated by the price at the beginning of the year (EPS_t/P_{t-1}) for the return model and change in earnings per share (ΔEPS_t) for the differenced model.

The price and return variable is P_t (Price model), P_t/P_{t-1} (Return model) and ΔP_t (Differenced model). Annual stock returns, exclusive of dividends, are used and the earnings and price data are adjusted for stock splits, stock dividends, stock issues and the fact of currency translation (drachma to euro). No further trimming of the sample was conducted since we did not want to lose observations and affect the final results.

Table 1 illustrates the descriptive statistics of the price and earnings variables for all the years under investigation. As we can see the mean of all variables is higher than the median which indicates that data are skewed to the right side of the distribution. Also the mean is below the standard deviation for all the variables and the minimum and maximum values indicate a wide range of observations in the sample. Finally the mean EPS is 37 cents while the mean stock price is 6.3 euros providing a mean stock return of up to 11 per cent for the whole period of investigation. Also the mean value of EPS is decreasing from 1994 to 2002 but on the other hand mean stock prices increase and reach their peak in 1999 (annual descriptive statistics are not presented here- they can be provided upon request).

INSERT TABLE 1 HERE

3.2 Methodological framework

3.2.1 Earnings and Return specification

The rationale for the economic interpretation of return-earnings relation originates from a fixed valuation model in which stock prices are the discounted present values of future net cash flows (Kothari and Zimmerman, 1995; Watts and Zimmerman, 1986; Kormendi and Lipe, 1987; Ohlson, 1991). Since the expectations for the market's future cash flows cannot be measured, we use current earnings as an approximation for the market's expectation in our valuation models.

These models assume that the ex-dividend stock price at the end of a time period t is a function of the published accounting earnings for the same period. There are four model specifications that we will use in our analysis, which all of them produce equal earnings response coefficients of 1/r, where r is the firm's expected rate of return (Kothari and Zimmerman, 1995).

$$\textbf{Price Model: } P_t = a + bX_t + e_t \tag{1}$$

$$\textbf{Return model: } P_t/P_{t-1} = a + bX_t / P_{t-1} + e_t \tag{2}$$

$$\textbf{Differenced Model: } \Delta P_t/P_{t-1} = a + bX_t / P_{t-1} + e_t \tag{3}$$

$$\textbf{Deflated Model: } P_t/EPS_t = a1/X_t + b + e_t \tag{4}$$

Where P_t is the ex-dividend price at the end of the fiscal year t and X_t is earnings per share excluding taxes and discontinued operations for period t, a and b are the intercept and the earnings response coefficient respectively and e is the error term. Two important assumptions are made in order to evaluate these models, earnings follow a random walk and prices do not lead earnings (prices are set by the market only using the current and past time series of earnings).

The differenced (changes) model is used because we can produce a stationary series by differencing the price and earnings variables, so the econometric problems that appear in the estimation of the price model could be mitigated by taking the first differences (Christie, 1987). Additionally, the deflated specification was used in order to reduce the heteroscedasticity of the error terms and the cross-sectional variation of the slope coefficient (which is assumed to be constant). So we estimated the price model using earnings per share as the deflator where b is the estimated ERC.

Finally, the above four models were estimated for each year separately and as a pooled sample as well. Also, in order to take into consideration the 1999-2000 stock market crash, we divided our sample into two sub periods 1994-1998 and 1999-2004 in order to examine the relation of returns and earnings before and after the market's crash.

3.2.2 Cross-sectional and time aggregated data

There are several studies that divide reported earnings into two basic components, the value relevant component and the price irrelevant component. The first includes the information needed to explain stock prices movements, while the second is irrelevant in explaining stock returns (see also Beaver et.al, 1980; Lev, 1989; Ramakrishnan and Thomas, 1998; Dumontier and Labelle, 1998). Under this framework the cross-sectional aggregation of the firms could reduce the variance of the irrelevant component compared to the value relevant component and improve the overall results. Consequently all four models were re-estimated using a portfolio aggregation procedure. Stocks in each year were clustered into 15 portfolios where each portfolio included 7 stocks. The intuition behind this refinement is that investors are supposed not to be influenced by the variability in reported earnings, thus returns will be uncorrelated with the measurement error in earnings variables.

Beyond the value irrelevant noise in earnings that can affect the returns-earnings relation, we must consider the fact that earnings events are not recognized as fast as stock price events. This often delays the incorporation of important events into the financial statements thus current earnings do not capture all the important economic events (see also Basu, 1997; Raonic et.al, 2004). As a result the low value relevancy (as measured by the ERC and the coefficient of determination) of the price and return models could be attributed to the timeliness of earnings recognition. If this situation actually exists, we can improve the return-earnings relation by aggregating earnings and returns over long time periods, thus the longer the time horizon the higher the recognition of economic events in both variables.

Under this framework, stock returns and annual earnings per share were aggregated over two and five year periods and models 1-4 were re-estimated using two and five year returns over two and five year earnings. Finally in order to examine the impact of both biases on earnings and returns (value-irrelevant noise and time recognition lag) we combined the two previous methodologies by clustering the five-year aggregated variables into 15 portfolios with 7 stocks each and following the same procedure as previous.

4. Empirical results

The regression results, of the four models specifications, for the one year period are presented in Table 2. The first column contains the results of the price model (equation 1), the second column includes results from the return model and columns 3 and 4 contain results from the change and deflated models respectively. The evidence indicates that securities returns are significantly associated with the earnings level (Price model) and the earnings level deflated by the previous year's price (Return model). Regarding the changes and the deflated models the only significant results were in 1995 and 2004. On the contrary the price model provided significant results for all the years under investigation while the return model only for the half (1995, 1997, 1999, 2000 and 2004).

INSERT TABLE 2 HERE

More specifically the most significant ERC in the price model was in 2003 (4.82) and the smallest in 2002 (0.015) which was also insignificant. Moreover, another fact that was documented by Kothari and Zimmerman (1995), is that all models yield significant intercept coefficient which is inconsistent with the theory. It is obvious that 2002 was the only year that did not provide any significant ERC in all four models. This fact can be attributed to the 1999-2000 stock market crash which affected the stock prices movements and the overall stability of the market. In our case the price and return models seem to perform and describe better the return-earnings relation compared to the remaining two specifications. This result is consistent to Kothari and Zimmerman (1995), who found that the price model provides unbiased ERCs but the return model yields ERC which are biased downward, and with other studies that document the ability of price and return models to describe better the aforementioned relation (Bowen, 1981; Olsen, 1985; Landsman, 1986; Barth et.al, 1990 and 1992; Barth, 1991; Harris et.al, 1994).

Additionally, in order to examine the impact of the 1999-2000 stock market crash, we divided the sample into to sub-periods the first from 1994 to 1998 and the second from 1999 to 2004 and estimated again the price and return models (results are not reported here but they can be provided upon request). The price model yields an ERC of 2.72 (the t statistic is 11.27) for the pre-1999 period and an ERC of 0.25 (t statistic is 1.62) for the post-1999 period, while the return model yield a higher ERC for the pre-1999 period but not statistically significant. Overall we can say that in the Greek case, the price model describes better the return-earnings relation compared to the other three specifications but we must not forget that the four models may suffer from value irrelevant noise and time recognition lag problems.

Consequently, in order to deal with the value irrelevant noise in the earnings variables, we clustered each year's stocks into 15 portfolios with 7 stocks in each one and re-estimated the four models specifications. The mean values of both coefficients and R^2 are portrayed in Table 3.

INSERT TABLE 3 HERE

The above results provide evidence for the existence of noise in the earnings variables that causes a downward bias in the estimates of individual stocks. The results concerning the price model are overwhelming. The ERC and R^2 have increased significantly for all years under investigation (including 2002). As for the return model, it yields significant ERCs for 8 out of 11 years and of course the overall coefficient of determination has increased considerably compared to Table 1. This result is consistent to Dumontier and Labelle (1998) who also found irrelevant noise in the French accounting earnings and experienced a rather significant improvement in their results after controlling for this bias. Their overall results indicated the superiority of the price and return models over the differenced model.

Also the low informational ability of the price and return models could be attributed to the recognition lag. That is when earnings do not incorporate all the important economic events as fast as prices do. For this reason we increased the returns and earnings measurement intervals in order to include more economic events in the association between earnings and returns. Two and five year returns were summed and regressed over two and five year earnings. The overall results were sensibly improved as the period increases from two to five year measurement window. The two year variables produced ERCs higher than the ERCs of the one year period for all years under consideration. In the pooled sample the ERC is 0.537 (t=3.68 significant at a=1 per cent) while the ERC for the one year measurement interval is 0.31 (t=2.52 significant at a=1 per cent) and the R^2 increased from 0.5 per cent to 1.2 per cent (results are not reported here but it can be provided upon request).

The increase in the earnings response coefficients was even more severe when we increased the measurement window up to five years. Results presented in Table 4 show that the ERC for the pooled sample is 1.35 (5.69 significant at a=1 per cent) and the R^2 is up to 4.1 per cent. These results support the main hypothesis that annual earnings do not provide an efficient measure for all the value relevant events captured by the investors and the use of other instrumental variables could improve the return-earnings relation which is based on short period data.

INSERT TABLE 4 HERE

Finally the last sensitivity test aims at correcting for both biases due to irrelevant noise in earnings and the timeliness of earnings recognition as they were depicted in the last two tables. For this reason we constructed 15 portfolios for each period (5 year measurement interval) of 7 stocks each and re-estimated models 1 to 4. The results, as they are reported in Table 5 show a relative improvement of both ERC and R^2.

In the price model column the ERC for the pooled sample increased to 4.65 (t=4.24 significant at a=1 per cent) from 1.35 an overall increase of 344 per cent. As for the R^2 has increased up to 26.6 from 4.1 or alternatively 648 per cent. Regarding the return model the ERC now became significant at a=5 per cent (t=1.94 from -0.249) and the R^2 increased from 0.1 per cent to 19.1 per cent.

Summarizing we can say that when temporal and cross-sectional aggregation of data is taken under consideration, it reinforces the returns-earnings relation by diversifying the value irrelevant noise in earnings and the timeliness between the period that a value relevant event occurs and the period when the specific event is incorporated in earnings (Basu, 1997).

INSERT TABLE 5 HERE

5. Conclusions, limitations and future research

The present paper analyses the association between stock returns and earnings in the context of the Hellenic capital market, using four model specifications, the price, return, differenced and deflated models. Data of published earnings and stock prices of 105 companies (all listed in the A.S.E.) for 11 years (1994-2004) were extracted from the Athens Stock Exchange database.

The initial results indicated the increased ability of the price and return models to explain better the aforementioned relationship by providing highly significant earnings response coefficients. The changes and deflated models on the contrary proved to be insufficient compared to the previous two models. These results are consistent to Kothari and Zimmerman (1995), Martikainen et.al (1997) and Dumontier and Labelle (1998).

Additionally, in order to improve the previous results we took under consideration the value irrelevant noise in the earnings variable and recognition lag of value relevant events in earnings. Consequently we re-estimated the previous four models using two refinements. The first deals with the problem of noise in earnings and for that we clustered our sample into 15 portfolios for each year and the second includes the increase of the return and earnings measurement interval (from 1 year to 2 and 5 years) in order to take under consideration more value relevant events.

The final results indicated a high value relevancy in the Hellenic capital market (of course smaller compared to US, France and UK) expressed by the large increase of the ERC and R^2 of the price and return models after controlling for the two biases mentioned above. Thus controlling for value irrelevant events and by increasing the variables measurement window we achieve to improve the low association of earnings and returns which was documented by many studies in the past.

However, we must mention that our analysis is based on two limitations. The first is that we do not consider the implications of deviations from the random walk property of annual earnings and the second refers to the fact that we did not examined foundational economic reasons that grounds the previously mentioned relation, as well as alternative econometric techniques that could improve the final result. A fruitful path for future research is to examine the effect of losses (negative earnings) in the earnings-returns relation and the inclusion of leading periods in earnings and returns as measures of increasing the earnings response coefficients and the explanatory power of the models.

Acknowledgements

We wish to thank Professor Costas Syriopoulos and Dr. Evangelos Koumanakos from the University of Patras, for providing us all the necessary data for the completion of this paper, participants on the 1st International Conference in Accounting and Finance of the University of Macedonia-Thessaloniki and the editor Amy King as well as two anonymous reviewers for useful comments and suggestions. Their contribution is highly acknowledged. Needless to say that the responsibility for any errors and omissions rests entirely with us.

References

Alford A., Jones J., Leftwich R. and Zmijewski M. (1993). The relative informativeness of accounting disclosures in different countries. *Journal of Accounting Research*, 31 (Suppl.), 183-223.

Ball, R. and Brown, P. (1968). An empirical evaluation of accounting income numbers, *Journal of Accounting Research,* 6, 159–78.

Barth, Mary E, William H. Beaver, and Mark A. Wolfson, (1990). Components of bank earnings and the structure of bank share prices. *Financial Analysts Journal*, 46 May/June, 53-60.

Barth, Mary E., (1991). Relative measarement errors among alternative pension asset and liability measures. *The Accounting Review,* 66, 433-463.

Barth, Mary E., William H. Beaver, avid Wayne R. Landsman, (1992). The market valuation implications of net periodic pension cost components. *Journal of Accounting and Economics*, 15, 2-62.

Basu, Sudipta, (1997). The conservatism principle and the asymmetric timeliness of earnings. *Journal of Accounting and Economics*, 24, 3-37.

Beaver, W. H., R. Lambert, and D. Morse (1980). The information content of security prices. *Journal of Accounting and Economics,* 2, 3-28.

Beaver, William H., (1989). *Financial accounting: An accounting revolution.* Prentice-Hall, Engle-wood Cliffs, N J.

Board, J. and Day, J. (1989). The information content of cash flow figures. *Accounting and Business Research*, 20, 3–11.

Bowen R. M. (1981). Valuation of earnings components in the electric utility industry. *The Accounting Review,* 56, 1-22.

Christie A. A., (1987). On cross-sectional analysis in accounting research. *Journal of Accounting and Economics,* 9, 231-258.

Collins, D. W., S.P. Kothari, J. Shanken, and R. G. Sloan, (1994). Lack of timeliness versus noise as explanations for low contemporaneous return-earnings association. *Journal of Accounting and Economics,* 18, 289-324.

Dumontier P. and Labelle R. (1998). Accounting earnings and firm valuation: the French case. *The European Accounting Review*, 7 (2), 163-183.

Gonedes N. J. and N. Dopuch, (1974). Capital market equilibrium, information production, and selecting accounting techniques: Theoretical framework and review of empirical work. *Journal of Accounting Research*, 12 (Suppl.), 48-130.

Harris, T. S. M. Lang. and H. P. Moiler, (1994). The value-relevance of German accounting measures: An empirical analysis. *Journal of Accounting Research*, 24, 187-209.

Kormendi R. and R. C. Lipe, (1987). Earnings innovations, earnings persistence and stock returns. *Journal of Business*, 60, 323-345.

Kothari, S. P., Zimmerman, J. L (1995). Price and return models. *Journal of Accounting and Economics*, 20, 155-192.

Landsman, W. R., (1986). An empirical investigation of pension fund property rights. *The Accounting Review*, 61, 662-691.

Lev B. and J. A. Ohlson, (1982). Market based empirical research in accounting: A review, interpretations, and extensions. *Journal of Accounting Research,* 20 (Suppl.), 249-322.

Lev B. (1989). On the usefulness of earnings: Lessons and directions from two decades of empirical research. *Journal of Accounting Research*, 27(Suppl.), 153-192.

Martikainen T. and Perttunen J. (1991). Return intervals, systematic risk and firm size: Evidence form a thin security market. *Economics Letters*, 36, 311-315.

Martikainen, T., Kallunki, J. and Perttunen, J. (1997). Finnish earnings response coefficients: the information content of losses. *The European Accounting Review*, 6 (1), 69-81.

Ohlson, J. A., (1991). The theory and value of earnings and an introduction to the Ball-Brown analysis. *Contemporary Accounting Research*, 8, 1-19.

Ohlson, J. A. and P. K. Shroff, (1992). Changes versus levels in earnings as explanatory variables for returns: Some theoretical considerations. *Journal of Accounting Research*, 30, 210-226.

Olsen, C., (1985). Valuation implication of SFAS no. 33 data for electric utility investors. *Journal of Accounting Research,* 23 (Suppl.), 28-47.

Pritchard N. J. (2002). The Relationship between Accounting Numbers and Returns in the Baltic Stock Markets. *Discussion Paper 2002/06/May 2002,* Heriot-Watt University.

Ramakrishnan, R. and Thomas, J., (1998). Valuation of permanent, transitory and price-irrelevant components of reported earnings. *Journal of Accounting Auditing and Finance*, 13 (3), 301-336.

Raonic I., S. McLeay and I. Asimakopoulos (2004). The Timeliness of Income Recognition by Interlisted European Companies: An Analysis of Institutional and Market Complexity. *Journal of Business Finance and Accounting*, 31 (1-2), 115-148.

Vafeas N., Trigeorgis L. and Georgiou X. (1998). The usefulness of earnings in explaining stock returns in an emerging market: the case of Cyprus. *The European Accounting Review*, 7 (1), 105-124.

Watts R. L. and J. L. Zimmerman (1986). *Positive accounting theory.* Prentice-Hall, Engle-wood Cliffs, N J.

White H. (1980). A heteroscedasticity consistent covariance matrix estimator and a direct test for heteroscedasticity. *Econometrica,* 48, 817-838.

Table 1. Descriptive statistics of earnings and price variables used in the valuation of the price, return and differenced price models specifications. Sample consists of annual data from 1994-2004.

Variable	**Mean**	**St.Deviation**	**Minimum**	**Median**	**Maximum**
EPS_t	0.37	2.72	-9.39	0.17	87.8
P_t	6.26	1.,3	0.08	2.83	128.7
EPS_t/P_{t-1}	0.11	0.96	-13.2	0.045	24.5
P_t/P_{t-1}	1.82	3.77	0.05	0.928	51.42
$\Delta EPS_t/ P_{t-1}$	-0.022	1.73	-47.4	-0.0024	24.52
$\Delta P_t/P_{t-1}$	0.76	3.84	-18.62	-0.072	50.42

Sample: The overall sample consists of 105 firms with full earnings and price data for 1994-2004. EPS_t is annual earnings per share excluding taxes, extraordinary items and discontinued operations. Price variables are measured over the fiscal year. Only firms with December fiscal year-end are included. Earnings and prices are adjusted for stock splits and stock dividends in order to obtain the differenced variables.

Table 2. Regression results of stock returns on earnings using four model specifications for one-year period.

	Price Model (1)			Return Model (2)			Changes Model (3)			Deflated Model (4)		
Years	a	b	R^2 adj	a	b	R^2 adj	a	b	R^2 adj	a	b	R^2 adj
1994	1.44	2.81	33.2%							1.182	8.65	98.1%
	(5.31)*	(7.25)*	[2.08]							(71.8)*	(1.21)	[2.02]
1995	1.69	2.53	23.2%	1.11	0.0597	1%	0.123	0.273	2.7%	2.77	-27.8	72%
	(5.45)*	(5.69)*	[1.99]	(28.68)*	(1.41)***	[1.88]	(3.25)*	(1.96)**	[1.9]	(16.45)*	(-1.96)**	[2.05]
1996	1.49	1.86	19.3%	0.835	-0.0095	0%	-0.166	-0.0029	0%	1.52	-2.73	97.9%
	(5.43)*	(5.08)*	[2.03]	(28.08)*	(-0.43)	[2.22]	(-5.58)*	(-0.03)	[2.23]	(70.03)*	(-1.02)	[2.00]
1997	1.84	2.76	20.4%	1.18	0.488	12.7%	0.287	0.0266	0%	3.52	3.412	100%
	(4.89)*	(5.26)*	[2.07]	(18.13)*	(4.01)*	[1.83]	(4.48)*	(0.50)	[1.82]	(500.6)*	(0.041)	[2.04]
1998	3.38	4.12	25%	2.13	-0.360	0%	1.09	-0.144	0%	3.70	-8.1	95.7%
	(7.17)*	(5.98)*	[1.77]	(12.97)*	(-0.93)	[2.13]	(6.84)*	(-0.33)	[2.14]	(48.1)*	(-0.253)	[1.97]
1999	21.5	10.7	5.5%	8.19	8.62	2.1%	8	2.72	0%	22.07	-101.98	81.6%
	(7.99)*	(2.66)*	[1.77]	(8.33)*	(1.79)**	[1.73]	(9.13)*	(0.58)	[1.71]	(21.55)*	(-1.27)	[2.06]
2000	7.4512	6.3742	28.9%	0.331	2.62	10.7%	-0.641	0.389	0%	2.20	-254.2	100%
	(8.82)*	(6.57)*	[1.69]	(19.64)*	(3.66)*	[1.86]	(-38.46)*	(0.70)	[1.91]	(6879.11)*	(-1.14)	[2.02]
2001	5.476	7.3112	36.5%	0.812	-0.302	0%	-0.194	-0.082	0%	2.86	-5.94	64.8%
	(7.50)*	(7.80)*	[2.03]	(24.10)*	(-0.88)	[2.22]	(-5.88)*	(-0.24)	[2.24]	(13.9)*	(-0.37)	[2.06]
2002	4.1738	0.0155	0%	0.616	-0.0034	0%	-0.384	-0.0036	0%	2.37	2.14	68.9%
	(7.31)*	(0.24)	[2.08]	(21.91)*	(-0.29)	[1.99]	(-13.7)*	(-0.31)	[1.99]	(14.61)*	(0.47)	[1.99]
2003	3.7761	4.8289	39.2%	1.24	0.025	0%	0.239	0.00398	0%	7.98	446.2	100%
	(7.13)*	(8.25)*	[2.17]	(30.07)*	(0.23)	[2.13]	(5.92)*	(0.46)	[2.14]	(51427)*	(1.16)	[2.04]
2004	3.6197	2.6015	29.6%	0.782	0.220	4.7%	-0.712	1.17	8%	1.025	7.85	89.8%
	(6.09)*	(6.69)*	[2.31]	(23.28)*	(2.47)*	[1.79]	(-3.79)*	(3.17)*	[2.03]	(30.21)*	(1.36)***	[2.13]
Pooled	6.1479	0.3074	0.5%	1.81	0.042	0%	0.7611	0.020	0%	7.56	24180	96.2%
	(18.36)*	(2.52)*	[1.29]	(15.44)*	(0.34)	[0.99]	(6.41)*	(0.30)	[1.01]	(170)*	(0.70)	[2.00]

White (1980) T statistics are in the parenthesis (* significant at a = 1%, **significant at a =5%, *** significant at a = 10%), numbers in brackets is Durbin-Watson Statistic

Price Model: $P_t = a + bEPS_t + e_t$ - **Return model**: $P_t/P_{t-1} = a + bEPS_t / P_{t-1} + e_t$

Differenced Model: $\Delta P_t/P_{t-1} = a + b\Delta EPS_t / P_{t-1} + e_t$ - **Deflated Model**: $P_t/EPS_t = a1/EPS_t + b + e_t$

Table 3. Regression results of portfolio returns on earnings using four model specifications for one-year period.

	Price Model (1)			**Return Model (2)**			**Changes Model (3)**			**Deflated Model (4)**		
Years	a	b	R^2 adj	a	b	R^2 adj	a	b	R^2 adj	a	b	R^2 adj
1994	1.46	3.23	36.3%							2.46	-1.39	83%
	(1.52)***	(2.53)*								(78.76)*	(0.27)	
1995	1.30	4.43	42.12%	0.79	1.07	37.2%	0.211	1.57	16.1%	1.49	3.34	76.8%
	(1.51)***	(2.57)*		(9.5)*	(2.77)*		(1.28)***	(1.47)***		(16.7)*	(1.19)	
1996	1.15	2.75	34%	0.695	0.448	31%	-0.253	1.07	24.9%	1.15	0.68	65.4%
	(1.52)***	(2.44)*		(6.14)*	(0.424)		(-3.54)*	(1.21)		(9.13)*	(1.15)	
1997	1.24	2.81	47%	0.903	0.78	44.6%	0.183	1.21	63.9%	1.49	85	80.6%
	(2.10)**	(3.19)*		(5.31)*	(2.36)*		(0.71)	(0.44)		(36.27)*	(0.79)	
1998	1.34	7.14	58.8%	1.42	0.978	46.6%	0.89	-1.44	55.6%	2.35	5.24	72.8%
	(2.27)**	(3.19)*		(6.45)*	(0.707)		(3.3)*	(1.25)		(32.6)*	(1.14)	
1999	19.4	27.8	55.9%	6.1	39.8	54.6%	7.35	-5.17	42.6%	20.8	-127.7	94.6%
	(3.6)*	(2.13)**		(2.68)*	(3.55)*		(4.19)*	(1.18)		(42.8)*	(-0.35)	
2000	4.95	17.7	42.9%	0.231	11.5	55%	0.69	4.57	40.9%	6.24	-28.7	74.7%
	(2.14)**	(3.34)*		(4.43)*	(3.31)*		(17.59)*	(0.95)		(6791.3)*	(0.80)	
2001	2.27	17.74	66.5%	0.729	2.196	45.4%	0.25	0.96	34.1%	5.1	-13.3	79.8%
	(2.58)*	(5.25)*		(8.52)*	(1.69)**		(2.85)*	(0.62)		(12.2)*	(0.48)	
2002	2.51	9.33	48.4%	0.54	1.94	40%	-0.40	1.49	27.7%	2.77	3.71	77.5%
	(2.55)*	(3.28)*		(9.65)*	(2.53)*		(-7.08)*	(1.33)***		(8.89)*	(0.72)	
2003	2.68	9.87	50.3%	0.989	2.61	34.3%	0.298	3.76	25.9%	3.5	-9.16	79.2%
	(2.52)*	(3.11)*		(10.7)*	(2.24)**		(2.96)*	(1.76)**		(195480)*	(0.248)	
2004	2.12	6.54	58%	0.761	1.65	41.9%	-0.141	0.68	40.9%	1.90	8.16	68.1%
	(2.28)**	(4.45)*		(8.59)*	(1.95)**		(-1.31)***	(-0.69)		(13.2)*	(0.37)	
Pooled	4.78	4.38	17.2%	1.53	1.89	2%	0.567	-1.89	8.13%	5.04	-25.3	70.7%
	(4.57)*	(4)*	[1.04]	(4.29)*	(1.55)***	[1.01]	(1.94)**	(1.12)	[1.02]	(47141)*	(0.36)	[1.98]

White (1980) T statistics are in the parenthesis (* significant at a = 1%, **significant at a =5%, *** significant at a = 10%), numbers in brackets is Durbin-Watson Statistic.

Price Model: $P_t = a + bEPS_t + e_t$ - **Return model**: $P_t/P_{t-1} = a + bEPS_t / P_{t-1} + e_t$

Differenced Model: $\Delta P_t/P_{t-1} = a + b\Delta EPS_t / P_{t-1} + e_t$ - **Deflated Model:** $P_t/EPS_t = a1/EPS_t + b + e_t$

Table 4. Regression results of stock returns on earnings using four model specifications for five-year periods.

	Price Model (1)			Return Model (2)			Changes Model (3)			Deflated Model (4)		
Years	a	b	R^2 adj	a	b	R^2 adj	a	b	R^2 adj	a	b	R^2 adj
1994-98	8.39	3.47	30.9%							3.51	-172	99.9%
	(5.21)*	(6.89)*	[2.02]							(305.73)*	(-1.28)***	[2.09]
1995-99	28.4	4.8	12.8%	15.7	-0.626	1.2%	9.36	0.8	0%	3.52	112	99.8%
	(7.64)*	(4.03)*	[1.68]	(13.72)*	(-1.52)***	[.,77]	(11.02)*	(0.63)	[1.77]	(246.74)*	(0.66)	[2.06]
1996-00	32.5	6.49	18.1%	15	-0.871	1.1%	8.58	1.59	1.1%	2.2	-1518	100%
	(7.82)*	(4.89)*	[1.62]	(13.06)*	(-1.47)***	[1.78]	(10.29)*	(1.46)***	[1.82]	(986.3)*	(-0.98)	[2.07]
1997-01	35.5	8.36	25.6%	14.5	-0.06	0%	8.53	1.61	2.9%	2.2	-1518	100%
	(8.18)*	(6.06)*	[1.58]	(11.57)*	(-0.05)	[1.79]	(10.31)*	(2.02)**	[1.82]	(982.6)*	(-0.97)	[2.07]
1998-02	47.6	0.649	0.5%	13.9	-0.178	0%	7.96	-0.282	0%	2.2	-21	100%
	(9.89)*	(1.25)	[1.88]	(12.13)*	(-0.39)	[1.82]	(9.21)*	(-0.79)	[1.80]	(5943)*	(-0.08)	[2.02]
1999-03	47.6	0.809	1.2%	13.1	-0.073	0%	7.07	0.163	0%	7.57	265669	96.1%
	(9.48)*	(1.50)***	[1.92]	(11.28)*	(-0.16)	[1.79]	(8.11)*	(0.43)	[1.7]	(50.65)*	(0.69)	[2.01]
2000-04	26.7	1.09	7.4%	3.81	-0.0022	0%	-1.73	0.131	0.9%	7.57	265517	96.1%
	(7.83)*	(3.06)*	[2.11]	(53.69)*	(-0.08)	[2.07]	(8.09)*	(1.41)***	[1.88]	(50.65)*	(0.69)	[2.01]
Pooled	36.1	1.35	4.1%	12.6	-0.249	0.1%	6.63	0.274	0.1%	7.03	171476	91.4%
	(23.04)*	(5.69)*	[1.72]	(27.78)*	(-1.25)	[1.57]	(18.84)*	(1.37)***	[1.43]	(88.44)*	(1.49)***	[2.01]

White (1980) T statistics are in the parenthesis (* significant at a = 1%, **significant at a =5%, *** significant at a = 10%), numbers in brackets is Durbin-Watson Statistic.

Stock returns and earnings variables were summed over five year periods.

Price Model: $P_t = a + bEPS_t + e_t$ - **Return model**: $P_t/P_{t-1} = a + bEPS_t / P_{t-1} + e_t$

Differenced Model: $\Delta P_t/P_{t-1} = a + b\Delta EPS_t / P_{t-1} + e_t$ - **Deflated Model**: $P_t/EPS_t = a1/EPS_t + b + e_t$

Table 5. Regression results of portfolio returns on earnings using four model specifications for five-year periods.

	Price Model (1)			Return Model (2)			Changes Model (3)			Deflated Model (4)		
Years	a	b	R^2 adj	a	b	R^2 adj	a	b	R^2 adj	a	b	R^2 adj
1994-98	7.59	3.62	40.7%							2.06	109.5	85.2%
	(1.63)***	(2.62)*								(45.2)*	(0.67)	
1995-99	24.8	4.09	41.7%	11.26	0.27	28.8%	8.67	-21.34	39.8%	8.62	-122.7	85.4%
	(2.7)*	(2.34)**		(11.8)*	(-0.69)		(4.41)*	(-1.05)		(30.22)*	(0.09)	
1996-00	28.62	7.57	40.4%	10	-0.40	43.5%	9.33	-19.16	33.6%	6.56	-230.4	82.5%
	(2.46)*	(2.44)*		(16.05)*	(-1.74)**		(3.49)*	(0.30)		(635.7)*	(0.62)	
1997-01	32.38	12.18	45.7%	10.35	0.1	45%	6.99	-20.06	46.3%	6.09	97.5	80.7%
	(2.38)**	(3.05)*		(14.6)*	(-1.11)		(3.29)*	(-1.58)***		(630)*	(0.9)	
1998-02	28.5	12.3	55.6%	12.5	-0.91	32.9%	5.17	-9.10	32.1%	7.16	-195.8	81.6%
	(2.73)*	(3.33)*		(5.74)*	(-0.79)		(3.73)*	(-0.29)		(572.8)*	(0.275)	
1999-03	28.28	14.9	63.5%	16.45	-8.3	46.9%	5.06	5.36	56.9%	9.78	-80.23	82.6%
	(3.16)*	(4.19)*		(3.99)*	(0.95)		(4.07)*	(0.46)		(17345)*	(0.16)	
2000-04	14.11	12.52	57.6%	3.71	1.57	45.9%	-1.41	5.44	47.2%	5.76	-22.75	80.1%
	(2.15)**	(5.35)*		(18.05)*	(2.29)**		(-4.52)*	(1.53)***		(23435)*	(0.82)	
Pooled	31.9	4.65	26.6%	10.2	1.94	19.1%	5.63	-4.71	11.2%	5.88	-165.3	77.2%
	(6.36)*	(4.24)*	[1.51]	(11.38)*	(2.01)**	[1.55]	(6.06)*	(-0.51)	[1.34]	(28748)*	(0.40)	[2.13]

White (1980) T statistics are in the parenthesis (* significant at a = 1%, **significant at a =5%, *** significant at a = 10%), numbers in brackets is Durbin-Watson Statistic.

Stock returns and earnings are summed over five year periods. In each period 15 portfolios were created with 7 stocks each and models 1-4 were re-estimated for all the portfolios in each period.

Price Model: $P_t = a + bEPS_t + e_t$ - **Return model**: $P_t/P_{t-1} = a + bEPS_t / P_{t-1} + e_t$

Differenced Model: $\Delta P_t/P_{t-1} = a + b\Delta EPS_t / P_{t-1} + e_t$ - **Deflated Model**: $P_t/EPS_t = a1/EPS_t + b + e_t$

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Order Aggressiveness on the ASX Market

Ying Xu
Faculty of Economics
Humboldt-Universität zu Berlin
Unter den Linden 6, 10117, Berlin, Germany
E-mail: muhualanxi@hotmail.com

Abstract

This paper investigates how traders on the Australian Stock Exchange (ASX) market make their joint decision on price aggressiveness and quantity when they submit an order, based on information from limit order book and stock price movement. We use a simultaneous-equation system including an Ordered Probit model to account for the discrete nature of price aggressiveness and an OLS model to fit the continuous quantity. The results suggest a negative trade-off between price aggressiveness and quantity. Factors such as depths at the best quotes, bid-ask spread, volatility and price changes are major determinants of the traders' order submission decision. This paper also provides evidence for a U-shaped intraday pattern of order flows and positive serial correlation in order type.

Keywords: Limit order book, Price aggressiveness, Ordered probit model, Order flow

1. Introduction

Many stock exchanges around the world, such as the Australian Stock Exchange (ASX), have adopted a fully computerized order-driven trading system. In such a market, liquidity is provided by participants who submit orders to buy or sell electronically to Stock Exchange Automated Trading System (SEATS). The electronic limit order market gains its popularity because of the greater transparency offered by this system compared to dealer market setting. One critical part of study on the limit order trading process is how the trader makes the decision to trade. Following the paper of Lo and Sapp (2007), traders are assumed to confront with a simultaneous choice on both price aggressiveness and quantity. Price aggressiveness is defined by the execution priority: markets orders, those to buy or sell a pre-specified quantity at the best available price and thus being executed immediately, is regarded as the most aggressive orders. In contrast, limit orders are orders that buy or sell a pre-specific quantity at a pre-specific price. A trade results only when the incoming bid or ask order matches or overlaps the price of the best opposite quote (the lowest limit sell or highest limit buy), otherwise, the limit order is stored on the limit order book followed strict price and time priority waiting for execution. We suppose that traders collect information from limit order book, for example, depths at best quotes, bid-ask spread, together with information from market price movement to decide an order submission. Therefore, order flow has information content and traders behave strategically using limit orders. This paper is designed to empirically investigate some features of order flow and the determinants for traders to make the joint decision on price aggressiveness and quantity.

In order to test the joint decision on price aggressiveness and quantity based on trader's information set, we set up a simultaneous-equation system, including a price aggressiveness equation estimated via Ordered Probit model and a quantity equation fitted via OLS method. Examining order submission and trading data for five most traded stocks on the ASX market during July and August 2002, the results indicate that: First, there exist a negative trade-off between price aggressiveness and quantity, i.e. traders tend to employ a strategy to combine more aggressive orders with smaller size, aiming to increase the probability of execution and meanwhile reduce transaction costs and minimize potential loss. Second, factors like depths at best quotes, bid-ask spread, volatility and price changes explicitly affect the decision criterion used by market participants. Third, the analysis of order flows on the ASX market supports the "diagonal effect" and shows a positive autocorrelation in order type; orders exhibit a U-shaped pattern during normal trading periods. In most cases, the empirical results we find are consistent with theory and literature for other market structures.

The remainder of this paper is organized as follows. Section 2 reviews relevant theoretical and empirical literatures. Section 3 introduces economic framework and formulates hypotheses. Section 4 explains the econometric model. Section 5 presents a detailed description of the ASX trading system and data set. Section 6 reports and analyzes the empirical findings on limit order submission. Section 7 provides some concluding remarks.

2. Theoretical and Empirical Literature

One stream of the theoretical literature makes restrictive assumptions about the behavior of informed traders, while some ignore such traders completely. When studying the limit order market, they focus on the tradeoff between the immediate execution of placing the market order versus the better price, and uncertain execution, of submitting a limit order. As elaborated by Hasbrouck (2007), a key determinant of order submission choice is the probability that a particular limit order will be executed, while the execution probabilities are determined by the order choice decisions of other agents. An ask order may be executed only when it meets an offering price at the other side, and the latent reservation values may trigger market orders. Therefore, there is an obvious relationship between an agent's own order choice problem and the others' order choice in the past and in the future. A set of traders are assumed to follow the trading strategy by the former traders, and the prerequisite that the solution to order choice problem should be consistent with the assumed price and execution processes is necessary for the market to be in dynamic equilibrium. The early study by Cohen et al. (1981) develops a gravitational pull model to explain the trading decision in limit order markets by a dynamic balance of the relative costs of price improvement and execution risk. They find that the end-of-period wealth of participants who choose to trade via a limit order is an increasing function of the order execution probability. The larger the expected execution probability, the shorter the expected waiting time and thus the smaller the expected adverse selection cost. They further indicate that, as spread narrows, the benefits of the better price available to limit order traders decrease, causing more traders to prefer the certain execution of the market order. Some researchers like Foucault et al. (2005) tend to classify investors as patient and impatient traders, and argue that prices and times-to-execution are jointly determined in equilibrium and there exist intertemporal trade-offs involved with limit orders. Hollifield et al. (2003) build a structural model of a pure limit order market, which captures the trade-off between order price and probability of execution. They estimate their model nonparametrically and derive implications for trader's order submission strategies.

In some theoretical literature on price discovery and limit order book's informativeness, researchers believe that limit orders are not as informative as market orders. Glosten (1994) and Rock (1996) explicitly incorporate informed traders into their models, but assume that they always use market orders instead of limit orders. Their common belief is that with short-lived private information, informed traders will be impatient and prefer market orders because these orders guarantee immediate execution. On the other hand, without private information on the direction of price movements, limit order submitters will face an adverse selection risk, as such orders will be more likely to execute when they generate a loss to the uninformed submitters. For instance, if the informed trader knows that the current market price is too high and the price will decrease in the near future, only a bid limit order with relatively higher price could be executed. Since the submitter purchases at a price over the security's true value, the execution of this submission is a loss for him, and this phenomenon is the winner's curse.

The second stream of the theoretical literature takes market characteristics such as state of limit order book as exogenous variables and examines how both liquidity and informed traders choose between limit and market orders. Angel (1994) derives an analytical expression for the probability of limit order execution, conditional upon an investor's information set. His result indicates that informed traders are less likely to use limit order than liquidity traders, especially when the realized asset value is further away from its expected value and therefore they have the opportunity to capitalize on their private information. Handa and Schwartz (1996) model the effects of liquidity traders and informed traders on the choice between market and limit orders, and find that limit order submitters lose from trading with informed traders but gain from accentuated short-run volatility induced by liquidity trades. Harris (1998) predicts that liquidity traders who need to meet a target will start by using limit orders and then switch to market orders as the end of trading or their deadline approaches. Similarly, the probability of submitting a limit order by the informed traders decreases with time until the end of trading, when their information is revealed. Using an experimental market setting to investigate the evolution of liquidity in an electronic limit order market, Bloomfield et al. (2005) confirm Harris' suggestion that large liquidity traders prefer limit order early in the trading period, but their preferences shifts over time toward using more market orders. Contradicting the theoretical prediction, informed traders start more likely to employ market orders and then shift to trading mostly using limit orders. This may reflect that informed traders search for a source of profit, earning the bid-ask spread via limit order submission. They further conclude that the informed traders' behavior results in a provision of liquidity to market and allows the market to endogenously create liquidity even in the presence of information asymmetry.

The third stream of literature is consistent of dynamic equilibrium models in which key market attributes arise endogenously. A trader's optimal strategy depends on conjectures of other traders' strategies.

Due to Parlour's (1998) equilibrium model, depth on either side of the market has an impact on traders' choice between market and limit orders. The increase in book depth on the bid side decreases the probability that a buyer will submit a limit order; while a depth's increase on the offer side increases the probability that a buyer will submit a limit order. As a consequence of limit order queuing, when adding an order to a book which is already with large depth, i.e. many more orders standing in front of this order, the trader receives a lower probability of execution. Although there is no uncertainty about the asset's cash flows, nor is there price determination in Parlour's model, it illustrates a key equilibrium interaction between a trader's order choice and his beliefs about similar choices made by others in the future.

The Foucault's (1999) equilibrium model provides an analysis of individual order choice, when the opposite side depth is deep, a trader is more likely to use a limit order. If the fundamental risk of a security increases then a given limit order faces a higher pick-off risk, causing limit order trader place less aggressive prices and then the spread widens. Later, market orders become more expensive, leading traders to favor limit orders. Additionally, Foucault (1999) argues that there is a direct relationship between bid-ask spread and volatility, and an inverse relationship between market depth and volatility. The reason might lie on: when the market is volatile, the probability of trading against informed investors increases; therefore, the expected loss for uformed traders is larger. To protect themselves from trading disadvantage, limit buy (sell) order traders have to post lower (higher) bid (ask) price or reduce their order sizes. Thus, the price volatility determines the trader's decision between market and limit orders through spread and depth.

On the empirical side, numerous studies analyze the order book, order flow and especially examine the intraday variation in bid-ask spread. A U-shaped intraday pattern is widely found in the spreads of NYSE stocks. However, Brock and Kleidon (1992), Chan, Chung and Johnson (1995) attribute this pattern to specialists' market making behavior, while Chung and Van Ness (1999) show this pattern largely reflects the intraday behavior of the spread established by limit order traders. Al-Suhaibani and Kryzanowski (2000) survey the behavior of market participants in a pure order-driven market without market makers, Saudi Stock Market, and find that its intraday patterns are surprisingly similar to those found in other markets with different structures, including the persistent U-shaped patterns in number of shares traded, volume, and bid-ask spread. In their paper, the call market is taken as a contributing factor to the concentration at the opening, whereas the high level of limit orders at the end of every trading session could be a result of limit price adjusting.

Other empirical studies study the question whether and to what extent the state of limit order book influences traders' order submission strategies. Biais et al. (1995) are among the first to study this problem, using data on order flow on the Paris Bourse. They report that investors tend to place limit (market) orders when the spread is wide (narrow) or the depth is low (high). To compete for price and time priority, traders quickly submit limit orders within the quotes when the depth declines or the spread widens. And the unexecuted orders are less competitive orders that at prices away from the quotes. Griffiths et al. (2000) examine the costs and determinants of order aggressiveness through the limit order submissions on the Toronto Stock Exchange during June of 1997, and report that aggressive orders have large price impacts but smaller opportunity costs than passive orders. Additionally, aggressive buy (sell) orders are observed to tend to follow other aggressive buy (sell) orders and occur when bid-ask spreads are narrow and depth on the same (opposite) side of the limit book is large (small), resulting in improved execution probabilities in the more competitive market states. Finally, an analysis of longer-term returns to holding stock subsequent to the order suggests that aggressive buyers are more likely to be motivated by information, whereas aggressive sellers tend to be motivated by liquidity. Applying the similar method proposed by Biais et al. (1995) and Griffiths et al. (2000) to the data of 15 Swiss Stock Exchange issued during March and April 1997, Ronaldo (2004) shows that patient traders become more aggressive when the own (opposite) side book is thicker (thinner), which are consistent with earlier studies. He argues that this result demonstrates that the depth of limit order book is a proxy for the execution probability of an incoming trade and strengthens or weakens trading aggressiveness. Besides, transient volatility and a wider spread are found to encourage limit order placement and discourage market order submission. Similar investigation on NYSE market is conducted by Beber and Caglio (2005), whose results indicate that order aggressiveness depends on the state of the order book and on the asset dynamic, and the most important determinants of order submission strategies are the depth on the same side of the book and a momentum indicator. They further support their conjecture that in the case of information-based trading, orders are less aggressive, by a different response to changes in the investor's information set and by a stronger price impact of less aggressive orders.

According to the empirical studies on the Australian Stock Exchange (ASX) market, Verhoeven et al. (2004) examine factors that affect the decision used by market participants for the stocks CML and BHP, using logit regressions. Their results indicate that the bid-ask spread, depth at the best price, price changes in the last five minutes and order imbalance are major determinants for traders to choose between market and limit orders. Cao et al. (2004) employ a more comprehensive dataset, covering 21 constituent stocks from the ASX-20 index to study the order submission strategies of traders using an Ordered Probit modeling approach. They provide evidence to confirm that traders use the available information on the state of the limit order book, including spread, depths on the same and opposite sides, when developing their order submission strategies. More recently, Hall and Hautsch (2006) conduct a dynamic investigation on five most traded stocks from ASX market using a six-dimensional autoregressive conditional intensity (ACI) model, in order to study the determinants of order aggressiveness and traders' order submission strategy. They find that market depth, the queued volume, bid-ask spread, recent volatility, as well as recent changes in order flow and stock price are important determinants of order aggressiveness, consistent with theoretical predictions.

3. Economic framework and test design

3.1 Order submission strategies

When making the decision of buy or sell a stock, the investor's order submission strategy is not just a choice between a market and a limit order. A trader can submit very different limit orders, trading off probabilities of execution with transaction costs. Moreover, the order placement is a joint decision of price and quantity: large orders can be split into smaller ones. Thus, the order aggressiveness of submission is defined as the investor's preference for an immediate and certain execution, coupled with his indifference for the transaction costs and risks. Greater price aggressive orders have a price priority in execution, but suffer from price risk, originating from the bid-ask spread and market depth. On the other hand, limit orders confront with execution risk and also adverse selection risk, because these orders are more likely to be executed when the value of the asset has changed adversely. The simultaneous choice of order size is also important to traders. For limit orders, both the execution risk and adverse selection risk increases with the ordered quantity, since that a large-sized order may not be executed entirely and a large loss may occur when the value of underling stock changes. Lo and Sapp (2007) provide an empirical study of the joint nature of order price and quantity in the limit order market and they find a significantly negative relationship between price aggressiveness and order size through explicitly examining the trade-offs between these two dimensions, conditioning on the state of the limit order book. They offer two explanations why more aggressive orders are smaller: order splitting or order consolidating (Lo and Sapp, 2007, p.14). Since large orders may reveal more information, as suggested by Easley and O'Hara (1992), informed investors split their orders into smaller amounts to trade without information disclose. Besides, breading up large orders may also reduce transaction cost and picking-off loss, because a large, aggressive limit orders make a trader to suffer from a great loss if the value of the security moves against the trader and he cannot cancel the order in time. Consequently, traders prefer to place more aggressive but smaller order, which have priority in execution and in the meantime minimize the pick-off risk. In the case of order consolidation, traders submit small orders to fulfill immediate liquidity needs. Because liquidity-oriented traders require an immediate execution, they usually place aggressive orders to obtain the priority, and then we could also observe a negative relation between price aggressiveness and order size. Accounting for this idea, we assume that:

Trade-off between price aggressiveness and quantity hypothesis: There is a negative relationship between price aggressiveness and quantity.

Previous studies document a conditional order flow pattern called a diagonal effect, which refers to a higher likelihood to observe a given type of order after an event just occurred than it would be unconditionally. When observing order flow pattern, both Biais et al. (1995) from Paris Bourse and Al-Suhaibani and Kryzanowski (2000) from the Saudi Stock Market find that: after the arrival of a limit buy order, an incoming order is most likely to be the same order type. Biais et al. (1995) offer three explanations to account for the correlation of market orders: strategic order splitting, trade imitation, or similar reaction to information. First, traders may strategically split large orders to reduce price impact. Lo and Sapp (2007) test the autocorrelation of market orders and aggressive limit orders, and their findings suggest that dealers split their orders to either hide their information or to decrease the execution costs. Second, if certain market participants can observe the orders of others whom they believe to be informed, then they may imitate their orders. However, Al-Suhaibani and Kryzanowski (2000) employ a dataset which allow them to refer a subset of orders belonging to the same trader, and find that the proportion of the same trader involved in two consecutive orders is 28.94% of all order flow events. Their findings support strategic splitting rather than an imitation. Third, news events may lead various traders to trade in a similar manner, i.e. the diagonal effect may be caused by a similar reaction of market participants to a particular stock, an information event, or the economy as a whole. In line with this interpretation, Wermers (1999) shows that institutional investors buy (sell) more stocks in the US after it experiences extreme positive (negative) returns in the previous quarter.

The diagonal effect of limit orders within the best quotes is explained by the undercutting and overbidding behavior of traders competing to supply liquidity to the market (Biais et al. 1995). When the number of limit orders placed inside the quote increases, the spread narrows and the cost of submitting a market order reduces, so that it is more likely to place more aggressive orders. The succession of less aggressive orders such as cancellations is explained as trading imitation or similar reaction to the same events. The diagonal effect leads to the hypothesis:

Price aggressiveness autocorrelation hypothesis: The higher is the price aggressiveness of the previous orders, the higher is the aggressiveness of the new orders placed.

3.2 Exogenous explanatory variables

In this part, a number of hypotheses are developed to embody the intuitive idea that investors optimally choose their order submission strategy, including price aggressiveness and order size, so as to maximize their expected profits on the basis of their own information sets, which can be distinguished with limit order book-related variables, market movement-related variables and time-related variables.

3.2.1 Limit order book-related variables

According to theoretical work on order choice, Parlour's (1998) finds that depth on either side of the market has an impact on traders' choice between market and limit orders. Harris (1998) and Foucault (1999) predict that market orders become less and limit orders become more attractive as the spread increases. Foucault et al. (2005) further point out that limit order aggressiveness either increases or decreases with the inside spread, depending on whether patient or impatient traders dominate the trading population. The general empirical findings indicate that the rate of limit order submissions increases with the size of the spread, and that depth at the top of both sides of the book affects order choice, consistent with the theoretical literature.

In order to examine whether and to what extent the state of the limit order book influences traders' order submission strategies, order book statistics which proxy for depths on both sides and bid-ask spread will be included as explanatory variables.

Depth

Depth, which usually refers to the total volume (number of shares) of limit orders on one side of the order book, affects the probability of limit order execution because time priority is enforced after price priority by ASX trading rules. As to Parlour (1998), the execution probability depends on the size of the book and on the trader's belief about further order arrivals, i.e. the aggressiveness of orders arriving over the remainder of the trading day. Therefore, the thicker is the buy side of the order book, the more possible that the incoming buyer submits a market order, since the lower the probability of a new limit bid order being executed. Besides the direct competition effect, sellers rationally anticipate a crowding-out effect of buy limit orders and are thus better off submitting less aggressive sell orders (the potential strategic effect). Both of these effects reduce the incentive to submit a buy limit order when the depth is thicker at the bid. This logic can also hold for the seller's submission decision and the crowding-out effect is symmetric. Thus, we hypothesize that:

Hypothesis 1: The greater the depth of the bid (ask) side, the stronger the order aggressiveness of the incoming buyer (seller).

Hypothesis 2: The greater the depth of the ask (bid) side, the weaker the order aggressiveness of the incoming buyer (seller).

According to the potential strategic effect, which is tested through hypothesis 2, we need to take the characteristic of ASX market into consideration. On the ASX market, market orders are allowed to walk up the book without any market intervention. When there is a bid order imbalance, i.e. greater depth on the bid side than on the ask side, investors perceive this security owning a high valuation and thus increase their demand for this security. A larger demand compared with supply makes the security's price more likely to rise. The increasing price leads to a higher risk of non-execution of a new limit bid, causes a more serious competition among all bids and then increases bid prices to approach to the best ask price. Because the gravitational pull effect of the best ask becomes stronger, as stated by Handa et al. (2000), the execution probability of a limit buy order becomes more uncertain and a buy market order becomes more attractive. As such, we hypothesize that:

Hypothesis 3: The larger the depth at the ask side relative to the bid side, the stronger (weaker) the order aggressiveness of an incoming seller (buyer).

Bid-ask spread

In an order driven marketplace, the spread represents the margin that a holder (purchaser) of shares is required to surrender (pay) to a prospective purchaser (seller) away from the equilibrium price to trade immediately (Note 1). The bid-ask spread is confirmed by empirical research as a function of proxies for the cost of liquidity provision and competition. On one side, a limit order placed within the spread gains the highest price priority and thus has the highest probability of execution. On the other side, the cost of submitting a market order, especially paying the spread, increases as the spread widens. In addition, the larger the spread, the lower the gravitational pull effect exerted by the opposing side (Cohen et al., 1981), and the less possible that the price change, which attributes to a loss suffered by limit order submitter due to the winner's curse. This consideration leads to the following hypothesis:

Hypothesis 4: The wider the spread, the weaker the order aggressiveness.

3.2.2 Market movement-related variables

This group of variables focuses on the investor's information set describing stock price evolution, including the factors volatility and price changes. Grammig, Heinen and Rengifo (2004) assert that the market movement-related variables could be used to predict the future order submission process. When it indicates "bad news", the number of aggressive sell limit and market sell orders increase while buyer activity decreases.

Volatility

The literature has identified various effects that volatility may have to influence order submission behaviors. For instance, Cohen et al. (1981) point out that as price uncertainty increases, risk-averse traders reward a premium on a certain outcome regarding the execution of their trades. As a consequence, more aggressive orders like market orders are placed by risk-averse traders since higher volatility increases the dispersion of outcomes for a limit order submission.

On the contrary, Copeland and Galai (1983) suggest that posting an order is more costly when volatility is higher, limit order traders are therefore less aggressive, because of the risk of being picked off by informed traders. Consistent with this conjecture, Handa et al. (1998) assert that limit order traders always lose from trading with informed traders who have superior information, but they gain from trading with liquidity traders who must meet their liquidity need. However, regardless of whether volatility is caused by liquidity traders or a difference of opinions as to the true price of a security, higher price volatility indicates that a trader faces with a greater opportunity of executing his order at a better price. In this sense he could prefer a limit order rather than a market order. Besides, Lo et al. (2002) show that with higher volatility, the probability of the stock price hitting a limit price barrier increases and hence reduce order aggressiveness, assuming that investors prefer an expected immediate execution. As a result, we can formulate the following hypothesis:

Hypothesis 5: The higher the stock price volatility, the weaker the order aggressiveness.

Although the above hypothesis is based on a series of theories, we should notice that it may be challenged when considering an equilibrium setting, such as the model by Foucault (1999), where the actions of some traders affect others' behaviors. If the uncertainty due to an increase in the asymmetry of information across traders rises, and induces a shift to limit order trading, the market order submission in equilibrium decreases and the probability of execution drops. Therefore, trading will shift back to market order and partially offset the initial shift. This logic leads to an ambiguous relationship between volatility and order aggressiveness, which is confirmed by the mixed results of empirical studies by Hasbrouck and Saar (2004) as well as Bae et al. (2003).

Price changes

Information-motivated trades may occur with the arrival of information and the presence of informed traders in the market. Price changes could be used as a proxy for the arrival of information, where a rising price infers that there is good news being assimilated into the market and a falling price implies bad news. Informed traders tend to use market orders which offer them the immediacy to profit from their private informational advantage. Because soon after, uninformed traders will eventually learn about this information either through observing the behavior of informed traders or a public announcement by the firm, and thereby the potential of profitable gain for the informed will disappear. Kaniel and Liu (1998) find that market orders convey more information than limit orders about the value of the underlying security and suggest that informed traders are more likely to employ markets orders instead of limit orders. Therefore, we hypothesize that:

Hypothesis 6: When prices are increasing, the order aggressiveness rises at the bid side, but drops at the ask side.

3.2.3 Time of the day

Hourly dummy variables are used to capture the changing features of the market during the trading day. There are six hourly-dummies running from 10:00 to 16:00, which allow us to capture the opening and closing effects. Bloomfield et al. (2005) point out that during market openings there is a large quantity of new information arriving at the market, and therefore less aggressive orders are submitted at market openings because of the uncertainty. Furthermore, according to the call market, all qualifying orders are executed at a single price at the open and traders benefit from their orders being executed at a price better than their quotes. Thus, limit order traders lose less to informed traders if they trade during the call market, and we expect less aggressive orders during openings and more aggressive orders over the trading day. As the end of the trading session approaches, liquidity traders are likely to place more aggressive orders since they must meet their liquidity need. A high level of limit orders could also be anticipated, resulting from price adjusting: canceling an existing order and submitting a new one. Greater proportions of large orders are executed at the end of session, partially due to liquidity trading and also because small trades at the opening contribute to price discovery and large trades tend to occur after price discovery has already occurred (Biais et al., 1995). Applying the theoretical prediction we can formulate hypothesis as follows:

Hypothesis 7: Traders submit less aggressive order during the earlier hours, but more aggressive orders when the end of the trading day approaches.

Table 1 summarizes the abbreviation and description for each exogenous explanatory variable analyzed in this study, and also shows the seven hypotheses tested thereafter and their related empirical studies.

4. Econometric models

The empirical analysis of price aggressiveness is conducted using an Ordered Probit model, a method introduced in the market microstructure literature by Hausman et al. (1992) to study the nature of discrete price changes, and widely employed by empirical investigations such as Ranaldo (2004) and Lo and Sapp (2007). The Ordered Probit model fits the price aggressiveness equation well because the different actions available to traders are inherently ordered in terms of their aggressiveness. A market order with no price limit can be considered as the most aggressive type of order as its submitter prepares to transact at any price. Limit orders are less aggressive and can be ranked in terms of their limit price. Finally, limit order cancellations, by which traders remove liquidity from the order book, are the least aggressive.

The method of categorizing price aggressiveness applied by Cao, Hansch and Wang (2004) is based on the priority in execution and described as follows: (1) market orders (immediately executed at the best available price) or marketable limit orders (limit orders at prices better than the best price standing on the opposite side of the market) which consume liquidity on the opposite side of the market; (2) best limit orders placed at prices better than the existing best price and therefore improve the current best bid or offer (3) best limit orders placed at the existing best price; (4) off-best orders placed within three ticks of the best price; (5) off-best orders placed more than three ticks away from the best price; and (6) the cancellation of a limit order that removes liquidity from the book without a transaction. The descriptive statistics of price aggressiveness on the ASX market are documented in section 5 in detail. The Ordered Probit model links the discrete observable choice of order type Z_t with the latent variable Z_t^*, which is continuous and whose domain is the set of real numbers, that is:

$$Z_t = \begin{cases} 1, & if\ Z_t^* \in (-\infty, \mu_1] \\ 2, & if\ Z_t^* \in (\mu_1, \mu_2] \\ 3, & if\ Z_t^* \in (\mu_2, \mu_3] \\ 4, & if\ Z_t^* \in (\mu_3, \mu_4] \\ 5, & if\ Z_t^* \in (\mu_4, \mu_5] \\ 6, & if\ Z_t^* \in (\mu_5, \infty) \end{cases} \tag{1}$$

where μ_j, $j = 1, 2, ..., 5,$ are additional parameters such that $\mu_1 < \mu_2 < ... < \mu_5$. Thus the range of Z_t^* is partitioned in 6 mutually exclusive and exhaustive intervals and the variable Z_t indicates the interval into which a particular observation falls. The dependent variable Z_t is ordinal and μ_j is treated as parameters to be estimated.

4.1 Estimation of Ordered Probit model

In order to explain how the Ordered Probit model is estimated, we first assume that the order quantity is exogenous, and therefore we have a single equation for price aggressiveness as follows:

$$Z_t^* = \gamma_1 Q_t + \beta' X_{t-1}^Z + \varepsilon_t^Z \tag{2}$$

where Z_t^* is the latent price aggressiveness, Q_t is the exogenous order quantity, X_{t-1}^Z is the lagged values of exogenous explanatory variables, and ε_t^Z is the error term, conditionally normally distributed with mean zero and variance σ_t^2. The conditional probability of observing a particular price aggressiveness category depends on the location of the conditional means of the underlying variables $\beta' X_{t-1}^Z$ and Q_t, relative to the partitions μ_j. With this formulation, the conditional probability of a particular observed price aggressiveness, for $2 \le j \le 5$, is given by:

$$\begin{aligned} \Pr(Z_t = j) &= \Pr(\mu_{j-1} \le Z_t^* \le \mu_j) \\ &= \Pr(\mu_{j-1} - \gamma_1 Q_t - \beta' X_{t-1}^Z \le \varepsilon_t^Z \le \mu_j - \gamma_1 Q_t - \beta' X_{t-1}^Z) \\ &= F(\mu_j - \gamma_1 Q_t - \beta' X_{t-1}^Z) - F(\mu_{j-1} - \gamma_1 Q_t - \beta' X_{t-1}^Z) \end{aligned} \tag{3}$$

where F is the cumulative distribution function of ε_t^Z and is assumed to contain no additional unknown parameters, so that, for instance, the variance of ε_t^Z is known. This assumption fixes the scale of the measurement of Z_t^*, but not the origin. In most cases, the identification could be achieved by assuming a zero intercept, i.e. X_{t-1}^Z does not contain a constant term. Therefore, the full set of probabilities of the possible price aggressiveness is:

$$\begin{aligned} &\Pr(Z_t = 1) = F(\mu_1 - \gamma_1 Q_t - \beta' X_{t-1}^Z), \\ &\Pr(Z_t = j) = F(\mu_j - \gamma_1 Q_t - \beta' X_{t-1}^Z) - F(\mu_{j-1} - \gamma_1 Q_t - \beta' X_{t-1}^Z), j = 2,3,4,5, \\ &\Pr(Z_t = 6) = 1 - F(\mu_5 - \gamma_1 Q_t - \beta' X_{t-1}^Z). \end{aligned} \tag{4}$$

By adopting an additional notation that $\mu_0 = -\infty$ and $\mu_6 = +\infty$, we can obtain a compact version as:

$$\Pr(Z_t = j) = F(\mu_j - \gamma_1 Q_t - \beta' X_{t-1}^Z) - F(\mu_{j-1} - \gamma_1 Q_i - \beta' X_{t-1}^Z) \tag{5}$$

for $j \in [1, 6]$. By this definition, the set of cumulative probabilities can be transformed into a linear function that depends on both Q_t and X_{t-1}^Z, and only the intercepts in this function differ across the six categories:

$$F^{-1}[\Pr(Z_t \le j)] = \mu_j - \gamma_1 Q_t - \beta' X_{t-1}^Z \tag{6}$$

Through the Maximum Likelihood estimation we can obtain a natural estimator for this model. Define: $Z_{tj} = \begin{cases} 1 & if\, Z_t = j \\ 0 & else \end{cases}$, then the log-likelihood for the price aggressiveness is given by:

$$\log L = \sum_{t=1}^{T} \sum_{j=1}^{6} Z_{tj} \log[F(\mu_j - \gamma_1 Q_t - \beta' X_{t-1}^Z) - F(\mu_{j-1} - \gamma_1 Q_t - \beta' X_{t-1}^Z)] \tag{7}$$

The "Likelihood" estimation refers to a probability of prediction, and reflects how likely it is that the observed value of dependent variables is predicted by some analytical function of the observed values of independent variables and parameters. This likelihood is maximized with respect to $(\gamma_1,\ \beta,\ \mu_1, ...,\ \mu_5)$, i.e. a vector of *M*+5 parameters, where *M* is the number of exogenous explanatory variables. The standard Ordered Probit model takes *F* to be Normal and assumes that $\varepsilon_t^Z \sim N(0, \sigma^2)$, therefore, by adopting the scale normalization $\sigma = 1$ and imposing a zero intercept for identification, the probabilities are given by:

$$\Pr(Z_t = j) = \Phi(\mu_j - \gamma_1 Q_t - \beta' X_{t-1}^Z) - \Phi(\mu_{j-1} - \gamma_1 Q_t - \beta' X_{t-1}^Z) \tag{8}$$

where Φ is the cumulative distribution function of a standard Normal and the log-likelihood by equation (7) with *F* is replaced by Φ.

As pointed out by Ranaldo (2004), employing the results obtained from the Ordered Probit model, one can estimate the cumulative probabilities that any of the six events will occur and also the probability that a specific order is likely to be submitted. However, the marginal effects of the regressors Q_t and X_{t-1}^Z on the outcomes' probabilities are not equal to the estimated coefficients $\hat{\gamma}_1$ or $\hat{\beta}$, because of the ambiguous effect on the middle partitions. Therefore, in order to understand in what direction and magnitude a change in the explanatory variables affects, the marginal effects for each outcome are calculated by differentiating the probabilities in equation (8):

$$\begin{aligned}
&\frac{\partial \Pr(Z_t = 1)}{\partial X_{t-1}^Z} = -\phi(\hat{\mu}_1 - \hat{\gamma}_1 E[Q_t] - \hat{\beta}' E[X_{t-1}^Z])\hat{\beta}, \\
&\frac{\partial \Pr(Z_t = j)}{\partial X_{t-1}^Z} = -[\phi(\hat{\mu}_j - \hat{\gamma}_1 E[Q_t] - \hat{\beta}' E[X_{t-1}^Z]) - \phi(\hat{\mu}_{j-i} - \hat{\gamma}_1 E[Q_t] - \hat{\beta}' E[X_{t-1}^Z])]\hat{\beta}, \quad j = 2,3,4,5, \\
&\frac{\partial \Pr(Z_t = 6)}{\partial X_{t-1}^Z} = \phi(\hat{\mu}_5 - \hat{\gamma}_1 E[Q_t] - \hat{\beta}' E[X_{t-1}^Z])\hat{\beta}.
\end{aligned} \tag{9}$$

where ϕ is the normal probability density function, $\hat{\mu}_j$ are the estimated thresholds and $\hat{\gamma}_1$ and $\hat{\beta}$ represent the estimated coefficients resulting from Eqs (1) and (2). The unconditional mean of instruments of quantity and each independent variable over the entire sample is used as the estimate of $E[Q_t]$ and $E[X_{t-1}^Z]$, respectively. Thus, the marginal effect, induced by an incremental variation in one of the order flow components, on the order choice probabilities can be calculated by using equation (9). For instance, we could estimate how the probability of submitting orders at each level of aggressiveness changes marginally when the spread increases by one tick, avoiding difficulties in interpreting the coefficients in the Ordered Probit model.

As criticized by Hall and Hautsch (2006) and Lo and Sapp (2007), the above Ordered Probit model suffers from several disadvantages: First, it is set up based on the assumption that all the observations are independent, however, orders for the same stock might share similarities, in which case the model will be misspecified and the usual standard errors will be incorrect. Second, the Ordered Probit model is set up in a static environment, however, ignoring the dynamics within the individual processes may also induce misspecification and bias. Third, the model treats the cancellation of limit order as the least aggressive type of order submission. Since this behavior happens only conditional upon an earlier submission, but not a choice available to all traders, and moreover, the aggressiveness of cancelled orders is ignored, treating cancellation same as other order submissions is controversial. Finally, the Ordered Probit model ignores the exact time when the order is submitted given the current state of market.

4.2 Joint decision on price aggressiveness and quantity

Explicitly allowing traders to simultaneously select their limit price and order size, we transform equation (2) to a set of simultaneous equations for each stock:

$$Z_t^* = \gamma_1 Q_t + \beta' X_{t-1}^Z + u_t^Z \tag{10a}$$

$$Q_t = a + \gamma_2 Z_t^* + b' X_{t-1}^Q + u_t^Q \tag{10b}$$

where Z_t^* is the latent price aggressiveness, Q_t is the endogenous quantity, X_{t-1}^Z and X_{t-1}^Q consist of a series of predetermined variables in price aggressiveness equation and quantity equation, respectively, while u_t^Z and u_t^Q are disturbance terms. For identification purpose, X_{t-1}^Z contains the variable of price change Δp, which is a predetermined variable that not in the quantity equation, i.e. $X_{t-1}^Z = [X_{t-1}^Q, \Delta p]$ (Note 2). As to the identification of simultaneous equations with Probit model, Nelson and Olson (1978) suggest that the number of independent restrictions on the coefficients of each equation should be not smaller than the number of structural equations minus one. In order to exactly identify each of the tow equations in our model, there must be one independent restriction in equation (10a) and (10b). Intercept equal to zero and $\Delta p = 0$ are the restrictions for each equation respectively, and thus both equations in the simultaneous system are just identified. These two simultaneous equations are estimated adopting a method analogous to the two-stage least squares method of Lo and Sapp (2007) (Note 3). In the first stage, we calculate the reduced-form of price aggressiveness and quantity equations:

$$\begin{aligned} Z_t^* &= \frac{\gamma_1 a}{1-\gamma_1\gamma_2} + \frac{\gamma_1}{1-\gamma_1\gamma_2} b' X_{t-1}^Q + \frac{1}{1-\gamma_1\gamma_2} \beta' X_{t-1}^Z + \frac{\gamma_1}{1-\gamma_1\gamma_2} u_t^Q + \frac{1}{1-\gamma_1\gamma_2} u_t^Z \\ Q_t &= \frac{a}{1-\gamma_1\gamma_2} + \frac{1}{1-\gamma_1\gamma_2} b' X_{t-1}^Q + \frac{\gamma_2}{1-\gamma_1\gamma_2} \beta' X_{t-1}^Z + \frac{1}{1-\gamma_1\gamma_2} u_t^Q + \frac{\gamma_2}{1-\gamma_1\gamma_2} u_t^Z \end{aligned} \tag{11}$$

Equation (11) can be written compactly as:

$$\begin{aligned} Z_t^* &= \pi_1' X_{t-1} + \upsilon_{1t} \\ Q_t &= \pi_2' X_{t-1} + \upsilon_{2t} \end{aligned} \tag{12}$$

where X_{t-1} is a vector of explanatory variables, i.e. $\begin{pmatrix} 1 & X_{t-1}^Q & X_{t-1}^Z \end{pmatrix}$. In order to estimate the parameters in the structural forms, we first estimate the coefficients in the reduced form of the simultaneous equations separately. Specifically, π_1 is estimated using an Ordered Probit model and π_2 is estimated using OLS method. With the estimated coefficients, we could form the instruments for price aggressiveness $\hat{Z}_t^* = \hat{\pi}_1' X_{t-1}$ and for quantity $\hat{Q}_t = \hat{\pi}_2' X_{t-1}$, which are at least asymptotically uncorrelated with disturbance terms. In the second stage, $\hat{Z}_t^*$ replaces its counterpart in right hand side of equation (10b). Treating the instrument as fixed regressor and the resulting equation as single equation model, we estimate the structural parameters in equation (10b) via OLS model. Similarly, $\hat{Q}_t$ is substituted for Q_t in equation (10a) and the price aggressiveness is estimated via an Ordered Probit model.

The order submission analyses for the buy and sell sides are performed separately, indicating that for any one stock, the entire time series of the order flow are broken into two sub-samples. Each of these sub-samples contains six order types submitted on one side of the book, their order sizes and dynamic explanatory variables (volatility and price change) at time *t*, which is calculated already based on lagged values, as well as the static explainable variable data (ask depth, bid depth, relative depth and actual spread) immediately before, i.e. at time *t-1*. Therefore, these explanatory variables are predetermined and exogenous.

4.3. Hausman specification test

In order to correctly build the econometric model, it is necessary to judge whether the regressor quantity is endogenous or exogenous. For the liquidity-oriented trading, quantity is predetermined and should be regarded as exogenous, however, for informed traders, they simultaneously decide order price and quantity on the available information, and thus both are endogenous. A disadvantage in recent study is that we cannot distinguish whether an order submission is liquidity- or information- motivated. As respect to the econometric respective, the simultaneity problem arises because some regressors are endogenous and are therefore likely to be correlated with the disturbance. The Hausman specification error test is a test of simultaneity, essentially a test of whether regressor is correlated with the error term. If the simultaneity problem exists, the usual least squares estimator or Probit estimator is inconsistent (Rivers and Vuong, 1984), instead, with the valid instrumental variables, the two-stage estimator is consistent and should be used. In the above setting, the exogeneity condition is stated in terms of the correlation coefficient ρ, which can be interpreted as the correlation between Q_t and u_t^Z for equation (10a) and (10b). When $\rho = 0$, Q_t and u_t^Z are uncorrelated, Q_t is exogenous for equation (10a) and we can directly estimate this equation. On the contrary, $\rho \neq 0$ implies that Q_t is correlated with u_t^Z, and therefore endogenous, in such case we have to use instrumental variables for the estimation. The Hausman test to test whether Q_t is exogenous involves two steps: First, estimate the reduced form of the quantity equation by OLS, as shown in the second line of equation (12), to obtain the residual $\hat{\upsilon}_{2t} = Q_t - \hat{\pi}_2{}' X_{t-1}$. Second, include the residual computed in step 1 as an explanatory variable in the regression:

$$Z_t^* = \gamma_1 \hat{Q}_t + \beta' X_{t-1}^Z + \delta \hat{\upsilon}_{2t} + u_t^Z \tag{13}$$

Estimate the artificial regression, as stated by equation (13), by ML and perform the z-test for the hypothesis of the coefficient δ:

H_0: $\delta = 0$ (Q_i is exogenous)

H_1: $\delta \neq 0$ (Q_i is endogenous)

If the z-test is significant, do not reject the hypothesis of simultaneity; otherwise, reject it.

5. Description of the market and data set

The data used in this paper is provided by the Australian Stock Exchange (ASX), which employs the fully computerized Stock Exchange Automated Trading System (SEATS), modeled on Toronto's Computer Assisted Trading System (CATS). The ASX data is particularly suitable for the purpose of the study, because: First, ASX market relies almost solely on liquidity provision by investors. There are no dealers or designated market makers on the ASX, although brokers may act as informal market makers and facilitate their client's trades by placing opposing orders on their own accounts (Note 4). In contrast, the New York Stock Exchange where the specialist plays an important role in the provision of liquidity and maintenance of market integrity, and the Paris Bourse where exists a number of markets makers assigned to its stocks, confront with a principal-agency problem and higher transaction costs that might influence the test results. Second, ASX has an almost instantaneously updated limit order book, containing detailed information about each order. Third, ASX strictly enforces price before time priority on limit order execution. Finally, it is considered to be a highly transparent market, and only orders with a total value exceeding $200,000 can be entered with a hidden volume. Therefore, on the ASX market, traders are likely to act strategically, taking into account the beliefs, reactions and order submission strategies of other traders.

5.1 Institutional details of the ASX trading system

Several modes of operations in SEATS may affect the information available to traders and their order submission behavior. The various phases in a typical day on the ASX are illustrated in Table 2. The pre-opening period takes place between 7:30am and 10:00am, during which orders can be entered into the system but they cannot be matched at this time. The ASX opens at 10:00am with a call auction procedure aimed principally at maximizing traded volume at the chosen opening price. After that, the normal trading takes place continuously until 4:00pm. Orders entered during normal trading hours are matched, resulting in trades for market orders, or stored in the order book automatically for limit orders. A 5-minute pre-closing period begins at 4:00pm and then is followed by the official single-price closing auction. The closing phase between 4:05pm and 5:00pm facilitates late trading, during which the system accept new orders but does not match them. Once the order has been matched, the selling broker is responsible to apply manual procedures, which execute orders in the same manner as though SEATS is fully active, reporting and subsequently canceling their order. Afterwards, the after hours adjust phase operates between 5:00pm and 7:00pm, aiming to adjust to take account of capitalization and other changes which come into effect on the following day. After a half-hour period of housekeeping to routinely maintain the system, from 7:30pm to 11:30 pm and again from 3:30am to 7:30am the system is in enquiry mode only providing information regarding the current orders for stock. None can log into the system between 11:30pm to 3:30am when SEATS is off-line.

The order book enforces strict price and time priority, and limit orders are stored in the buy and sell queues. During normal trading, orders can be modified by either changing the limit price or the number of shares or cancellation. While a size reduction does not influence the order's priority, a price change leads an order to be relocated in the queue at the new price level. The minimum tick rules stipulate the minimum distance in cents at which an order can be placed from another at a different price. According to the amendments to the minimum bid rule in 1996, which is outlined in Table 3, a minimum bid would be $0.001 for equity securities with the market price up to $0.10, and a minimum bid of $0.05 for those with market price from $0.10 to $0.50. Limit orders for at least $200,000 can include an undisclosed reserve volume which replenishes the visible order size upon partial execution. However, order book information is sufficiently transparent, and traders can view details of individual buy and sell orders along the book and aggregate depth at multiple prices through SEATS Trading Screen.

5.2 Descriptive statistics

The investigated sample consists of the order book data of five most liquid stocks on the ASX market, specifically, Broken Hill Proprietary Limited (BHP), National Australia Band (NAB), Telstra (TLS), Woolworths (WOW) and Mount Isa Mines (MIM), during the period 1 July to 30 August 2002. The ASX Intra-Day dataset provides historical details on order submission of the five stocks placed on SEATS as well as the resulting trades. Each order and trade record includes information on the price, volume and direction, time-stamped to the nearest one hundredth of a second. In order to focus on the information content of the limit order book and avoid confounding effects from the opening and closing procedures, we restrict our attention within the normal trading period from 10:09am to 4:00pm, and remove data from the opening and closing call auctions as well as all market crossings (Note 5) and off-market trades. The resulting samples consist of 133,059, 102,366, 94,089, 52,493 and 30,559 observations for BHP, NAB, TLS, WOW and MIM, respectively.

Table 4 shows summary statistics of limit order book, including the explanatory variables across the sample period. Depth at the best ask price is larger than the depth on the bid side, however, depth at off-best prices on the ask side seems smaller than that on the opposite side. This suggests that generally the bid side of the market is more liquid, but it is not so liquid to market order buyers. Given that the minimum tick size is 1 cent for all the observed stocks and the actual spread is one or two ticks, the stocks can be regarded quite liquid. These results confirm the work by Aitken et al. (1995), which provides evidence that bid-ask spreads are constant and flat during the day, but contrasts with data from US markets (Note 6), which suggests that spreads follow a U-shaped pattern. The reason may be that Australian market, unlike US market, has a fully competitive market-making system, without specialists or market-makers who may be exercising their monopoly power. Information models (such as Easley and O'Hara, 1987) predict that specialist may keep spread wider enough so that the profits from trading with liquidity traders sufficiently compensate for the losses from trading with informed traders. According to this theory, without these specialists, the intraday pattern of spread could be flat and keep narrow. The relative spread (Note 7) in our sample (0.271%) is lower compared with the results of Angel (1997), in which the relative bid-ask spread median is calculated to equal 0.65% for major market indices of fifteen countries. The mean of volatility for five stocks on the ASX market is 0.747, while the value on Swiss market is 0.482, as surveyed by Ranaldo (2004). Generally, "a market is liquid if traders can buy or sell large number of shares when they want and at low transaction costs. Liquidity is the willingness of some traders to take the opposite side of a trade that is initiated by someone else, at low cost" argues Harris (1990) when defining a liquid market. According to this definition, market liquidity can be measured from four dimensions: width, referring to the spread for a given number of shares; depth, measured by the number of shares that can be traded at given quotes; immediacy, referring to how quickly trades of a given size can be done at a given cost and resiliency, defined as how quickly prices revert to former levels after they change in response to large order flow imbalance initiated by uninformed traders. Because of its relatively narrow spread and high depth, the ASX market is regarded quite liquid. In addition, Vo (2007) proves that the price volatility is negatively related to market liquidity, which measured by both spread and depth. Intuitively, if price is high volatile, the probability that a limit order is executed is higher even though spread is wide. As a result, limit order submitters decrease limit price when they buy or increase limit price when they sell and thus lead to a narrow spread. The coexistence of a high volatility and narrow spread in our sample confirms this suggestion.

5.3 Order aggressiveness on the ASX market

Table 5 breaks down the price aggressiveness of order submissions into six categories in our sample, and shows the frequency of order submissions ranked by price aggressiveness: market orders or marketable limit orders (23.7%--40.8%), best limit orders improving existing best prices (0.1%--3.7%), best limit orders placed at the best prices (30.9%--43.1%), off-best limit orders within 3 ticks of the best prices (8.4%--16.7%), off-best limit orders more than 3 ticks of the best prices (4.8%--10.6%) and cancellations (3.3%--14.6%). These results indicate that the most frequent order types are market or marketable limit orders and limit orders placed at the best prices, consistent with the observations provided by Lo and Sapp (2007) as well as Beber and Caglio(2005). The high number of orders placed at the best quote implies a good average performance for these orders, as suggested by Harris and Hasbrouck (1996). The percentage of best limit orders which improve the present best prices in our sample is quite small, compared with the findings (14%) of Lo and Sapp (2007). One of the main attributes of this small proportion is that the most liquid securities have actual bid-ask spread equal to the tick size, thus giving less opportunities for traders to better the market. The proportion distributions are generally similar for both ask and bid sides of the market, however, for some stocks, there is a slightly higher aggressiveness for buy orders with respect to those for sell orders. This finding may occur because the market is perceived as a rising market, considering that market ask orders confront with a potential lost when stock prices rise, sellers would like to adopt less aggressive orders instead of a market order.

Table 6 shows the likelihood of order arrival conditional on the level of the last price aggressiveness on both ask and bid sides. The columns correspond to an event at time *t*, and the rows to events at time *t-1*. The diagonal effect in the conditional order arrival can be obviously observed through that most bold numbers, which represent the maximum proportion of order type based on each previous price aggressiveness category, locate along the diagonal of the frequency table. This result indicates an autocorrelation between order sequences, for instance, after a limit order at best price is submitted, the probability that the same limit order type will arrive is higher then other order types, consistent with Biais et al. (1995), who report that an incoming order is likely to be at the same aggressiveness level as the previous order.

Table 7 illustrates the total number of shares submitted and average order sizes according to every price aggressiveness category. A general observation on the order quantity is that order size increases gradually as the price aggressiveness decreases from category 1 to category 3, and the order size for the category of cancellations is largest, except for MIM. This phenomenon results of the joint decision of order price and quantity by investors, who have a negative trade-off between price aggressiveness and size. The possible loss of a price-aggressive order could be reduced by decreasing its quantity. Another interesting finding is that the order sizes are relatively small for limit orders off-best prices. Since high trading volume is considered to be an indication of the advent of informed traders. Uninformed traders are more inclined to submit these small orders, in order to lower their risks from trading with informed traders who usually expropriate their private information by placing aggressive orders with great volumes. For all the stocks, the total bid volume is higher then the ask volume, indicating that there is a higher demanding market during the period of analysis. As to a cross-section analysis, we find that the higher the stock price, the smaller the average quantity for each price aggressiveness classification. The negative relationship between stock price and order size reflects traders' budget constraint and the intention to minimize the potential loss if the stock price walks to the opposite side from their expectation.

6. Empirical findings

This section presents the results from the analysis of the price aggressiveness and the corresponding quantity for the order submission on the ASX market. According to Table 8, which documents the main results from Hausman specification test, all coefficients for error terms are significantly not equal to zero, therefore, we reject the null hypothesis that Q_t is exogenous. The Hausman test confirms that traders simultaneously decide their order price aggressiveness and quantity, and we should adopt a simultaneous system, as shown in equation (10a) and (10b) to model traders' behavior. Table 9 reports coefficient estimates for the sample on demand and supply side events separately. The price aggressiveness equation and quantity equation are estimated using a method analogous to the two-stage OLS approach as introduced in section 4. When interpreting the results, a positive coefficient in the price aggressiveness equation means that the factor increases the price aggressiveness of the order submitted, since the ascending ranking from 1 to 6 reflects increasing price aggressiveness.

6.1 Trade-off between price aggressiveness and quantity

Table 9 shows that the estimated coefficients of quantity on price aggressiveness are negative but not significant on the ask side, whereas the quantity has a slight positive influence on price aggressiveness on the bid side. For the quantity equation, price aggressiveness is found to be significantly positively associated with quantity on both ask and bid sides. These results are not totally consistent with the findings of Lo and Sapp (2007), which document a negative relationship between price aggressiveness and quantity after controlling for the state of the limit order book for both two equations, although their estimated coefficient on quantity in the price aggressiveness equation is also not statistically significant. Our results are also not consistent with the findings of Bessembinder et al. (2006), which indicate a slightly positive effect of order size in the price aggressiveness equation and a negative effect of price aggressiveness in the quantity equation, using two-stage Least Squares estimation that allows for endogeneity. The marginal probability analysis in Table 10 illustrates that, buyers placing orders at more aggressive prices are more likely to submit smaller orders, suggesting that a submission strategy that combines aggressive price with small order size. However, sellers for stocks except NAB behave in an opposite pattern.

In the price aggressiveness equation, quantity is found to be negatively related with price aggressiveness for sellers. This result is quite reasonable if we assume that sellers are dominated by liquidity-motivated traders, who have to first meet their trading target. Under this assumption, the negative correlation between price aggressiveness and quantity implies that larger liquidity traders are propone to place less price-aggressive orders. This result confirms the prediction of Harris (1998) about the behavior of large liquidity traders, that traders needing to meet a target by a certain deadline would start a greater propensity for using limit orders. One might conjecture that large traders would submit aggressive orders to transact early to avoid any possibility of incurring trading penalties. However, the empirical finding suggests that liquidity traders are more concerned with the execution quality of their orders, a result in line with the behavior of traders in a simulated market of Bloomfield et al. (2005). For buyers, if we suppose that they are dominated with informed traders, we can interpret that investors owning private information are incline to use price aggressive orders to guarantee an immediate execution, as well as higher order quantity to increase their profit. Besides, after a transaction, if the transacted price does not reflect a stock's true value and future stock price moves against the trader, the trading generate a real loss for sellers, but only a potential loss for buyers, because stock price may latterly move in favor of them. Therefore, buyers tend to behave more aggressive than sellers, and their behavior may attribute to the positive relationship between price aggressiveness and quantity.

On the other hand, the estimated coefficients on the price aggressiveness in the quantity equation are significantly positive, except for the stock of MIM on the bid side, which is significantly negative. An optimal order size is obtained through investors' consideration on the balance between execution risk and potential profit. By submitting a small order, trader face low execution risk, since his order is more possible to be executed totally. However, a small order directly reduces the trader's profit, especially when he has valuable information on the movement of stock price. Easley and O'Hara (1992) provide a possible reason why more aggressive orders are smaller, i.e. order splitting, and their assumption is latterly supported by Lo and Sapp (2007). This explanation asserts that investors slip their orders to either hide their private information or to reduce the execution costs. Nevertheless, if we consider that this market is mainly consist of informed traders, whose strategies most likely stem from the changing value of their private information associated with the dynamic adjustments of prices, we could understand why there exists a positive relationship between price aggressiveness and quantity. When the market prices differ from the true value, informed traders have an incentive to take liquidity by trading via market orders with large quantity, as long as the investment value is not beyond their budget constrains. After the market price adjustment and the true value's reveal, the informed traders change their strategy. In an asymmetric information setting, the informed traders ultimately have the advantage because while all traders placing limit orders face execution risk, they do not have the potential risk of adverse selection , which confronting the liquidity traders. Thus, informed traders can price the limit orders more aggressively and be at the inside quote to make profit. Without uncertainty and trade-off between different risks but only considering the stock's market price dynamics, informed traders have more possible combinations of decisions on order price and size, and this is perhaps the reason why we observe different correlations between price aggressiveness and quantity among stocks.

6.2 Limit order book-related variables

Depth

The coefficients of depths at the best quote on the same side for sellers of all stocks except NAB and for buyers of stocks TLS, WOW and MIM in the price aggressiveness equation are significantly positive, implying that an increase in the depth at the best prevailing price on the same side of the market encourages aggressive orders. This is consistent with the empirical findings of Ranaldo (2004), beber and Caglio (2005) and the crowding-out hypothesis of Parlour (1998), and therefore basically confirms our hypothesis 1. The crowing-out effect forces a trader to compete for execution, since high market depth on the same submission side indicates a long queue for execution. The increased competition provides newly arriving traders with incentives to place more aggressive orders in order to jump the queue. This effect is similar on both sides of the market. Table 9 shows that the marginal probabilities of market orders and aggressive limit orders being submitted drop, meanwhile, the marginal probabilities of less aggressive limit orders and cancellations increase, when the same side depth rises, for stocks TLS, WOW and MIM. Moreover, depths on the same side are significantly positively correlated with submitted order size for stocks NAB and MIM.

Considering changes in the depth on the opposite side of the market, the results show that the opposite market depth is inversely associated with price aggressiveness for the stocks BHP, TLS and MIM, consistent with Ranaldo (2004) and Verhoeven et al. (2004). A large opposite market depth indicates a higher chance of execution, which shortens the order queue, and thus traders need not to adopt a price-aggressive strategy. However, for the stocks NAB and WOW, the results are not in accordance with theoretical prediction, where orders become more aggressive when the market liquidity supply on the other side of the book is greater. For the quantity equation, the results are more ambiguous. Generally the findings do not provide enough evidence to support hypothesis 2. Table 9 reveals that for changes in the opposite depth, bid and ask sides of the market react symmetrically in terms of price aggressiveness, but not in the case of stock BHP. The asymmetry in buyers and sellers trading attitude may be due to the impatience of sellers compared with buyers. A slow execution when the market moves down leaves the sellers a realized loss, while a slowly execution when price goes up represents only an opportunity cost for buyers. Besides, as sellers are considered more likely to be liquidity-motivated than buyers, a higher market depth on the bid side implies a lower marginal cost of submitting market orders, especially when orders can walk up the order book in the markets. Thus, when opposite depth increases, sellers' marginal probability of placing aggressive orders rises at a greater degree than buyers'.

As *reldep* is defined as shares at the best five ask prices divided by total shares at the best five bid and ask prices, the significantly negative coefficients on the ask side except for BHP implies that the larger the depth at the ask side relative to the bid side, the lower the probability of submitting a more aggressive ask order, contradicting hypothesis 3. On the bid side of the market, price aggressiveness is significantly positively associated with the relative depth for stocks NAB and TLS. A possible explanation provided by Verhoeven et al. (2004) is that a large depth at the off-best ask price indicates to the buyer that the true price is likely to be at this off-best ask, which promotes the buyer to submit a market order to buy at the best ask price to make profits, because the stock is perceived to be under-priced. Due to the same logic, once the sellers find depth further from the market is high, they are less likely to fill in the schedule in front of those off-best prices. Table 10 indicates that the marginal probability to place more aggressive sell orders increases and the marginal probability to submit less aggressive sell orders drops, with an increment in relative depth, for stocks NAB, TLS, WOW and MIM. On the bid side, when relative depth rises, the marginal probability to observe a more aggressive order decreases, and that to observe a less aggressive order increases for stocks NAB, TLS and WOW. The quantity of orders reacts quite differently to a change in relative depth for a cross-section analysis. On the ask side, order sizes of stocks NAB, TLS and WOW rise significantly as the relative depth increases, whereas for BHP and MIM, the submitted quantity drops. On the bid side, relative depth significantly negatively affects order sides of stocks NAB, TLS and MIM, but is found to be positively correlated with quantities for BHP and WOW.

Bid-ask spread

The bid-ask spread is significantly negatively associated with the degree of price aggressiveness for both buyers and sellers, confirming hypothesis 4 and in accordance with Biais et al. (1995), Hollifield et al. (2004) and Ranaldo (2004). However, this result is not in accordance with Lo and Sapp (2007), who find that orders become more aggressive and larger in size when the spread widens. Since a large spread is costly for the market order, and it indicates high asymmetric information, traders may consider it optimal to submit less aggressive orders. On the contrary, when the spread decreases, i.e. the cost of trading by market orders or the price of liquidity reduces, traders reasonably prefer to demand liquidity and place aggressive orders. An analysis of the marginal effects in Table 10 demonstrates that an increase in the bid-ask spread increases the probability of order submission in some aggressive order categories, whereas the likelihood of observing less aggressive orders and cancellations reduces. This is contrary to the intuition that market orders become more expensive as the spread widens. La and Sapp (2007) propose that this phenomenon suggest that investors become more concerned about ensuring their orders are executed in an uncertain market. Furthermore, we could also notice that a change in the bid-ask spread affects buy price aggressiveness slightly more than sell price aggressiveness, similarly with the results of Beber and Caglio (2005). As to the quantity equation, spread is found to be positively related with order size for stocks NAB, TLS and WOW, in line with Lo and Sapp (2007). The positive relationship is probably because that, in an order-driven market, limit order traders are more inclined to supply additional liquidity if the potential profitability increases as the spread widens.

6.3 Market movement-related variables

Volatility

As illustrated in Table 9, an increase in volatility drives more passive orders for buy side but has ambiguous influence on ask side, partially confirming hypothesis 5. A negative relationship between volatility and price aggressiveness is predicated by the model of Foucault (1999) and is also documented in Ranaldo (2004) and Beber and Caglio (2005). Because volatility is considered as a proxy of information-driven trading: when volatility rises, the trader concerns that his order may be picked off by a better-informed trader on the other side of the market during times of greater uncertainty. Therefore, the expected cost of market order trading increases and the submission of less aggressive orders turns out to be the optimal strategy. On the quantity dimension, order sizes of both bid and ask orders increase statistically significantly as volatility rises. This result signifies that when market condition is turbulent, the information is more valuable and it is easier to pick off other traders, and thus informed traders try to expropriate their information by submitting large orders, and earn more profits from trading with uninformed traders.

Price changes

We find that an increment in the proportion of positive price changes makes sellers to submit orders at less aggressive prices, however, the effect on bid orders are still not perceptible. This result partially supports hypothesis 6. Earlier empirical literature on the effect of price also shows different evidence. For instance, Beber and Caglio (2005) find that when the stock price moves up, buy orders become more aggressive and sell orders less aggressive. The reason is, as they point out, that traders perceive that a positive price trend reduces the probability of execution of passive buy orders, but increases the probability of execution of sell limit orders. Nevertheless, Lo and Sapp observe a completely different result, that an increase in the proportion of positive price changes makes ask (bid) orders more (less) aggressive. The analysis of marginal probabilities in Table 10 shows that price change is one of the most important variables in affecting price aggressiveness. For all stocks, we find that with the increasing of the proportion of positive price change, it is more likely to observe more aggressive asks.

6.4 Time of the day

The signs of coefficients of hourly dummies for most stocks in the price aggressiveness equation of Table 9 turn from negative to positive gradually on both ask and bid sides during normal trading period, suggesting that traders submit less aggressive orders after openings and more aggressive orders when the trading end approaches. This result confirms our hypothesis 7 and is consistent with Beber and Caglio (2005), who document a monotonic increment for the time-of-the-day dummy variables and argue that the proportion of market orders increases at day’s end. The finding proves the existence of “deadline effect”, as suggested by Bloomfield et al. (2005), where liquidity traders who need to buy or sell a large number of shares tend to use more limit orders early on, but they switch to market orders to meet their targets as the end of the trading period approaches. This deadline effect is observed widely in earlier literature, for instance, Roth et al. (1988) conduct experimental test and find many deals occur just before deadlines, indicating that traders become more aggressive as the close of trading approaches. Similar effect is also posited by Harris (1998) and Hollifield et al. (2003). According to the quantity equation, the result indicates that larger orders are more likely being submitted after opening and just before the end, and shows a U-shape pattern in order submission activity over the trading day. The cluster of order submission early in the day confirms a 24-hour pattern in information arrival and disclosure. Greene and Watts (1996) find that two-thirds of sample earnings announcements take place after the stock exchange close. Because it is possible for the traders to accumulate information overnight, larger orders could be placed early in the day. On the other hand, many traders face an end-of-day deadline for either liquidity requirement or exploiting their information, and therefore larger orders are also submitted at the end of the day.

Table 11 summarizes the main findings for price aggressiveness equation and shows that most of our hypotheses are supported by empirical results, as indicated by the matching of the expected and empirical signs of the coefficients in the regressions. Therefore, we can conclude that limit order traders do make their decisions about order submission, with respect to the information on the limit order book as well as market price movement. Furthermore, traders' submission behavior follows some specific time-of-the-day pattern. However, some results such as relative depth do not support our hypothesis, and the reason may lie on that traders believe that limit orders outside the best bid and ask are not so informative and simply ignore them. Other explanations include that monitoring and analyzing the book is costly or traders are overconfident and give little weight to order book information and other traders' strategy, relatively to their own assessment of the stock's true value, as stated by Cao et al. (2004). Besides, there exists a difference of the effects of explanatory variables on a cross-sectional level to some degree, indicating that some economic relations may hold individually.

7. Conclusion

The limit order book promptly updates the accumulation of limit orders after market orders or marketable limit orders matching, and therefore reflects order submission and trading information, liquidity change, as well as traders' expectation. Using data from five most traded stocks on the ASX market during July and August 2002, this study analyzes how the information from limit order book and stock market determines the trader's join choice on price aggressiveness and quantity. The empirical results suggest that quantity is negatively traded off with price aggressiveness, because traders have to balance execution risk, transaction costs and potential loss. Besides, the depths at the best quotes are informative: order placement becomes aggressive during the period of thick market depths on the same side and thin market depths on the opposite side, consistent to the crowding out mechanism. Off-best depths have ambiguous effect on order submission, due to traders' overconfidence or the inconveniency to observe. Additionally, bid-ask spread, a proxy for the cost to place a market order, is found to significantly negatively related with price aggressiveness and positively related with quantity. Furthermore, high volatility encourages less aggressive limit order placement and discourages market order submission. Price change is also proved as am important factor to affect the price aggressiveness decision. Finally, the order sequences exhibit the diagonal effect, in which the probability that the same order type is more likely to occur after this order type has just occurred than it would be unconditionally. And order flows distribute following a U-shape intraday pattern; the intensity of order submission happens just after opening and as the trading end approaches.

This empirical study offers several practical suggestions for market regulators and participants: First, public release of complete and in time information on the evolution of the limit order book is essential for traders to provide liquidity in an order-driven market. Thus, guaranteeing the limit order book market enough transparency could make it more liquid. Second, because the autocorrelation of price aggressiveness on the ASX market reflects that informed traders strategically slip their orders to hide private information, it seems impossible for uninformed traders to identify informed traders and to imitate their submission strategy. Third, the findings support the practical feasibility of sophisticated order submission strategy, that trades off price aggressiveness and quantity, since a large aggressive order reveal more information and generate high transaction costs and potential loss. The specific order submission strategies and their profitability are not investigated directly in this paper, and we could leave this topic in future research.

References

Admati, A. and Pfleiderer, P. (1988). A theory of intraday patterns: volume and price variability, *Review of Financial Studies* 1, 3-40.

Aitken, M.J., Brown, P. and Walter, T. (1995). The intraday behavior of the Bid/ask Schedule under an automated trading regime, *Working Paper*.

Aitken, M., Frino, A., Jarnecic, E., McCorry, M., Segara, R. and Winn, R. (1997). The microstructure of the Australian stock exchange: An introduction, SIRCA.

Al-Suhaibani, M. and Kryzanowski, L. (2000). An exploratory analysis of the order book, and order flow and execution on the Saudi stock market, *Journal of Banking and Finance* 24, 1323-1357.

Angel, J. (1994). Limit order versus Market orders, *Working Paper*, School of Business Administration, Georgetown

University.

Angel, J. (1997). Tick size, share prices, and stock splits, *Journal of Finance* 52, 655-681.

Bae, K., Jand, H. and Park, K. (2003). Traders' choice between limit and market orders: Evidence from NYSE stocks, *Journal of Financial Markets* 6, 517-538.

Beber, A. and Caglio, C. (2005). Order submission strategies and information: Empirical evidence from NYSE, *FAME Research Paper*, No.146.

Bessembinder, H., Panayides, M. and Venkataraman, K. (2006). In search of liquidity: An analysis of order submission strategies in automated markets, *NBER Working Paper*.

Bloomfield, R., O'Hara, M. and Saar, G. (2005). The "make or take" decision in an electronic market: Evidence on the evolution of liquidity, *Journal of Financial Economics* 75, 165-199.

Biais, B., Hillion, P. and Spatt, C. (1995). An empirical analysis of the limit order book and the order flow in the Paris Bourse, *Journal of Finance* 50, 1655-1689.

Brock, W. and Kleidon, A. (1992). Periodic market closure and trading volume: a model of intraday bids and asks, *Journal of Financial Dynamics and Control* 16, 451-489.

Chan, K. C., Chung, Y. P. and Johnson H. (1995). The intraday behavior of bid- ask spreads for NYSE stocks and CBOE options, *Journal of Financial and Quantitative Analysis* 30, 329 – 346.

Cao, C., Hansch, O. and Wang, X. (2004). The information content of an open limit order book, *EFA 2004 Maastricht Meetings Paper* No. 4311.

Chung, K. and Van Ness, R. (1999). Limit orders and the bid-ask spread, *Journal of Financial Economics* 53, 255-287.

Cohen, K. J., Maier, S. F., Schwartz, R. A. and Whitcomb, D. K. (1981). Transaction costs, order placement strategy, and existence of the bid-ask spread, *Journal of Political Economics* 89(2), 287-305.

Copeland, T. and Galai, D. (1983). Information effects on the bid-ask spread, *Journal of Finance* 38, 1457-1469.

Easley, D. and O'Hara, M. (1987). Price, trade size and information in securities markets, *Journal of Financial Economics* 19, 69-90.

Easley, D. and O'Hara, M. (1992). Time and process of security price adjustment, *Journal of Finance* 47, 577-605.

Foucault, T. (1999). Order flow composition and trading costs in a dynamic limit order market, *Journal of Financial Markets* 2, 99-134.

Foucault, T., Kadan, O. and Kandel, E. (2005). Limit Order Book as a Market for Liquidity, *Review of Financial Studies* 18 (4), 1171-1217.

Glosten, L. (1994). Is the electronic limit order book inevitable?, *Journal of Finance* 49, 1127-1161.

Grammig, J., Heinen, A. and Rengifo, E. (2003). Trading activity and liquidity supply in a pure limit order book market: An empirical analysis using a multivariate count data model, *CORE Discussion Paper* No.2004/58.

Greene, J. and Watts, S. (1996). Price discovery on the NYSE and the Nasdaq: The case of overnight and daytime news releases, *Financial Management* 25 (1), 19-42.

Griffiths, M., Smith, B., Turnbull, A. and White, R. (2000). The costs and determinants of order aggressiveness, *Journal of Financial Economics* 56, 65-88.

Hall, A. D. and Hautsch, N. (2006). Order aggressive and order book dynamics, *Empirical Economics* 30, 973-1005.

Handa, P. and Schwartz, R. A. (1996). Limit order trading, *Journal of Finance* 51, 1835-1862.

Handa, P., Schwartz, R. A. and Tiwari, A. (2000). Quote setting and price information in an order driven market, *Working Paper*, Iowa University.

Harris, F. H., MacInish, T. H. and Chakravarty, R. (1995). Bids and asks in disequilibrium market microstructure: The case of IBM, *Journal of Banking & Finance* 19 (2), 323-345/

Harris, L. (1994). Minimum price variations, discrete bid-ask spreads and quotation sizes, *Review of Financial Studies* 7, 149-178.

Harris, L. (1998). Optimal dynamic order submission strategies in some stylized trading problems, *Financial Markets, Institutions, and Instruments* 7, 26-74.

Harris, L. and Hasbrouck, J. (1996). Market vs. limit orders: the SuperDOT evidence on order submission strategy, *Journal of Financial and Quantitative Analysis* 31, 213-231.

Hasbrouck, J. (2007) Empirical market microstructure, *Oxford University Press*.

Hasbrouck, K. and Saar, G. (2004). Limit orders and volatility in a hybrid market: The island ECN, *Working Paper*, NYU Stern School of Business.

Hausman, J., Lo, A. W. and MacKinlay, C. (1992). An ordered probit analysis of transaction stock prices, *Journal of Financial Economics* 31, 319-379.

Hollifield, B., Miller, R. and Sandås, P. (2003). Empirical analysis of limit order markets, *Working Paper*, Carnegie Mellon University.

Kaniel, R. and Liu, H. (1998). Are transactions and market orders more important than limit orders in the quote updating process?, *Working Paper* No. 16-98, The Wharton School, University of Pennsylvania.

Lo, A. W., MacKinlay, A. C. and Zhang, J. (2002). Econometric models of limit order executions, *Journal of Financial Economics* 65, 31-71.

Lo, I. and Sapp, S. G. (2007). Order aggressiveness and quantity: How are they determined in a limit order market?, *Working Paper*, Bank of Canada 2007-23.

Nelson, F. and Olson, L. (1978). Specification and estimation of a simultaneous-equation model with limited dependent variables, *International Economic Review* 19, No.3 695-709.

Parlour, C. (1998). Price dynamics in limit order markets, *Review of Financial Studies* 11, 789-816.

Ranaldo, A. (2004). Order aggressiveness in limit order book markets, *Journal of Financial Markets* 7, 53-74.

Rock, K. (1996). The specialist's order book and price anomalies, *Working Paper*, Harvard University.

Roth, A., Murnigham, K. and Schoumaker, F. (1988). The deadline effect in bargaining: some experimental evidence, *American Economic Review* 78, 806-823.

Verhoeven, P., Ching, S. and Ng, H. G. (2004). Determinants of the decision to submit market or limit orders on the ASX, *Pacific-Basin Finance Journal* 12, 1-18.

Vo, M. T. (2007). Limit orders and the intraday behavior of market liquidity: Evidence from the Toronto stock exchange, *Global Finance Journal* 17, 379-396.

Wermer, R. (1999). Mutual fund herding and the impact on stock prices, *Journal of Finance* 54, 580-622.

Table 1. Definitions of explanatory variables, hypotheses to test and related empirical studies

Abbreviation	Explanatory variables	Measurements
askdep	Depth at the best ask	Number of shares at the best ask divided by 10,000
biddep	Depth at the best bid	Number of shares at the best bid divided by 10,000
reldep	Relative depth	Total number of shares at the best five ask prices divided by total number of shares at the best five bid and ask prices
spread	Bid-ask spread	Difference between the lowest ask and the highest bid quotes
volat	Midquote volatility	Standard deviation of the last 20 midquotes
Δp	Proportion of positive price change	Number of positive price changes within the past two minutes divided by the total number of quotes submitted within the past two minutes
hdumm	Hourly dummy	Hourly dummy variables from 10:00 to 16:00

Hypotheses	Predictions	Related Literature
Hypothesis 1	*biddep* (*askdep*) positively related to order aggressiveness of buyers (sellers)	Parlour (1998) and Ranaldo (2004)
Hypothesis 2	*askdep* (*biddep*) negatively related to order aggressiveness of buyers (sellers)	Parlour (1998) and Ranaldo (2004)
Hypothesis 3	*reldep* positively (negatively) related to order aggressiveness of sellers (buyers)	Verhoeven et al. (2004)
Hypothesis 4	*spread* negatively related to order aggressiveness	Foucault (1999) and Ranaldo (2004)
Hypothesis 5	*volat* negatively related to order aggressiveness	Foucault (1999) and Ranaldo (2004)
Hypothesis 6	Δp positively (negatively) related to order aggressiveness of buyers (sellers)	Verhoeven et al. (2004) and Lo and Sapp (2007)
Hypothesis 7	Less aggressive orders during earlier hours but more aggressive orders during ending	Hollifield et al. (2004) and Bloomfield et al. (2005)

Table 2. Trading phases on SEATS

Time*	SEATS Market Phases
3:30am-11:30pm	Enquiry Only (Saturday)
3:30am-7:30am	Enquiry Only (Monday to Friday)
7:30am-10:00am	Pre-Opening
10:00am-10:09am	Opening
10:00am-4:00pm	Normal Trading
4:00pm	End of Trading
4:00pm-4:04pm	Pre-Close
4:05pm-5:00pm	Closing
5:00pm-7:00pm	After Hours Adjust
Shortly after 7:00pm	Shut Down
7:00pm-7:30pm	End of Day Processing
7:30pm-11:30pm	Enquiry Only
11:30pm-3:30am	System off-line

*all times in Sydney time

Source: Aitken et al. (1997) The microstructure of the Australian stock exchange: An introduction. Chapter 2, Exhibit 3.

Table 3. Minimum ticks for various stock price interval from 8 December, 1996

Market Price	Minimum Bid
Up to \$0.10	\$0.001
Over \$0.10 to \$0.50	\$0.005
Over \$0.50 to \$998.99	\$0.01
\$999.00 or greater	\$1.00

Source: Aitken et al. (1997) The microstructure of the Australian stock exchange: An introduction. Chapter 4, Exhibit 6.

Table 4. Descriptive statistics for explanatory variables

Stock	BHP		NAB		TLS		WOW		MIM		Mean
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	
Bid depth	9.62	9.20	0.82	0.78	47.78	47.30	1.07	1.08	84.32	95.47	29.74
	(22.05)	(21.74)	(2.06)	(1.94)	(36.77)	(37.48)	(1.30)	(1.29)	(75.29)	(71.39)	
Ask depth	11.74	11.28	1.20	1.09	47.22	47.71	1.06	1.09	81.79	93.60	29.78
	(31.00)	(28.95)	(4.30)	(3.90)	(36.79)	(36.84)	(1.15)	(1.21)	(70.53)	(84.65)	
Relative depth	0.493	0.475	0.516	0.484	0.502	0.488	0.505	0.466	0.534	0.518	0.498
	(0.153)	(0.154)	(0.225)	(0.227)	(0.139)	(0.139)	(0.206)	(0.205)	(0.126)	(0.123)	
Midquote	950.3	943.8	3382.6	3365.2	481.9	481.4	1220.4	1217.6	121.5	118.8	1228.4
	(51.74)	(52.13)	(97.42)	(105.8)	(6.08)	(6.13)	(37.44)	(38.48)	(10.18)	(10.05)	
Actual spread	1.068	1.063	2.176	2.199	1.006	1.006	1.502	1.512	1.005	1.005	1.354
	(0.277)	(0.272)	(1.504)	(1.545)	(0.080)	(0.078)	(0.948)	(0.953)	(0.069)	(0.073)	
Relative spread	0.113	0.113	0.064	0.065	0.209	0.209	0.123	0.124	0.833	0.852	0.271
	(0.030)	(0.030)	(0.045)	(0.046)	(0.017)	(0.016)	(0.078)	(0.078)	(0.087)	(0.090)	
Volatility	0.604	0.497	1.711	1.650	0.245	0.238	1.122	1.017	0.211	0.170	0.747
	(0.971)	(0.845)	(1.689)	(1.635)	(0.299)	(0.289)	(1.116)	(1.041)	(0.297)	(0.272)	
Price change	0.334	0.346	0.348	0.385	0.342	0.334	0.336	0.373	0.338	0.320	0.346
	(0.121)	(0.115)	(0.124)	(0.121)	(0.118)	(0.116)	(0.124)	(0.120)	(0.121)	(0.114)	

This table reports the sample statistics of the explanatory variables for each stock defined in section 3 over the sample period, July and August 2002. The mean and standard deviation values are presented. *Bid (ask) depth* is the number of shares available at the highest (lowest) bid (ask) quote (in 10,000). *Relative depth* is the total number of shares at the best five ask prices divided by total number of shares at the best five bid and ask prices. *Midquote* is the mid-price in Australian cents. The *actual spread* is the difference between the prevailing best ask and bid quotes in Australian cents. The *relative spread* is the actual spread divided by the midquote times 100. *Volatility* is the standard deviation of the last 20 midquotes. *Price change* is the percentage of the positive price changes during the last 10 order submissions.

Table 5. The frequency of price aggressiveness

Price aggressiveness	BHP		NAB		TLS		WOW		MIM	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
1.Market/marketable limit order	15950	28884	11695	13285	18329	19933	6747	8852	3918	5462
	(0.288)	(0.372)	(0.237)	(0.250)	(0.408)	(0.405)	(0.282)	(0.310)	(0.299)	(0.313)
2.Limit order with improved price	555	515	1841	1752	71	60	539	618	9	10
	(0.010)	(0.007)	(0.037)	(0.033)	(0.002)	(0.001)	(0.023)	(0.022)	(0.001)	(0.001)
3.Limit order at best price	21836	25263	21233	22509	14431	15199	10268	10943	4599	6210
	(0.394)	(0.326)	(0.431)	(0.424)	(0.321)	(0.309)	(0.430)	(0.383)	(0.351)	(0.356)
4.Limit order within 3 ticks from best price	6510	9196	4282	4843	6032	5587	1999	2890	2194	2383
	(0.117)	(0.119)	(0.087)	(0.091)	(0.134)	(0.114)	(0.084)	(0.101)	(0.167)	(0.137)
5.Limit order more 3 ticks from best price	5208	7649	5228	5332	4564	3715	2401	2823	921	839
	(0.094)	(0.099)	(0.106)	(0.100)	(0.102)	(0.076)	(0.100)	(0.099)	(0.070)	(0.048)
6.Cancellation	5409	6084	4974	5392	1487	4681	1948	2465	1475	2539
	(0.098)	(0.078)	(0.101)	(0.102)	(0.033)	(0.095)	(0.081)	(0.086)	(0.112)	(0.146)
Total	55468	77591	49253	53113	44914	49175	23902	28591	13116	17443

This table shows the price aggressiveness frequency of the stocks BHP, NAB, TLS, WOW and MIM traded on the ASX market during July and August 2002, including the number of orders in the individual categories as well as their corresponding percentages with respect to the sub-sample of sell and buy markets for each individual stock. The price aggressiveness categories are defined in section 4.

Table 6. Conditional probability of price aggressiveness

BHP	Ask						Bid					
t-1 \ t	1.	2.	3.	4.	5.	6.	1.	2.	3.	4.	5.	6.
1.	**36.68**	30.27	27.98	20.43	25.04	21.96	**44.18**	35.73	34.33	29.57	35.18	30.51
2.	0.82	**4.68**	1.32	0.55	0.83	0.59	0.37	**2.52**	1.06	0.52	0.61	0.49
3.	34.87	35.32	**43.57**	41.27	33.83	39.12	27.60	31.46	**37.85**	33.75	30.42	35.12
4.	10.28	7.39	10.98	**15.61**	12.90	13.75	10.88	11.46	11.02	**16.18**	12.01	13.17
5.	8.65	13.33	7.80	10.65	**16.61**	9.10	10.07	11.46	8.69	10.19	**13.35**	8.65
6.	8.71	9.01	8.35	11.49	10.79	**15.47**	6.89	7.38	7.04	9.79	8.42	**12.06**

Table 6. Conditional probability of price aggressiveness (continued)

NAB	Ask						Bid					
t-1 \ t	1.	2.	3.	4.	5.	6.	1.	2.	3.	4.	5.	6.
1.	**30.67**	21.84	23.10	19.36	20.73	17.85	**32.91**	23.12	23.35	20.51	22.36	19.79
2.	2.90	**9.89**	4.44	2.34	2.54	2.92	2.61	**8.56**	3.87	2.07	2.51	2.76
3.	38.88	40.25	45.97	**46.64**	39.42	42.74	36.53	40.53	45.23	**45.99**	39.95	44.64
4.	8.83	8.53	8.18	**12.75**	8.24	7.60	8.78	8.73	8.72	**12.70**	8.74	8.88
5.	9.71	9.34	9.33	9.69	**17.08**	12.67	10.33	9.36	8.86	9.56	**16.86**	8.12
6.	9.00	10.16	8.98	9.22	11.97	**16.22**	8.83	9.70	9.96	9.17	9.58	**15.80**

Table 6. Conditional probability of price aggressiveness (continued)

TLS	Ask						Bid					
t-1 \ t	1.	2.	3.	4.	5.	6.	1.	2.	3.	4.	5.	6.
1.	**47.45**	42.25	36.64	34.04	37.69	36.38	**47.54**	46.67	35.75	33.55	39.27	35.51
2.	0.11	0.00	0.25	0.12	0.07	**0.27**	0.04	0.00	**0.27**	0.14	0.03	0.06
3.	27.32	26.76	**38.10**	34.38	30.06	30.93	25.94	30.00	**37.16**	33.19	29.64	30.04
4.	12.29	16.90	12.94	**18.30**	14.22	9.89	10.07	8.33	11.05	**16.47**	12.62	10.83
5.	9.83	11.27	8.93	10.69	**15.21**	8.54	7.90	8.33	6.79	7.55	**9.93**	6.67
6.	3.00	2.82	3.13	2.47	2.76	**13.99**	8.51	6.67	8.99	9.09	8.51	**16.90**

Table 6. Conditional probability of price aggressiveness (continued)

WOW	Ask						Bid					
t-1 \ t	1.	2.	3.	4.	5.	6.	1.	2.	3.	4.	5.	6.
1.	**35.78**	25.60	26.16	22.11	25.07	23.82	**38.26**	27.51	27.78	27.27	30.07	25.11
2.	1.51	**8.72**	2.66	1.85	1.50	2.26	1.21	**7.61**	2.70	1.66	1.88	2.72
3.	38.34	39.89	**46.03**	45.92	40.61	43.48	33.00	39.00	42.45	39.13	33.23	**43.25**
4.	7.75	8.16	7.94	**12.71**	8.95	7.60	9.55	11.49	9.38	**14.26**	11.16	8.97
5.	9.41	9.65	9.34	10.26	**15.58**	9.03	10.38	7.77	8.56	9.65	**15.87**	7.79
6.	7.20	7.98	7.87	7.15	8.29	**13.81**	7.60	6.63	9.13	8.03	7.79	**12.17**

Table 6. Conditional probability of price aggressiveness (continued)

MIM	Ask						Bid					
t-1 \ t	1.	2.	3.	4.	5.	6.	1.	2.	3.	4.	5.	6.
1.	40.20	**44.44**	27.59	22.44	26.60	22.58	**40.39**	40.00	28.63	24.60	27.41	25.88
2.	0.08	0.00	**0.13**	0.00	0.00	0.00	0.02	0.00	**0.13**	0.04	0.00	0.00
3.	29.76	22.22	**39.60**	35.07	31.38	37.36	30.67	30.00	38.89	36.57	30.63	**38.95**
4.	14.27	11.11	15.55	**24.21**	19.98	13.83	11.74	10.00	13.29	**19.82**	18.12	11.50
5.	6.28	11.11	6.83	7.66	**11.94**	5.56	4.06	0.00	4.44	5.16	**12.63**	4.41
6.	9.42	11.11	10.31	10.62	10.10	**20.68**	13.13	**20.00**	14.62	13.81	11.20	19.26

This table shows the frequency of each order category at time t conditional on the degree of the last price aggressiveness, on the ask side and the bid side separately, for stocks BHP, NAB, TLS, WOW and MIM. Orders range in a sequence of price aggressiveness, i.e. category 1 is the most aggressive order, market/marketable limit order and category 6 is cancellation. Each row and column represents the order flow event at time t-1 and t respectively. Each column is an analogous probability vector and adds up to 100%. The maximum proportion in each row is in bold.

Table 7. Order Size for each order aggressiveness category

Order aggressiveness	BHP		NAB		TLS		WOW		MIM	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
1.Market/marketable limit order	144,449	164,024	35,696	42,790	341,197	381,855	30,578	34,549	179,929	210,823
	(9056)	(5679)	(3052)	(3221)	(18615)	(19157)	(4532)	(3903)	(45924)	(38598)
2.Limit order with improved price	6,934	6,044	9,019	8,878	3,912	2,880	3,144	3,913	982	1,074
	(12601)	(11737)	(4899)	(5068)	(55098)	(47994)	(5832)	(6332)	(109091)	(107421)
3.Limit order at best price	319,292	297,337	115,736	115,715	807,444	781,504	77,797	77,773	335,590	389,001
	(14622)	(11770)	(5451)	(5141)	(55952)	(51418)	(7577)	(7107)	(72970)	(62641)
4.Limit order within 3 ticks from best price	71,410	71,674	16,561	17,263	311,446	264,190	11,003	12,012	93,275	88,235
	(10969)	(7794)	(3868)	(3564)	(51632)	(47287)	(5504)	(4156)	(42513)	(37027)
5.Limit order more 3 ticks from best price	29,051	24,123	15,766	15,374	46,614	22,615	7,900	6,922	12,560	12,730
	(5578)	(3154)	(3016)	(2883)	(10213)	(6088)	(3290)	(2452)	(13638)	(15173)
6.Cancellation	73,359	76,718	25,682	26,787	116,815	310,510	13,803	16,898	73,731	143,431
	(13562)	(12610)	(5163)	(4968)	(78557)	(66334)	(7086)	(6855)	(49987)	(56491)
Total	644,555	639,921	218,461	226,807	1,627,427	1,763,554	144,225	152,068	696,066	845,293
	(11620)	(8247)	(4435)	(4270)	(36234)	(35863)	(6034)	(5319)	(53070)	(48460)

The table shows the total number of shares (in 1,000) submitted and the average number of shares of each order (in brackets) during the sample period, according to every price aggressiveness category.

Table 8. Hausman specification test

	BHP		NAB		TLS		WOW		MIM	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
residual	-0.0022**	-0.0007**	-0.0111**	-0.0037**	-0.0018**	-0.0024**	-0.0061**	-0.0057**	0.0002**	-0.0001**

This table shows the estimated coefficients of residuals in the following equation of Hausman specification test:

$$Z_t^* = \gamma_1 \hat{Q}_t + \beta' X_{t-1}^Z + \delta \hat{\upsilon}_{2t} + u_t^Z$$

This equation is estimated by ML and a z-test is performed for the coefficient δ . If the z-test is significant, we do not reject the hypothesis of simultaneity; otherwise, we reject it. The estimated coefficients of other explanatory variables are not reported here, for brevity.

Table 9. Order submission regressions

Order submission: Price aggressiveness

	BHP		NAB		TLS		WOW		MIM	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
volume	-0.0122	0.0335	0.0078	0.1789*	-0.0048*	0.0026	-0.0496	0.0579	-0.0052	0.0033
askdep	0.0006	0.0419	-0.0221**	0.0251**	0.0030**	-0.0020**	0.0765*	0.0086	0.0016**	-0.0010**
biddep	-0.0027**	-0.0617	0.0545**	-0.0600**	-0.0016**	0.0039**	0.0422	0.0081	-0.0017**	0.0017**
reldep	0.1237**	-0.4562	-0.0851*	0.1347**	-0.1643**	0.2726**	-0.1032*	-0.0291	-0.3356*	0.0918
spread	-0.0704**	-0.1184**	-0.0408**	-0.0410**	-0.1118	-0.2675**	-0.0702**	-0.1058**	-0.0479	-0.2299
volat	-0.0049	-0.0240	0.0007	-0.0207**	-0.0766	-0.1371**	0.0074	-0.0291	-0.0408	-0.1173
Δp	-0.3965**	0.0243	-0.2679**	-0.0767	-0.2826**	-0.0173	-0.2099*	-0.1321	-0.5429**	0.3426**
h10	-0.0230	-0.0276	-0.0469	-0.0343	-0.1122	-0.0460	-0.0929	-0.0968	0.0590	0.2163
h11	-0.0184	-0.0059	-0.0502	-0.0137	-0.0537	-0.0030	-0.0401	-0.0237	0.0444	0.1696
h12	-0.1014	-0.0110	-0.0460	-0.0214	0.0095	-0.0225	-0.0005	-0.0245	0.1222	0.1194
h13	0.0031	0.0262	0.0074	0.0483	0.0283	-0.0092	0.0177	-0.0265	0.1409	0.1231
h14	0.0057	-0.0223	0.0019	-0.0104	0.0483	0.0074	0.0536	0.0451	0.1212	0.1976
h15	0.0742	0.0440	0.1006	0.0459	0.1134	0.0657	0.0788	0.0905	0.2791	0.3302

Table 9. Order submission regressions (continued)

Order Submission: Quantity										
	BHP		NAB		TLS		WOW		MIM	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
intercept	14.1520**	-0.2397	4.5304**	4.7215**	31.1207**	39.4181**	5.6098**	4.7513**	50.9637**	55.9695**
aggre	17.6050**	38.2906	8.0472**	4.3889**	333.6641**	419.6289**	47.9554**	9.8143**	102.1842**	-204.5744**
askdep	-0.0061	-1.2376**	0.0934**	-0.1179**	-1.0206**	0.8382**	-1.0313**	0.1081	0.0205*	-0.2700**
biddep	0.0607**	1.8380**	-0.2250**	0.2975**	0.6260**	-1.6355**	-0.2867	0.1499	0.1862**	0.4335**
reldep	-2.1205**	13.2597**	1.5729**	-0.6885**	83.0317**	-113.5445**	2.7016**	0.2201	-10.3149	-61.7074**
spread	-0.0364	4.7110	0.2635**	0.1643*	44.2413**	112.2703**	3.1927**	1.0278**	26.4232*	-9.0178
volat	0.2006**	0.7760**	0.0847**	0.1175**	63.7451**	54.0213**	0.3754**	0.4586**	34.7240**	7.2812
h10	11.2522**	7.7268**	4.5452**	4.4253**	39.2476**	36.4712**	5.7327**	5.5103**	54.5288**	50.0137**
h11	11.2270**	7.5708**	4.0912**	4.0238**	37.6526**	34.9134**	5.6518**	5.0448**	51.4126**	42.3759**
h12	10.1150**	8.5449**	3.9438**	3.6178**	26.5998**	30.0016**	5.2446**	4.3570**	44.8105**	39.7190**
h13	7.8262**	4.5211**	2.9002**	2.7239**	20.9048**	21.3192**	3.9821**	3.0749**	37.4060**	25.4867**
h14	12.5039**	8.7159**	4.4627**	4.4794**	42.8137**	40.3389**	6.5143**	5.3291**	59.3629**	55.3145**
h15	13.4288**	9.9956**	5.2076**	4.8147**	38.4823**	42.1040**	7.1975**	6.4840**	58.5360**	59.5938**

This table shows the results from estimations of

$Z_t^* = \gamma_1 Q_t + \beta' X_{t-1}^Z + u_t^Z$ for price aggressiveness using Ordered Probit model, and

$Q_t = a + \gamma_2 Z_t^* + b' X_{t-1}^Q + u_t^Q$ for quantity by two-stage OLS model.

Where Z_t^* is the latent price aggressiveness, Q_t is the quantity, X_{t-1}^Z and X_{t-1}^Q consist of explanatory variables for price aggressiveness and quantity equations. The hourly dummies in price aggressiveness equation are analyzed in a separate regression. Both equations are estimated for the order submission on ask and bid sides of market separately, using data of five stocks on the ASX market during July and August, 2002. ** and * denote statistical significance at 1 and 5 percent levels based on the z-statistics for price aggressiveness equation and **t**-statistics for quantity equation, respectively.

Table 10. Marginal probability

Marginal probability: BHP (%)

	1.Market/ marketable limit order		2.Limit order with improved price		3.Limit order at best price		4.Limit order within 3 ticks from best price		5.Limit order more 3 ticks from best price		6.Cancellation	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
volume	0.210	-0.489	0.543	-1.357	0.096	0.048	-0.009	0.059	-0.007	0.007	-0.417	1.266
askdep	-0.010	-0.613	-0.025	-1.699	-0.004	0.061	0.000	0.074	0.000	0.009	0.019	1.584
biddep	0.046	0.903	0.118	2.504	0.021	-0.089	-0.002	-0.109	-0.001	-0.013	-0.091	-2.335
reldep	-2.121	6.673	-5.484	18.51	-0.969	-0.661	0.087	-0.806	0.068	-0.096	4.212	-17.26
spread	1.207	1.732	3.120	4.802	0.552	-0.172	-0.049	-0.209	-0.039	-0.025	-2.397	-4.478
volat	0.084	0.351	0.216	0.972	0.038	-0.035	-0.003	-0.042	-0.003	-0.005	-0.166	-0.907
Δp	6.801	-0.356	17.58	-0.986	3.108	0.035	-0.277	0.043	-0.217	0.005	-13.51	0.920

Table 10. Marginal probability (continued)

Marginal probability: NAB (%)

	1.Market/ marketable limit order		2.Limit order with improved price		3.Limit order at best price		4.Limit order within 3 ticks from best price		5.Limit order more 3 ticks from best price		6.Cancellation	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
volume	-0.137	-3.156	-0.358	-8.178	-0.050	-1.005	0.004	0.183	0.019	0.372	0.239	5.676
askdep	0.389	-0.443	1.020	-1.148	0.144	-0.141	-0.012	0.026	-0.055	0.052	-0.683	0.797
biddep	-0.959	1.058	-2.513	2.742	-0.355	0.337	0.030	-0.062	0.135	-0.125	1.682	-1.903
reldep	1.498	-2.377	3.925	-6.160	0.554	-0.757	-0.047	0.138	-0.211	0.280	-2.626	4.275
spread	0.718	0.723	1.880	1.873	0.265	0.230	-0.023	-0.042	-0.101	-0.085	-1.258	-1.300
volat	-0.012	0.365	-0.032	0.947	-0.005	0.116	0.000	-0.021	0.002	-0.043	0.022	-0.657
Δp	4.715	1.354	12.355	3.507	1.744	0.431	-0.149	-0.079	-0.663	-0.160	-8.267	-2.434

Table 10. Marginal probability (continued)

Marginal probability: TLS (%)

	1.Market/ marketable limit order		2.Limit order with improved price		3.Limit order at best price		4.Limit order within 3 ticks from best price		5.Limit order more 3 ticks from best price		6.Cancellation	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
volume	0.035	-0.044	0.138	-0.110	0.038	-0.029	0.016	-0.019	0.000	0.000	-0.186	0.102
askdep	-0.022	0.033	-0.088	0.084	-0.024	0.022	-0.010	0.014	0.000	0.000	0.118	-0.078
biddep	0.012	-0.065	0.046	-0.162	0.013	-0.042	0.005	-0.028	0.000	0.000	-0.062	0.150
reldep	1.198	-4.544	4.725	-11.38	1.296	-2.974	0.548	-1.935	-0.006	0.008	-6.382	10.57
spread	0.815	4.460	3.215	11.17	0.882	2.919	0.373	1.900	-0.004	-0.008	-4.343	-10.37
volat	0.559	2.285	2.202	5.722	0.604	1.496	0.255	0.973	-0.003	-0.004	-2.975	-5.315
Δp	2.060	0.288	8.125	0.722	2.229	0.189	0.942	0.123	-0.010	-0.001	-10.97	-0.671

Table 10. Marginal probability (continued)

Marginal probability: WOW (%)

	1.Market/ marketable limit order		2.Limit order with improved price		3.Limit order at best price		4.Limit order within 3 ticks from best price		5.Limit order more 3 ticks from best price		6.Cancellation	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
volume	0.744	-0.901	2.050	-2.435	0.221	-0.413	0.033	-0.127	-0.061	0.060	-1.674	2.038
askdep	-1.148	-0.133	-3.162	-0.360	-0.341	-0.061	-0.051	-0.019	0.095	0.009	2.583	0.301
biddep	-0.633	-0.126	-1.743	-0.341	-0.188	-0.058	-0.028	-0.018	0.052	0.008	1.424	0.286
reldep	1.548	0.454	4.265	1.226	0.460	0.208	0.069	0.064	-0.128	-0.030	-3.484	-1.026
spread	1.052	1.647	2.900	4.451	0.313	0.755	0.047	0.232	-0.087	-0.109	-2.369	-3.726
volat	-0.112	0.453	-0.308	1.225	-0.033	0.208	-0.005	0.064	0.009	-0.030	0.251	-1.025
Δp	3.149	2.055	8.675	5.555	0.936	0.943	0.141	0.290	-0.259	-0.136	-7.086	-4.650

Table 10. Marginal probability (continued)

Marginal probability: MIM (%)

	1.Market/ marketable limit order		2.Limit order with improved price		3.Limit order at best price		4.Limit order within 3 ticks from best price		5.Limit order more 3 ticks from best price		6.Cancellation	
	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid	Ask	Bid
volume	0.099	-0.075	0.236	-0.164	0.039	-0.036	-0.020	-0.003	0.000	0.000	-0.181	0.116
askdep	-0.031	0.022	-0.074	0.048	-0.012	0.011	0.006	0.001	0.000	0.000	0.057	-0.034
biddep	0.033	-0.039	0.078	-0.086	0.013	-0.019	-0.007	-0.002	0.000	0.000	-0.060	0.061
reldep	6.350	-2.084	15.18	-4.590	2.514	-1.009	-1.311	-0.082	-0.012	0.003	-11.63	3.250
spread	0.907	5.216	2.169	11.49	0.359	2.527	-0.187	0.206	-0.002	-0.007	-1.661	-8.135
volat	0.771	2.661	1.844	5.862	0.305	1.289	-0.159	0.105	-0.002	-0.003	-1.412	-4.151
Δp	10.27	-7.775	24.56	-17.13	4.068	-3.766	-2.121	-0.306	-0.020	0.010	-18.81	12.13

This table shows the marginal reactions of price aggressiveness in each category to a change in the explanatory variables. Orders are divided into six price aggressiveness category as discussed in section 4. The values are obtained using equation (9). In order to calculate the marginal probability, the estimated coefficients resulting from the Ordered Probit regressions, together with the unconditional mean of quantity and explanatory variables are used. The marginal probability of hourly dummies is not reported, for simplicity.

Table 11. Comparison between predicted and empirical signs of regression coefficients

Hypothesis	Variable	Ask orders		Bid orders	
		Predicted	Empirical	Predicted	Empirical
Trade-off	*volume*	−	−	−	+
H1	*askdep*	+	+*	−	?
H2	*biddep*	−	?	+	?
H3	*reldep*	+	−*	−	?
H4	*spread*	−	−**	−	−**
H5	*volat*	−	?	−	−
H6	Δp	−	−**	+	?
H7	*hourdummies*	− → +	− → +	− → +	− → +

This table summarizes the empirical findings, compared with tested hypotheses through matching the expected and empirical signs of the coefficients in the price aggressiveness regressions for stocks BHP, NAB, TLS, WOW and MIM. Empirical signs are given when as least the results for four stocks keep the same, otherwise, a symbol (?) is marked, representing ambiguous results. The signs of predicted and empirical coefficients for hourly dummies change from negative to positive.

International Journal of Economics and Finance
Vol. 1, No. 1
February 2009
www.ccsenet.org/journal.html

Cognitive Pattern Analysis Employing Neural Networks: Evidence from the Australian Capital Markets

Edward Sek Khin Wong
Faculty of Business & Accountancy
University of Malaya
50603, Kuala Lumpur, Malaysia
E-mail: edwardwong@graduate.uwa.edu.au

Abstract

An artificial neural network is an intelligent system using computers that allows users to improve performance through pattern recognition. Neural networks benchmark their predictions with actual results and constantly revise their predictions, improving forecasting capability. The purpose of this paper is to support the use of neural networks as a detection mechanism tool to discover market inefficiencies in financial markets. Using the Australian capital market as an example, this study investigates the question of the existence of market inefficiencies using artificial neural networks as the investigative tool. This study also focuses on whether additional publicly sourced information when used as input through a neural network, can provide investors with a trading advantage over traditional financial models. In finance, any forecasting advantage obtained through the use of publicly available information even if internal, external or both indicate some form of inefficiency in the financial markets. In this paper, we explore the efficiency of the four capital markets, the United States, Japan, Hong Kong and Australia, but focusing on the Australian capital market using the Australian Stock Exchange's 200 (ASX 200) index. Our research demonstrates how the inclusion of external information to our neural network model can provide a significant trading advantage. Although our preliminary analysis suggests the Australian market is as efficient as the US market, other results suggest the Australian market may be less efficient, when we take into account of external information from other financial markets. Our results show that accounting for external market signals can significantly improve forecasts on the ASX200 index but show little benefit on forecasts for the Dow Jones Industrial Average (DJIA) index. Using external market signals, our neural network model's prediction accuracy of the ASX200 index increases by an additional 10 percent from 50 to 60 percent accuracy. This suggests the inclusion of publicly available external signals can significantly improve neural network forecasts. By exploiting this additional information, this method can improve returns for investors who incorporate such signals in their neural network models.

Keywords: Neural networks, Cognitive engineering, Intelligence systems, Australian capital market, Market inefficiency

1. Introduction

In finance literature, the issue of capital market efficiency is a most important topic and factor for decision making. Any ability by any investor to improve their returns by exploiting publicly available information constitutes market inefficiency (Walczak 1999). In an efficient market, information when released to the markets, are also assumed to be priced into the capital market, as they react to the new knowledge and adjust their value instantaneously to reflect their "true" values. Hence, there is no trading advantage, when publicly available information is used to trade. (Note 1). Since publicly available information is expected to be quickly reflected in financial markets when it is released, one would expect the more established capital markets, such as New York Stock Exchange (Dow Jones Industrial Average, DJIA) and Tokyo Stock Exchange (Nikkei 225 average) to be close to efficiency due to the fluidity of these markets. The ability to forecast movements in these market using publicly available external signals would be inconsistent to market efficiency. For the purposes of this paper, an efficient market is defined as a market that is random and unpredictable, and inefficiency is any variation from this definition. According to Levich (1981), 50 percent prediction accuracy is reasonable, when compared against human trading performance in the capital markets. This suggests any model that can predict future values with accuracy greater than 50 percent violates this concept of an efficient market.

1.1 Research Objectives

The purpose of this paper is to test the proposition of whether market inefficiencies exist in the Australian market using neural networks. As neural networks are capable of predicting future values of data, neural networks can simultaneously test market inefficiencies and evaluate these to show the trader if there is any trading advantage using these inefficiencies for trading advantages. This study also investigates whether the inclusion of additional external market indicators such as information from other markets can improve forecasts of the Australian market

2. Literature Review

2.1 Australian Stock Exchange

The Australian stock exchange is one of the largest exchanges with the largest volumes of daily trades in southern hemisphere. Despite its relative size, several factors are known to influence movements on the Australian stock exchange. Yoda (1994) finds the Australian markets are influenced by changes in anticipated production levels and interest rates of more established markets such as the United States (US), Europe, and Japan, with the U.S. commanding the highest influence.

Besides the influence anticipated foreign production levels, other factors also influence the Australian market. A study by McGee (1997) finds a strong correlation among these established markets when the global economy world faced the Asia Crisis around the October 1997 crashed. Another study by Kwok (1995) also finds movements in these markets such as the US, Hong Kong, London, and Japan influences movements in the Australian financial market. The co-integration between these markets shows some nonlinear dependencies of daily stock returns for periods lasting one to five days (Yadav, 1996). This suggests neural networks may be appropriate for predicting market performance of periods lasting one to five days.

2.2 Neural Networks in Market Forecasting

Neural Networks mimics the human brain's ability to classify patterns or to make predictions or decisions based upon past experience using data sets (Specht, D. and Shapiro, P, 1991). This enables us to build sophisticated custom problem solving applications without programming (Burrascano 1991).

Past studies have used neural networks to forecast individual stock market indices. Due to the popularity of neural networks, there is an overabundance of studies using neural networks in financial market research. Most of these studies focus on predicting established markets such as the US, London, Hong Kong, and the Japan markets (Bosarge, 1993, Yoda, 1994 and Tsibouris and Zeidenberg, 1995). Although many studies have been done on these markets, most past studies continue to focus on using only internal market information. This includes moving averages, trading volume, and interest rates of the particular market for predicting future values. According to Bosarge (1993) and Yoda (1994), these studies have a typical performance rate of 50 to 53 percent accuracy. We believe by accounting for additional external information produced by markets which are co-integrated with the Australian market, our model can improve forecasts. For the purpose of this study, the 50 percent prediction accuracy is assumed to typical neural network capability, or random based behaviour.

3. Data and Methodology

In this study, the data is sourced from DataStream International where the closing values of the market index of four countries; the ASX200, DJIA (Dow Jones Industrial Average), Nikkei 225, and Hang Seng index are selected. Data for each index is sourced between the periods of Jan 3, 2006 through October 16, one-day, two-day, and five-day simple returns for each index are calculated. Figure 1 shows the ASX200 index value for the entire period.

Insert Figure1 Here

Other input variables over this period are also added as inputs into our neural network. These include simple averages and trade volume for each financial market, which has been consistently used in past studies.

3.1 Development of Model

From the literature review, it is well established that the presence of inefficient behaviour exist in financial markets. (Note 2). This is particularly highlighted in Lo and MacKinlay (1988) using a variance ratio test. (Note 3) However, in this study we only focus on the performance of our neural network in predicting the index values of the four financial markets; (1) ASX200, (2) DJIA, (3) Nikkei 225, and (4) Hang Seng. For each of the four markets, our neural network is conditioned to predict the five-day future value of each index for each respective market. An example of a neural network using the ASX 200 is shown below in Figure 2.

Insert Figure 2 Here

Insert Figure 3 Here

In this study, we use a neural network with one hidden layer back propagation network for each input variable set. For this type of neural network, the decision nodes for each market are developed in three main stages. The first stage involves feeding inputs into the model to develop its predictive capability through the pattern recognition by its nodes (January 3, 2005 to December 30, 2005). The second stage is often known as the training stage January 3, 2006 to June 30, 2006) and involves creating a back-propagative or feedback function into the neural network algorithm. In this stage, the available data of the model is tested by predicting future values and these predicted values are not produced as output to the user, instead this data allows the algorithm to adjust for predictive errors encountered, increasing its accuracy in predicting future market index values. In the backward-propagation stage, the training data is vital as it improves the predictive capability by accounting for errors in the network's predictive process. In the final stage (July 3, 2006 to October 15, 2006), from which our research results are based, is the process of simply predicting future five day values that are not in the earlier two stages.

For each financial market the inputs of one-day, two-day, and five-day lags of closing index values, along with the corresponding one-day, two-day, and five-day average trading volumes are calculated for respective markets.

The predictive capability of our neural network is also evaluated in two stages, with both using the nonparametric runs test. This test is specifically chosen as it compares the frequency of a run (two consecutive periods of financial data moving in the same direction) against the frequency of instances where they do not move in the same direction. This test basically suggests the data tested should be random and hence approximately half (50%) of the observations should occur in runs and half that do not (Urrutia 1995). This statistic is calculated by multiplying the value of 1 - *MAE* (Mean Absolute Error) by 100, where the error is 0 if the correct direction of change is predicted and 1 if not so predicted. For each market, the null hypothesis of 50 percent prediction accuracy is assumed based on the review of past studies. The *p* value indicates the confidence of rejecting the null hypothesis of 50%.

In the first stage, our neural network model's ability is evaluated using only internal data for each financial market. The second stage focuses upon investigating if adding external data from other markets as inputs can improve our model's predictive capability.

4. Results

Stage One:

Insert Table 1 Here

From Table 1, we can see at a 0.10 confidence level, it is sufficient to reject our hypothesis for both the ASX200 and Hang Seng market index. At the higher 0.05 confidence level, the null hypothesis can only be rejected for ASX200 market index. This suggests that there may be some inefficiency to indicate the presence of non-random market behaviour and this factor can provide a significant trading advantage to investors. Having identified the potential for the Australian market to be inefficient, we next focus on evaluating whether the use of information from different markets effects can improve our forecasts.

Stage Two:

Focusing only on the Australian market, six different sets of variables are implemented in our neural network model to investigate if the inclusion of external signals has an effect in the apparently inefficient markets of Australia. The various combinations of external impacts are used to determine the effect of specific signals to the Australia markets. Prediction accuracy results for the single hidden layer networks are displayed in Table 2.

Insert Table 2 Here.

From our results in Table 2, we find that when additional external signals, publicly available information, are taken into account, our nonlinear neural network model predictive performance improves from 49.71% to 59.57%. Our result suggests that external publicly available signals when used as inputs can be used to predict the market more reliably; this indicates some inefficiency in the Australia market.

5. Conclusion

Neural networks provide a tool for evaluating the presence of an inefficient market and exploiting that inefficiency. In this study, it is shown that neural networks can evaluate inefficiencies in the Australian capital market. It is also able to predict future values of the market quite reliably.

Our results show the Australian capital market appears the most inefficient market against the other three markets the Hang Seng, Nikkei 225, and DJIA. This suggests there may be hidden investment rules that allow traders to "beat the market" make positive abnormal profits on the Australia stock market.

Although our results indicate some form of inefficiency in the Australian market, it must be noted that our results are limited to using a simple neural network as a predictive model. The neural network used had only one hidden layer back propagation network for each input variable set. The use of a single hidden layer back propagation was assumed to be sufficient to validate the nonlinear function.

Although our approach is geared towards the results of past studies, the development of neural networks in providing an intelligent decision aid for determining the relative efficiency in Australia market is shown. Further work may be conducted by re-investigating our results using a more complex layer or identifying the optimal neural network architecture for use in forecasting, to show if trading advantages do exist in other equity markets. Neural networks as a tool can aid investors in identifying investment strategies to capitalise on any potential market inefficiencies.

References

Bosarge, W. E. (1993). *Adaptive Processes to Exploit the Nonlinear Structure of Financial Market*.

In: R. R. Trippi & E. Turban (eds.). Neural Networks in Finance and Investing. Probus Publishing, pp. 371--402.

Burrascano, P. (1991). Learning Vector Quantization for the Probabilistic Neural Network. *Proceedings of the IEEE International Conference on Neural Networks,* July 1991, Vol. 2, pp. 458-461.

Kwok, R.H.F. (1995). Market Integration in the Four Newly Industrialised Economies of Asia, *Advances in Pacific Basin Financial Markets,* Vol. 1, pp. 199-209.

Levich, R. & L.R. Thomas. (1993). The Significance of Technical Trading-Rule Profits in the Foreign Exchange Market: A Bootstrap Approach, *Journal of International Money and Finance,* Vol12, pp. 451-474.

Lo, A. W. & MacKinlay, A. C. (1988). Stock Market Prices Do Not Follow Random Walks: Evidence from a Simple Specification Test, *Review of Financial Studies* Vol.1, pp. 41-66.

Specht, D. and Shapiro, P. (1991). Generalization Accuracy of Probabilistic Neural Networks Compared with Back-Propagation Networks. *Proceedings of the International Joint Conference on Neural Networks*, July 8-1, 1991, Vol 1, pp. 887-892.

Steven Walczak. (1999). Quality of Neural Network Financial Models as a Function of Time. *Decision Sciences Institute 5th International Conference Proceedings*, pp. 537-539.

Steven Walczak. (2001). An Empirical Analysis of Data Requirements for Financial Forecasting with Neural Networks. *Journal of Management Information Systems*, *17 (4)*, 203-222.

Tsibouris, G. and Zeidenberg, M. (1995). Testing the efficient markets hypothesis with gradient descent algorithms. In A.P. Refenes, editor, *Neural Networks in the Capital Markets,* chapter 8, pages 127--136. John Wiley & Sons, Chichester, England, 1995.

Urrutia, J.L. (1995). Tests of random walk and market efficiency for Latin American emerging equity markets. *Journal of Financial Research,* 18, 3 (Fall 1995), pp. 299-309.

Yadav, P.K.; Paudyal, K.; and Pope P.F. (1996). Non-linear Dependence in Daily Stock Returns: Evidence from Pacific Basin Markets - *Advances in Pacific Basin Financial Markets,* Vol 2 Part B (1996) pp 349-377.

Yoda, M. (1994). Predicting the Tokyo stock market. In G.J. Deboeck (ed.), T*rading on the Edge: Neural, Genetic, and Fuzzy Systems for Chaotic Financial Markets.* New York: Wiley, 1994, pp. 66-79.

Notes

Note 1. In finance, this is commonly known as semi-strong form efficiency of financial markets.

Note 2. See Bosarge 1993, Yoda 1994 and Tsibouris and Zeidenberg 1995 for some evidence.

Note 3. The variance ratio test is also performed in the markets by the author, but not reported. For these results please contact the author.

Table 1. Single Intramarket index set Neural Network Model

Index	Nonparametric Runs Prediction accuracy	*P* value (confidence level)
ASX200	49.81	0.082**
DJIA	56.87	0.069
Nikkei	47.29	0.956
Hang Seng	54.97	0.102*

* indicates significant at the 0.10 level

** indicates significant at the 0.05 level

Table 2. Neural Networks to Predict ASX200 Five-day Futures.

Indexes	Prediction accuracy	P value
ASX200	49.81	0.082
ASX200 and Hang Seng	51.79	0.366**
ASX200 and Nikkei	52.59	0.219**
ASX200 and DJIA	54.27	0.1841**
ASX200, Hang Seng, Nikkei and DJIA	59.57	0.003**

** indicates significant at the 0.05 level

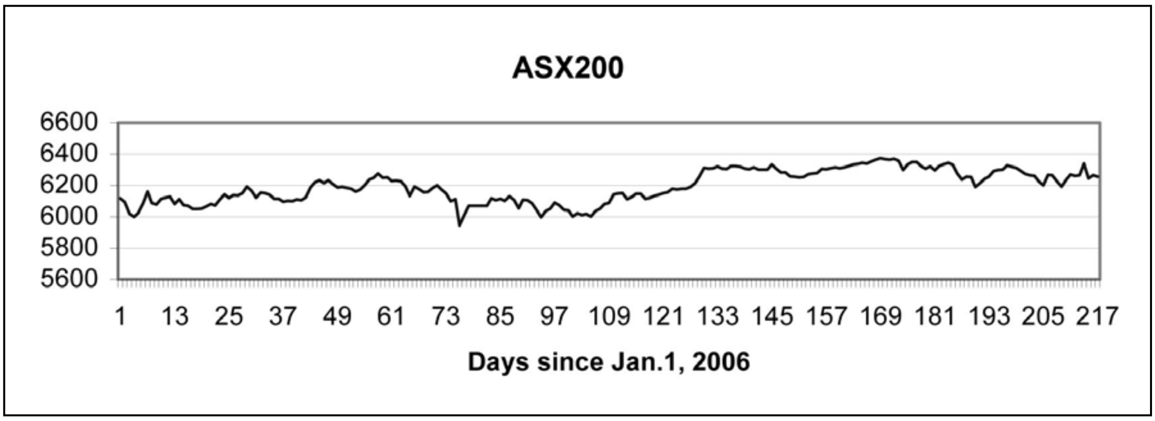

Figure 1. ASX200 Market Index Values, January 3, 2006 – October 16, 2006

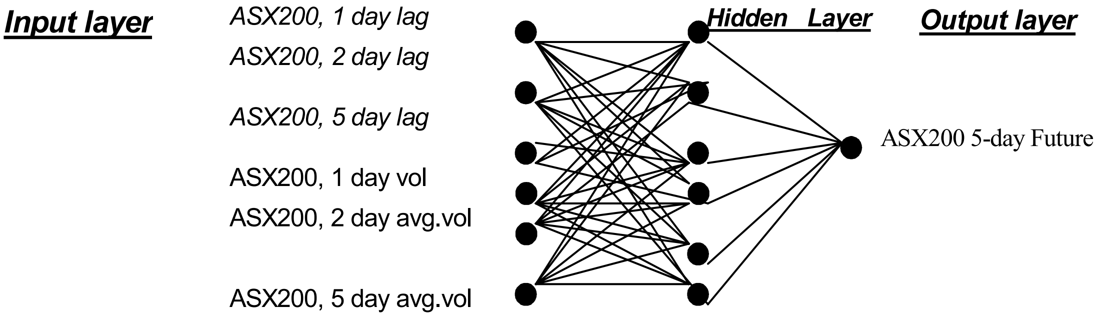
Input layer
ASX200, 1 day lag
ASX200, 2 day lag
ASX200, 5 day lag
ASX200, 1 day vol
ASX200, 2 day avg.vol
ASX200, 5 day avg.vol
Hidden Layer
Output layer
ASX200 5-day Future

Figure 2. Sample Neural Network for forecasting ASX200 Five-day Future Values

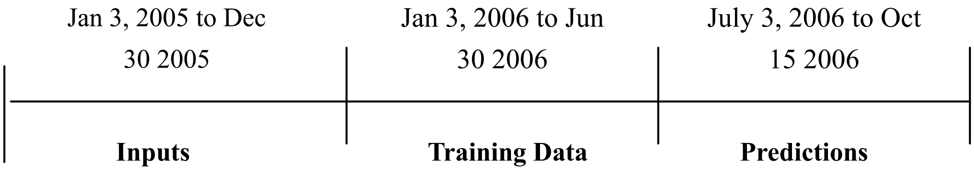
Jan 3, 2005 to Dec 30 2005
Jan 3, 2006 to Jun 30 2006
July 3, 2006 to Oct 15 2006
Inputs
Training Data
Predictions

Figure 3. Timeline of data used for each market for forecasting five-day future values

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Study on the Transnational Venture Capital Investments:
A Theoretical Framework

Yijun Zeng
School of Business Administration
Henan University of Finance & Economics
Zhengzhou 450002, China
E-mail: zengyijun63@163.com

Abstract

Starting from reviewing the transnational venture capital investment phenomena, this article points out the necessity to study the transnational venture capital investment, and establish a theoretical framework about the research of the transnational venture capital investment. The framework is composed by the entry system that both investors and attractors find the values of the venture capital investment, the operation system of the venture capital investment and the exit system realizing the investment value, and these three systems are interdependent as well as relatively independent, and they reflect the generality and the particularity of the transnational venture capital investment, and they are also the uniform integer.

Keywords: Transnational venture capital investment, Investment motivation, Operation mechanism, Theoretical framework

1. Research backgrounds and introduction

Since 1990s, with the development of the global economic integration, the internationalization tendency of the venture capital investment is more and more obvious, and the transnational venture capital investment has been the important part of the international investment activity.

As we all know, the venture capital investment has important drive function for the economic growth and technical advancement of one country. From 1995 to 2000, the financial amount of the global venture capital added from 44 billion dollars to 250 billion dollars, and the venture capital investment industry has been one of the industries with the quickest growth in the world, and the yearly average venture capital investment growth rate in seven most developed countries has achieved 42%. The venture capital investment has been the important channel of the national economic adjustment, especially for the high and new technology industry. The flow of the venture capitals in the national market drives the development of the transnational venture capital investment. In US where the venture capital investment industry develops most, from 1980s to the middle of 1990s, the proportion of the foreign venture capitals in the US venture capital investment amount was about 10%, and in 2000, this number achieved 26%. After 1990s, the amount that the venture capitals of developed countries flow to developing countries continually increased, for example, the foreign venture capitals flowing to India, China and Latin America presented the quick development tendency. In 2002, the proportion of the foreign venture capitals flowed to US is 20% of the whole venture capitals of US, and in the whole Europe, the venture capital flow among countries and foreign flow occupied 27% of the amount of the venture capitals, and in Asia, the transnational flow of venture capitals among nations and regions and foreign flow occupied 51% of the amount. These data indicated that the transnational venture investment had formed the scale through the quick development in the late of 1990s, and it has been a sort of new tendency for the international direct investment.

The practice of the transnational venture capital investment promotes the economic development and the technical advancement of various countries. This sort of investment activity has been concerned by some scholars who had issued some articles about the research of venture capital investment internationalization. But most of these researches are only limited in the factor analysis and development tendency about the transnational venture capital investment, and there are few foreign and domestic scholars to systematically study the transnational venture capital investment. The amount the international venture capital investment increases steadily and the process of the transnational venture capital investment is quickening gradually, so we should quietly emphasize the research about the transnational venture capital investment theory. For a series of problems about the transnational venture capital investment such as investment motivation, investment mode, operation mechanism, exit mode and the influences on the host countries, we urgently need to explain them in theory and summarize them in practice. So, aiming at above problems, this article tries to establish a basic theoretical framework to systematically study the transnational venture capital investment behaviors.

2. Research view and literatures summarization

The venture capital investment is a sort of equity investment, and it seeks the long-term investment incomes of the capitals, and it doesn't require the fluidity of the capitals, so it belongs to the direct investments. And in his article, we define the "transnational venture capital investment" as one sort of special form of the international direct investment. For the research about the transnational venture capital investment, we can start from the supply and demand relationship of the international venture capitals, and systematically study the capital operation process of the transnational venture capital investment based on the international direct investment theory.

About the research of venture capital investment, there are many systematic and special research results through tens years' developments, and they are relatively completed. But for a series of problems brought by the transnational factors, the traditional theory is hard to offer reasonable explanations, because the transnational venture capital investment possesses many characters such as high uncertainty and high risk. First, the transnational factors will bring added risks such as politic risk, exchange rate risk and high agency costs. Second, in the operation process of the venture capitals, the two-class agent problem exists, i.e. the agent relationships between investors and venture investors and between venture investors and venture enterprises, so the converse selection and moral risk will exist in the operations of venture enterprises because of asymmetric information. Third, as a sort of equity investment, the venture capital investment realizes the capital increment through the stock transfer, so the exit problem of the venture capitals possesses special characters. In fact, the research about the transnational venture capital investment is just rising in recent several years. Western scholars studies the transnational venture capital investment from the environment and mode of the venture capital investment, the cultivation about the supply and demand of the venture capitals, the decision factors and future tendency of the transnational venture capital investment, and the representative research results include following aspects.

Martin Kenney, Kyonghee Han and Shoko Tanaka (2002) summarized some characters of the transnational venture capital investment through reviewing the historical process of the international investment of the venture capitals. They thought that it depended on an acceptable environment which mainly included the commercial opportunity that allowed large capital to profit for the venture capitals to successfully implement transnational extension. In some countries which venture capital investment industry develops successfully, all governments have offered necessary assistances, which is the important factor to start the growth of the venture capital investment. Of course, the cultural traditions, social concepts and governmental regulations in some countries will baffle the developments of the venture capital investments, and they all need the government creates an advantaged environment. The environmental difference among countries may make the investment mode of the international venture capitals mix with the venture capital investment mode of the host country, and form the mode with the character of the host country. Therefore, the internationalization mode of the venture capital investment is mainly decided by the bilateral relationship.

Gompers and Lerner (2001) studied the internationalization of the venture capital investment from the supply and demand of the venture capitals. They thought that for each country which wanted to participate in the venture capital investment internationalization, it should emphasize the infrastructure construction of the venture capital investment, i.e. the cultivation about the supply and demand of the venture capitals. The supply of the venture capitals depends on three factors. (1) Investors' confidences. The venture capital investment will produce more return than other investments. (2) The fashion to support the venture capital investment. (3) The amount of the experienced venture investors who are in the business domain. The demand of the venture capitals depends on following conditions. (1) Regulated investment environment and legal structure. They should sweep the obstacles for the new enterprises and offer sufficient salaries for entrepreneurs. (2) Good investment basic structure, technical conditions and employee training to support the new enterprises and markets. (3) The cultural atmosphere to respect entrepreneurs' innovational spirits, especially for the venture spirits. (4) Healthy economic situation including thrifty IPO market. Gompers and Lerner advised that to promote the development of the venture capital investment, the government and national leaders have to emphasize the main factors influencing the supply and demand of the venture capitals and establish corresponding policy measures, and encourage the investment and attraction of the venture capitals.

Martin Haemmig (2003) firstly systematically studied the decision factors of the transnational venture capital investment on the layer of the enterprises. Haemmig analyzed the motivation and the strategy of the enterprise which engaged in the transnational venture capital investment, and described the decision factor of the venture capital investment internationalization. He thought that the "transferability" of employee, technology and capital decided that the new enterprise can be founded at any place in the world, and for the transnational venture capital investments, venture investors' experiences and talents are the key factors to decide whether the investment could be successful. When establishing the concrete strategies, then venture capital investment enterprises must consider many problems such as the industrial selection, the geographical location, the interior organization form of the company, the venture funds structure of the company, entrepreneur's salary and execution network because of the complexity of transnational investment and the fixed risk of venture capital investment.

William L. Megginson (2004) studied the internationalization mode and the future development tendency of the venture capital investment. The modern venture capital investment industry rooting in US has been the important tool to drive the economic growth and technical advancement for various countries, but Megginson thought that the venture capital investment mode of US can not be simply copied to other countries, and it could not be the "international" venture capital investment mode. Though in recent years, the financing ability, investment mode and investment return of the venture capitals in the world presented the centralized tendency, but for whether the future venture capital investment industry could be go to real market integration, Megginson doubt that. He thought because of the differences of the investment environment such as the legal environment in different countries, the special "international" venture capital investment market composed by part of countries might occur.

From above introductions, we can see that the research about the transnational venture capital investment is mainly centralized in the summarization and the development tendency about the influencing factors of the venture capital investment internationalization. These results can help us to understand the development process of the transnational venture capital investment in theory, but the systematic research about the transnational venture capital investment still lacks. We think that the systematic characters of the transnational venture capital investment are mainly embodied in the investment motivation, the operation mechanism, the exit mode and the influences on the host country. And this sort of research method is consistent with the influencing degree of the development history of the venture capital investment to the society and the economy. At present, under the macro background of the global economic integration, the transnational flow of the venture capitals is gradually rising, and its influences are more and more obvious. But the relative researches are still trailed the practice of the transnational venture capital investment. First, the development history of the transnational venture capital investment is very short and its rule is hard to grasp. Second, the data about the transnational venture capital investment is difficult to be acquired, and its scientific research is limited. However, in theory, the systematic research is necessary, and it will offer theoretical guidance for the development of the transnational venture capital investment and bring important references for the academic references.

3. Theoretical framework about the transnational venture capital investment research

As a sort of special investment mode of the international direct investment, the history of the transnational venture capital investment is very short, and it is only developed to be scaled investment in recent ten years. That is decided by the changes of the exterior conditions and interior factors of the venture capital investment. First, the development of the global economic integration process in recent ten years increasingly strengthens the international investment activities, which drives the transnational flow of the venture capitals. Second, the high-speed development of the scientific technology drives the demand for the international venture capitals and promotes the supply of the international venture capitals. So, to systematically study this economic phenomenon, we must base on the international direct investment theory.

In the international direct investment theory system, the monopolistic advantage theory, the internalization theory and the eclectic theory of international production occupy very important status. These theories mainly explain three problems about the international direct investment. The first one is the reason of the direct investment in foreign countries, i.e. why we invest in foreign countries. The second one is the condition of the direct investment, i.e. why we can invest in foreign countries. The third one is the object of the direct investment in foreign countries, i.e. where we invest in foreign countries. For the research about the transnational venture capital investment, we still emphasize the reason, condition and object of the transnational venture capital investment. However, the international direct investment theory still can not explain all peculiarities of the transnational venture capital investment, and we have to utilize other economics theories to study the problem. Based on that, we can primarily establish a theoretical framework about the research of the transnational venture capital investment.

First, the transnational venture capital investment theory should explain following three problems. The first problem is the reason of the transnational venture capital investment. The transnational flow of the venture capital investment requires that investor's investment motivation is consistent with the attractor's attraction motivation, and find the potential values of the invested venture enterprise and produce the transnational venture capital investment behavior through the link with the venture entrepreneur. The second problem is the operation condition of the transnational venture capital investment. Through eliminating the limitation of the transnational investment and reducing the entry bulwark, investors and attractors can establish effective operation mode according to both conditions, and create values for the invested enterprise through the service increment offered by venture entrepreneur. The third problem is the exit environment of the transnational venture capital investment. The exterior environment of the investment should ensure the venture capitals can exit at proper moment to realize their values, so the investor realizes the increment of the capital and the attractor realizes the extension of the capital.

Second, to systematically study the transnational venture capital investment, we must establish organic association between the transnational investment system of the investment country and the transnational attraction system of the invested country. As viewed from the capital operation characters of the transnational venture capital investment, above three problems basically contain the basic process of the venture capital investment, i.e. the investment, the operation and the exit of the venture capital. So we can divide the total framework of the transnational venture capital investment theory into three systems. The first system is the value discovery process of the venture capital investment object, i.e. the process translating the investment (attraction motivation) into behaviors, and we call it as the entry system. The second system is the capital increment process of the venture capital investment, and it creates values for the operation and management for the venture enterprise through the investors and the attractors, and we call it as the operation system. The third system is the value implementation process of the venture capital investment, and it realizes the high capital increment and income through the exit of the venture capitals in the special environment, and we call it as the exit system. The theoretical framework system about the research of the transnational venture capital investment is seen in Figure 1.

In the framework, the entry system, the operation system and the exit system embody the generality and the particularity of the transnational venture capital investment. The final aim of the transnational investment of the venture capitals is to realize the increment of the capital. And to achieve this aim, it must experience the basic process including discovering investment value, creating investment value and realizing investment value, and the process embodies the generality of the transnational venture capital investment. Because the transnational venture capital investment possesses many characters such as high uncertainty and high risk, so we should adopt more scientific method to make right decisions in many aspects such as the investment object selection, the item risk evaluation, the management team reputation investigation and the location selection. The entry mode and the investment mode of the transnational venture capital investment have their own characters, and especially the management increment service offered by the venture entrepreneur makes the operation mechanism of the transnational venture capital investment be differing from general international direct investment. The transnational venture capital investment needs realizing the high capital increment income through the capital exit, so the venture enterprise has special requirement for the exit environment of the capitals, which also reflects the particularity of the transnational venture capital investment. The entry system introduces the increment channel for the venture capitals, and the operation system offers the increment condition for the venture capitals, and the exit system ensures the increment aim for the venture capitals, so these three systems are interdependent as well as relatively independent.

4. Conclusions

Based on the international direct investment theory, we study the basic problems about the transnational venture capital investment, seek the combination point of the international direct investment theory and the venture capital investment starting from analyzing the characters of the venture capital investment, and establish the theoretical research framework of the transnational venture capital investment. To systematically research the transnational venture capital investment, we should depend on the analysis and method application of key problems in every subsystem such as the investor’s motivation and the attractor’s motivation research, the project risk evaluation and the decision-making behavior research in the entry system, the contract arrangement research of the venture capital investment and the function research of venture entrepreneur in the operation system, and the exit environment research and the exit mode research in the exit system. For the concrete research methods, we should further explore.

In a word, the theoretical research framework of the transnational venture capital investment establishes a complete system on the theoretical layer and on the operation layer. This system starts from the internationalization phenomenon of the supply and demand relation of the venture capitals and takes the transnational investment (attraction) motivation, operation and exit as factors to establish the mutual association mechanism between transnational investors and transnational attractors, takes the value discovery, value creation and value implementation of the venture capitals as the routes to establish the transnational venture capital investment research theoretical system through the organic unification of the three systems including the entry system, the operation system and the exit system, and it offers a new idea for us to study the transnational venture capital investment.

References

Martin Haemmig. (2003). *The Globalization of Venture Capital: A Management Study of International Venture Capital Firms.* Switzerland: Haupt Press. P.2-36.

Martin Kenney, Kyonghee Han and Shoko Tanaka. (2002). *The Globalization of Venture Capital: The Cases of Taiwan and Japan.* Paper presented at the International Conference on Financial systems, corporate investment in innovation and venture capital, Brussels, November 7-8, 2002.

Paul A. Gompers and Josh Lerner. (2001). *The Money of Invention: How Venture Capital Creates New Wealth.* Boston, Massachusetts: Harvard Business School Press.

Paul A. Gompers and Josh Lerner. (2001). The Venture Capital Revolution. *Journal of Economic Perspectives.* No.15(2). P.145-168.

William L. Megginson. (2004). Toward A Global Model of Venture Capital. *Journal of Applied Corporate Finance*. No.16. P.8-26.

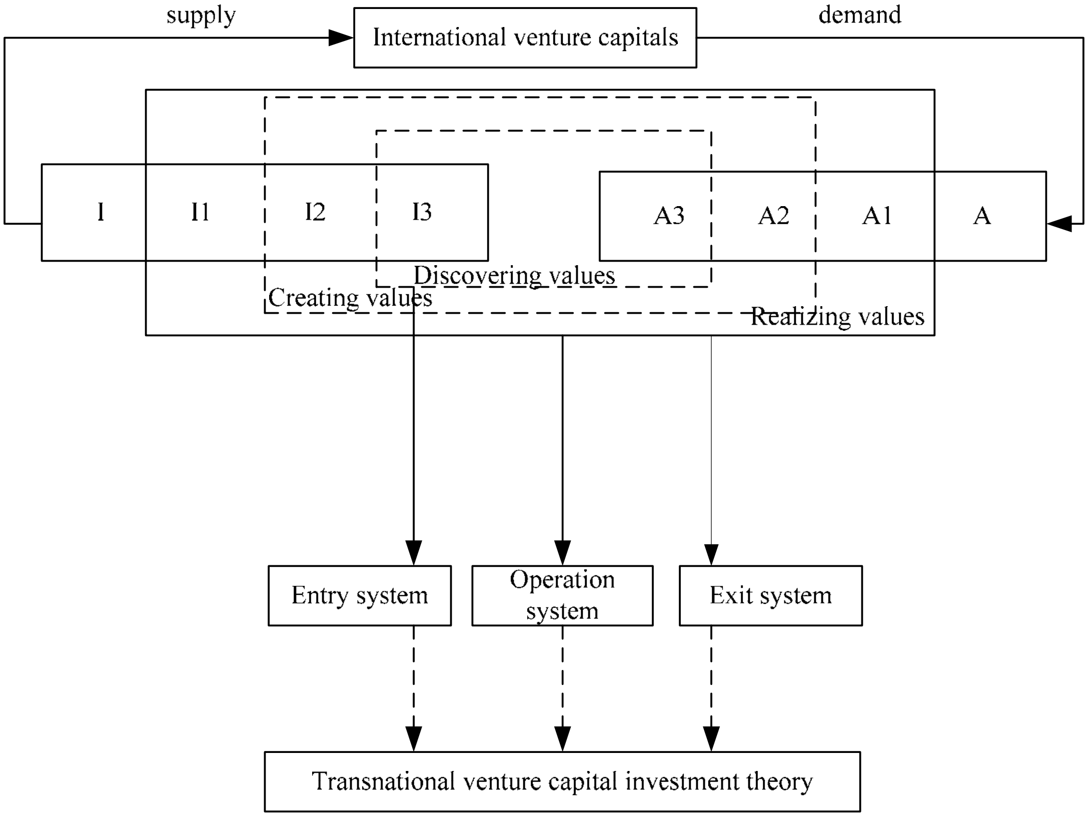
supply
International venture capitals
demand
I
I1
I2
I3
A3
A2
A1
A
Discovering values
Creating values
Realizing values
Entry system
Operation system
Exit system
Transnational venture capital investment theory

Figure 1. Theoretical Framework of the Transnational Venture Capital Investment Research

Note: I: Transnational investment system, I1: Investment exit, I2: Investment operation, I3: Investment motivation, A: Transnational attraction system, A1: Attraction extension, A2: Attraction operation, A3: Attraction motivation.

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Globalization of the Malaysian Manufacturing Sector:

An Agenda for Social Protection Enhancement

Evelyn S. Devadason (corresponding author)
Faculty of Economics & Administration
University of Malaya, 50603 Kuala Lumpur, Malaysia
Tel: 60-3-7967-3726 E-mail: evelyns@um.edu.my

Chan Wai Meng
Faculty of Business & Accountancy
University of Malaya, 50603 Kuala Lumpur, Malaysia
Tel: 60-3-7967-3890 E-mail: chanwm@um.edu.my

Abstract

Globalization offers opportunities for workers to achieve higher levels of economic prosperity. Unfortunately, it also causes job insecurity for specific groups in certain sectors. This is particularly so in the Malaysian manufacturing sector which is actively fragmenting its production structure. There are prior concerns that some workers in certain labour-intensive segments of the production process will be displaced. The issue is whether the country's current social protection systems are adequate to protect the disadvantaged. The paper first addresses the effects of one perspective of fragmentation on relative labour demand in Malaysian manufacturing. Subsequently, the relevant legislations and schemes enacted and implemented to protect workers are examined. The key conclusions of the study are first, unskilled workers are more likely to lose out from the international fragmentation of imports; and second, the existing social protection systems are inadequate to meet the needs of the unskilled workers.

Keywords: International production fragmentation, Imports of parts and components, Labour demand, Social protection

1. Introduction

Trade developments in the East Asian region in particular, point to a rapid expansion in international production fragmentation (Note 1) (Ng and Yeats, 2003; Athukorala and Yamashita, 2005) throughout the 1990s (Kimura and Ando, 2003). This translates into the rising importance of trade in parts and components. In fact, trade in *parts and components* ('middle products,' 'intermediates' or 'fragments of final products') has grown at a faster pace than trade in final manufactured goods (Athukorala and Yamashita, 2005; Jones and *et al*., 2005).

Imports of parts and components are essential to Malaysia as she is considered largely as an assembler. In 2003, parts and components accounted for 56 per cent of total Malaysian imports. Component trade is found to be heavily concentrated in electronics and electrical industries, whereby semiconductors and other electronic components (SITC 776) accounted for 68 per cent of parts and component imports (Athukorala and Yamashita, 2005). Figure 1 caricatures the development in fragmentation for selected industries in Malaysia. The share of parts and components in total imports had declined marginally from 37 per cent in 1983 to 34 per cent in 2000. However, parts and components recorded a higher average annual growth rate of 22 per cent as opposed to 19 per cent for total imports for the entire period. The higher import growth of parts and components vis-à-vis that for total imports signify the importance of fragmented imports in the Malaysian manufacturing sector.

The international procurement of parts and components is believed to have a greater impact on the labour market (Note 2) than trade in final goods (see Sakurai and Moriizumi 2000; Egger and Egger 2003). Most countries experienced a shift in demand towards skilled labour. The Malaysian manufacturing sector is no exception in this respect. Figure 2 presents skill differentials (the ratio of skilled to unskilled labour, S/U) and skill intensity (the ratio of skilled labour to total employment, SIE) for selected industries.

The 1980s saw a marginal decline in skill differentials (from 17 per cent in 1983 to 16 per cent in 1990) in manufacturing. The 1990s however was characterized by a reversal in the relative quantity decline of the 1980s.

Explanations based on the increases in the relative supplies of unskilled labour are compatible with the skill differential trends observed in the 1980s since unskilled labour increased by 9 per cent per annum while skilled labour increased merely by 8 per cent per annum. In the 1990s, the reverse occurred with a higher growth rate of 10 per cent per annum for skilled labour vis-à-vis 6 per cent per annum for unskilled labour. The rise in skill intensity in the 1990s was thus considered demand-driven relative to that of the 1980s.

Malaysia is thus an interesting case to analyze given the increase in relative labour demand (share of skilled to unskilled labour) and the evidence of greater component trade. The study combines both features of the manufacturing sector, fragmented imports and labour demand, to examine the impact of the former on the latter. The study approaches one perspective of fragmentation (focusing on the import side) by estimating the effects on relative demand for labour. Further, given that there will always be winners and losers as a result of globalization of the manufacturing sector, it is important to ensure that there are policies in place to minimize the negative impact on those who are most badly affected. This may include adequate social safety nets. The crucial questions are: What safety nets exist to support workers who are displaced or retrenched? Are the existing legislations and schemes adequate?

This paper is structured in the following manner. Section 2 discusses the data employed for the study. Section 3 provides the background for the empirical work, details the econometric analysis and presents the results. Section 4 examines the relevant safety nets pertaining to retrenchment and termination benefits and discusses the adequacy of these systems. Finally, Section 5 concludes with some recommendations to enhance social protection in Malaysia.

2. Data

The labour data is drawn from manufacturing surveys conducted annually by the Department of Statistics (DOS) Malaysia. The study only considers full-time paid employees (N), which excludes working proprietors, active business partners, unpaid family workers and part-time paid employees. Similarly, only the wages and salaries of full-time employees are considered for the study. The wage variable (W) refers to the average yearly earnings per full-time employee in each industry. All wage variables are deflated by the Malaysian consumer price index (at constant 1980 prices).

The definition of skills used for the study is solely based on occupational groupings governed by the availability of data from the manufacturing surveys. Skilled workers (S) refer to the number of employees in the managerial, professional, technical and supervisory categories. Unskilled workers (U) comprise production/operative workers. The real average wages for skilled and unskilled workers (SW and USW respectively) are constructed based on their average yearly earnings. Other industry measures employed comprise real value-added (Q), the share of foreign direct investment in total capital investment (FDI/CI) and the share of foreign workers in total employment (FW/N).

The data on exports (X) and imports (M) are derived from the *Malaysia: External Trade Statistics* publications. The data is compiled for industries at the 3-digit Standard International Trade Classification (SITC) level. Exports are valued f.o.b. while imports c.i.f. Total manufacturing exports and imports are deflated with the export price and import price index (1980 =100) for the entire economy respectively.

Since imports (Note 3) based on the SITC scheme do not separate fragmentation trade from final goods, the study adopts Ng and Yeats (1999) classification of intermediate goods inferred from trade statistics for a total of six industries (furniture and fixtures, machinery manufacturing, electrical and electronic products, transport equipment, scientific and measuring equipment and miscellaneous goods) in sections SITC 7 and 8 (industries in which fragmentation is relevant). For the study, only items termed as "parts and accessories" at the 4-5 digit level SITC are counted as fragmented products while others are treated as final goods. Thus the fragmentation intensity variable (F) in this study is defined as the share of component imports in total imports.

Integrating trade, labour market and industrial statistics, the empirical analysis involves a small panel data set of 6 major industrial groups, spanning the period 1983 to 2000. The data is a balanced panel of 108 observations.

3. Methodology and Results

3.1 The Theory

The issue of trade in goods that belongs to the same sector, intra-industry trade or IIT (Grubel and Lloyd, 1975), has recently received much attention. The assumptions of the models explaining IIT are that consumers love variety and there are increasing returns to scale in the production of the differentiated good. The original impression seemed to be that IIT does not affect the relative demand for skilled labour. However recent contributions to trade literature have showed that this type of trade can lead to increased differentials within sectors. Assuming skilled workers determine the quality of final goods produced and that the opportunities for greater trade rests with industries that are basically producing high quality products (differentiated in a vertical and horizontal way), the demand for high skilled labour would increase much faster (see Manasse and Turrini, 1999).

Duranton (1999) however explains labour demand effects based on trade in *intermediate goods* (fragmentation) instead of final goods. Final good producers who desire for advanced production technology will resort to trade to acquire the *high quality intermediates* abroad due to the scarcity of skilled labour locally. Based on this argument, the attraction of fragmentation at an international level may be reflected in different requirements for labour skills thereby causing a new international division of labour. One country may contain labour skills more appropriate to one fragment and another labour-abundant country may be relatively more productive in the other fragment.

Theoretically it is conceived that unskilled labour intensive stages of production are shifted to unskilled labour abundant countries while more technologically advanced stages remain in skilled abundant countries. The argument is that fragmentation will lower the demand for unskilled labour in developed countries (see Hijzen *et al.* 2004), leading to a fall in the employment of unskilled. However Feenstra (1998) points out that the reduction in demand for unskilled labour is also possible for developing countries if the fragmented activity received is relatively more skilled intensive than the home country. Thus skill differentials may increase in both developed and developing countries.

However, Geishecker and Gorg (2005) point out that the consequences of fragmentation for local labour markets are not clear-cut, particularly when there is labour mobility between industries. Fragmentation in one industry may affect labour in other industries as workers move from the affected industries. Thus the effects of fragmentation may not just be confined to changing demand between industries but relative demand within industries (see also Feenstra, 1998 and Hijzen *et al*., 2004). Geishecker and Gorg (2005) also point out that even without labour mobility, fragmentation effects on labour depend, among other things, on whether it complements or substitutes which type of labour.

A few recent papers that examined the impact of fragmentation on relative demand for skilled labour find significant results attesting to the importance of trade of such a nature for the latter. Most of these studies are based on European countries, such as the United Kingdom (Anderton and Brenton, 1999; Hijzen *et al.,* 2004), Austria (Egger and Egger 2001; 2003), Italy (Helg and Tajoli, 2004) and Germany (Geishecker and Gorg, 2005).

The above discussion ultimately rests on the fact that industries that depend on imported parts and components are likely to incur some adjustments in the relative labour demand *within* industries. Such a relationship is put to test for the Malaysian case.

3.2 Econometric Specification

The translog function is commonly used in the literature and is considered appealing in that it provides a second order approximation to any cost function and it does not impose any restrictions on the substitutability of imports. The variable cost function in translog form that assumes capital to be a fixed factor of production is as follows:

$$\ln C_i = \alpha_0 + \alpha_q \ln Q_i + \tfrac{1}{2}\alpha_{qq}\ln(Q_i)^2 + \beta_k \ln K_i + \tfrac{1}{2}\beta_{kk}\ln(K_i)^2 + \Sigma_j\gamma_j \ln W_{ij} + \tfrac{1}{2}\Sigma_j\Sigma_k\gamma_{jk}\ln W_{ij}\ln W_{ik} + \Sigma_j\delta Q_j \ln Q_i \ln W_{ij} +$$
$$\Sigma_j\delta_{qj}K_j \ln K_i \ln W_{ij} + \Sigma_j\delta_{Kj}\ln K_i \ln W_{ij} + \rho \ln Q_i \ln K_i + \lambda_T T_i + \tfrac{1}{2}\lambda_{TT}(T_i)^2 + \lambda_{QT}T_i \ln Q_i + \lambda_{KT}T_i \ln K_i + \Sigma_j\varphi_{iWj}T_i \ln W_{ij}$$

where

C_i = variable costs in industry i

Q_i = output in industry i

K_i = capital stock in industry i

W_{ij} = price of variable factor j

T_i = technology in industry i

Cost minimization of the above generates the following linear equations for the factor shares (L):

$$L_{ij} = \alpha_j + \delta_{Qj}\ln Q_i + \delta_{Kj}\ln K_i + \Sigma_k\gamma_{jk}\ln W_{ik} + \varphi_{iWj}T_i \qquad (1)$$

Differencing the above generates:

$$\delta L_{ij} = \varphi_{iWj}dT_i + \delta_{Qj}d\ln Q_i + \delta_{Kj}d\ln K_i + \Sigma_k\gamma_{jk}\ln W_{ik} \qquad (2)$$

Assuming homogeneity of degree one in prices imposes:

$$\Sigma_k\gamma_{ik} = \Sigma_j\gamma_{ik} = \Sigma_j\delta_{iK} = \Sigma_j\delta_{Qj} = 0 \qquad (3)$$

this generates

$$dL_{ij} = \varphi_{iWj}dT_i + \delta_{Qj}d\ln Q_i + \delta_{Kj}d\ln K_i + \gamma d\ln(W_j/W_k) \qquad (4)$$

with two variable factors j and k.

Machin *et al.* (1996) and Anderton *et al.* (2001) define the two variable factors of production as skilled (S) and unskilled (U). The relative labour demand equation is examined with the inclusion of trade variables, which are exports (X) and imports (M). The study adopts the more conventional factor demand equation of estimating "relative" labour demand as the share of skilled to unskilled (hereafter referred to as skill differentials or S/U), since Machin *et al.* (1996) do acknowledge that the theoretical foundation for estimating relative labour demand regression using the share of skilled to total employment as a dependent variable is weak. Following the earlier discussion on fragmentation, total imports (M) are disaggregated into imports of parts and components (Mpc) and other final imports (Mo). Mpc is also used interchangeably with the fragmentation intensity variable (F).

In addition to trade variables, foreign labour (FW) and foreign direct investment (FDI) are also introduced in the equation, due to the importance of foreign presence and foreign participation in the Malaysian manufacturing respectively. Foreign labour is distinguished by skills to capture the differential impact on inequality. Thus foreign migrant components are entered into equations as the share of skilled foreign workers in total employment (FWs/N) and the share of unskilled foreign workers in total employment (FWu/N), while FDI is entered into the inequality equations as the share in total capital investment (FDI/CI).

The relative labour demand (skill differential) equations that are estimated are as follows:

$$\ln(S/U)_{it} = \Omega + \varphi_1 \ln(SW/USW)_{it} + \varphi_2 \ln Q_{it} + \lambda K_{it} + \mu_1 \ln X_{it} + \mu_2 \ln M_{it} + \mu_3 (FDI/CI)_{it} + \mu_4 (FWs/N)_{it} + \mu_5 (FWu/N)_{it} + \nu_{it} \quad (5a)$$

$$\ln(S/U)_{it} = \Omega + \varphi_1 \ln(SW/USW)_{it} + \varphi_2 \ln Q_{it} + \lambda K_{it} + \mu_1 \ln X_{it} + \mu_2 \ln Mo_{it} + \mu_3 \ln Mpc_{it} + \mu_4 (FDI/CI)_{it} + \mu_5 (FWs/N)_{it} + \mu_6 (FWu/N)_{it} + \nu_{it} \quad (5b)$$

$$\ln(S/U)_{it} = \Omega + \varphi_1 \ln(SW/USW)_{it} + \varphi_2 \ln Q_{it} + \lambda K_{it} + \mu_1 \ln X_{it} + \mu_2 F_{it} + \mu_3 (FDI/CI)_{it} + \mu_4 (FWs/N)_{it} + \mu_5 (FWu/N)_{it} + \nu_{it} \quad (5c)$$

where

Ω = constant

S/U = ratio of skilled to unskilled labour

SW/USW = real relative wages (average skilled wages relative to average unskilled wages)

Q = output measured as real value-added

K = capital intensity measured as total fixed assets per unit of output (in per cent)

X = real exports

M = real imports

Mo = other imports, in real terms (final imports, excluding parts and components)

Mpc = real imports of parts and components

F = fragmentation intensity measured as share of imports of parts and components in total imports

FDI/CI = share of foreign direct investment in capital investment is a proxy for technology (in per cent)

FWs/N = share of skilled foreign workers in total employment (in per cent)

FWu/N = share of unskilled foreign workers in total employment (in per cent)

ν represent error terms that pick up random measurement errors in skill differential and the effects of labour demand shocks on skill differential, which are not picked up by the included independent variables.

3.3 Findings

The unit root panel test on the levels and first differences are investigated prior to estimating the relative labour demand equation. The panel unit root test of Im, Pesaran and Shin (IPS, 1997) is performed with the following conditions: (a) constant but no trend, and (b) one lag are assumed for all cases. Table 1 presents the results of the panel unit root tests in levels and first differences. The results confirm that the null of a unit root cannot be rejected for all variables in levels. The panel unit root tests are then further investigated for the variables in first differences. All variables are found to be of I (1) process, which is stationary in first differences. Thus, the variables in equations (5a) to 5(c) are first differenced.

The fixed effects (FE) and the random effects (RE) models are both employed and the choice of the model is determined by consistency properties. The Hausman (1978) specification test is performed to check if the coefficients estimated by the fixed-effects estimator and the same coefficients estimated by the random effects estimator do not differ statistically. Given that the Hausman test reveals that the coefficients of the FE and RE models do not differ statistically for all specifications [5(a) – 5(c)] only the FE estimates are reported in Table 2. Equation (5a) of Table 2 reports the skill differential specification including total imports, while equation (5b) presents the same specification but distinguishes total imports into imports of parts and components (Mpc) and imports of final goods (Mo). Equation (5c) introduces the fragmented intensity variable. All specifications in Table 2 are corrected for AR(1) disturbances.

In terms of the individual specific variables, statistically significant and negative estimates are found between relative wages and skill differentials, as expected. Similarly, output estimates are significantly negative. The negative sign on the output variable that controls for the scale of production simply means that the employment of unskilled labour increases more rapidly than the increase in skilled labour as output increases. Conversely the significant positive estimates on capital intensity signify complementarity effects between capital and skilled labour as factors of production.

In contrast to the significance of the above variables, FDI does not significantly affect relative labour demand. The evidence seems to suggest that FDI has not brought in skilled labour using technology. Mahadevan (2002) in her study on Malaysian manufacturing between 1981 and 1996 agrees that Malaysia has obtained better technology and equipment *via* FDI, but has undoubtedly failed to learn to use it adaptively. Conversely, another study on the Malaysian manufacturing industry by Oguchi *et al.* (2002) for the period 1992 to 1996 shows that FDI did not come with more efficient technology, as domestic firms were as efficient as foreign firms. The lack of sufficient absorptive capacity (or even efficient FDI-related technology) probably explains why FDI inflows are not translated into higher demand for skills.

Though, foreign participation in the form of FDI inflows do not matter, the presence of foreign labour has contributed to growing skill differentials. The quality of migrant labour is found to be important for skill differentials. Higher presence of skilled migrants significantly increases relative labour demand while unskilled migrants have an opposite impact, albeit insignificant.

Though both trade terms are insignificant in equation (5a), exports are deemed to have employment effects in favour of unskilled labour. However, prior evidence indicates a significant shakeout of unskilled labour relative to skilled labour owing to exports due to labour rationalization. As for imports, the negative impact on skill differentials implies slower growth in skilled vis-à-vis unskilled as imports increase (prior evidence indicates that imports into Malaysia increases both skilled and unskilled labour demand).

It is important to note that the signs of the estimates on other imports differ from imports of parts and components [see equation 5(b) of Table 2]. Though both Mo and Mpc remain insignificant, Mpc is positive while Mo remains negative. Since the variable of most interest in the study is the measure of fragmentation, the Mpc is substituted with the fragmentation-intensity variable (F) in equation (5c). F is found to be positive and significant in equation (5c).

Fragmentation on the import side appears to be skill-biased, thereby widening skill differentials in the Malaysian case. The views that render support to the widening impact of fragmentation on skill differentials in unskilled abundant countries center on the fact that "knife-edge comparative advantage" of fragmented trade necessitates skill upgrading of industries (Rajan, 2004). The factor bias towards skills inherent in imports of parts and components (Note 4) is of more relevance to growing skill differentials as opposed to imports of final goods. The group that is likely to be redundant is the unskilled, and this is already apparent with high retrenchments during the period 1996 to 2000 owing to the introduction of labour saving techniques (7th Malaysia Plan).

In total, though fragmented imports in manufacturing are significantly important in explaining the rise in skill differentials, it cannot explain all of the increase. The empirical results of the study should be taken as suggestive instead of conclusive. However, the plight of the unskilled can no longer be ignored with growing trade in parts and components that beckon a higher demand for skilled labour. The next section expounds the existing protection available for these workers in the event of retrenchments or job losses. This is a critical issue since the unskilled are basically lowly paid workers who encounter difficult job transitions in the labour market.

4. Safety Nets

4.1 Existing Legislations

The statutory provisions in Malaysia pertaining to the retrenchment of a worker are found in the Employment Act 1955 and its subsidiary legislation, the Employment (Termination and Lay-Off Benefits) Regulations 1980, the Labour Ordinance of Sabah, the Labour Ordinance of Sarawak and the Industrial Relations Act 1967.

The Employment Act 1955 prescribes the minimum benefits for a worker who comes within its scope, which includes a person who is employed to perform manual labour. Manual labour is defined as "duties which are purely physical in nature with very little or no mental effort" (Colgate Palmolive (M) Sdn Bhd v. Cheong Foo Weng and 12 others, [2002] 2 AMR 2107, at 2136). An unskilled worker is therefore covered by this Act and is entitled to the benefits prescribed by the Employment (Termination and Lay-Off Benefits) Regulations 1980, if he is retrenched. The entitlement of the retrenched worker depends on his tenure of service and wages. The tenure of service must be continuous for a period of at least 12 months whilst what amounts to wages (Note 5) is prescribed in the Employment Act 1955.

There are three key issues regarding the Act for the protection of the unskilled. First, it does not apply to all who are employed in Malaysia. The Employment Act 1955 applies to West Malaysia and the Federal Territory of Labuan only.

Thus, it does not protect an unskilled worker in Sabah or Sarawak. Second, a worker who was provided with housing accommodation loses not only his job and source of income, but also a roof over his head when he is retrenched. As his retrenchment benefits are calculated based on his basic wages which does not include the value of accommodation, he will not be compensated for the loss of accommodation. He may be rendered homeless upon his retrenchment. He thus suffers a double blow of being both jobless and homeless. Third, the minimum benefit which is prescribed in Regulation 6 of the Employment (Termination and Lay-Off Benefit) Regulations 1980 is as follows: (a) ten days' wages for every year of employment if he has been employed by that employer for a period between 12 months to less than two years; (b) fifteen days' wages for every year of employment if he has been employed by the employer for two years or more but less than 5 years; and (c) twenty days' wages for every year of employment if he has been employed for a period of five years or more. The amount of compensation as prescribed by the regulation is pittance particularly where the worker was in the employer's employment for less than two years. The corresponding entitlement of less than one month's wages may not be sufficient to sustain the worker even for a short period. It is unfortunate that an important social security legislation such as the Employment Act 1955 is limited in its scope of application.

As discussed in Section 3.3, the group that is likely to be made redundant due to fragmentation of the manufacturing sector is the unskilled. Unskilled workers comprise both local and foreign workers in the Malaysian case. It is noteworthy to mention here that an unskilled worker who is a Malaysian enjoys better protection under this Act when there is redundancy for the following reason. The Act provides that where there is a redundancy, the employer may terminate the employment of a local worker only after it has first terminated the employment of all foreign workers employed in a similar capacity as that of the local worker.

Currently, Sabah and Sarawak have their own respective Labour Ordinances. The Ordinances were amended in October 2005. Though the Ordinances cover unskilled workers, these persons do not enjoy the protection conferred by the Employment (Termination and Lay-Off Benefits) Regulations 1980. However, the provisions in the Ordinances confer better rights on a worker compared to that in the existing Employment Act 1955 in one aspect, namely the employer has to pay for the costs of the repatriation of the retrenched worker and his dependants to their place of origin. Neither the Employment Act 1955 nor the Employment (Termination and Lay-Off Benefits) Regulations 1980 confers this benefit on an unskilled worker in West Malaysia or the Federal Territory of Labuan.

Finally, the Industrial Relations Act 1967 plays a significant role in that it provides for the workers' trade union to negotiate a better package for the retrenched workers. Unlike the Employment Act 1955, the Industrial Relations Act 1967 applies throughout Malaysia. Under the Industrial Relations Act 1967, the trade union is prohibited from raising the issue pertaining to the terms of retrenchment benefits when the union is negotiating for a collective agreement with the employer. However, the Act does not prohibit an employer who finds retrenchment of its workers inevitable, from inviting the trade union to discuss the terms of the retrenchment scheme. Once the employer and the trade union have agreed on the terms, they may enter into a collective agreement and deposit it with the Industrial Court. The terms cannot be less favourable than those prescribed in the Employment Act 1955 and the Employment (Termination and Lay-Off Benefits) Regulations 1980.

Once the Industrial Court takes cognizance of the agreement, the agreement is binding on the employer and all workers to which the agreement relates. If the employer fails to pay the termination benefits as agreed to a retrenched worker, the worker may lodge a complaint with the Industrial Court. The Act provides that the employer shall be guilty of an offence if it fails to comply with the order.

4.2 Social Security and Statutory Training Schemes

The current statutory retrenchment schemes discussed above are lacking in many aspects. In view thereof, it is important to study the relevance of the social security schemes that are available to the unskilled. The schemes include the compulsory savings scheme under the Employees Provident Fund Act 1991 and the statutory training scheme implemented by the Ministry of Human Resources, namely, the Human Resources Development Fund (HRDF).

Apart from relying on his retrenchment benefits, an unskilled worker may also have to rely on his savings to sustain him during the period of his unemployment. In this connection, it may be pertinent to review the compulsory savings scheme established under the Employees Provident Fund Act 1991, the Employees Provident Fund (EPF). It must be emphasized however that this scheme is not available to all persons who are employed. Under the EPF scheme, at least 11 per cent of a worker's wages will be deducted and paid into his account with the EPF, whereas his employer is required to contribute at least a sum which is not less than 12 per cent of the worker's wages into the worker's account with the EPF. It is to be noted that the definition of the term "wages" (Note 6) under the Employees Provident Fund Act 1991 is different from that given in the Employment Act 1955. It is wider and encompasses even the worker's bonus, commission and allowance.

It is noteworthy that the Employees Provident Fund Act 1991 has a few safeguards to protect a worker. For example, the employer is prohibited from deducting or otherwise recovering its contribution from the worker. Further, where the employer is a company, its directors will be jointly and severally liable with the employer for any unpaid contribution. Furthermore, the worker cannot assign or transfer the amount in his EPF account. In addition, the amount cannot be subject to any attachment or levies. And in the event the worker becomes a bankrupt, the Official Assignee is not entitled to nor have any claim on the moneys in the worker's account with the EPF. All these are to uphold the spirit and purpose of the EPF scheme which is best reflected in the EPF's tagline "Savings for Old Age".

There are also several issues that prevail regarding the EPF scheme. First, the multipurpose withdrawals that are allowed for housing, education and medical purposes may render the retirement benefits insufficient. Second, the current withdrawal scheme does not allow a worker to make any withdrawal in the event he is retrenched. Third, the retrenched unskilled worker who becomes a self-paid employee (private worker) will not be covered by any social security retirement programme. His account with the EPF remains but there will be no further contributions with his new status as a self-paid employee.

The second issue is crucial as many of the unskilled workers may face difficulties to be re-employed in the formal sector. This is despite the existing training schemes that are offered to them such as the HRDF established pursuant to the Pembangunan Sumber Manusia Berhad Act 2001. One of the many purposes of the Fund is to undertake activities or projects to train those retrenched. As the termination benefits, if any, may not be able to sustain the worker for the period of training, it is important to review the withdrawal scheme under the EPF scheme. Further, the retrenched worker, armed with new skills but without a permanent job, may need funds to start a small business. The EPF scheme should support his needs to earn a honest living.

5. Conclusion

The results produce important implications concerning the increase in fragmented imports for the domestic labour market. First, the concerns that imports displace skilled labour for an unskilled abundant country like Malaysia may be misplaced, given the role of imports of parts and components. Instead, the study points out that unskilled labour stands to lose out from increasing fragmentation in imports. With increasing trade in parts and components following further globalization of the Malaysian manufacturing sector, the unskilled may witness more job losses. Second, the retrenched unskilled workers may end up in dire straits given the inadequacies in the current protection schemes. The plight of these workers becomes even more critical as there is no welfare programme for the unemployed in Malaysia.

The focus therefore should be on the coverage and adequacy of the current support mechanisms. The existing social protections systems need to be progressively reviewed to ensure that its provisions are relevant to the prevailing needs of the labour market. First, some reforms are needed to the current statutory provisions and social security schemes to enhance the protection for the unskilled, whose contract of employment is terminated prematurely due to no fault of his. The application of the Employment (Termination and Lay-Off Benefits) Regulations 1980 should not exclude anyone who is gainfully employed. It is proposed that the Minister makes comprehensive regulations pertaining to the termination benefits of a redundant worker. Further, the formula on the termination benefits should be reviewed. The legislature should take cognizance of one of the purposes of the termination benefits, which is to compensate the retrenched worker for his hardship due to the loss of employment and possibly, the loss of his home.

Second, the EPF should cover all Malaysians who are gainfully employed. The scheme should not discriminate against ex-workers who become self-paid employees. As for training needs, there should be a different scheme for retrenched workers. In order not to overburden an employer who is currently contributing a percentage of his employees' monthly wages to the HRDF with an additional levy, it is proposed that a part of the current levy be channeled towards the training of retrenched workers. More effort should be given to retrain a retrenched worker. The new scheme should adopt the maxim "if you give a man a fish, he will eat for a day. If you teach a man *how* to fish, he will eat for the rest of his life".

Third, the Ministry of Human Resources should consider the proposal made by the Malaysian Trade Union Congress (MTUC) to establish a special fund, with a joint contribution from employer and workers, for the latter to receive a certain amount of monies for certain duration upon retrenchment. In this respect, the recent proposal to set up a National Retrenchment Fund to tide over workers during their period of unemployment is considered a positive step to enhance social protection in Malaysia. The proposal is currently under a 6-month review by the Malaysian government.

References

Anderton, B. & Brenton, P. (1999). Outsourcing and low skilled workers in the UK. *Bulletin of Economic Research*, 51(4), 267-285.

Anderton, B., Brenton, P. & Oscarsson, E. (2001). What's trade got to do with it? Relative demand for skills within Swedish manufacturing. Working Document No.162, Centre for European Policy Studies, Stockholm University, Sweden.

Ando, M. & Kimura, F. (2003). The formation of international production and distribution networks in East Asia. NBER Working Paper 10167*,* National Bureau of Economic Research, US.

Arndt, S.W. (2001). Globalization of production and the value-added chain. *North American Journal of Economics and Finance*, 12, 217-218.

Arndt, S.W. (2003). Global production networks and regional integration. Working Paper 2003-12, Claremont McKenna College, California, US.

Arndt, S.W. (2004). Trade integration and production networks in Asia: The role of China. Mimeo, Claremont McKenna College, California, US.

Athukorala, P. (2003). Production fragmentation and trade patterns in East Asia. Working Paper No. 2003/21, Research School of Pacific and Asian Studies, Australian National University, Australia.

Athukorala, P. & Yamashita, N. (2005). Production fragmentation and trade integration: East Asia in a global context. Paper presented at the Workshop on Regional Integration with Particular Reference to Asia, Claremont, US, 25 February.

Charles, A. (2004). *A-Z Guide to Employment Practice in Malaysia*, CCH, Kuala Lumpur.

Devadason, E. (2005). International production fragmentation and the implications for relative labour demand: The Malaysian case. Paper presented at the 2nd Global Labour Forum, New Delhi, India, 13-14 Dec.

Duranton, G. (1999). Trade, wage inequalities and disparities between countries: The technology connection. *Growth and Change*, 30, 455-478.

Egger, H. & Egger, P. (2001). Cross-border sourcing and outward processing in EU manufacturing. *North American Journal of Economics and Finance*, 12, 243-256.

Egger, H. & Egger, P. (2003). Outsourcing and skill-specific employment in a small economy: Austria after the fall of the iron curtain. *Oxford Economic Papers*, 55(4), 625-643.

Egger, P., Pfaffermayr, M. & Wolfmayr-Schnitzer, Y. (2001). The international fragmentation of Austrian manufacturing: The effects of outsourcing on productivity and wages. *North American Journal of Economics and Finance*, 12, 257-272.

Falk, M. & Koebel, B.M. (2002). Outsourcing, imports and labour demand. *Scandinavian Journal of Economics*, 104(4), 567-586.

Feenstra, R.C. (1998). Integration of trade and disintegration of production in the global economy. *Journal of Economic Perspectives*, 12(4), 31-50.

Feenstra, R.C. & Hanson, G.H. (1996). Globalization, outsourcing and wage inequality. *American Economic Review*, 86(2), 240-5.

Geishecker, I. & Gorg, H. (2004). Winners and losers: Fragmentation, trade and wages revisited. IZA Discussion Paper No.982, Institute for the Study of Labor (IZA), Germany.

Geishecker, I. & Gorg, H. (2005). Do unskilled workers always lose from fragmentation? *North American Journal of Economics and Finance*, 16, 81-92.

Grubel, H.G. & Lloyd, P.J. (1975). *Intra-Industry Trade: The Theory and Measurement of International Trade in Differentiated Products.* US: John Wiley & Sons.

Hausman, J. A. (1978). Specification tests in econometrics. Econometrica, 46(6), 1251-1271.

Helg, R. & Tajoli, L. (2004). Patterns of international fragmentation of production and implications for the labor markets. Discussion Paper No.503, Research Seminar in International Economics (RSIE), University of Michigan, US.

Hijzen, A., Gorg, H. & Hine, R.C. (2003). International fragmentation and relative wages in the UK. *IZA* Discussion Paper No.717, Institute for the Study of Labor (IZA), Germany.

Hijzen, A.., Gorg, H. & Hine, R.C. (2004). International outsourcing and the skill structure of labour demand in the United Kingdom. IZA Discussion Paper No.1249, Institute for the Study of Labor (IZA), Germany.

Hummels, D., Ishii, J. & Yi, K-M. (2001). The nature and growth of vertical specialization in world trade. *Journal of International Economics*, 54, 75-96.

Im, K.S., Pesaran, M. H. & Shin, Y. (1997). Testing for unit roots in heterogenous panels. Mimeo, University of Cambridge, UK.

Jones, R. & Kierzkowski, H. (2005). International fragmentation and the new economic geography. *North American Journal of Economics and Finance*, 16, 1-10.

Jones, R., Kierzkowski, H. & Lurong, C. (2005). What does evidence tell us about fragmentation and outsourcing? Paper presented at the UNU-WIDER Jubilee Conference WIDER Thinking Ahead: The Future of Development Economics, Geneva, 17-18 June.

Kaminski, B. & Ng, F. (2001). Trade and production fragmentation: Central European economies in European networks of production and marketing. Policy Research Working Paper 2611, World Bank, Washington DC.

Lall, S., Albaladejo, M. & Zhang, J. (2004). Mapping fragmentation: Electronics and automobiles in East Asia and Latin America. QEH Working Paper Series 115, Oxford University, UK.

Lovely, M.E. & Richardson, J.D. (1998). Trade flows and wage premiums: Does who or what matter? NBER Working Paper 6668, National Bureau of Economic Research, Cambridge.

Machin, S., Ryan, A. & Reenan, J.V. (1996). Technology and changes in skill structure: Evidence from an international panel of industries. Discussion Paper No.297, Centre for Economic Performance, London School of Economics & Political Science, UK.

Mahadevan, R. (2002). Is there a real TFP growth measure for Malaysia's manufacturing industries. *ASEAN Economic Bulletin*, 19(2), 178-190.

Maimunah, A. (2006). *Malaysian Industrial Relations and Employment Law*, (5th ed.) Kuala Lumpur: Mc-Graw-Hill.

Manasse, P. & Turrini, A. (1999). Trade, wages and superstars. Discussion Paper 2262, Centre for Economic Policy Research, UK.

Morrison-Paul, C.J. & Siegel, D. (2001). The impacts of technology, trade and outsourcing on employment and labour composition. *Scandinavian Journal of Economics*, 103(2), 241-264.

Ng, F. & Yeats, A. (1999). Production sharing in East Asia: Who does what for whom, and why? Mimeo, World Bank, Washington D.C.

Ng, F. & Yeats, A. (2003). Major trade trends in East Asia: What are their implications for regional cooperation and growth? World Bank Research Working Paper 3084, World Bank, Washington D.C.

Oguchi, N., Amdzah, N.A.M., Bakar, Z., Abidin, R.Z. & Shafii, M. (2002). Productivity of foreign and domestic firms in the Malaysian manufacturing industry. *Asian Economic Journal*, 16(3), 215-228.

Okamoto, Y. (2004). Agglomeration and international competitiveness: Can Malaysia's growth be sustainable? In A. Kuchiki & T. Masatsugu (Eds.), *Industrial Clusters in Asia: Analyses of their Competition and Cooperation*, (pp.255-275). Japan: IDE, JETRO.

Pathmanathan, N., Kanagasabai S.K. & Alagaratnam, S. (2003). *Law of Dismissal*, CCH: Singapore.

Rajan, R.S. (2004). FDI, trade and the internationalization of production in the Asia-Pacific region: Issues and policy conundrums. Mimeo, National University of Singapore, Singapore.

Robbins, D.J. (1996). HOS hit facts: Facts win – Evidence on trade and wages in developing countries. Development Discussion Paper No.557, Harvard Institute for International Development, Harvard University, US.

Sakurai, N. & Moriizumi, Y. (2000). Trade patterns and factor use: Evidence for Asia Pacific countries. Paper presented at the 13th International Conference on Input-Output Technique, 21-25 August, Macerata, Italy.

Tombazos, C.G. (1999). The role of imports in expanding the demand gap between skilled and unskilled labour in the US. *Applied Economics*, 31, 509-516.

Zebregs, H. (2004). Intraregional trade in emerging Asia. IMF Policy Discussion Paper 2004/1, International Monetary Fund, Washington D.C.

Notes

Note 1. Alternative terms have been given to reflect the same concept, such as segmentation, integrated production, outward processing (see Lall et al., 2004), intra-product specialization, super-specialization (see Helg and Tajoli, 2004), multi-stage production (Hummels et al., 2001), de-localization, disintegration (see Hijzen et al., 2003), production sharing, vertical specialization, slicing the value chain and outsourcing (see Feenstra and Hanson, 1996; Athukorala and Yamashita, 2005), kaleidoscope comparative advantage and intra-mediate trade (Rajan, 2004). The term 'fragmentation' refers to the splitting up of production processes into separate components that can be produced in different locations.

Note 2. Fragmentation differs importantly from import penetration in final goods in the sense that it explicitly takes into account the extent to which firms move production activities abroad. Moreover labour demand is not only affected in import-competing industries but in all industries that use foreign inputs.

Note 3. Imports as pointed out by Lovely and Richardson (1998) may take the following form: (a) finished goods; (b) import of parts and components.

Note 4. Robbins (1996) also finds a positive association between the relative demand for skilled labour and imports for capital goods for Malaysia.

Note 5. The term "wages" is defined in s.2(1) of the Employment Act 1955 to mean the basic wages and all other payments in cash payable to an employee for work done in respect of this contract of service but does not include: the value of any house accommodation or the supply of any food, fuel, light or water or medical attendance, or of any approval amenity or approved service; any contribution paid by the employer on his account to any pension fund, provident fund, superannuation scheme, retrenchment, termination, lay-off or retirement scheme, thrift scheme or any other fund or scheme established for the benefit or welfare of the employee; any traveling allowance or the value of any traveling concession; any sum payable to the employee to defray special expenses entailed on him by the nature of his employment; any gratuity payable on discharge or retirement; or any annual bonus or any part of any annual bonus.

Note 6. The definition is given in s.2 of the Employees Provident Fund Act 1991 to mean all remuneration in money, due to an employee under his contract of service or apprenticeship whether agreed to be paid monthly, weekly, daily or otherwise and includes any bonus, commission or allowance payable by the employer to the employee whether such bonus, commission or allowance is payable under his contract of service, apprenticeship or otherwise, but does not include - service charge; overtime payment; gratuity; retirement benefit; retrenchment, lay-off or termination benefits; any traveling allowance or the value of any traveling concession; or any other remuneration or payment as may be exempted by the Minister.

Table 1. Unit Root Panel Tests (Levels and First Differences)

	IPS Test			
	Levels		First Differences	
Variable	t-bar	Psi(t-bar)	t-bar	Psi(t-bar)
S/U	-1.474	0.107	-3.310	-4.595
SW/USW	-1.661	-0.389	-3.880	-6.054
Q	-1.903	-1.026	-3.961	-6.260
K (%)	-1.903	-1.026	-3.706	-5.608
FDI/CI (%)	-2.241**	-1.920**	-4.075	-6.554
FWs/N (%)	-2.853**	-3.533**	-4.449	-7.511
FWu/N (%)	-0.245	3.349	-2.707	-3.049
X	-1.681	-0.441	-3.347	-4.688
M	-1.554	-0.104	-3.505	-5.094
Mo	-1.566	-0.138	-3.836	-5.942
Mpc	-1.362	0.401	-3.574	-5.271
F (%)	-1.805	-0.739	-5.085	-9.140

Note:

1. The above tests assume a constant but without trend. One lag is assumed for all cases.
2. The values marked with an ** signify that the variables are stationary in levels.
3. All variables in levels are in logarithmic values except for those in (%).

Table 2. Relative Labour Demand Equations Across Manufacturing Industries, by Fixed Effects (with AR1 disturbances)

Variables	(5a)	(5b)	(5c)
dln(SW/USW)	-0.018**	-0.018**	-0.017**
	0.006	0.006	0.006
dln(Q)	-2.044**	-1.963**	-2.028**
	0.622	0.619	0.608
dK	0.104**	0.103**	0.103**
	0.031	0.031	0.031
d(FDI/CI)	-0.0002	0.001	0.001
	0.006	0.006	0.006
d(FWs/N)	0.799**	0.861**	0.829**
	0.409	0.397	0.39
d(FWu/N)	-0.130	-0.140	-0.144
	0.101	0.102	0.101
dln(X)	-0.009	-0.228	-0.335
	0.847	0.845	0.846
dln(M)	-0.499	-	-
	0.973		
dln(Mo)	-	-0.874	-
		0.745	
dln(Mpc)	-	0.303	-
		0.464	
dF	-	-	**0.028***
			0.016
cons	0.778**	0.827**	0.783**
	0.258	0.257	0.236
Hausman test	χ^2 (8) = 5.44 (Prob > χ^2 = 0.710)	χ^2 (8) = 2.79 (Prob > χ^2 = 0.972)	χ^2 (8) = 3.66 (Prob > χ^2 = 0.889)
R-sq	0.411	0.415	0.415
DW	1.736	1.689	1.648

Note: 1. The dependent variable is dln(S/U).

2. Total number of observations in 96.

3. DW refers to the Bhargava *et al.* (1982) Durbin-Watson statistics.

4. Figures below coefficient estimates are standard errors.

**significant at 5% and *significant at 10%

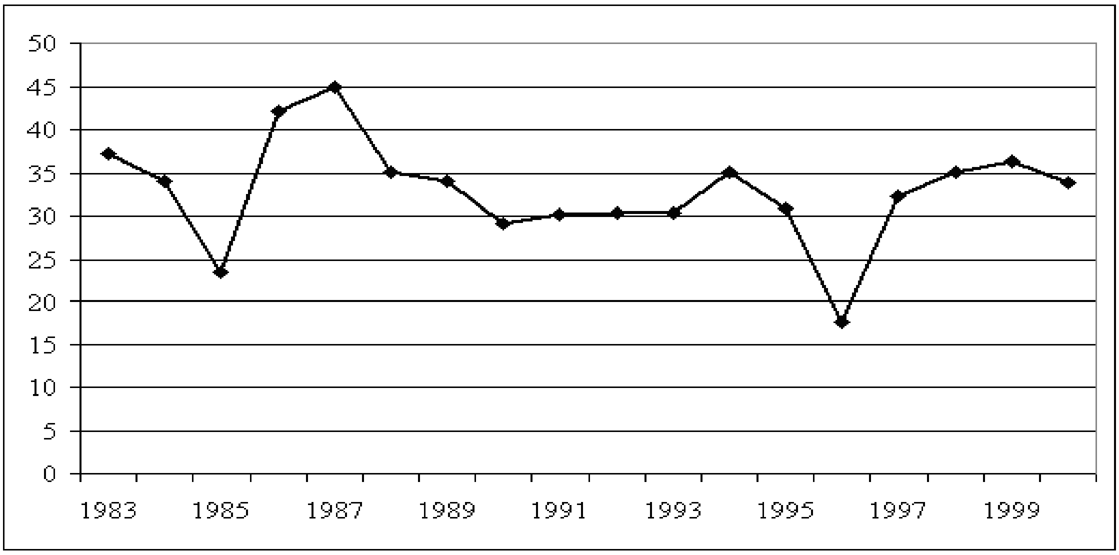

Figure 1. Share of Parts and Components in Total Imports (in per cent)

Source: Calculated from the Malaysia: External Trade Statistics, various issues.

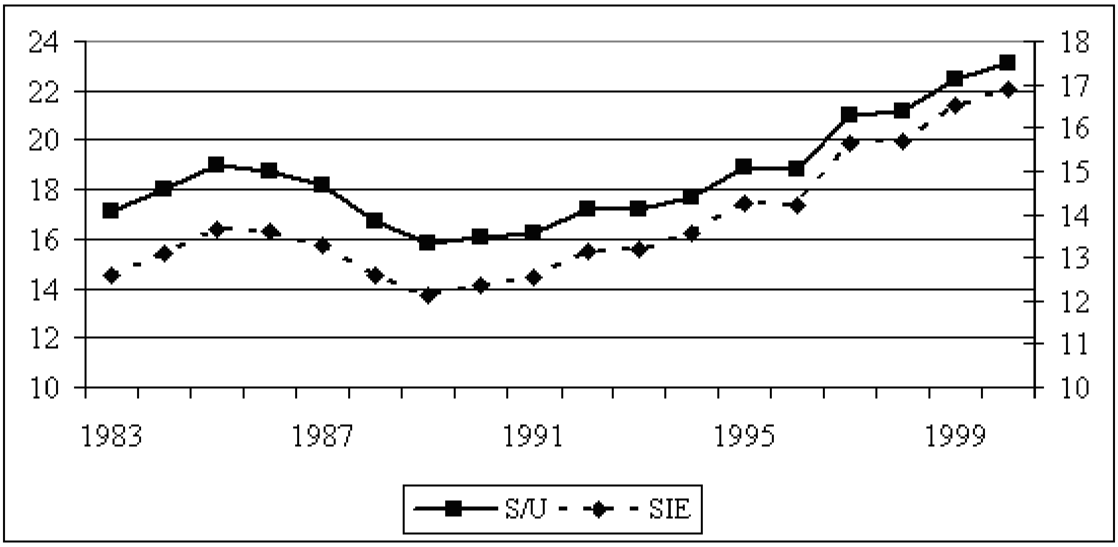

Figure 2. Skill Differentials and Skill Intensity in Manufacturing (in per cent)

Note: The left axis measures skill differentials (S/U) while the right axis measures skill intensity (SIE).

Source: Calculated from unpublished data obtained from the Department of Statistics, Malaysia.

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Study on the Attributes and Correlations of Production Factors

Zuohua Yue
College of Economic Management
Henan University of Science and Technology
Luoyang 471003, China
Tel: 86-379-6428-2038 E-mail: yzh5701@126.com

The research is financed by the"Henan Provincial Modern Agricultural Innovation System and Support System" of Henan Provincial Soft Science Project (No. 072400450470). (Sponsoring information)

Abstract

With the advancement of the society and the development of the economy, human cognitions about the production factors are more and more deepened. However, as viewed from the attribute structure of production factor, the production factor only contains two attributes such as hard factor and soft factor, and the soft factor is composed by knowledge soft factor and society soft factor, which form three capital forms, i.e. material capital, human capital and social capital, and these three sorts of capital drive the increase of the economy and the development of the society. In the background of modern economy, the human capital is in the center position and its advance could enhance the energy and efficiency of other factors.

Keywords: Production factors, Attributes, Correlation, Research

As the material base and drive dependence, the production factor has been concerned by the academe all along. The product factors mean various economic resources and conditions which are necessary for human material production, and they are necessary conditions for people to engage in production or economic activities. Without the existence of production factors, people can not engage in production or economic activities. From relative research histories, the production factors are dynamic, and in different economic development stages, human cognitions about production factors are developing and changing, and with the advance of social civilization and the enhancement of economic level and the change of the production pattern, the cognitions are continually fulfilled. So under the present economic background, it possesses theoretical and practical meanings to study the attributes and correlations of production factors fro the drive dependence of Chinese economic development.

1. Human cognitions about the development of production factors and the researches in recent years

According to traditional western economics, since Say's Law, the production factors were divided into three basic forms including land, capital and labor, which occurred in the classic book, "Economic", by Samuelson (Samuelson, 1999, P.9). With the advance of human civilization, especially after WW II, the world economy entered into the stage of quick development, and in the middle of 1960s, Solow R., Swan T., Denison E. and other western economists found the "increasing residual" in the research about the economic growth mechanism, and as an independent production faction, the technology has gradually been realized by people.

The function of system to the economic growth firstly occurred in some economists' (such as Veblon, T.B) research and analysis about the system from the end of 1800s and the 1930s-1940s. But at that time, they only emphasized the importance of the system, and they didn't confirm the connotation of the system, and they only said "demand and desire, result and objective, method and measure, the change and tendency of individual behaviors all are the function of the variable of the system, but the system possesses highly complexity and total instability (Thorstein, 1993, P.166)". So the opinion didn't enter into the traditional economic theory and until the new system school where Coese R, Williamson O and North D were main representative economists who studied the problems including the property of the enterprise, the trade cost and asset specificity, the function of the system to the economic growth begun to be accepted by people. The occurrence of the system factor further developed the human view for studying the production factors. The research about the economic growth was not only limited in the investment of the production factors, and the economists also begun to study the environment of the production factors, especially they could more scientifically explain many macro-economic problems such as the difference of economic development in different countries where labor, capital, land and human resource were basically same. Just as North said, "the system offers a sort of economic stimulation structure, and with the evolvement of the structure, it plans the direction of the economy toward growth, stagnancy or decay (North, 1991, P.235)". Thus, the system had been normally brought into the study of production factors.

As one of production factors, the human capital also occurred in certain stage of economic development through continual cognition deepening. In the beginning of 1900s, in Marshall's Politics Principle, he regarded the entrepreneurships as the production factor. And he thought that with the development of productivity and the extension of production scale, the entrepreneurships became into the new production factor. In 1960s, Theodore W. Scholtz put forward the theory of human capital after he studied the economic developments of undeveloped countries such as India. And in 1980s, when Romer P. and Lucas R. studied the micro layers such as knowledge overflow and externality of human capital of economic development, the human capital had been emphasized as an independent production factor. At present, Marshall's "entrepreneurships" should belong to the range of human capital because the "entrepreneurships" is formed when the human capitals are accumulated in "entrepreneur", and it is one of representation form of human capitals. So as an independent production factor, the "entrepreneurships" is just one part of the human capital.

In recent years, about the research of production factor, some scholars put forward new opinion. Chinese Academy of Social Sciences researcher, Xue Yongying comprehensively studied the composing of production factors from various aspects such as factor, combination, operation and optimization of productivity in his book of "Productivity Economics" (Zhejiang People's Press, 1986), and he thought except for three basic factors, energy, information, education, management and scientific technology all should be regarded as production factors. After 1990, with the rise of information economy and network economy, domestic scholar Wu Jiapei deeply studied the relation between information and economy in his book of "Information and Economy (1993)", and he put ward the opinion that "information is the most important soft factor of productivity" (Wu, 1993, P.27-30). Yu Rengang though the complexity degree of the labor process decided the connotation of production factor, and in the simple labor, the production factor only included human and material, and with the continual development of social production, the production process would become more and more complex, so many new production factors would continually enter into the production process. In the modern large production condition, the basic production factor includes four sorts, i.e. natural resources (including land), labor (mainly including simple labor), capital and knowledge (mainly including management knowledge, technical knowledge and information). At the same time, Yu Rengang thought that in different economic terms, each production factor possessed different function in the combination of various factors, so the correlation existed among production factors, and this sort of relation was decided by the status and function of production factor in the production. The research about the connotation of the production factor doesn't stop, and with the advance of social civilization, new production factor will certainly occur (Xu, 2006, P.15-18).

About the theoretical research of production factor, Xu Bin et al put forward the concepts of generalized production factor and special production factor, and put forward the hierarchy theory of production factor based on the generalized concept, i.e. the generalized production factor system was composed by four hierarchies including basic production factor, unit production factor, industrial production factor and department production factor (Xu, 2006, P.15-18).

From the human research history and actuality about the production factor, we can see that there are not deep researches about the attributes and correlation of production factors, and the relative research can not only promote the theory of production factors, but can improve the practice for the dependence of production factors in modern economic development.

2. Attributes of production factors

About the research of production factors, when people study the dependence factors of economic growth from different domains of economics, they will continually find new production factor, which will be inconvenient for studying and cognizing the essential of production factor. First, when some scholars study the relationship between economic growth and production factors and find some factors influence the economic growth, they will regard them as the production factors, so the quantity of production factors will become more and more and it become very complex to establish the production function model through taking various factors as the independent variables and taking the output as the attributive variables, and it goes against to look for the key factor influencing the output level. Second, confusions and "repetition" will occur when people cognize the production factors from different views, for example, for "energy", some economists describe it as an independent production factor, and it should be contained in the factor of "land" in fact. So it is very necessary to describe the essential attribute of production factor by the clustering method through getting rid of the concrete form of production factors. It not only embodies the scientific research opinion of "induction and deduction", but possesses meanings to study the property of the production factor and the relationship between production factor and economic growth, and simplify the production function.

Based on above analysis, after rescanning the set of production factors, the author thinks that the production factors only can be divided into two sorts from attributes, i.e. hard factors and soft factors. The so-called hard factors mean the factors which can be materialized, and they have concrete or actual images, and they are objective, such as traditional land, capital and the labors with material form. In the process of economic growth, their function can be measured. The soft factors mean the factors which can not be materialized, and they are abstract and they have not concrete material forms, and they are not easy to be measured, and they are "created" by human through the accumulation of knowledge and the advance of socialization, and they are subjective on consciousness. The soft factors include knowledge soft factor and social soft factor, and the knowledge soft factors mean the human capitals such as technology, talent, knowledge and intelligence of people, and the social factors mean the social capitals such as system and regulation formed by the social power guaranteeing the economic activity order. Thus, we have the total cognition to the property, composing and the relation with economic development of production factors (seen in Figure 1).

Convenient for studying the relationship between production factors and economic growth, we analyze and classify the production factors, which can avoid the blur and superposition of the connotations among factors. At the same time, we also definitely confirm the natural labor and human capital. Through the attribute analysis of production factors, the connotation and the extension of the expression of production function would be further abundant and extensive. According to the Cobb-Douglas Production function, the production function containing soft factors and hard factors can be denoted as

$Q=f(H, S)$

Where, Q denotes the output, H denotes the investment of hard factors and S denotes the investment of soft factors.

Thus, we can understand the production function as the social total output which is the function of the hard factors and soft factors investments. If denoting the hard factors and soft factors by the capital forms in the economic activity process, the social total output is the function about material capital, human capital and social capital investments.

$Q=f(M, H, S)$

Where, Q denotes the output, M denotes the material capital, H denotes the human capital and S denotes the social capital.

The hard factors and soft factors are contrary uniform entity in the process of economic growth. And they are different and associated. As viewed from the form, the hard factors are concrete, and they have material forms and they can be measured. And the acquisition of hard factors is easy because most of them exist in the nature and society. The process of acquisition is simple only by the mode of trade. The acquisition of soft factors is complex and needs more costs, and they can be acquired through education, training and long-term "studying from working", and the process of the acquisition is long. Some soft factors also are related with human culture and social development level, so the acquisition of soft factors is much difficult than the acquisition of hard factors. Though both have large difference, but they are closed related in the economic activities. The soft factors can exert the function through the hard factors in the economic growth process, and the existence of hard factors is only the necessary condition of economic growth, and it is not the sufficient condition. If the hard factors can development the original economic activities in the original economic stage, but in the modern society, without the supports of soft factors, it is not feasible to develop the economic activities only depending on the hard factors. The hard factors and soft factors supplement each other.

3. Correlations among production factors

The production factor is the base to develop the economic activity, and different production factors play different roles in the economic activities, and they have different function mechanisms to economic growth and economic development in the process of economic activity. It can better allocate the production factors and promote the economic growth and development to study the relationship between different factors with economic growth and development.

As above analysis, according to the attributes, the production factors can be divided into two clusters, i.e. hard factors and soft factors, and the soft factors are composed by knowledge soft factors and social soft factors. In the process of human economic activity, the hard factors composed by land, labor and capital offer material capitals in the economic activities, and the knowledge soft factors composed by knowledge, technology and intelligence offer human capitals in the economic activities. The social soft factors composed by system, regulation and law offer social capitals in the economic activities. Material capital, human capital and social capital are necessary capital factor set for the economic growth and development in the economic activities. But their functions are different and the hierarchies exist among them.

Through the analysis of production factors in economic activities, we find that the hard factors composed by land, labor and capital are the direct participators in the economic activities, and they directly act on the objects of economic activities and complete the attention of economic activities, and their values are coagulated in the new material products in the economic activities to complete the value conversion and realize the economic growth. Because of the direct property that they participate in the economic activity, we call them as the direct production factors. For the knowledge soft factors composed by knowledge, technology and intelligence, they don't directly act on the objects of economic activity in the economic activity, and they influence the economic growth through enhancing the allocation benefits of land and capital and enhance the efficiency cost of labor. So we can see that human capital is the headspring of land, labor and capital to the economic growth, and the land, labor and capital can be enhanced their economic activity because of the existence of human capitals. Therefore, we call human resources as the source production factors. The social soft factors composed by system, law and regulations are the set of behavior rules between people and people and between human and materials, and their existence regulate the rules that the economic bodies should follow in the economic activities, and these rules not only make each economic body know its behavior rule in the economic activity, but make each economic body know other economic bodies, so the social honest sources are extended and the trade costs are reduced. In nature, the system is the guarantee that the economic activities run in high efficiency, so we call it as the guarantee factor. The direct production factor, source production factor and guarantee factor act on each other and perform their own responsibility in economic activities, and they make the economic activity go along orderly and complete the cycle, and make the economic activity develop from low level to high level and realize the continual growth and development of economy. The hierarchy and correlation among production factors in the economic activities are seen in Figure 2.

From Figure 2, in human economic activity, labor, capital and land directly participate in the economic activity and complete the conversion of value form, and they are the basic dependence of economic growth. The human capital doesn't directly act on the economic activity, but it is directly associated with labor, capital and land, and it connects with the economic activities through them, and it forms the relationships of condition guarantee and information feedback with the factor of system, and it is in the central position in all production factors. Through the functions to land, labor and capital, human capital can enhance the economic activity and increase the contribution degree in the economic growth. The correlation of human capital and system forms a positive feedback function, i.e. the function of human capital to the system will make the system more adapt to the accumulation and development of human capitals, and the function of system to human capital will make the human capital form larger productivity in the frame of system. The factor of system ensures the condition in the economic activity, and it acts on the economic activity to make labor, capital and land to be allocated optimally, and it also acts on the human capital to make human capital be released to the largest extent.

The study about the hierarchy and correlation of production factors can explain why the human capital is the core factor in modern economic development.

References

North, D. (1991). Institutions. *Journal of Economics Perspective.* No.1. quoting from a secondary source of Tan Chongtai's "New Development of Development Economics" of Wuhan University Press in 1999. P.235.

Samuelson. (1999). *Economics (16th edition).* Beijing: Huaxia Press. P.9.

Thorstein B Veblen quoting from a secondary source of Geoffrey M. Hodgson. (1993). *Pronouncement of Modern System Economics.* Beijing: Beijing University Press. P.166.

Wu, Jiapei. (1993). *Information and Economy.* Beijing: Tsinghua University Press. P.27-30.

Xu, Bin & Xu, Shoubo. (2006). Hierarchical Theory on Production Factors. *Journal of Beijing Jiaotong University (Social Sciences Edition).* No.4. P.15-18.

Yu, Rengang. (1999). Meaning, Function and Correlation of Production Factors. *Contemporary Economic Research.* No.5. P.30-32.

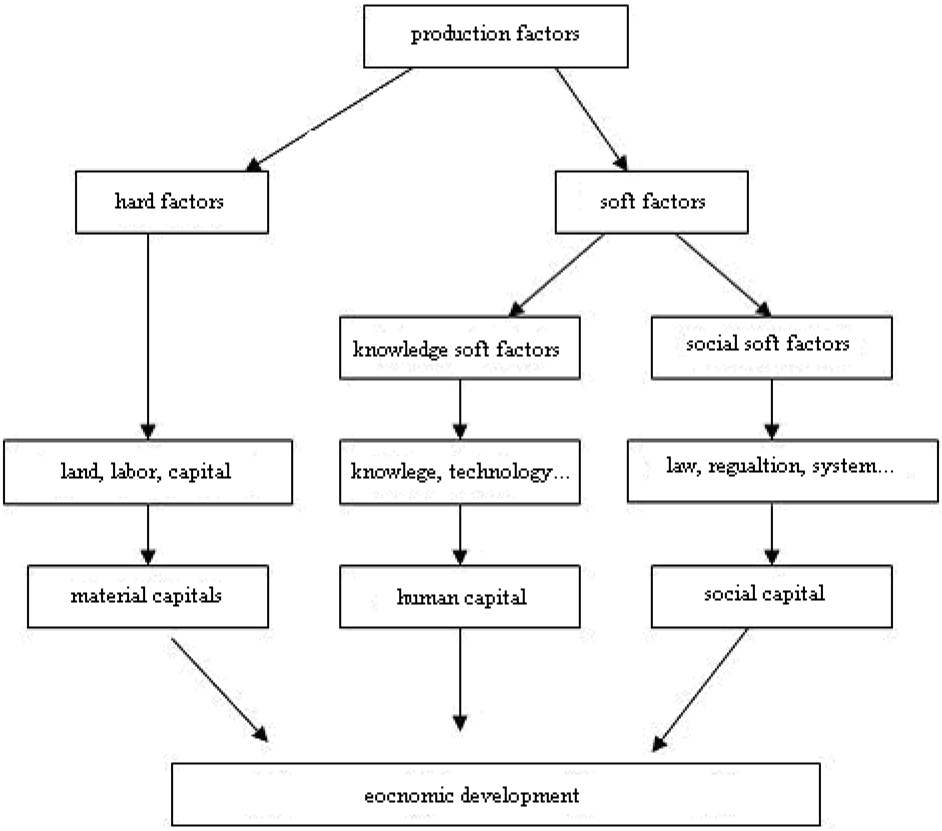
production factors
hard factors
soft factors
knowledge soft factors
social soft factors
land, labor, capital
knowlege, technology…
law, regualtion, system…
material capitals
human capital
social capital
eocnomic development

Figure 1. Attributes and Capital Forms of Production Factors

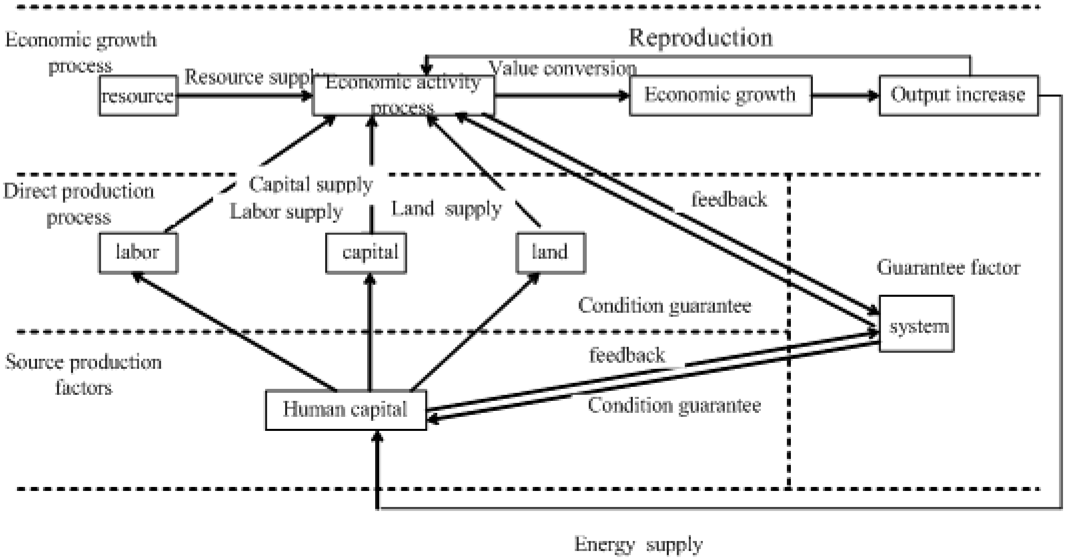
Reproduction
Economic growth process
resource
Resource supply
Economic activity process
Value conversion
Economic growth
Output increase
Direct production process
Capital supply
Labor supply
Land supply
feedback
labor
capital
land
Guarantee factor
Condition guarantee
system
Source production factors
feedback
Human capital
Condition guarantee
Energy supply

Figure 2. Hierarchy and Correlation of Production Factors in the Economic Growth Process

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

The Relationship between Human Resource Management and Firm Performance in Malaysia

May-Chiun Lo (Corresponding author)
Universiti Malaysia Sarawak, Faculty of Economics and Business, 94300 Sarawak, Malaysia
Tel: 60-82-582-360; Fax: 60 82 671 794
E-mail: mclo@feb.unimas.my

Abang Azlan Mohamad
Universiti Malaysia Sarawak
Faculty of Economics and Business, 94300 Sarawak, Malaysia
Tel: 60-82-582-459; Fax: 60 82 671 794
E-mail: maazlan@feb.unimas.my

Maw King La
Universiti Malaysia Sarawak
Faculty of Economics and Business, 94300 Sarawak, Malaysia

Abstract

Drawing upon resource based theory which explains human capital as the key resource for the firms' development, this paper hypothesized that the human resource management (HRM) was anchored on firm performance. HRM were conceptualized as 4- and unidimensional constructs, respectively. The main thrust of the paper is to assess the impact of HRM on firm performance. The research uses a sample of 85 firms surveyed in Sarawak, Malaysia. The findings suggested that incentives and information technology are positively related the firms' performance. Implications of the findings, potential limitations of the study, and directions for future research were further discussed.

Keywords: Firm performance, Training, Incentives, Performance appraisal, Information technology

1. Introduction

Human resource management faces challenges of bringing better fitted workers into the organizations and meeting the workers' needs and expectations. Thus, there is a compelling demand to develop better ideas, strategies to improve the interface between employees and employers, and to elaborate comprehensive insight that can help human resource managers get better results and improved performance (Vigoda & Cohen, 2003). The relationship between human resource management (HRM) and firm performance has received extensive considerable attention from researchers in recent years (e.g., Li, Zhao, & Liu, 2006; Sanchez, Jimenez, Carnicer, & Perez, 2007; Lin & Chen, 2007).

A vast amount of research has proved the positive relationship between HRM and firms' performances in terms of sales revenue, profitability, net asset ROI, and market share (Huselid, 1995; Schuler & Jackson, 1987; Hill & Rothaermel, 2003). However, limited researches on HRM have been observed to empirically examine its impact on the firms' performance. Hence, this study will add to the growing body of research by linking HRM to firms' performance and expands the domain of this relationship. Considering the potential cascading effect that employees' power could have on firms' performance, previous researches may have underestimated the impact of HRM on firms' performance.

HRM and firm performance have become the foundation of a new era of managing a diversified workforce in the advent of globalised world. Sarawak, a state located in East Malaysia, has gone through rapid transformation from primary sector such as agriculture and mining to a highly skilled manufacturing sector in the 1990s. Thus, understanding the relationship between HRM and firm performance would serve as a platform for companies' managers, to assist them in achieving better performance. Hence the main purpose of this study is to investigate whether certain components of HRM such as training, incentives, performance appraisal and information technology influence firms' performance. This study assumes that answers to such questions would have relevant implications for the business growth and further enhance a firms' market value.

2. Theoretical Background

2.1 Firms Performance

Performance can be viewed in many aspects and connotations depend on the application. Derek, Torrington and Laura (1995) attributed performance as bottom line profit, doing better than competitors, maximum organization effectiveness and achieving specific organization objectives. In fact, Laitinen (2002) defined performance as the ability of an object to produce results in a dimension determined a priori, in relations to a target. Recent study conducted by Cho and Punick (2005) have confirmed that top managers increasingly relate quality to firm performance and hence it has been viewed as one of the important key variables in achieving long-term competitive advantage. This is further supported by Yun and Good (2007) that loyalty is positively related to the company's profitability and the long-term growth. Indeed, renowned researcher posted there is something about the way that decisions were made in successful organizations that shows the seeds of eventual failure (Christiansen, 1997).

2.2 Employee Training and Firm Performance

HRM is known as the central business concern, as matters related to policies, practices, and systems that shape employees' behavior, attitudes, and performance will be referred to (Noe, Hollenbeck, Gerhart & Wright, 2000). One important role that HRM plays is training. Training is a planned and systematic effort to modify and develop knowledge, skill, competency and attitude toward learning experience to achieve effective performance through work productivity (Buckley & Caple, 2004). Past researchers such as Drummond (2000) revealed that training provides adequate criteria to an individual to perform better in a given task and subsequently contributes to the firm performance (Rothwell & Sullivan, 1994). However, Drucker (1995) commented that training is an expensive way of attempting to enhance human productivity.

On the other hand, Robert, Alan, Compton and McCarthy (1999) were of the opinion that effective training would not only equip employee with most of the knowledge and skills needed to accomplish jobs, it would also help to achieve overall organization objectives by contributing to the satisfaction and productivity of employee. Therefore, we state the following hypothesis:

H1: Employee training will be positively correlated to firms' performance.

2.3 Employee Motivation and Firm Performance

Theorist such as Harsanyi (1969) indicated that people's behavior can largely be categorized into two components: economic gain and social acceptance. These two dominant interests have actually created incentive for the firms' employees. Incentive pays are part of a complex arrangement to express and to maintain the working relationship between the employers and employee. It demonstrates not only what the management is trying to achieve but in circumstances which contributed to the overall firms' performance (Lupton & Bowey, 1975). Rajkumar (2004) posited that the concept of incentives practiced by organizations like stock options and bonuses particularly should help to reinforce employee productivity.

Research by Armstrong (2001) linked incentives to the achievement of previously set targets which are designed to motivate people to be more productive to achieve high level of firm performance. In this further supported by Ian, Jim and Haper (2004) who noted that incentives should be incorporated to organization strategies as seen as a technique which organization can apply in order to achieve higher productivity in accordance with goals. In view of the above, the following hypothesis was proposed.

H2: The more emphasis on employees' incentives, the most possibility of better firms' performance.

2.4 Information Technology and Firm Performance

Technological innovation was found to have strong impact and influence on firm performance (Nohria & Gulati, 1996). Hitt, Hoskisson, and Kim (1997) further elucidated that the technology capabilities of the firms has vital influence on long-term performance of the firms. In addition, Dave and Wayne (2005) concluded that human resources regularly find new application of technology to improve their efficiency and their effectiveness in an effort to influence firm performance.

Nonetheless, past researchers (Mcloughlin & Harris, 1997) found out that technology account on business is minimal as many firms which incorporated technology to do transaction work, surprisingly, has a relatively low impact on performance. As stated by Mumford (2000), if firms emphasize too much on outcomes, they will tend to develop low-level technological innovation in order to avoid high uncertainty. Hence, we offer the following hypothesis.

H3: Information technology will be positively correlated to firms' performance.

2.5 Performance Appraisal and Firm Performance

Comprehensive performance appraisal system form basis yardstick for assessing individual's performance, highlight potential for future career advancement, most importantly, to improve the performance (Mullins, 2002). Lecky (1999) defined performance appraisal system as a benchmark which is set against specific task performance, define and evaluating current performance. It requires the input and output where criteria like remuneration, pay rise, level of expectation, promotion and managerial planning. In addition, it is a merit rating which should be benefit to both parties and must be constantly reviewed to suit the requirement. The system explicitly mentioned the individual's needs and thus has far reaching effect of improving productivity. Dave and Wayne (2005) argued that performance appraisai is an instrument whereby an individual was retaliated by the assessment due to certain personal disgruntled, and it has adversely affected future performance. Hence, the following hypothesis has been stated.

H4: Performance appraisal will be positively correlated to firms' performance.

3. Methodology

3.1 Sample

With an aim to generalize on firms in Sarawak, the population of the present study consists of manufacturing companies located in Sarawak, East Malaysia. Currently, the manufacturing sector is considered as one of the cornerstone of Malaysia's economic diversification strategy. Two hundred sets of questionnaires were distributed to executives working at manufacturing companies in Sarawak, however only 85 copies of questionnaires were usable for analysis.

3.2 Measures

The questionnaire used in this study consists of three parts. Section 1 required the respondents to rate a total of 20 items on the four components of HRM namely, training, incentives, information technology, and performance appraisal which were extracted from past researches such as Snell and Lau (1994), Kuratko, Hornsby, and Naffziger (1997), and Zahra, Neubaum, and Huse (2000). Section 2 contained 5 items of questions pertaining to firm performance based on the research of Daily and Johnson (1997). Finally Section 3 contains items regarding the demographic of the respondents such as gender, age, education background, working experiences, monthly gross salary, etc. The respondents were asked to describe on a 7-point Likert scale with: *7 = strongly agree, 6 = agree, 5 = slightly agree, 4 = neutral, 3 = slightly disagree, 2 = disagree, and 1 = strongly disagree.*

4. Findings

The Cronbach's coefficients alphas for HRM factors ranged from .75 to .80, respectively. The firm performance retained all the 5 items which accounted for its Cronbach's coefficients alpha of .83. Generally, the values indicated good internal consistency estimate of reliability of the grouped items for both factors. The findings of the reliability analysis are summarized in Table 2.

Table 3 illustrated the intercorrelations among the subscales were obtained from the Pearson Correlations Matrix to determine whether the subscales were independent measure of the same concept. Generally, the values indicating intercorrelations among the predictors variables were low, ranging from .29 to .46 ($p<.01$), thus indicating the independence of the scales used for measuring the predictors. In addition to that, a number of the predictor variables were noted to be significantly correlated to the criterion variables ranging from .35 to .52, which were considered as low intercorrelation values.

Multiple regression analysis was carried out to test the hypotheses that comprised the direct effects of HRM on firm performance. Table 4 presents the results of the analyses.

In analyzing firm performance, the main effect of the predictors explained a total of 40.4% of the variability in firm performance. Two dimensions in HRM, namely incentives and information technology were found to be significantly predictive of the firm performance at .35 ($p<.01$) and .23 ($p<.05$), respectively.

A confirmatory factor analysis (CFA) with AMOS 6.0 was conducted to test the four-dimensional structure of HRM-firm performance. As shown in table 5, the results has clearly indicated that only two dimensions of HRM were found to have significant impact on firm performance.

5. Discusssion and Conclusion

A number of researchers have revealed that there is significant relationship develops between HRM and firm performance (Huselid, 1995; Li, Zhao, & Liu, 2007; Schuler & Jackson, 1987) and this is also confirmed by resource-base theory. In this particular study, the outcomes of the analysis indicated that two dimensions of HRM have significant relationship with firm performance based on the data sample covering manufacturing companies in Sarawak.

As hypothesized, incentives were found to have a statistical significant relationship with firm performance. The result is not surprising in view of the fact that incentives are a vital part of remuneration which practically influenced a firm's performance.. This is also consistent with the research findings by Ian et al. (2004) and Nelson (1994) that it is inevitable that incentive rewards is a technique which must be applied in forming organizational strategy if companies strive to achieve better performances.

On the other hand, information technology is also found to be instrumental in achieving better firm performance. As commented by Dave and Wayne (2005), information technology was not taken lightly by global firms nowadays. In the same vein, Sorge et al. (1995) had discovered unbreakable link between information technology and firm performance. They concluded that the advent of technology has indeed helping many of the firm to turnaround in the aspect of performance. As stated by Preece (2000) and concurred by other researchers (e.g., Abernathy & Utterback, 1978; foster, 1986; Hill & Rothaermel, 2003) information technology increases effectiveness and outputs by cutting short transaction time for tasks.

Interestingly, the findings have demonstrated that employee training and performance appraisal did not correlate with better firm performance. The results contradicted with the study by Mullin (2002) that training is the key element in influencing the performance of a firm. This could probably due to the fact that manufacturing companies do not pay much attention on employees trainings and appraisal as much as the high-technology firms which need broader expertise and knowledge in technology innovation (Mumford, 2000).

The study has proven where the strong positive correlation values for HRM determinants such as incentives and information technology with firm performance. This denotes that, employees value incentives and technology information more importantly than trainings or appraisals given by the companies. These findings were consistent with previous research results that allowance of self-growth and independence in the workplace can enhance performances of companies (Dudeck & Hall, 1991; Gruber, 1996; Mumford, 2000).

6. Limitation and Implication

The major limitations of this study revolve around sampling issues as the small sample size reported here may have affected the current results. However, the small sample size coupled with the significant results reported does suggest that the current findings are found to be reliable. Secondly, it could be conceivable that issues related to organizational climate might have affected some of the relations studied. Clearly, a longitudinal approach would have placed researcher in a better position to draw causal conclusions. Therefore, only conclusions or discussions of the general relationships between the variables of interest could be drawn.

References

Abernathy, W. J., & Utterback, J. M. (1978). Patterns of innovation in technology. *Technology Review, 80,* 7, 40-47.

Appiah-Adu, K., & Singh, S. 91998). Customer orientation and performance: A study of SMEs. *Management Decision, 36,* 6, 385-399.

Armstrong, M. (2001). *A handbook of human resource management practice*, 8th Ed. Kogan Page Limited.

Buckley, R., Caple, T. (2004). *The Theory and Practice of Training*. Kogan Page, London.

Cho, H., & Pucik, V. (2005). Relationship between innovativeness, quality, growth, profitability, and market value. *Strategic Management Journal, 26,* 555-575.

Christensen, C. M. (1997). The innovators dilemma: When new technologies cause great firms to fail. *Harvard Business School Press,* New York, N.Y.

Daily, C. M., & Johnson, J. L. (1997). Sources of CEO power and firm financial performance: A longitudinal assessment. *Journal of Management, 23,* 2, 97-117.

Dave, U., & Wayne, B. (2005). *HRM value of proposition,* Boston, Harvard Business School.

Drucker, P. (1995). *Managing in time of great change*, New York, Truman Talley Books.

Drummond, H. (2000). Effective Management: *A hand book for manager.* Kogan Page.

Dudeck, S.Z., & Hall, W.B. (1991). Personality consistency: eminent architects 25 years later. *Creativity Research Journal, 4,* .213-32.

Foster, R. N. (1986). *Innovation: the attacker's advantage.* Summit Books, New York, NY.

Harsanyi, J. (1969). Rational-choice models of political behavior vs functionalist and comformist theorist. *World Politics, 21,* 21, 513-538.

Hill, C. W. L., & Rothaermel, F. T. (2003). The performance of incumbent firms in the face of radical technological innovation. *Academy of Management Review, 28,* 2, 257-274.

Hill, C. W. L., & Rothaermel, F. T. (2003). The performance of incumbent firms in the face of radical technological innovation. *Academy of Management Review, 28,* 2, 257-274.

Huselid, M. A. (1995). The impact of human resource management practices on turnover, productivity and corporate financial performance. *Academy of Management Journal, 38,* 3, 635-672.

Ian, B. Jim, J., & Will, H. (2004). *Human Resource Management.* New York Prentice Hall.

Kuratko, D. F., Hornsby, J. S., & Naffziger, D. W. (1997). An examination of owner's goals in sustaining entrepreneurship. *Journal of Small Business Management, 35,* 1, 24-33.

Laitinen, E. K. (2002). A dynamic performance measurement system: Evidence from small Finnish technology companies. *Scandinavian Journal of Management, 18,* 65-69.

Lecky, R. (1999). *Constructive appraisal.* Thompson Marshall Publishing, London.

Li, Y., Zhao, Y., & Liu (2006). The relationship between HRM, technology innovation and performance in China. *International Journal of Manpower, 27,* 7, 679-697.

Lin, C. Y. Y., & Chen, M. Y. C. (2007). Does innovation lead to performance? An empirical study of SMEs in Taiwan. *Management Research News, 30,* 2, 115-132.

Lupton, T. & Bowey, A. (1975). *Handbook of salary and wage systems.* Gower Press.

McLoughlin, I., & Harris, M. (1997). *Innovation, organizational change and technology.* London: International Thomson Business Press.

Mullins, J.L. (2002). *Management and organizational behaviour*, Prentice Hall.

Mumford, M. D. (2000). Managing creative people: Strategies and tactics for innovation. *Human Resource Management Review,10,* 3, 313-351.

Nankervis, A.R, Compton, R. L., & McCarthy, T. E. (1999). *Strategic human resource management,* (3rd Ed.), Nelson ITP, Melbourne.

Noe, R. A., Hollenbeck, J. R., Gerhart, B., & Wright, P. M. (2000). *Human resource management.* Chicago, IL: Irwin.

Nohria, N., & Gulati, R. (1996). Is slack good or bad for innovation? *Academy of Management Journal, 39,* 245-264.

Preece, J. (2000). *Online communities: Designing usability, supporting sociability.* Chichester, UK: John Wiley & Sons

Rothwell, W., Sullivan, R., & McLean, G. (1995). *Practicing organization development: A guide for consultants.* Jossey-Bass Pfeiffer, San Francisco, CA.

Sanchez, A. M., Jimenez, J. V., Carnicer, P. D. L., & Perez, M. P. (2007). Managerial perceptions of workplace flexibility and firm performance. *International Journal of Operations & Production Management, 27,* 7, 714-734.

Schuler, R., & Jackson, S. (1987). Linking competitive strategies and human resource management practices. *Academy of Management Executive, 1,* 202-219.

Snell, R., & Lau, A. (1994). Exploring local competences salient for expanding small businesses. *Journal of Management Development, 13,* 4, 4-15.

Torrington, D., & Hall, L. (1995). *Personnel Management: Hrm in Action.*
London, United Kingdom, Prentice Hall.

Vigoda, E., & Cohen, A. (2003). Work congruence and excellence in human resource management. Empirical evidence from the Israeli nonprofit sector. *Review of Public Personnel Administration, 23,* 192-216.

Yun, Z. S., & Good, L. K. (2007). Developing customer loyalty from e-tail store image attributes. *Managing Service Quality, 17,* 1, 4-22.

Zahra, S. A., Neubaum, D. O., Neubaum, D. O., & Huse, M. (2000). Entrepreneurship in medium-size companies: Exploring the effects of ownership and governance system. *Journal of Management, 26,* 947-976.

Table 1. Demographic Profile of the Respondents

Demographic Characteristics		**Frequency**	**Percentage**
Age	<20	3	3.5
	21-30	20	23.5
	31-40	39	45.9
	41-50	14	16.5
	Above 50	9	10.6
Gender	Male	40	47.1
	Female	45	52.9
Race	Malay	20	23.5
	Chinese	49	57.6
	Indian	0	0
	Others	16	18.8
Supervisor Gender	Male	50	58.8
	Female	35	41.2
Year with present organization	<10	36	42.4
	11 to 20	27	31.8
	21 to 30	15	17.6
	31 to 40	7	8.2
Position in the company	Clerical	28	32.9
	Lower level of management	20	23.5
	Middle level of management	26	30.6
	Top level of management	5	5.9
	Others	6	7.1
Sector	Consumer products	17	20
	Industrial products	21	24.7
	Construction	9	10.6
	Trading or services	29	34.1
	Others	9	10.6

Table 2. Results of Reliability Analysis

Variables	Cronbach's alpha
HRM factors	
Employee training	.75
Incentives	.78
Information technology	.86
Performance appraisal	.92
Firm performance	.83

Note. *N* = 85

Table 3. Pearson Correlations for HRM and Firm Performance Measures

Variables	1	2	3	4	5
Incentive	1.00				
Information technology	.31**	1.00			
Training	.44**	.34**	1.00		
Performance appraisal	.20*	.46**	.29**	1.00	
Firm performance	.52**	.45**	.44**	.35**	1.00

Note. *N* = 85 *p<.05, **p<.01.

Table 4. Regression Results: The Relationship between HRM Factors and Firm Performance

Independent Variable	Std Beta
Model Variables	
Incentives	.350**
Information Technology	.225*
Employee Training	.172
Performance Appraisal	.127
R^2	.404
Adj R^2	.380
R^2 Change	.404
F Value	17.09**

$^{*}p < 0.05$, $^{**}p < 0.01$

Table 5. Results of the testing of the Hypotheses

				Estimate	S.E.	C.R.	P
H1	firmPer	<---	training	.166	.084	1.970	.049
H2	firmPer	<---	Incentive	.507	.124	4.099	***
H3	firmPer	<---	info	.233	.092	2.525	.012
H4	firmPer	<---	appraisal	.104	.070	1.474	.141

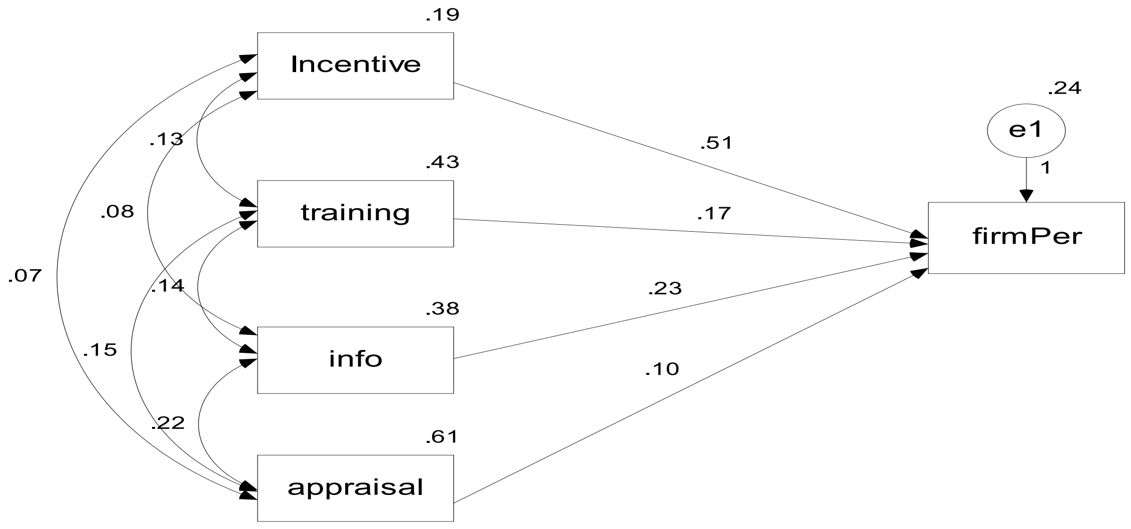
.19
Incentive
.24
e1
1
.13
.51
.43
training
.17
firmPer
.08
.07
.14
.23
.38
info
.15
.10
.22
.61
appraisal

Figure 1. Hypothesized model of HRM and Firm Performance

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

An Overview Study on Dornbusch Overshooting Hypothesis

Wenwen Tu
Department of Economics
School of Business, Nanjing University
Nanjing 210093, China
E-mail: winnie1225_tu@hotmail.com

Junwen Feng
School of Economics and Management
Nanjing University of Science and Technology
Nanjing 210094, China
E-mail: Fengjunwen8@hotmail.com

Abstract

Thirty years have passed since Dornbusch first published his overshooting hypothesis on “Expectations and Exchange Rate Dynamics”. Reviewing and appraising the advantages and disadvantages of this elegant model in international economics is of great importance to the consideration of establishing exchange rate and macro-economic policy in the future. In this paper, the author provides an analysis and empirical evidence for and against this influential model. Some conclusions are drawn from an all-sided discussion.

Keywords: Dornbusch overshooting hypothesis, Exchange rate, Macro-economic policy

1. Outline of Dornbusch Overshooting Hypothesis (DOH)

The Dornbusch overshooting hypothesis (DOH) was first written as “Expectations and Exchange Rate Dynamics” by economist Rudiger Dornbusch and was published in the Journal of Political Economy in 1976. Everyone calls this masterpiece marks the birth of modern international macroeconomics (Rogoff, 2002).

The hypothesis develops a simple macroeconomic framework for the study of exchange rate movements. The purpose is to develop a theory that is suggestive of the observed large fluctuations in exchange rates while at the same time establishing that such exchange rate movements are consistent with rational expectations formation (Dornbusch, 1976).

The model assumes a country that is small in the world capital market so that it faces a given interest rate. Perfect capital mobility will ensure the equalization of expected net yields so that the domestic interest rate, less the expected rate of depreciation, will equal the world rate. In the goods market we will assume that the world price of imports is given. Domestic output is an imperfect substitute of imports, and aggregate demand for domestic goods, therefore, will determine their absolute and relative price (Dornbusch, 1976). Both the goods market and money market are in equilibrium. Purchasing Power Parity (PPP) holds in the long-run. Moreover, while commodity prices are sticky (slow to adjust), assets prices (i.e. exchange rate) adjust instantaneously in response to new information.

Three relationships lie at the heart of the overshooting result. The equation (1) below, where r is the domestic interest rate, r* is the given world rate of interest and x is the expected rate of depreciation of the domestic currency, is a representation of perfect capital mobility:

$$r = r^* + x = r^* + \theta\ (e' - e) \tag{1}$$

This also states that the expected rate of depreciation of the spot rate in proportional to the discrepancy between the long-run rate and the current spot rate, where θ is the coefficient of adjustment, e’ and e are the long-run rate and the short-run rate respectively.

The second core equation is the condition of equilibrium in the domestic money market. The demand for real money balances is assumed to depend on the domestic interest rate and real income, which equals the real money supply (Dornbusch, 1976). Assuming a conventional demand for money, the log of which is linear in the log of real income and in interest rates, we have:

$$m - p = -\alpha r + \varphi y \qquad (2)$$

Where m, p, and y denote the logs of the nominal quantity of money, the price level and real income; both α and φ are positive parameters and they are interest elasticity of demand for money and income elasticity of supply for money respectively.

The third equation is the demand function for domestic output which depends on the relative price of domestic goods, e – p, interest rates and real income (Dornbusch, 1976). It is derived from aggregate demand function and is assumed to have the form

$$\ln D = u + Yy + a(e - p) - br \qquad (3)$$

Where D denotes the demand of domestic output and u is a shift parameter.

Now, an increase in the supply for money that is expected to persist will cause a goods and asset market disequilibrium at the initial exchange rate and price (Figure 1). The 45°line R is drawn through the origin on the assumption that PPP holds in the long-run. The QQ schedule reflects that money market equilibrium and international arbitrage of net expected yields. The positively sloped schedule P. =0 shows combinations in price levels and exchange rates for which the goods market and money market are in equilibrium. The economy in initial full equilibrium at point A, with a long-run price level P°and a corresponding long-run exchange rate e°where the level of prices is determined.

Insert Figure 1 Here

In the short –run, to maintain money market equilibrium, the increased quantity of money would have to be matched by the money market schedule shifting out to Q’Q’. Due to the assumption that exchange rates and asset markets adjust fast relative to goods market, the nominal money supply raises but the price level is temporarily fixed. Therefore, the short-run equilibrium is at point B.

But in the long run, the price level could go up in order to maintain money market equilibrium. According to equation (2), the demand for real balances will go down. Therefore, the domestic interest rate r could fall less then the short-run. Thus the new long-run equilibrium is at point C, where both goods and money markets clear and exchange rate and price changes exactly reflect the increasing in money. As a result, the short-run initial depreciation of the exchange rate on impact must be larger than the long-run depreciation, i.e. the new short-run exchange rate e is larger than the new long-run exchange rate e’. That is, a monetary expansion results in an overshooting exchange rate.

2. The Strength of the Dornbusch Overshooting Hypothesis

Firstly, even with the inevitable onslaught of more modern approaches, the Dornbusch overshooting hypothesis remains a central tool for policy analysis. It is not only a framework for considering international monetary policy but also a model for understanding exchange rates. The hypothesis shows that with sticky prices in goods market, exchange rate volatility is needed to temporarily equilibrate the economy in response to monetary policies. The general approach has been applied to different problems, including the “Dutch disease”, the choice of exchange rate regime, commodity price volatility and analysis of disinflation in developing countries (Rogoff, 2002). Sichei et al. (2005) suggest that the Dornbusch overshooting model appears to underlie the movement of the nominal Rand-USD exchange rate in the period 1994 to 2004 in South Africa (Figure 2).

Insert Figure 2 Here

Source: Sichei et al. (2005)

This model fits the data well and prices in South Africa are sticky which is derived from the high-income elasticity of demand. Thus increasing interest rates with a view to strengthening the Rand does the opposite in the short-run, i.e. depreciate the Rand in the first 8 quarters.

Secondly, the “Expectations and Exchange Rate Dynamics” was the first paper in international economics to marry sticky prices with rational expectations, both central features of today’s mainstream. Rational expectations here imply that private agents must form exchange rate expectations in a way that is consistent with the model itself (Rogoff, 2002). In other words, the paper combined the Keynesian short-run analysis and the Monetarism long-run analysis in order to be more suitable to reality.

Thirdly, the Dornbusch overshooting hypothesis is used as a fundamental model to analyze open economy in international macroeconomics until now, precisely because it is theoretical simplicity, clarity and elegance. Rogoff (2002) pointed out that if one is in a pinch and needs a quick response to a question about how monetary policy might affect the exchange rate, most of people will still want to check any answer against Dornbusch’s model.

3. The Weakness of the Dornbusch Overshooting Hypothesis

Although the Dornbusch overshooting hypothesis has keeping itself in the forefront of practical policy analysis for thirty years, it has been under severe attack since the model was published. Various version of critical evaluation of the overshooting phenomenon argue that undershooting may happen under some given circumstances. In general, the hypothesis can be problematic for those following critical evaluations.

3.1 The Interest Elasticity of the Demand of Money

In the extreme case, however, when α, the interest elasticity of the demand for money, is infinity, as in the Dornbusch hypothesis, a different conclusion exists (Figure 3).

Insert Figure 3 Here

To analyze the effects of a monetary expansion, consider an initial long-run equilibrium at point A where money market curve QQ is vertical. A rise in the money supply causes this schedule shift out to Q'Q' which is also vertical. The exchange rate reaches immediately its long-run value, i.e. the short-run equilibrium at pint B is on the same line with the long-run equilibrium at point C. Therefore, the adjustment in short-run coincides with the adjustment in long-run, i.e. the new short-run exchange rate e equals the new long-run exchange rate e'. Neither overshooting nor undershooting happens under this circumstance.

3.2 The Imperfect Capital Mobility

Dornbusch assumed that money markets clear instantaneously while adjustment in goods markets is sluggish. The hypothesis emphasizes that the key factor determining whether the exchange rate overshoots or undershoots its equilibrium value is the speeds of adjustment in goods and money markets. The hypothesis assumes the perfect capital mobility and the complete flexibility of exchange rate. However, in real markets imperfect or low capital mobility happens frequently. At this moment the QQ schedule is positively sloped and the current account of balance of payment dynamics can have large medium term impacts on real exchange rates (Figure 4).

Insert Figure 4 Here

A trade deficit of the current account will require a depreciation of domestic currencies. One often sees this pattern in practice, particularly for countries that are forced to quickly reverse current account balances (Rogoff, 2002). It offers a typical instance that the correlation of real exchange rate and current account imbalances in Thailand, Korea and Indonesia in the 1990s (see Figures 5 to 7). Since the capital mobility is imperfect, a monetary expansion causes the QQ schedule shift out to Q'Q'. It is shown in Figure 4 that the new short-run exchange rate e is smaller than the new long-run exchange rate e'. In other words, a monetary expansion results in exchange rate undershooting.

Insert Figure 5 Here

Insert Figure 6 Here

Insert Figure 7 Here

Frenkel and Rodriguez (1982) modified the Dornbusch's model and allowed a finite speed of adjustment in money markets. It shown that under this circumstance, the short-run effects of a monetary expansion depend on the degree of capital mobility. When capital is highly mobile, the exchange rate must overshoot its long-run value, but when capital is relatively immobile, the exchange rate undershoots its long-run value.

3.3 The Long-run Expenditure Function

The Dornbusch overshooting hypothesis implies that the expenditure function is given. But in practice it will be illustrated that the effects of a monetary expansion on the dynamics of exchange rates and in particular on whether exchange rates overshooting or undershooting their equilibrium path depend critically on the variance of expenditure function.

For instance, the expenditure function of a small net importer of tradable goods country will be assumed as equation (4) below.

$$P = \alpha Pt + (1 - \alpha) Pnt \tag{4}$$

Where P, Pt and Pnt denote the price level, prices of tradable goods and prices of non-tradable goods; α is the percentage coefficient of trade portfolio. Let α equal 3, i.e. (1 – α) will be 7. In the short-run, the equation (4) is fixed. Thus a 10% monetary expansion causes a 3% increase in prices of tradable goods and a 7% increase in prices of non-tradable goods respectively. Moreover, the domestic currency depreciates. But during the long-run, the expenditure function can be changed according to new information. Since the prices of non-tradable goods have increased more than tradable goods, the country will shift its expenditure to the tradable goods. Therefore, α will be changed into 5, i.e. (1 – α) will also equal to 5. Under this circumstance, the exchange rates move only imposed by the prices of tradable goods. As a result, the adjustment of exchange rate in short-run less than the adjustment in long-run, i.e. the exchange rate must be undershooting.

3.4 The Empirical Data

In the first place, Rogoff (2002) points out that the basic problem with the Dornbusch overshooting hypothesis is that whereas it seems to capture major turning points in monetary policy quite well, the model does not capture all the other huge exchange rate swings that regularly take place. Thus formally testing the Dornbusch model is easier said than done. To take the model to the data, one needs to resolve many issues. This drawback can be described with the help of a graph of the USD rate versus the German Mark (Figure 8).

Insert Figure 8 Here

The solid line gives real exchange rate and the dashed line is the one-year real interest differential. A rise in the real exchange rate represents a depreciation of the Mark. A rise in the real interest differential represents a rise in German real interest rates relative to those prevailing in the United States. As one can see, the model does seem to say something about major turning points, though we will not press to see if it robustly passes regression tests (Rogoff, 2002).

In the second place, another piece of evidence comes from the co-movements of forward and spot exchange rates. Dornbusch's thought experiments of a one-time unanticipated change in the level of the money supply, the spot rate should move by more than the forward rate at any horizon. This excess movement is precisely the overshooting. However, the hypothesis has come under pressure from real data (Table 1).

Insert Table 1 Here

Indeed, as Table 1 illustrates, forward rate volatility is slightly lower than spot rate volatility over the sample period. The data are inconsistent with the Dornbusch overshooting hypothesis. As Flood (1981) highlighted that, in most cases; the overshooting model predicts that forward rates and spot rates will not move one for one in general. The exact co-movements may depend on the nature of the shock (real versus nominal, temporary versus permanent) and on the horizon of the forward rate.

4. Conclusions

This paper has analyzed and examined the Dornbusch overshooting hypothesis in models and empirical data. From what have been discussed above, we may safely draw those following conclusions.

First of all, the overshooting concept is an extraordinary framework for reflecting the relationship between international macroeconomics policy and exchange rate. It remained a central workhorse model in its original from for policy analysis over the past quarter century. In addition, the "Expectations and Exchange Rate Dynamics" was the first paper in international finance to marry sticky prices with rational expectations, both central features of today's mainstream. Moreover, the Dornbusch model defines a high-water mark of theoretical simplicity and elegance in international economics.

Secondly, it has been seen that exchange rate overshooting is not an intrinsic characteristic of the foreign exchange rate market and that it depends on a set of specific assumptions. As long as the interest elasticity of demand for money is infinity, the adjustment of exchange rate in short-run is equal to the long-run adjustment. In addition, when capital mobility is imperfect, i.e. the speed of adjustment in money markets is relatively immobile; the exchange rate undershoots its long-run value. Moreover, when the expenditure function can be changed during the long-run and where a country is a small net importer of tradable goods, the exchange rate must undershoot.

The last but not the least, the Dornbusch overshooting hypothesis about why exchange rates overshoot has proven of relatively limited value empirically, although a plausible case can be made that it captures the effects of major turning points in monetary policy.

References

Dornbusch. (1976). Expectations and Exchange Rate Dynamics. *Journal of Political Economy,* Vol. 84 (December 1976), Page 1161-1176.

Flood. (1980). Explanations of Exchange-Rate Volatility and Other Empirical Regularities in Some Popular Models of the Foreign Exchange Market. *Carnegie-Rochester Conference Series on Public Policy,* Vol. 15, Page 219-249.

Krugman. (2006). *International Economics: Theory and Policy.* Boston: Pearson Education.

Mussa. (1980). *The Role of the Current Account in Exchange Rate Dynamics.* Chicago: University of Chicago.

Mussa. (1982). A model of Exchange Rate Dynamics. *Journal of Political Economy,* Vol. 90 (February 1982), Page 74-104.

Rogoff. (2002). Dornbusch's Overshooting Model after Twenty-Five Years. Imf.org, Available URL:http://www.imf.org/external/np/speeches/2001/112901.htm.

Sichei, Gebreselasie and Akanbi. (2005). *An Econometric Model of the Rand-US Dollar Nominal Exchange Rate.* Web.up.ac.za. Available URL:http://www.up.ac.za/dspace/bitstream/2263/4509/1/Sichei_Econometric%282005%29.pdf.

Table 1. Variance of the Forward Rate Divided by Variance of Spot Rate

COUNTRY	THREE MONTHS	ONE YEAR
Canada	0.88	0.89
France	1.08	1.07
Germany	0.88	0.87
Italy	0.92	0.91
Japan	0.99	0.99
United Kingdom	1.00	1.00

Source: IMF, International Financial Statistics; IMF staff estimates.

Note: Quarterly data for 1979 Q-Ⅰ ~ 2000 Q-IV.

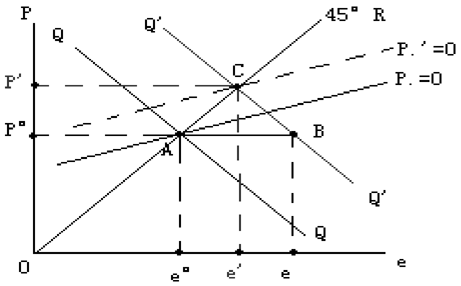

Figure 1. The Dornbusch Overshooting Hypothesis Display

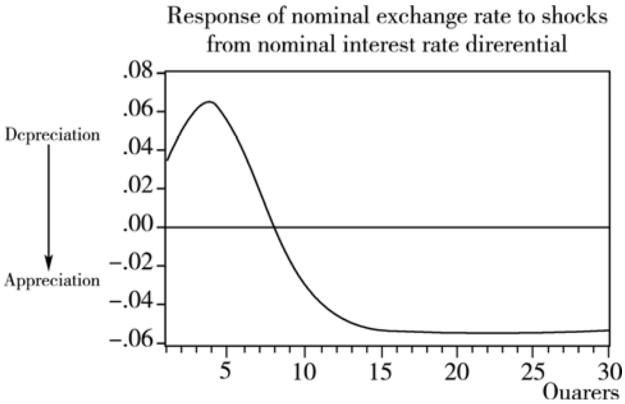

Figure 2. Response of Nominal exchange rate to Stocks

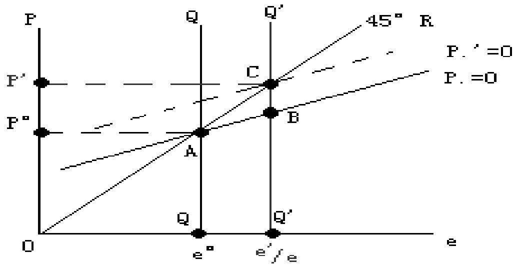

Figure 3. The Interest Elasticity of the Demand of Money

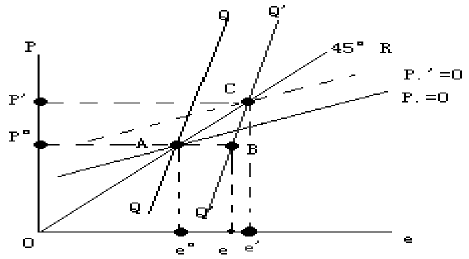

Figure 4. The Imperfect Capital Mobility

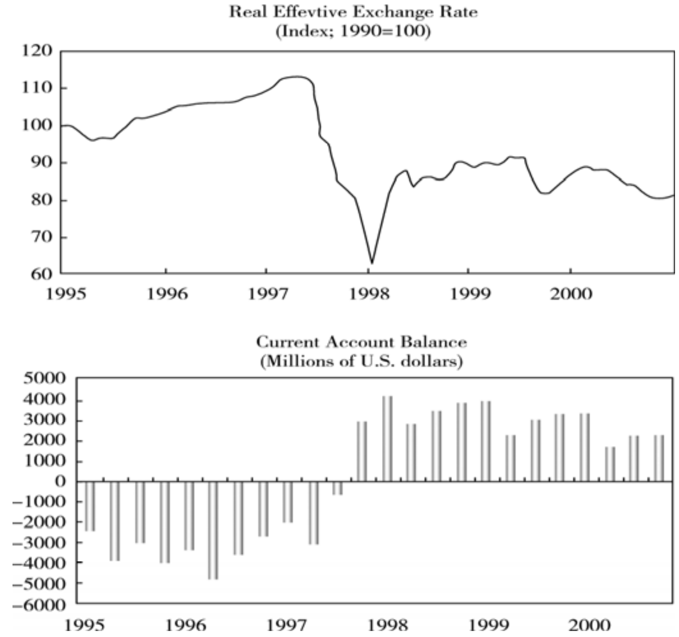

Figure 5. Real Exchange Rate and Current Account Imbalance in Thailand

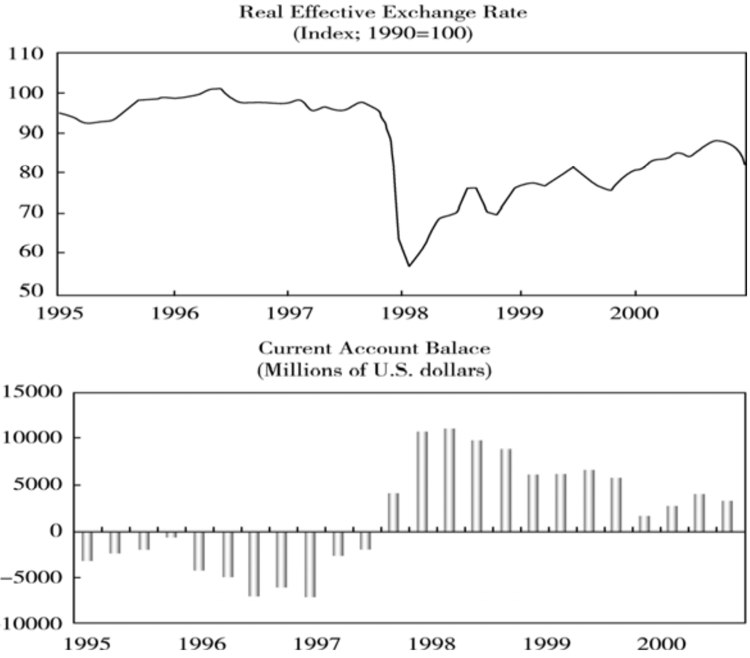

Figure 6. Real Exchange Rate and Current Account Imbalance in Korea

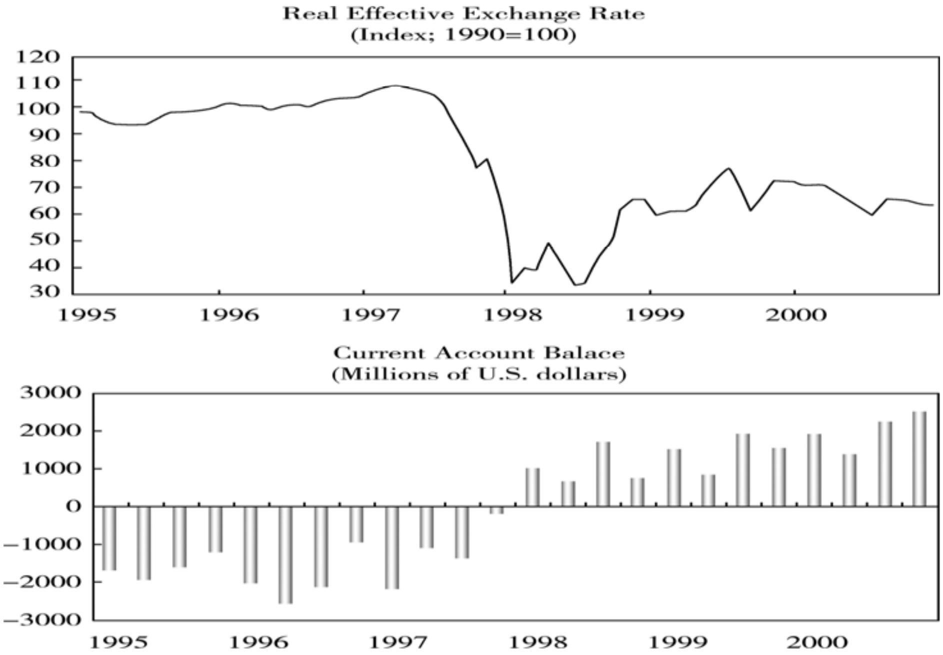

Figure 7. Real Exchange Rate and Current Account Imbalance in Thailand

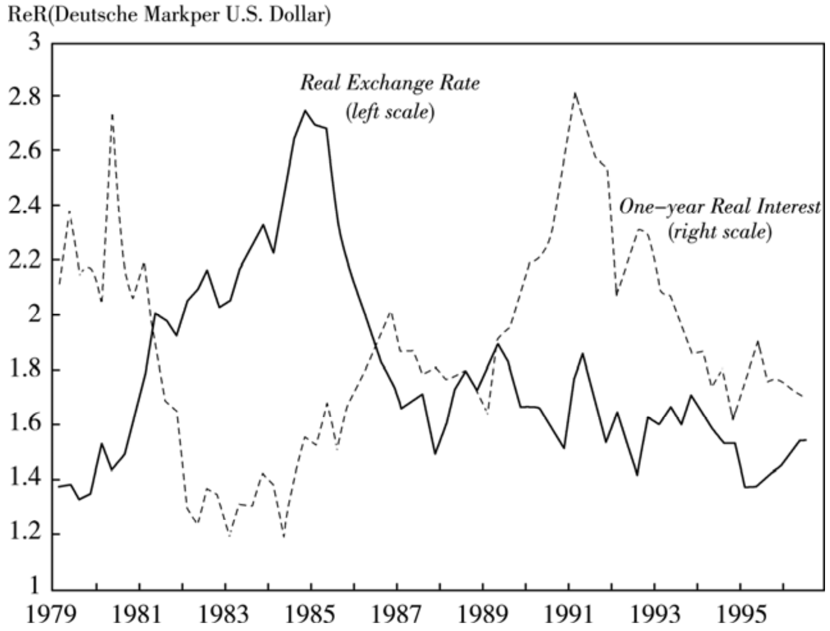

Figure 8. USD rate Versus German Mark

International Journal of Economics and Finance
Vol. 1, No. 1
February 2009
www.ccsenet.org/journal.html

AFTA and the Intra-Trade Patterns among ASEAN-5 Economies: Trade-Enhancing or Trade-Inhibiting?

Kim-Lan Siah
Department of Economics, Faculty of Economics and Management
Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
E-mail: audreysiah@yahoo.com

Chee-Keong Choong (Corresponding author)
Department of Economics and Finance, Faculty of Business and Finance
Universiti Tunku Abdul Rahman (Perak Campus)
Jalan Universiti, Bandar Barat, 31900 Kampar Perak Darul Ridzuan, Malaysia
Tel: 60-5-466-2323 E-mail: choongck@utar.edu.my

Zulkornain Yusop
Department of Economics, Faculty of Economics and Management
Universiti Putra Malaysia, 43400 UPM Serdang, Selangor, Malaysia.
Tel: 60-3-8946-7631 E-mail: zolyusop@econ.upm.edu.my

Abstract

This paper examines ASEAN economic integration and its ability to promote intra-ASEAN trade, namely, Indonesia, Malaysia, Philippines, Singapore and Thailand. In order to achieve this goal, a modified gravity model is estimated within autoregressive distributed lag (ARDL) framework, or bounds testing approach for each of the five ASEAN countries based on data from year 1970 to year 2001. The empirical results indicate that the effects of the size of economy on bilateral trade flow in ASEAN are either trade-enhancing or trade-inhibiting dependent on the country-specific. There is evidence that AFTA preferential arrangements are important and prevalent in enhancing intra-ASEAN trade. However, ASEAN countries may not as a whole benefit from the formation of AFTA as the trade deflection might occur in the regional market.

Keywords: Free trade area, Trade creation, Trade diversion, Trade deflection, Trade-enhancing, Preferential trade arrangement

1. Introduction

The economic emergence of the Association of Southeast Asian Nations (ASEAN) brings not only opportunities but also challenges to each of the member countries. Although ASEAN countries occupy similar geographical area, they are extremely diverse in terms of economic structure and development, political orientation, factor endowments, culture, history and religion. The divergence in income and background are the obvious feature of the ASEAN Free Trade Area (AFTA) member countries. The divergence in background is reflected by the readiness of each nation approaching AFTA dateline. An indicator of readiness can be seen from the members' proposal of the average import tariff with the smaller figures implying the higher degree of readiness while larger figures implying the lower degree of readiness. As observed by Naya and Plummer (1997: pp. 118): "ASEAN economic co-operation had not developed significantly and many scholars felt it couldn't (because of the dominance of politics in the Association) or that it shouldn't (due to the potential effects trade diversion in a region characterized by low level of intra-regional trade)."

From the theoretical standpoint, regional trade agreements' benefits and costs can be portrayed from two perspectives. First, the static effects of economic integration that measured in terms of increasing the effectiveness of production and consumer welfare (Note 1). Second, the dynamic effects of economic integration, affecting the overall economic growth rate of member countries in the long term (Note 2). The static effects include both the trade creation and trade diversion effects, in which they occur directly on the formation of the integration project (Note 3). According to Viner (1950), trade creation taking place whenever economic integration leads to a shift in product origin from a domestic producer whose resource costs are higher to a member producer whose resource costs are lower. Whereas, Plumber (1997) and Hassan (2001) view net trade creation as an improvement in the allocation of resources. Trade creation is beneficial to a country because it leads to the contraction of inefficient resources that could be used more effectively elsewhere in the economy. On the contrary, Plummer (1997) and Hassan (2001) argued that trade diversion would worsen the international allocation of resources and shifts production away from comparative advantage. However, the static welfare effect of economic integration that involves trade diversion is ambiguous, depending on the relative strength of these two opposing forces such as innovation of technology (Beaulieua, *et al.*, 2004). Viner (1950) concludes that regional trading arrangements can either increase or reduce world welfare depending upon the relative magnitudes of the trade-creation and trade-diversion effects.

Plummers (1997) relates dynamic effects as a medium term, and are long-term, implications of regional integration and, as such, are likely to be more important than the static effects that they may lead to the economic structure and performance of participating countries may evolve differently than if they had not integrated economically. He considers four possible areas that define to be "dynamic" in nature and derive directly from regional integration: "economies of scale", "X-efficiency" improvements, changes in investment flows, and the potential for industrial expansion in a developing country context. Hassan (2001) provides a similar view of point and claims that dynamic effects can be viewed as the factors that lead to faster economic growth.

There is a model of bilateral trade suggests that trade diversion or trade creation can be justified misallocation of resource reasons. Empirically, Ramasamy (1995) estimates the negative effect of trade diversion on resource allocation among members of ASEAN and reveals that trade creation is greater than trade diversion for Indonesia, Philippines and Thailand, whereas the reverse is true for Malaysia and Singapore. He concludes that resources will be misallocated in Malaysia as a result of Free Trade Area. However, AFTA will cause a positive gain in the allocation of resources for Indonesia. These are static effects that measuring the impact in the short term only. In the long term, it is possible that the positive dynamic effects of increased competition, economies of scale and the benefits of intra-industry trade far outweigh the short term effects.

Endoh (1999) attributes the effects of bilateral trade to the feature and the transition of trade relations in the Asia-Pacific region during the post-World War II period. He concludes that ASEAN has had no effect of its own on promoting trade among its member countries; the volume of trade among EAEC has been at a high level as compared to the hypothetical trade level since 1960; the amount of trade between EAEC economies and other APEC countries has been growing throughout the post-war period and close trade relations among APEC economies and some other Asian countries.

Besides, there is a broad literature that applies the gravity approach to study the international bilateral trade. Thornton and Goglio (2002), for example, confirm the importance of economic size, geographic distance and common language in intra-region bilateral trade by applying the gravity model. The finding shows that re-exports and membership of ASEAN have been important factors promoting intra-regional trade. In line with this research, Clarete, *et al*. (2003) also use gravity model to estimate the coefficients of the basic determinants such as the GDP, distance between capitals of trading partners, population, and physical area to explain cross-country trade flows. They reveal that Preferential Trade Areas (PTAs) have augmented trade in Asia. The new AFTA members are less integrated with the world economy as the founding members of this PTA. The PTAs in this analysis have contributed significantly to trade expansion both at the global and regional (Asia and Pacific) level. There is strong evidence that PTAs can create rather than divert trade. The results obtained suggest that PTAs offer a next-best path towards expanding world trade if negotiations for multilateral trade liberalization take a longer time to get completed.

This empirical result is also supported by Frankel (1998), who also used gravity model to detect and quantify a possible intra-regional trade bias by establishing a "norm" of bilateral trade volume based on economic, geographic and cultural factors (Note 4). Frankel concludes that trade is indeed concentrated regionally in formal RTAs such as ASEAN and NAFTA, or less formal grouping such as APEC. Although East Asia has a high level of intra-regional trade concentration, there has been no increase over time. Hence, the rapidly increased in intra-Asian trade can be entirely explained by the rapid growth in the region. On the other hand, the rapidly increased in trade among the APEC countries cannot be entirely explained by the growth of the member countries. APEC has operated as an intensifying trade bloc and its high level of trade is all the more impressive in that the distances are so large.

In their recent study, Elliott and Ikemoto (2004) have investigated the effect of AFTA on world and regional trade patterns, concerning on data for the five founding ASEAN countries by applying a modified gravity equation. They found that trade flows were not significantly affected in the years immediately following the signing of the AFTA agreement in 1993. This reinforces the findings of previous studies. Moreover, they argued that the degree of trade creation was lower than the preceding period (1988-1992). When the gravity equation is re-estimated for intra-ASEAN trade only, nevertheless, they did find some evidence of a positive AFTA effect that, although limited at the beginning stage, gradually increased. They argued that the fall in trade creation immediately following 1993 was the emergence of credible competition for market share from the new industrial and exporting powers of China, South American and Eastern Europe.

From the previous literature, obviously, co-operation through economic integration such as free trade areas or preferential trade agreements would lead to an increase (decrease) in trade efficiency (inefficiency) and secure a more (less) economical use of resources in order to increase (decrease) intra-ASEAN trade in the long run. Recently, nevertheless, as AFTA involves different background of economies and being more heterogeneous, thus it faces problem in securing internal implementation though it has obtained the consensus. Each member countries are clearer of their own large external seeking than it is about its internal objectives as well as retains the right to establish its own separate protection against imports from the rest of the world. Obviously, within the current framework of AFTA, each nation is acted as individualistic rather than collaborative or complementary with one another. This non-collective planning and strategies among ASEAN members raises the possibilities of a particularly dark scenario – trade deflection to take place (Note 5).

With the persistence of this unhealthy trend not only makes ASEAN as a whole might not gain collectively from AFTA, indeed all of them might be worse off. Whether ASEAN countries will, as a whole, benefit from the formation of AFTA is still an open question. It is therefore important to ask: Do they really promote trade among members? What are the significant influences on trade linkages or interactions among ASEAN economies – Indonesia, Malaysia, Philippines, Singapore, and Thailand – before and after AFTA? How much do trade patterns change over time among members? In this paper, we re-examine bilateral trade relationship between each ASEAN-Five members in order to discover what determinants of intra-trade that can explain the changing trade patterns.

The first contribution of the paper is to examine and prove a proposition that has existed for many years in the trade-policy literature: there can exist gains (trade creation) from trade due to the economic integration of participating countries, but there will be a loss (trade diversion) resulting from the heterogeneous background of the participating countries. The second contribution is to provide explanations to these issues by exploring the effects of AFTA and crisis on intra-ASEAN besides identifying the factors that explaining the intra-ASEAN trade via a modified gravity model within autoregressive distributed lag (ARDL) framework for each of the five ASEAN countries based on data from year 1970 to year 2001. To the best of our knowledge, there does not exist a formal ARDL modeling framework in which to estimate the intra-trade relationship between ASEAN members.

This paper is organized as follows: section 2 provides some historical background to ASEAN regional co-operation. Section 3 discusses the modified gravity model and its essential features. This section also presents the testing framework of the modified gravity approach within ARDL model. Section 4 discusses the empirical results. Section 5 provides some concluding comments and policy implications on the future of regional integration in ASEAN.

2. Historical overview in the ASEAN economies

It could hardly be asserted that the economic co-operation programmes implemented through the establishment of ASEAN were a full success. ASEAN remains to be a loose regional grouping. Each member country still insists on its own independent law and policies, its legal system, and its sovereign right to control and regulate internal activities as well as conduct external relationships, except those mutually agreed in the economic, social and political cooperation programs. Every programme implemented in ASEAN has been agreed among the member countries on a consensus basis. It was only in 1992 that ASEAN began to develop the idea of establishing the ASEAN Free Trade Area, a scheme to strengthen intra-regional economic cooperation to respond to global change. Through Free Trade Arrangement (FTA) member countries are expected to enjoy more positive outcomes from an economic integration. Resulting from the Fourth Summit Meeting, the ASEAN heads of government formally agreed and signed the Singapore Declaration, marking the commitment to intensify economic co-operation in the entire region. Basically, the Declaration is the creation of the ASEAN Free Trade Area (AFTA) in 15 years beginning 1st January 1993.

The main purpose of Free Trade Area in ASEAN is to remove the obstacles to a freer trade among member states. Under AFTA, high tariffs or taxes on traded goods and the quantitative restrictions and other non-tariffs barriers that limit the entry of imports had been abolished. Meanwhile, each ASEAN member country has the right to set its own level of tariffs on imports from non-members. The ultimate objective of AFTA is to increase the competitiveness of ASEAN as a globally competitive production base. AFTA will facilitate efficient utilisation of scarce resources and provides opportunities for the ASEAN member countries to further strengthening their competitive advantage. With the widespread, rapid and sustained economic development, AFTA is expected to narrow the economic gap among the ASEAN member countries through trade-creating effect, improve resource allocation and create greater employment opportunities. Besides, AFTA provides opportunities for the ASEAN producers to enjoy a bigger market of 526 million people and with the GDP capacity of US$555.3 billion. Furthermore, AFTA is created for the purpose of complementing the forward and outward-looking ASEAN countries in promoting economic growth through export and investment-promotion strategies. It will further improve the external competitiveness of the ASEAN on the back of substantial trade-creation effects and bigger flows of trade-related investment. As pointed out by Sharma and Chua

(2000), the complementarity in market size, resource endowment and income level provided enormous possibilities for interactions between countries in either goods or factor markets. For example, Resource-rich countries such as Malaysia and Indonesia can complement resource-poor countries like Brunei and Singapore. On the other hand, countries with high ratios of saving, such as Brunei and Singapore, can channel their investments through foreign investments to countries with low capital resources, like Thailand and Indonesia.

The main mechanism to realize AFTA is the Agreement of the Common Effective Preferential Tariff (CEPT) scheme. The CEPT Agreement allows ASEAN member Countries to reduce their tariffs to 0-5% on a MFN basis among ASEAN members. However, the ASEAN Economic Ministers Meeting Chiang-Mai, Thailand on 22nd and 23rd September 1994 had reached an agreement to shorten the time frame for the realisation of the AFTA from 15 to 10 years (Note 6). The CEPT scheme will cover nearly 98% of all tariff lines in ASEAN by the year of 2003 excluded those in the General Exceptions category and sensitive agricultural products. The Fifth ASEAN Summit held in Bangkok, Thailand in December 1995 reaffirmed the commitments of removing all quantitative restrictions and non-tariff barriers (NTBs). This further mandated greater transparency in standards and establishment of mutual recognition arrangements on a bilateral or multilateral basis to facilitate greater intra-regional trade.

The year 1997-1998 Asian crisis was a major turning point. The ASEAN economies and business dynamism in the region has been severely affected by the financial and economic crisis. In order to regain business confidence, enhance economic recovery and promote economic growth, the ASEAN leaders are oblige to be committed in the realization of AFTA as well as to enhance further economic integration of the region. At the 6th ASEAN Summit in year 1998, which was held in Hanoi, the ASEAN leaders adopted measures to accelerate rather than reversed the process of tariff reductions during the year 1997 Asian financial crisis. They agreed that the original signatories to the Agreement on the CEPT Scheme for the AFTA namely, Indonesia, Malaysia, Philippines, Singapore, Thailand and Brunei Darussalam would advance the implementation of AFTA by one year from year 2003 to year 2002. ASEAN accelerated rules permitting 100 per cent non-ASEAN ownership of regional companies has also brought about a clear message that FDI from any region was welcomed in ASEAN.

3. Data and Methodology

The core methodology adopted in this study is based upon the gravity model (Note 7) used by Sharma and Chua (2000). This study follows the modify gravity model by using the gross domestic product (GDP) instead of gross national product (GNP) that has been used by Frankel (1993). In this study, the gravity model is conducted by using the time series techniques among ASEAN-5. The modified gravity equation used in this study is in natural logs as shown below:

$$\begin{aligned}X_{ij} = \beta_0 + \beta_1(RGDP_i * RGDP_j) + \beta_2(PGDP_i * PGDP_j) + \beta_3 TC_{ij} + \beta_4 D1 \\ + \beta_5 D2 + \varepsilon_{ij}\end{aligned} \tag{1}$$

Where $(RGDP_i * RGDP_j)$ is the product form of RGDPs for both countries *i* and *j* (to take into account the size of the economy), $(PGDP_i * PGDP_j)$ is the product form of GDP per capita for both countries *i* and *j* (to proxy for the income level), TC_{ij} is the transportation cost to proxies the distance between a particular country *i* to a particular country *j*, *D1* representing the AFTA preferences, in which a value of one is assigned for period after forming of AFTA and zero is assigned for period before forming of AFTA; and *D2* representing the Asian financial crisis, in which a value of one is assigned for period after Asian financial crisis and zero is assigned for period before the crisis.

This study employs the autoregressive distributed (ARDL) bounds test by Pesaran, et al. (2001) to examine the cointegration relationship between trade (X) and a set of explanatory variables. The ARDL technique is selected based on the several considerations of its advantages. Firstly, unlike most of the conventional multivariate cointegration procedures, which are valid for large sample size, the bound test is suitable for a small sample size study (Pesaran, et al., 2001). Given that our sample size is limited with a total of 32 observations only, conducting bounds test will be appropriate. Secondly, the bound test does not impose restrictive assumption that all the variables under study must be integrated of the same order. The test statistic underlying the procedure is similar to the Wald test (or *F*-statistic) in a generalized Dickey–Fuller type regression used to test the significance of lagged levels of the variables under consideration in a conditional unrestricted equilibrium correction model (UECM). It is shown that the asymptotic distributions of both statistics are non-standard under the null hypothesis that there exists no long-run relationship in levels between the variables under concerned, irrespective of whether the regressors are purely I(0), purely I(1) or mutually cointegrated. As such, the order of integration is no more a sensitive issue and thus one could bypass the unit root tests (Note 8).

This procedure allows for testing of at most one long-run relationship and thus requires a zero restriction on one of the off diagonals of the λ matrix. The following autoregressive distributed lag, ARDL model will be estimated in order to test the cointegration relationship between the examined gravity model variables:

$$\begin{aligned}\Delta X_{ij,t} &= \beta_0 + \beta_1 X_{ij,t-1} + \beta_2 (RGDP_i * RGDP_j)_{t-1} + \beta_3 (PGDP_i * PGDP_j)_{t-1} + \beta_4 TC_{ij,t-1} \\ &+ \beta_5 D1 + \beta_6 D2 + + \sum_{m=1}^{p} \beta_7 \Delta X_{ij,t-m} + \sum_{m=0}^{p} \beta_8 \Delta (RGDP_i * RGDP_j)_{t-m} \\ &+ \sum_{m=0}^{p} \beta_9 \Delta (PGDP_i * PGDP_j)_{t-m} + \sum_{m=0}^{p} \beta_{10} \Delta TC_{ij,t-1} + \varepsilon_{ij,t}\end{aligned} \tag{2}$$

The ARDL procedure involves two stages. At the first stage the existence of the long run relation between the variables under investigation is tested by computing the *F*-statistic for testing the significant of the lagged levels of the variables in the error correction form of the underlying ARDL model. However, the (asymptotic) distribution of the *F*-statistic is non-standard, irrespective of whether the regressors are I (0) or I (1). First, we estimate Equation (1) by ordinary least square (OLS) technique. Second, the presence of cointegration can be traced by restricting all estimated coefficients of lagged level variables equal to zero. That is, the null hypothesis is $\beta_1=\beta_2=\beta_3=\beta_4=0$ against its alternative $\beta_1\neq\beta_2\neq\beta_3\neq\beta_4\neq0$. If the computed *F*-statistic is less than lower bound critical value, then we do not reject the null hypothesis of no long-run levels relationship. Conversely, if the computed *F*-statistic is greater than upper bound critical value, then we reject the null hypothesis and conclude that there exists steady state equilibrium between the variables. However, if the computed value falls within lower and upper bound critical values, then the result is inconclusive and knowledge of the order of the integration of the underlying variables is required before conclusive inferences can be made.

The most important trade data source for our analysis of ASEAN trade directions are collected from the *Direction of Trade Statistics Yearbook* of the International Monetary Fund (IMF). The transportation costs data to proxy the distance between the economies is from the Wan Hai Lines (M) Sdn. Bhd. Real GDP and GDP per capita data are from the *World Bank World Development Indicators Database* (*2002*).

4. Empirical Results and Discussions

From the results of the ARDL models of the computed *F*-statistics (Table 1), there exists a long-run relationship between the variables of the gravity equation are supported by all the bilateral trade flows among ASEAN-5, namely, Indonesia, Malaysia, Philippines, Singapore and Thailand data. For all countries, the computed *F*-statistic is greater than upper bound critical value, thus we reject the null hypothesis of no long-run levels relationship and conclude that there exists steady state equilibrium among the variables under study. For all the countries, we do not reject the null hypothesis of no residual correlation, ARCH and residual normality for each experiment from first order to fourth order. Meanwhile, we do not reject the null hypothesis of no Ramsey RESET problem for each experiment from first order to third order. The validity and reliability of the model, therefore, is confirmed as the model obeys the classical linear regression model (CLRM) namely, residuals are serially uncorrelated, homoscedasticity, normally distributed and the model is correctly specified.

The empirical results indicate that the overall effects of the size of economy – proxy by GDP – on bilateral trade flow are mixed and inconclusive. In the cases of bilateral trade flows from Malaysia to Indonesia, Malaysia to Thailand, Indonesia to Malaysia, Indonesia to Singapore, Philippines to Indonesia, Philippines to Thailand, Singapore to Thailand, and Singapore to Malaysia, the size of economy is obviously crucial in promoting the bilateral trade flows among the relevant countries significantly and negatively. Although the results are inconsistent with a priori, which implies a positive relationship between the economic size and trade flows, the findings support the Linneman's (1966) and Leamer and Stern's (1970) findings. According to their framework, an increase in economic size leads to less reliance on foreign trade because self sufficiency is achieved with an increasing economic size, thereby reducing trade flows with their trade partners.

On the other hand, the coefficient of GDP for Malaysia to Philippines, Indonesia to Philippines, Philippines to Singapore, Thailand to Indonesia, Thailand to Malaysia, Thailand to Singapore, and Singapore to Indonesia indicate significantly positive sign. This may support the hypothesis that trade increases with the size of the economy, as proven by Brada and Mendez's (1983) and Sharma and Chua (2000). This may suggests that a large domestic market promotes division of labour and thus creates opportunities for trade in a wide variety of goods. Although the coefficient for Malaysia to Singapore, Philippines to Malaysia, and Singapore to Philippines show positive sign, but are insignificant. These findings are in accord to Oguledo and Macphee (1994), in which indicated that the insignificance of the economic size variable might suggest that political and colonial ties are more important than economic sizes.

The income level (GDP per capita) has shown the significant and expected positive sign in the case of bilateral trade flows from Malaysia to Indonesia, Malaysia to Philippines, Malaysia to Singapore, Malaysia to Thailand, Indonesia to Malaysia, Indonesia to Singapore, Indonesia to Thailand, Philippines to Indonesia, Philippines to Malaysia, Philippines to Thailand, Singapore to Malaysia, Singapore to Philippines, and Singapore to Thailand. These results confirmed the idea that the higher the income level, the greater the production capacity and the larger the amount they can export as well as the greater the domestic consumption and import demand. Higher income level has also indicated a greater ability of the exporters to produce and export at a lower cost, all else remaining constant. The positive and significant effects of GDP per capita on trade flows reported in the study are in line with Linneman (1966), Aitken (1973), Geraci and Prewo (1977) and Brada and Mendez (1983).

With the exception, the coefficient of the GDP per capita for trade flows between Thailand to Indonesia, and Singapore to Indonesia show a significant negative sign, in which are not in line with theoretical expectation. This may be due to the collapse in the provision of credit caused by their poor balance sheets, weak levels of capitalization, and high levels of foreign debt exposure. With the increased cost of foreign debt service, firms were unable to service this debt or to obtain further credit. In addition, the heavy layoffs and increased unemployment have jeopardized the income level of labour. Thus, the widespread loss of purchasing power reduced the domestic consumption and import demand. Furthermore, the combination of higher Rupiah costs and the disappearance of short-term financing have deteriorated the import of the raw material. It has also brought about the fall in the exports of many industries. The political instability follows by the fall of the Suharto regime as well as the widespread social unrest during 1998, making the Indonesia's economy worse off. Whereas, the coefficient of GDP per capita from Thailand to Malaysia, and Thailand to Singapore that are also significantly negative sign may be explained by the Thailand economy that was seriously affected by the financial crisis impact.

The coefficient of distance variable (TC) for bilateral trade flows from Malaysia to Singapore, Malaysia to Thailand, Indonesia to Thailand, Philippines to Malaysia, Thailand to Indonesia, Thailand to Malaysia, Thailand to Philippines, and Singapore to Indonesia show significant and positive sign. This might be explained by the new technological advances in transportation and telecommunications, in which they have reduced the economic distances among the nations. The coefficient of distance for bilateral trade flows from Malaysia to Indonesia, Malaysia to Philippines, Indonesia to Philippines, Indonesia to Singapore, Philippines to Indonesia, Philippines to Singapore, Philippines to Thailand, Thailand to Singapore, and Singapore to Philippines are in the expected negative sign and significant. This means that distance is an important factor, which influent the volume of goods traded and is generally regarded as a significant barrier to trade. As Krugman (1991) revealed that distance tends to increase the cost of transacting international exchange of goods and services. The farther apart two potential trading partners are, the more costly their bilateral trade, in which it erodes possible gains from trade. The findings are expected because most of the countries in ASEAN are divided by water, and the port facilities are not that advance as compared to industrial countries.

The mixed results between transportation costs and trade, therefore, are not surprising. To some extent, these costs might be expected to depend on geographical factors and can be treated as an exogenous variable. Nonetheless, it is plausible to expect that they would also depend inversely on the development of transport and telecommunications infrastructure. Variations in transport costs across countries may be able to account for differences in their ability to compete in international markets. Thus, differences in the volume and quality of infrastructure across countries may be responsible for differences in transport costs, which in turn, may be able to account for differences in competitiveness (Bougheas, et al., 1999: pp. 170). Indeed, in a recent paper, Limao and Venables (2001) show that raising transport costs by 10% reduces trade volumes by more than 20%.

The significant and positive sign of dummy variable representing AFTA preferential (D1) of bilateral trade flows from Malaysia to Indonesia, Indonesia to Philippines, Indonesia to Thailand, Thailand to Indonesia, Thailand to Philippines Singapore to Indonesia, and Singapore to Philippines imply that the importance and prevalence of the formation of AFTA preferential arrangements in the intra-ASEAN trade flows. On the other hand, the significance negative sign of D1 of the bilateral trade flows from Malaysia to Philippines, Malaysia to Singapore, Malaysia to Thailand, Indonesia to Malaysia, Thailand to Malaysia, and Indonesia to Singapore may suggest that ASEAN as a whole does not gain collectively from the formation of AFTA. This may due to the existence of trade deflection in the regional market. Trade deflection will shrinks tariff revenue that those members would otherwise collect, in which tariff revenues are in effect redistributed in favour of low-tariff members as well as weakens the protective impact of high-tariff members' trade policies. It could be seen that the lower the average CEPT tariff rates, the higher the total trade amount except for the Philippines and Indonesia because of their serious macroeconomic difficulties, which were aggravated by political upheaval. For instance, by comparing the breakdown of intra-ASEAN and inter-ASEAN trade from the period of year 1981-2000, nearly half of ASEAN's intra- and inter-ASEAN trade originate in Singapore that has the lowest average CEPT tariff rates. Singapore bilateral pacts with non-ASEAN countries would result in massive trade deflection into the regional market. With the persistence of this unhealthy trend not only makes ASEAN as a whole might not gain collectively from AFTA, indeed all of them might be worse off.

The significant negative sign of the dummy variable representing Asian financial crisis (D2) of the bilateral trade flows from Malaysia to Indonesia, Malaysia to Singapore, Malaysia to Thailand, Indonesia to Philippines, Indonesia to Singapore, Philippines to Singapore, Singapore to Malaysia, and Singapore to Indonesia mean that crisis was reduced the bilateral trade flows. During the Asian Financial crisis, once Thailand floated the baht, countries lose competitiveness when their trading partners devalue. Therefore, they are more likely to be attacked and to devalue themselves. Currency devaluation increase competitiveness of a country's goods in international market, thus promoting exports and enhancing output growth. Whereas, the depreciation of domestic currency would increase the costs of imported inputs, eventually lead to contraction effects in the economy. Thus, during crisis, ASEAN countries increase their exports to developed countries while reduce their imports from neighbouring countries. Elliott and Ikemoto (2004) provide one explanation for the weakening of this negative import diversion effect after the Asian crisis. They argue that "… this situation may reflect the consensus that prior to this date that although ASEAN's success was based on its outward orientation, perceived problems of credibility and confidence in the region by the industrialized world meant that ASEAN countries were forced to turn inwards and to focus on their local markets." (pp. 17).

On the other hand, the dummy variable representing Asian financial crisis (D2) of the bilateral trade flows from Indonesia to Malaysia indicates a significant and positive sign. This may explain the upheavals of the Asian crisis did not affect the bilateral trade flows. However, the exceptional significant and positive sign of the coefficient D2 from Indonesia to Malaysia may be implied that the crisis has invoked the ASEAN sense of belonging among the member countries and thus, enhancing a greater level of economic co-operation. Furthermore, the contagion effect of the financial crisis has further demonstrated their high degree of economic inter-dependence, and therefore, highlighted the need for some policy co-ordination. For instance, the consolidation of Malaysia-Indonesia Trade and Economic relations through the revival of the Indonesia-Malaysia-Thailand Growth Triangle (IMT-GT) that have been stalled since the Asian financial crisis broke out in year 1997 by revisiting the basic regional co-operation strategies and redefining necessary support measures. Furthermore, there is an indication that Korea and Malaysia are among Asian countries experiencing the fastest path of recovery after crisis. Perhaps the significant positive sign might be due to the composition of trade (exports and imports) that are complimentarily rather than competitively to each other not affected by crisis.

5. Concluding Remarks and Policy Implications

Using a modified gravity equation, this paper examines ASEAN economic integration and its ability to promote intra-ASEAN trade, namely, Indonesia, Malaysia, Philippines, Singapore and Thailand within ARDL framework for each of the five ASEAN countries based on the annually data from 1970 to 2001. The analysis leads to the following conclusions. First, the effects of the size of economy on bilateral trade flow in ASEAN are mixed and inconclusive. Therefore, we could conclude that the size of economy can be either trade-enhancing or trade-inhibiting dependent on the country-specific. Second, an increase in the GDP per capita of the ASEAN country tends to increase the trade flows to the particular country. Third, the effects of the distance on bilateral trade flow in ASEAN are indeterminate. Distance can be either negatively or positively related to ASEAN trade flows, in which indicating the provision of improved accessibility that can result in reduced transport costs and thus reduced the economic distances among the nations. Fourth, although there exists importance and prevalence of the formation of AFTA preferential arrangements in enhancing intra-ASEAN trade, ASEAN countries may not as a whole benefit from the formation of AFTA. ASEAN countries as a whole will be worse off if trade deflection take place in the regional market. Finally, during crisis, ASEAN countries increase their exports to developed nation while reduce their imports from their neighbouring countries.

The outcomes of AFTA have not been as good as expected, the findings confirmed that policy makers of the ASEAN member countries should be more concerned in the effort of enhancing the co-operation in intra-ASEAN trade especially dealing with external shock collectively. With an understanding of the channel through which these factors influence trade flows, policy makers will be essential in suggesting and formulating appropriate policies in forming the future trade arrangement and the feasibility of a substantial expansion of regional trade. There are some issues derived from the findings should be concerned by relevant policy makers. First, countries should make efforts to reduce transaction costs between ASEAN regions to achieve a deeper economic integration since distance is among the main determinants of trade. Second, ASEAN countries need to improve their togetherness or strengthen their cooperative spirit through AFTA by identifying the strengths and weaknesses of each member country.

Third, if there is an increase in trade deflection as a result of the bilateral free trade agreement, it has to redesign and enforcement of AFTA's rules of origin requirement to make it difficult for imports from non-ASEAN members to be deflected through lowest tariff country. As suggested by Bagwell and Staiger (2004), the potential for opportunistic bilateral agreements is severe when there are no rules. They have proven that the welfare of a nonparticipating government is preserved in the presence of any bilateral agreement that respects the principles of reciprocity and nondiscrimination. As a consequence, they suggest that preferential tariff agreements, which are permitted under the special exception to most-favored-nation (MFN) granted by GATTs Article XXIV, represent a possible route to opportunistic bilateral agreements (pp. 24). Therefore, a more solid bilateral trade agreement should be recommended by the policy makers to ensure each ASEAN member will benefit from AFTA.

References

Aitken, N.D. (1973). The effect of the EEC and EFTA on European Trade: a temporal cross-section analysis. *American Economic Review*, 5, 881-892.

Anderson, J.E. (1979). A theoretical foundation for a gravity equation. *American Economic Review*, 69, 106-116.

Bagwell, K. & Staiger, R.W. (2004). Multilateral trade negotiations, bilateral opportunism and the rules of GATT/WTO. *Journal of International* Economics, 63, 1-29.

Beaulieua, E., Benarroch, M. & Gaisford, J. (2004). Trade barriers and wage inequality in a North-South Model with technology-driven intra-industry trade. *Journal of Development Economics*, 75, 113-136.

Beckerman, W. (1956). Distance and the pattern of Intra-European trade. *Review of Economics and Statistics*, 39, 31-38.

Bergstrand, J.H. (1985). The gravity equation in international trade: some microeconomic foundations and empirical evidence. *Review of Economics and Statistics*, 67, 474-481.

Bougheas, S., Demetriades, P.O. & Morgenroth, E.L.W. (1999). Infrastructure, transport costs and trade. *Journal of International Economics*, 47, 169-189.

Brada, J.C. & Mendez, J.A. (1983). Regional economic integration and the volume of intra-regional trade: a comparison of developed and the developing country experience. *KYKLOS*, 36, 92-105.

Clarete, R., Edmonds, C. & Wallack, J.S. (2003). Asian regionalism and its effects on trade in the 1980s and 1990s. *Journal of Asian Economics*, 14, 91-129.

Elbadawi, I. (1997). The impact of regional trade and monetary schemes on intra-subSaharan Africa trade. In A. Oyejide, I. Elbadawi and P. Collier (Eds.), *Regional integration and trade liberalization in SubSaharan Africa. Volume 1: Framework, Issues and Methodological Perspectives*, pp. 210-255, London: Macmillan Press.

Elliott, R.J.R. & Ikemoto, K. (2004). AFTA and the Asian crisis: help or hindrance to ASEAN intra-regional trade? *Asian Economic Journal*, 18, 1-23.

Endoh, M. (1999). International trade – economic aspects – Asia, Asia-Pacific Economic Cooperation (Organization). *Journal of Asian Economics,* 10, 571-589.

Engle, R.F. & Granger, C.W.J. (1987). Cointegration and error correction: representation, estimation and testing. *Econometrica*, 55, 251-276.

Evenett, S.J. & Keller, W. (2002). On theories explaining the success of the gravity equation. *Journal of Political Economy*, 110, 281-316.

Frankel, J. A. (1993). Is Japan creating a yen block in East Asia and the Pacific? In Jeffrey A. Frankel and Miles Kahler (Eds.), *Regionalism and rivalry: Japan and the United States in Pacific Asia*. pp. 53-85. University of Chicago Press, Chicago. (reprinted in R. Garnaut and P. Drysdale (Eds.), *Asia–Pacific Regionalism: readings in international economic relations*, pp. 227-249. Sydney: Harper Educational Publishers).

Geraci, V.J. & Prewo, W. (1977). Bilateral trade flows and transport costs. *Review of Economics and Statistics*, LIX, 67-74.

Granger, C.W.J. & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics,* 2, 111-120.

Gujarati, D.M. (2003). *Basic econometrics*. 4th Edition, Singapore: McGraw-Hill.

Hassan, M.K. (2001). Is SAARC a viable economic block? evidence from gravity model. *Journal of Asian Economics,* 12, 263-290.

Helpman, E. & Krugman, P.A. (1985). *Market structure and foreign trade*. Cambridge: MIT Press.

Krugman, P. (1991). The move toward free trade zones. In *Policy implications of trade and currency zones*: A symposium sponsored by the Federal Reserve Bank of Kansas City, Kansas City: Federal Reserve Bank of Kansas City, pp. 7-41.

Leamer, E.E. & Stern, R.M. (1970). *Quantitative international economics*. Boston: Allyn and Bacon.

Limao, N. & Venables, A.J. (2001). Infrastructure, geographical disadvantage and transport costs. *The World Bank Economic Review*, 15, 451-479.

Linnemann, H. (1966). *An econometric study of international trade flows.* Amsterdam: North-Holland.

Oguledo, V. I. & Macphee, C.R. (1994). Gravity models: a reformulation and an application to discrimination trade arrangements. *Applied Economics*, 26, 107-120.

Pesaran, M.H., Shin, Y.C. & Smith, R.J. (2001). Bound testing approaches to the analysis of level relationship. *Journal of Applied Econometrics*, 16, 289-326.

Plummer, M.G. (1997). ASEAN and the theory of regional economic integration. *ASEAN Economic Bulletin*, 14, 202-214.

Ramasamy, B. (1995). Trade diversion in an ASEAN free trade area. *ASEAN Economic Bulletin,* 12, 10-17.

Sanso, M., Cuairan, R. & Sanz, F. (1993). Bilateral trade flows, the gravity equation, and foundation form. *Review of Economics and Statistics*, 75, 266-275.

Sharma, S.C. & Chua, S.Y. (2000). ASEAN: economic integration and intra-regional trade. *Applied Economics Letters*, 7, 165-169.

Thornton, J. & Goglio, A. (2002). Regional bias and intra-regional trade in Southeast Asia. *Applied Economics Letters*, 9, 205-208.

Viner, J. (1950). The economics of customs unions. In *The customs union issue*, Chapter IV, New York: Carnegie Endowment for international Peace, pp. 41-81.

Notes

Note 1. The static effects of Preferential Trading Groups generally refer to onetime changes in the allocation of resources or efficiency.

Note 2. In order to estimate the overall national welfare effect, these two effects have to integrate together to obtain an overall picture of the economic impact of the regional bloc on a member country.

Note 3. Jacob Viner (1950) defined trade diversion as taking place whenever there is a shift in product origin from non-member producer whose resource costs are lower to a member country producer whose resource costs are higher. Trade creation is the net volume of new trade created by forming the trade bloc. It causes two kind of gain in an economy: gains on extra consumption of the product and gains on replacement of higher-cost production by lower-cost partner production.

Note 4. Other variables was entered into the equation include per capita incomes, and dummy variables indicating whether the pair share a common border or common language.

Note 5. Trade deflection refers to the import penetration of the ASEAN market by non-ASEAN countries through country that has the lowest tariff, and for subsequent re-exports to the rest of the ASEAN countries by utilizing the AFTA Common Effective Preferential Tariff (CEPT). Trade deflection will lead to greater inefficiency by stimulating production in the partner country in areas that has intra-regional comparative advantage but comparative disadvantage at the global level, in which will detriment the economic development prospects of its member states. If AFTA were agreed to be more collaborative but then led to greater trade deflection and inefficiency that are welfare reducing, this could threaten the development strategies of the ASEAN countries.

Note 6. The time frame for fully implementing AFTA has been set at 10 years for the more advanced economies of ASEAN and 15 years for the new ASEAN members.

Note 7. The gravity model has been using extensively to provide a natural framework and a useful multivariate approach for assessing the impact of regional trading blocs on the level and direction of bilateral trade flows. See Helpman and Krugman (1985) for the underlying theory of the model. Anderson (1979), Bergstrand (1985), Sanso, *et al*. (1993) and Evenett and Keller (2002) provide further discussion on the use of the gravity approach in trade literature. Moreover, Elbadawi (1997) reveals that gravity equation is superior in terms of its ability to incorporate the factors that cannot be accommodated by the conventional factor endowment theory such as intra-industry trade (by considering market size and scale economies) and the efficiency of resources reallocation following trade liberalisation (by introducing dummy variables).

Note 8. In order to avoid spurious regression, prior to conventional cointegration test such as the Engle and Granger (1987) bivariate test, one needs to conduct unit root tests to make sure that all system's variables are integrated of the order (Granger and Newbold, 1974; Gujarati, 2003).

Table 1. The results of bilateral trade between ASEAN-Five based on Equation (2)

Trade Flows from Malaysia

Variable	**Indonesia**	**Philippines**	**Singapore**	**Thailand**
RGDP*RGDP	-17.479***	0.410**	0.364	-18.468***
PGDP*PGDP	24.107***	2.143***	1.382*	28.215***
TC	-7.363***	-0.537***	0.645**	4.533**
D1	0.901**	-0.075**	-0.120***	-0.276*
D2	-0.209*	-0.002	-0.072*	-0.315*
Adjusted R^2	0.611	0.801	0.793	0.636
Computed F statistic	8.613***	17.537***	6.661***	4.704**

Trade Flows from Indonesia

Variable	**Malaysia**	**Philippines**	**Singapore**	**Thailand**
RGDP*RGDP	-11.405**	4.586**	-49.930***	-12.703
PGDP*PGDP	25.425***	1.716	64.292***	23.688*
TC	-2.416	-3.891***	-22.635	6.817*
D1	-3.125***	0.154***	-0.641**	0.998*
D2	2.755***	-0.112	-0.979***	0.077
Adjusted R^2	0.971	0.965	0.865	0.762
Computed F statistic	59.988***	95.498***	19.812***	6.844***

Trade Flows from Philippines

Variable	**Indonesia**	**Malaysia**	**Singapore**	**Thailand**
RGDP*RGDP	-20.712***	0.257	1.065***	-6.693**
PGDP*PGDP	29.546***	3.413***	0.247	13.416**
TC	-3.307**	0.896*	-1.885***	-2.203**
D1	0.018	-0.055	-0.001	-0.145
D2	0.009	-2.246	-0.251	-0.049
Adjusted R^2	0.904	0.737	0.900	0.766
Computed F statistic	11.073***	11.149***	15.890***	5.178**

Trade Flows from Thailand

Variable	**Indonesia**	**Malaysia**	**Philippines**	**Singapore**
RGDP*RGDP	27.904***	1.829	-2.128	6.334**
PGDP*PGDP	-38.603***	-0.794	-4.404	-7.672**
TC	28.297***	2.674**	2.602**	-1.285**
D1	5.955***	0.093	0.240*	-1.285
D2	-0.203	-0.254**	-0.120	0.005
Adjusted R^2	0.877	0.755	0.746	0.834
Computed F statistic	8.061***	5.310**	15.052***	11.793***

Trade Flows from Singapore

Variable	**Indonesia**	**Malaysia**	**Philippines**	**Thailand**
RGDP*RGDP	13.745**	-1.592***	0.401	-2.598**
PGDP*PGDP	-17.745**	3.842***	1.815**	5.157***
TC	2.489**	-0.012	-0.365*	0.529
D1	0.464**	0.001	0.132***	0.132***
D2	-0.390**	-0.029*	-0.034	-0.019
Adjusted R^2	0.786	0.865	0.889	0.774
Computed F statistic	14.333***	12.905***	17.557***	6.814***

Lower and Upper Bound Critical Value of F-Statistic in Bound Test for

Critical values (F test)	Lower bound	Upper bound
10%	2.72	3.77
5%	3.23	4.35
1%	4.29	5.61

Note: The LR Coeff. refers to the long-run coefficient . The next column contains the Jarque-Bera statistic for residual normality. The figure in the next column B-G is the Breusch Godfrey Serial Correlation LM test. The final two columns contain ARCH LM Test and Ramsey RESET Stability Test. The Figures in parenthesis beside the F-statistic are the probability of the test. The upper and lower bounds for critical values refers to the distribution of the F-statistic under the null hypothesis when the variables are individually I(1) and I(0) respectively.

Source: Pesaran et al. (2001, p.300)

Table CI (iii) Case III: Unrestricted intercept and no trend

The asterisks indicate the following levels of significance: *10%, **5%, ***1%.

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

A Literature Review on the Relationship between Foreign Trade and Economic Growth

Huan Chen
School of International Business
Southwest University of Finance and Economics
NO.55, Guanghuacun Street, Qingyang District
Chengdu 610074, China
E-mail: chenhuan_369@126.com

Abstract

Over the past years, the relationship between foreign trade and economic growth has been the debate of economic research in academe. On this issue, economists at home and abroad used the relative data of china and got different conclusions by different methods. At first, this thesis reviews the theories of the relationship between foreign trade and economic growth, and then sum up the main arguments of modern empirical economics. Finally, the author makes a brief comment and put forward some questions that should be explored in depth in this area. It provides a broader and more comprehensive perspective to the later researchers.

Keywords: Causality test, Technology spillover, Least squares, Regression analysis

The relationship between economic growth and foreign trade was focused by many economists when foreign trades come into being. With the development of foreign trade, it has been the debate of economic research in academe because of its impact on economic growth. To some extent, the emergence of foreign trade and its development were closely related with economic growth. In a way, foreign trade indeed promotes economic growth of a country.

1. The Review of the Theory on the Relationship

1.1 The Standpoint of Marx

In accordance with the Marxist point, the relationship between the two was the relationship between exchange and production in the final analysis. Marx said: "the depth, breadth and the way of exchange are decided by the development and structure of production. … We can see that all the elements of exchange are included in the production directly, or are decided by it." Essentially, production decides the exchange, but the exchange which is a stage of the exchange, is not merely decided by it and could react to produce under certain conditions. Sometimes, the counteractive of promoting or inhibiting is tremendous. On one hand, the expansion of production needs a growing market; on the other hand, the growing market will promote the expansion of production continuously. So, production and exchange, affect each other every seconds. This provides a very important revelation to many economists.

1.2 The theory of western schools on the relationship

1.2.1 The classical school of economics

They believed that foreign trade promoted economic growth in two ways. On the one hand, foreign trade improved the optimal distribution of resources and productivity consequentially and then stimulated the economic growth; on the other hand, one country could gain raw materials and equipments which it could not produce. Those provided the material basis for economic development. The most famous theories were exports of surplus of Adam Smith, comparative advantage of David Ricardo, the interests of the trade development of John Mueller and "trade is the engine of economic growth" of D • H • Robert Morrison. All these theories interpreted the relationship to some extent but ignored that the international environment is complex and ruleless.

1.2.2 The structure school

The representative is Lewis, who put forward dual economy model which parted a developing economy into capitalist part (the industry sector) and non-capitalist part (the traditional agricultural sector). The capitalist sector was bound to promote the growth of the economy through absorbing and accumulating surplus labor from non-capitalist sector. If the capitalist part produced the exporting goods and the traditional part produced the importing goods, foreign trade would undoubtedly expand the market and demand of products in capitalist part and reduce the wages of labor. Then it would further increase the profit and accumulation of the part and promote economic growth.

1.2.3 The effect school

The main point of Max Corden was that he analyzed foreign trade together with macro-economic variables and especially emphasized the impact of trade on the supply of production factors and productivity. Corden recognized that a country's foreign trade would affect macroeconomic from 5 aspects: the revenue effect, the effect of capital accumulation, the substitution effect, the income distribution effect and the effect of the weighted elements. All the above effects were cumulated which meant that the impact of trade on economic growth was strengthened gradually as the development of economy.

1.2.4 The new-growth school

Romer, Lucas and Svensson, the representatives of this school, took technology as the core factors to promote productivity. This theory pointed out that the growth of developed countries would be attributed to the improvement of productivity. Based on this fact, the theory made a series of models to study the relationship among international trade, technological progress and economic growth. They viewed that international trade could promote economic growth through technology spillover and external stimulation. On one hand, any technology had a spill-over process. The owners of advanced technologies, whether they had intention or no intention, would gradually make other countries learn these technologies through foreign trade; on the other hand, international trade provided a broader market, more frequent exchange of information and increased competition, which forced every country to develop new technologies and products. The mutual promotion relations between international trade and technical change could ensure a long-term economic growth.

1.2.5 The new- trade school

The new-trade theory economist, Haierpoman Paul Krugman believed there were two ways for international trade to promote economic growth. One was the effects of economies of scale brought by trade, and the other was that international trade could promote economic growth through improving the optimal allocation of resources between materials production sector and knowledge production sector.

2. The Main Arguments of Modern Empirical Economics

According to the general understanding of macroeconomics, import is often recognized as a leakage of revenue which will lead to unemployment rather than economic growth. Based on this assumption, the research on relationship between economic growth and foreign trade can be taken for the study of the relationship between growth and export. That means to prove the assumptions of export-led economic growth.

2.1 The main conclusions of western school

2.1.1 The conclusion of export having a single causal relationship with economic growth

Kwan and Cotsomitis (1991), Kwan and Kwok (1995) took use of Granger causality test to study Chinese growth and foreign trade. They come to the conclusion that the output was an exogenous variable and there was a one-way causal relationship between the two. Lee added other variables, such as the trend of time, FDI and the lagging investment and so on, and concluded that export promoted economic growth by using AD.HOC model and regression analysis. But he also found the result was affected by regional differences. Behzad Yaghmaian verified the assumption of export led economic growth using time-series data and regression analysis. He found that employment and output of manufacturing sector could promote export and economic growth. In the model of classical economics, if we took the statistics of population as labor force, export marked the leading role to economic growth, but it could not be reversed.

In these empirical studies, economists used ordinary least squares (OLS) to test these cross-regional or cross-section data, and the results generally support the promotion of export to growth. But its reliability is questionable: the results from OLS only showed the relevance between foreign trade and economic growth but could not explain the existence of a causal relationship between the two.

2.1.2 The conclusion of economic growth having a single causal relationship with export

In this problem, Ghartey (1993) made an interesting conclusion. He analyzed respectively the economic data of United States, Japan and Taiwan through cross-section data, and found that the American GDP promoted its export, but Taiwan is quite the opposite and there was a two-way causal relationship between the two in Japan. Sharma & Dhakal (1994) also reached the same conclusion.

2.1.3 The conclusion of there was a two-way causal relationship between the two

Jordan Shan and Fiona Sun (1998) took data from 1987 to 1996 as the samples and tested the two-way causal relationship between the two. They used 6 variables, which were output, export, import, investment, labor and energy consumption, to establish VAR (VAR) model. They make ADF single test and Granger causality test, and then found the result.

2.1.4 The conclusion of there was no causal relationship between the two

Jung and Marshall (1985) analyzed the relationship between GDP and export of 37 developing countries in 1950-1981, and found that there was no causal relationship except Israel.

Though these empirical studies have different conclusion, they are not repellent mutually, because different researchers used different methods. Moreover, so many researches have a great inspiration to scholars at home.

2.2 The main conclusions of domestic economic school

Chengxiang Shen (1999) used Granger causality test and co-integration test to test the hypothesis of "Chinese export-oriented economic growth" by the data of export and GDP in china from 1977 to 1998. He found that there was a two-way causal relationship between the two, but no long-term and stable relations. The studies took only the output and export into account, but import weighed with the export - output association (Grossman, G • M & E • Helpman, 1991), so the impact of import on economic growth should not be ignored. China is a developing country and its endowments elements vary considerably with developed countries. The developed countries mainly have advantage in capital and technology, but china mainly in natural resources and labor. This determines the major export products of china are agricultural products and low value-added products, and its import products are high-tech products. It can be said that the import is an important mean to break the bottleneck of economic development and promote economic growth. Meanwhile, the imported products will encourage domestic enterprises to improve product quality and production efficiency, and promote the upgrading of traditional industrial structure. As a result, the research on the relationship of import and economic growth is necessary. Jiadong Tong (1995) explored the relationship between economic growth and import, and he recognized that import at different times contributed to economy differently, but on the whole, there was a positive correlation between import and economic growth. Jingwen Li (1996) made an empirical analysis on economic growth model and pointed out that export boosted economic growth. FuWei Peng (1999) found that net exports had less relevance with economic growth. JiaQin Chen (1999) viewed that export had a great role in promoting economic growth. QuanFa Yang (1999) made Chinese data into Balassa model and found that export had a positive correlation with economy. XiaoPeng Liu (2001) started from the relevance of foreign trade and GDP growth rate and revealed that import had a strong role in the promotion of national economy by analyzing the data of china from 1980 to 1998. He also explained why Chinese export had weak correlation with economy from the angle of export structure. BoNai Fan, XiaoChi Mao, Shuang Wang (2005) studied the relationship through Granger causality model and broad difference approach, using Chinese statistics from 1952 to 2003. The result showed that Chinese GDP and export had a clear one-way causal relationship. That is to say, the export was an important factor to promote Chinese economic growth.

3. A Simple Comment on This Issue

3.1 Using econometric methods to explain the relationship between foreign trade and economic growth expands the vision of studying in this area, and has a profound impact on economics.

3.2 Taking foreign trade and other factors as independent variables to explore the relationship qualitatively between the two can not only explain the relationship accurately, but also make the affecting extent of independent variables on dependent variables clear. This is useful for a country to establish its foreign trade policy.

3.3 A few problems should be explore deeper

3.3.1 The general overlooked some important third variables, such as exports, investment and government spending.

3.3.2 The variables of many models is defined vaguely, and a possible common stochastic trend of variables has been overlooked or not been tested correctly.

3.3.3 The existing studies paid less attention to endogenous association of variables coming from the statistical process, for example, the GDP or GNP, on behalf of the economic growth, has already include the data of export.

3.3.4 The annual data was often chose by most scholars, but the quarterly data may be more suitable for measuring the causality of variables

3.3.5 The past studies over-relied on a specific study or a single method to deduct, and was lack of diversity and relativity.

References

Andy C. C & KwanJohn A. (1991). Cotsomitis. Economic Growth and the Expanding Export Sector: China 1952-1985, *International Economic Journal*, 5(1): 105 – 116.

Ghartey. E. E. (1993). Casual Relationship between Exports and Economic Growth: Some Empirical Evidence in Taiwan, Japan and the US, *Applied-Economics.* (9): 1145-1152.

Jordan Shan & Fiona Sun. (1998). On The Export-led Growth Hypothesis: The Econometric Evidence From China.

Applied Economics, (30).

Jung, S**.** W**.**, Marshall. Exports (1985). Growth and Causality in Developing Countries. *Journal of Development Economics,* (18):1-12.

Kwan A. C. C & Cotsomitis J. A. Kwok. (1996). Economic Growth and Exogeneity: Taiwan 1953-88. *The American Economist,* 28(4):467-471(5).

Lee, J. W. (1993). International trade, distortion and long-run economic growth, International monetary fund staff paper, 40(2):299-328.

Shan, J. & F. Sun. (1998b). On the Export-led Growth Hypothesis: The econometricEvidence from China. *Applied Economics,* 30: 1055-1065.

Shan, J. & F. Sun. (1998a). Export-led Growth Hypothesis for Australia: An Empirical Reinvestigation. *Applied Economics Letters,* 5:423-428.

Sharma, S. C. & Dhakal, D. (1994). Causal Analysis between Exports and Economic Growth in Developing Countries. *Applied Economics,* 26:1145- 1158.

Yaghmaian. B & Ghorashi. R. (1995). Export Performance and Economic Development: An Empirical Analysis. *The American Economist,* 37-45.

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Import Inflows of Bangladesh: the Gravity Model Approach

Md. Mahmudul Alam (Corresponding author)
CRM, Marketing Division, Grameenphone Ltd.
47 Shantinagar, Dhaka-1217, Bangladesh
Tel: 880-1711-503-782 E-mail: rony000@gmail.com

Md. Gazi Salah Uddin
School of Business, Presidency University
11/A, Road-92, Gulshan-2, Dhaka 1212, Bangladesh
Tel: 880-1715-201-727 E-mail: salahuddin@presidency.edu.bd

Khan Md. Raziuddin Taufique
School of Business, Presidency University
11/A, Road-92, Gulshan-2, Dhaka 1212, Bangladesh
Tel: 880-1711-348-223 E-mail: raziuddint@presidency.edu.bd

Abstract

Bangladesh suffers from a chronic deficit in her trade balance. The paper is an attempt to explore the imports of Bangladesh which is one of the most significant factors responsible for unfavorable trade balance of the country. The aim of the study is to intend some initiatives for an attempt to ultimately reshaping the trade balance of Bangladesh with her foreign trade partners. The paper examines the existence of the gravity theory for the imports of Bangladesh with its eight major trading partner countries- India, China, Singapore, Japan, Hong Kong, South Korea, USA and Malaysia. The data set consists of yearly data from 1985 to 2003 in a panel approach. The paper comes across with the findings that the gravity theory is consistent with the imports of Bangladesh. That is, the geographical distance of Bangladesh with its partner countries has significant impacts on its imports. But in near future this may change because of different factors such as profitability, easy trade procedures, product delivery time etc. that influence the imports decision more than does the geographical distance. This paper finds mixed relationship between the GDP and imports of Bangladesh. It also shows that the imports of Bangladesh influence the domestic production very little because Bangladesh mostly imports consumer goods rather than capital goods. Moreover, the population of Bangladesh has significant impacts on imports which in turn implies that Bangladesh is not capable of producing adequate consumer goods to meet the increased demand resulted from high population growth. It also shows that partner countries' GDP has significant positive impacts and partner countries' population has mixed impact on imports of Bangladesh. This paper concludes that it will be an alarming situation for trade balance of Bangladesh if the imports continue to increase in such a pattern that the rate increases five to eight times more in respect of population increases and at the same time the ratio of capital goods in proportion of total imports decreases.

Keywords: Gravity Model, Imports, Gross Domestic Product (GDP), Population, Geographical Distance, Panel

1. Introduction

In the past, national economies were relatively self-contained entities, isolated from each other by barriers to cross-border trade and investment; by distance, time-zones, and language; and by national differences in government regulations, culture, and business systems. As these barriers are being minimized by taking different initiatives, the contribution of cross-border trade to the development of national economy is becoming significant. The economic development of Bangladesh is also, to a significant extent, affected by the trade flows with partner countries. Although Bangladesh suffers from a chronic deficit in her trade balance (Table-1), scope might be found to explore the most significant factor(s) responsible for unfavorable trade balance with a view to take some initiatives for an attempt to ultimately reshaping the trend of Bangladesh's trade balance with her foreign trade partners.

Import is one of the moist important factors responsible for trade balance. Due to trade deficit where import is higher than export, the imports of Bangladesh need to look from different dimensions. While analyzing on imports, the import trading partners also get focus. In addition, recently the imports of Bangladesh also have started to move from border country to other countries. Though Bangladesh shares three of her borders with India and none with China, China has emerged as the largest source of import for Bangladesh replacing India for the first time in 2006 (Table-2). The value of imported amount of Bangladesh was 13.75b taka from China and 12.41b taka from India in 2005-2006 fiscal, where in 2004-2005 it was 12.33b taka from India and 6.94b taka from China.

The gravity model is a popular formulation describing the trade flows between different geographical entities. The theory says that there is negative relationship between geographical distance and trade volume which indicates that the lower distance influences for more trade and vice versa. As the imports of Bangladesh are geographically shifting, this paper aims at analyzing the gravity theory for Bangladesh. Moreover, this paper also finds out the impacts of home GDP, partner countries' GDP, home population, and partner countries' population on the imports of Bangladesh.

2. Literature Review

The empirical studies by Tinbergen (1962) and Linnemann (1966) showed trade flows follow the physical principles of gravity: two opposite forces determine the volume of bilateral trade between countries - the level of their economic activity and income, and the extent of impediments to trade. The latter includes, in particular, transportation costs, trade policies, uncertainty, cultural differences, geographical characteristics, limited overlap in consumer preference schemes, regulatory bottlenecks, etc. National borders are among these impediments, even for industrialized countries (Anderson and Van Wincoop, 2003).

Although trade potential primarily results from the matched export capacities and import demand (microeconomic), on a more aggregated level of analysis, proximity in demand, in per capita income, in space, and in culture, are the potential macroeconomic determinants of export. Thus various combinations of macroeconomic variables, such as gross domestic product (GDP) and population with geographic distance, are significant predictors of trade potentials. Hence, gravity equations have been used extensively in the empirical literature on international trade (Havrylyshin and Pritchett, 1991; Frankel and Wei, 1993; Bayoumi and Eichengreen, 1997; Evenett and Hutchinson, 2002).

Filippini (2003) used a gravity Equation model to analyze trade flows between East Asian industrializing countries (including China) and some developed countries in order to show the remarkable trade performance of East Asian countries. The study showed that all coefficient signs were consistent with model assumptions and found high propensity of Asian countries (including China but excluding Japan) to exchange high-tech manufactured products with Japan and USA. Another interesting result was that among the East Asian economies, China plays a very important role as an exporter and as an importer too in recent years.

Martinez-Zarzoso (2003) applied the gravity model to annual bilateral exports between 19 countries. His results indicated that incomes of exporter and importer, as expected, had positive influence on bilateral trade flows. Population had a large and positive impact on import, indicating that bigger countries import more than do small countries. Regarding transport infrastructure, he found that exporting countries' infrastructure fosters trade. Achay. (2006) investigated the determinants of trade flows between various countries of the world. He applied the gravity model on a sample of 146 countries for five-year sub-periods between 1970 and 2000. According to this study, GDP, distance and regional integration agreements are the determinants of trade. It is also found that all estimated coefficients were statistically significant and their signs were in conformity with expectations. The adjustment quality of the model as measured by determination coefficient (adjusted R^2) was quite high, standing at 71 percent. It was found that GDP, GDP per capita, common frontier, common official language, common currency or common colonial past have a positive impact on the volume of bilateral trade. On the other hand, the geographical distance had a negative impact on the volume of trade.

Very few studies on Bangladesh have been conducted so far in this area, which is perhaps due to unavailability of information, particularly regarding trade inflows and outflows with partner countries. Rahman (2004) applied a generalized gravity model to analyze Bangladesh trade flows with its trading partners using the panel data estimation techniques. The study estimated the gravity model of trade (sum of imports and exports). The results showed that Bangladesh's trade is positively determined by the size of the economies, per capita GNP (gross national product) differential of the countries involved and openness of the trading countries. The major determinants of Bangladesh's exports were found to be exchange rate, partner countries' total import demand, and openness of the economy. All these factors affected Bangladesh's exports positively whereas transportation cost was found to be a significant factor in influencing trade negatively.

From the extensive literature, gravity equations share common features that can be customized for different purposes: First, a gravity equation is bilateral; Secondly, gravity equations can be derived from various theoretical trade models

(Deardorff, 1995) and finally, a gravity equation may be used in order to estimate either determinants of the volume or determinants of the nature of trade flows.

3. Data and Model Specification

Classical gravity models generally use cross-section data to estimate trade effects and trade relationships for a particular time, for example one year. In reality, however, eight cross-section data observed over several periods (panel data methodology) result in more useful information than does cross-section data alone. There is couple of advantages of this method. First, panels can capture the relevant relationships among variables over time. Second, panels can monitor unobservable trading-partner-pairs' individual effects. If individual effects are correlated with the regressors, OLS estimates omitting individual effects will be biased. Therefore, the paper used panel data methodology for testing empirical gravity.

This panel study covers yearly data for Bangladesh and its eight major partner countries from 1985 to 2003; a balance panel of total 684 observations. The countries - India, China, Singapore, Japan, Hong Kong, South Korea, United States of America (USA) and Malaysia- are chosen based on the share of total import volumes of Bangladesh. These eight countries cover most of the imports of Bangladesh (Table-3). Bangladesh imported 64 percent in 1996, 60 percent in 2001 and 2002 of its total import from these eight countries.

<< Table 3 >>

Data for total import flows of Bangladesh (million US$) have been taken from Economic Review of Bangladesh. Then it has been converted into real value by adjusting with the US consumer price index (constant year 2000). GDP data are also in real term (constant 2000 million US$) that have been taken from World Development Indicators (WDI-2007) database. Population data (in million) were also taken from World Development Indicators (WDI-2007) database. Geographical distance in kilometers has been gathered from a web based database (www.timeanddate.com). All of these data have been converted into logarithmic (ln) format. For the statistical output, the SAS has been used and for the panel regression, the TSCSreg procedure has been used.

In 1687, Newton proposed the "Law of Universal Gravitation." It held that the attractive force between two objects *i* and *j* is given by,

$$Fij = G\,(MiMj / D^2ij) \tag{1}$$

Where notation is defined as follows, *Fij* is the attractive force, *Mi* and *Mj* are the masses, *Dij* is the distance between the two objects and *G* is a gravitational constant depending on the units of measurement for mass and force.

This model originates from the Newtonian physics notion. Newton's Gravity Law in mechanics states that two bodies attract each other proportionally to the product of each body's mass (in kilograms) divided by the square of the distance between their respective centers of gravity (in meters). The gravity model has been applied to a wide variety of goods and factors of production moving across regional and national boundaries under different circumstances since the early 1940s (Oguledo and Macphee, 1994).

The gravity model for trade is analogous to this law. The analogy is as follows: "the trade flow between two countries is proportional to the product of each country's 'economic mass', generally measured by GDP, each to the power of quantities to be determined, divided by the distance between the countries' respective 'economic centers of gravity', generally their capitals, raised to the power of another quantity to be determined" (Christie, 2002). This formulation can be generalized to,

$$Mij = K\, Yi^{\beta_1}\, Yj^{\beta_2}\, Dij^{\gamma}\, pop_i^{\delta_1} pop_j^{\delta_2} \tag{2}$$

The linear form of the model is as follows:

$$\ln(Mij) = K + \beta_1 \ln(Yi) + \beta_2 \ln(Yj) + \gamma \ln(Dij) + \delta_1 \ln(POPi) + \delta_2 \ln(POPj) \tag{3}$$

The basic statistical model is specified in the following logarithmic form in which five variables are assumed to determine the import flows to Bangladesh from its major trading partners:

$$\ln(Mij) = \alpha + \beta_1 \ln(Yi) + \beta_2 \ln(Yj) + \gamma \ln(Dij) + \delta_1 \ln(POPi) + \delta_2 \ln(POPj) + u_t \tag{4}$$

where, *Mij* is the flow of Bangladesh's *(i)* imports from partner country *(j)*, *Yi* is Bangladesh's GDP, *Yj* is the partner countries' GDP, *Dij* is the geographical distance between Bangladesh's *(i)* capital city to partner countries' *(j)* capital city, *POPi* is Bangladesh's population, *POPj* is the partner countries' population, *u* is the random error term and *β, γ, δ* are the structural parameters.

Here, this paper also observes the result for the basic model excluding the population variable from Equation no.-4, because there are few very highly populated countries such as China, India, and USA are available in partner countries' list that may impact the overall results. In this case the model is:

$$\ln(Mij) = \alpha + \beta_1 \ln(Yi) + \beta_2 \ln(Yj) + \gamma \ln(Dij) + u_t \tag{5}$$

The general hypothesis of the model is that the economic size of the importing countries, Mij, are usually measured with GDP. The estimated coefficients are usually close to the predicted value of one. If GDP of a country increases, the country is expected to import more from foreign countries. The same holds true for the partner countries that are exporters. So, both β_1 and β_2 are expected to be positive.

Distance is usually measured using the “great circle” formula. This formula approximates the shape of the earth as a sphere and calculates the minimum distance along the surface. As the distance between countries becomes larger, imports will fall. The distance is a factor, which is used as a proxy to consider the impact of transport costs and other transaction costs. Hummels, et al. (2007) have argued that shipping costs (freight charges and marine insurance) can go a long way towards explaining why distance matters. Distance indicates the time elapsed during shipment. For perishable goods, the probability of surviving intact is a decreasing function of time in transit. In gravity model, distance is a resistance factor and has a negative impact on volume of bilateral trade. As a result, γ is expected to be negative.

The coefficient of population of the exporters may have negative or positive sign depending on whether the country exports less when it is big (absorption capacity) or whether a big country exports more compared to a small country (economies of scale). For similar reasons, the coefficient of importer population may have negative or positive sign (Martinez-Zarzoso and Nowak-Lehmann, 2003). Therefore, the expected signs of $\delta 1$ and δ_1 may be negative or positive.

4. Results of Analysis

As the data has been taken from the countries of different regions and different economy sizes, here pooled OLS is not used. Moreover, geographical distance is fixed and not time sensitive. When working with constant data over time, time effect provides no output. Due to time constant, variable included in the equation-4 and equation-5, fixed effect model does not provide a good output. Moreover, *m*-value of Hausman test for equation-4 is insignificant, even in 90 percent confidence level which indicates to use fixed effect model. However, *m*-value of Hausman test for equation-5 is significant at 99 percent confidence level which indicates to use random effect model.

<< Table 4 >> << Table 5 >> << Table 6 >>< < Table 7 >>< < Table 8 >>

<< Table 9 >> << Table 10 >>< < Table 11 >>< < Table 12 >>< < Table 13 >>

Domestic GDP has no significant impact on the imports of Bangladesh both in random effect and fixed effect model for the qeutions-4. However, the GDP of Bangladesh shows significant positive relationships with imports of Bangladesh both in random effect and one way fixed effect model for the qeutions-5. The coefficients of determination (R^2) of random effect models for both cases indicate that 80 percent of the total variation in the dependent variable is accounted for by the independent variables that indicate good fit of the model. However, a qualitative focus can give an in depth idea about the study. Imports criteria show that it influences the domestic production very little because Bangladesh mostly imports consumer goods rather than capital goods. One third of total imports of Bangladesh were capital goods in 2000, but in recent years the rate declines to 9 percent in 2004-05, 11 percent in 2005-06, and 12 percent in 2006-07 (Table-14). As a developing and trade deficit country it is alarming for Bangladesh that the rate of total imports is increasing but the ratio of capital goods is decreasing.

<< Table 14>>

Due to time irresponsible fixed value, the fixed effect models do not provide any output for geographical distance variable. The random effect models show significant negative impacts of geographical distance on the imports of Bangladesh, which is accepted at the 99 percent confidence level for both equation-4 and equation-5. It is noted that here the sign of the coefficient is negative, means less geographical distance influence for higher import and vice versa. This result supports the theory of gravity model. However, recently Bangladesh has started to move its trade from lower distance country to higher distance country. The model of gravity mainly considers the geographical distance for trade, but the reality is more complex that might cause the inconsistency of this result in near future. There are lots of factors available that go beyond the assumption of gravity theory, such as, Bangladeshi importers are interested in China because it is offering goods with a wide price range, easy trade procedures, many duty free goods, shorter timing of shipment etc. China usually sends its products to Bangladesh within 25 days on an average by completing all the formalities while import from India takes 35 days. Many Chinese products, such as, food stuff, plastic goods, rubber goods, wooden products, raw hides, pulp, papers, electrical machinery and equipment and parts thereof, sound recorders and reproducers, television accessories of electronics articles, boilers, machinery and mechanical appliances, parts thereof, Cotton, yarn/thread and cotton fabrics etc., are cheap compared to the same quality products of India.

Both random effect and one-way fixed effect models show that domestic population has significant positive impacts on the imports of Bangladesh (equation-4). At 96 percent confidence level, one-way fixed effect model shows that 1 percent increases in domestic population will lead to 5.65 percent increases in imports of Bangladesh. And at 99 percent confidence level random effect models show that 1 percent increases in domestic population will lead to 7.19 percent increases in imports of Bangladesh. That means Bangladesh is not capable to produce enough in proportion of increasing demand due to population growth. As a result, demand for import increases five to eight times of the rate of its population increases.

Partner countries' GDP has significant positive impacts on the imports of Bangladesh both for equation-4 and equation-5. If Bangladesh imports more, the partner countries will produce more to meet the demand of Bangladesh. Fixed effect model shows no impact of Partner countries' population on Bangladeshi imports, where random effect models show significant negative impact.

5. Conclusion

Recently trade of Bangladesh is decreasing with its border countries and trade volume is increasing with far distance countries, such as, the case of China and India. Therefore, to find out the impacts of gravity model on the trade patterns of Bangladesh, its import is tested on home GDP, partner countries' GDP, geographical distance between home capital city and partner countries' capital city, home population, and partner countries' population.

The paper finds mixed results for the impact of Bangladeshi GDP on its imports. If population is not considered, GDP shows positive relationships with imports of Bangladesh. In a quantities focus, imports criteria show that it influences the domestic production very little because Bangladesh mostly imports consumer goods rather than capital goods. Further, population of home country has highly significant impact on imports of Bangladesh. It in turn implies that Bangladesh is not capable of producing enough in proportion of increasing demand due to population growth. Partner countries' GDP has significant positive impacts on the imports of Bangladesh and partner countries' population has mixed results on imports of Bangladesh.

Most importantly it is found that geographical distance has significant impact on imports of Bangladesh which means transport costs and other transaction costs, such as, the probability of surviving intact of perishable goods etc. still have significant impacts on its import. But in recent phenomenon, the imports are more influenced by profitability, easy trade procedures, product delivery time etc., rather than the geographical distance. As a consequence, policy makers need to conduct further serious studies to find out the relationship, if any, between trading pattern, geographical distance and trade deficits of Bangladesh. Policy makers also need to carefully consider the future alarming situation for trade balance of Bangladesh if the import increases continuously in such a pattern that the rate increases five to eight times more in respect of population increases and at the same time the ratio of capital goods in proportion of total imports decreases.

References

Achay, L. (2006). *Assessing Regional Integration in North Africa.* National Institute of Statistics and Applied economics, Rabat, Morocco.

Anderson, J. E. and Wincoop, E.v. (2003). Gravity with Gravitas: A Solution to the Border Puzzle. *American Economic Review*, 93(1), 170-192.

Bayoumi, T. and Eichengreen, B. J. (1997). Is regionalism Simply a Diversion? Evidence from the Evolution of the EC and EFTA. Takatoshi Ito and Anne O Krueger (eds.), *Regionalism vs. Multilateral Arrangements*. Chicago: The University of Chicago Press.

Bangladesh Bank (2004-2005, 2005-2006). *Import Payments (annual)*. Department of Public Relations & Publication, Dhaka, Bangladesh.

Bergstrand, J. H. (1989). The Generalized Gravity Equation, Monopolistic Competition, and the Factor-Proportions Theory of International Trade. *Review of Economics and Statistics*, 71, 143-153.

Christie, E. (2002). Potential Trade in Southeast Europe: A Gravity Model Approach. *Working Paper No. 21.* The Vienna Institute for International Economic Studies-WIIW.

Deardorff, A. V. (1995). Determinants of Bilateral Trade: Does Gravity Work in a Neoclassical World?. *Working Paper No. W5377.* National Bureau of Economic Research (NBER).

Evenett, S. J. and Hutchinson,W. K. (2002). The Gravity Equation in International Economics: Theory and Evidence. *Scottish Journal of Political Economy*, 49(5), 489-490.

Filippini, C. and Molini, V. (October 2003). The Determinants of East Asian Trade Flows: A Gravity Equation Approach. *Journal of Asian Economics*, 14(5), 695-711(17).

Frankel, J. A. and Wei, S. j. (1993). Trade Blocs and Currency Blocs. *Working Paper No. W4335.* National Bureau of Economic Research (NBER).

Havrylyshyn, O. and Pritchett, L. (1991). European Trade Patterns after the Transitions. Policy *Research Working Paper Series 748.* World Bank.

Hill, L. W. C. (2004). *Global Business Today* (pp. 5). New York: McGraw Hill Irwin.

Hummels, D., Volodymyr, L., and Skiba, A. (2007). The Trade Reducing Effects of Market Power in International Shipping. *Working paper no. W12914.* National Bureau of Economic Research (NBER).

Linnemann, H. (1966). *An Economic Study of International Trade Flows*. Amsterdam: North-Holland.

Martínez-Zarzoso, I., and Nowak-Lehmann, F. D. (2003). Augmented Gravity Model: An empirical application to Mercosur-European Union trade flows. *Journal of Applied Economics*, 6(2), 291-316.

Oguledo, V. I. and Macphee, C. R. (1994). Gravity Models: A Reformulation and an Application to Discriminatory Trade Arrangements. *Applied Economics*, 26(2), 107-120.

Rahman, M. (2004). The Determinants of Bangladesh's Trade: Evidence from the Generalised Gravity Model. *The Economic Society of Australia's 33 Conference of Economists,* University of Sydney, Australia.

Razzaque, A. and Raihan, S. (2007). WTO and Regional Trade Negotiation Outcomes: Quantitative Assessments of Potential Implications on Bangladesh. *Pathak Shamabesh*, 150 –152.

Tinbergen, J. (1962). *Shaping the World Economy: Suggestions for an International Economic Policy*. New York: The Twentieth Century Fund.

www.timeanddate.com

Table 1. Trade Balance of Bangladesh (Million US $)

Year	Goods Imports	Goods Exports	Trade Balance
1985	2,286	999	-1,287
1986	2,301	880	-1,421
1987	2,446	1,077	-1,369
1988	2,734	1,291	-1,443
1989	3,300	1,305	-1,995
1990	3,259	1,672	-1,587
1991	3,074	1,689	-1,386
1992	3,354	2,098	-1,256
1993	3,657	2,545	-1,113
1994	4,351	2,934	-1,416
1995	6,057	3,733	-2,324
1996	6,285	4,009	-2,275
1997	6,551	4,840	-1,711
1998	6,716	5,141	-1,574
1999	7,536	5,458	-2,077
2000	8,053	6,399	-1,654
2001	8,133	6,085	-2,049
2002	7,780	6,102	-1,678
2003	9,492	7,050	-2,442
2004	11,157	8,151	-3,006
2005	12,292	9,186	-3,105

Source: World Development Indicators, 2007

Table 2. Common Commodity Imports by Bangladesh from China and India

Group-wise Commodity	**China** *(TK in Millions)*		**India** *(TK in Millions)*	
	2005-2006	*2004-2005*	*2005-2006*	*2004-2005*
Cotton,(all types) cotton yarn/thread and cotton fabrics	2,987	2,072	2,284	1,246
Boilers, Machinery and mechanical appliances, parts thereof	2,094	1,595	994	1,147
Electrical machinery and equipment and parts thereof, sound recorders and reproducers, television image and sound recorders and reproducers and parts and accessories of such articles	1,334	636	323	302
Man-made staple fibers	1,016	884	286	267
Knitted or crocheted fabrics	857	665	91	93
Vehicles other than railway or tramway, rolling stock and parts and accessories thereof	214	200	704	504
Plastics and articles thereof	164	98	474	444
Cereals	21	20	1,425	2,961

Source: Bangladesh Bank (2004-2005, 2005-2006), Import Payments (annual), Department of Public Relations & Publication, Dhaka, Bangladesh.

Table 3. Goods Import Market Shares of the Eight Major Trading Partners of Bangladesh

Year	**India**	**China**	**Singapore**	**Japan**	**Hong-Kong**	**S. Korea**	**USA**	**Malaysia**	**Sub-Total (a)**	**Total Imports (b)**	**(a) as % of (b)**
1985	67	90	291	301	51	56	286	48	1,190	2,286	52.1%
1986	60	74	209	255	46	64	164	27	899	2,301	39.1%
1987	64	82	156	378	68	64	178	24	1,014	2,446	41.5%
1988	87	91	145	456	112	82	216	36	1,225	2,734	44.8%
1989	104	110	186	445	116	103	325	50	1,439	3,300	43.6%
1990	145	132	323	475	157	126	208	41	1,607	3,259	49.3%
1991	181	133	334	336	184	165	181	32	1,546	3,074	50.3%
1992	231	149	275	286	247	181	230	42	1,641	3,354	48.9%
1993	342	248	211	365	299	258	207	53	1,983	3,657	54.2%
1994	414	223	200	498	331	284	202	57	2,209	4,351	50.8%
1995	689	420	275	587	399	340	274	41	3,025	6,057	49.9%
1996	1,100	707	343	695	390	366	330	69	4,000	6,285	63.6%
1997	922	575	297	647	409	360	302	197	3,709	6,551	56.6%
1998	934	593	321	483	443	381	311	172	3,638	6,716	54.2%
1999	1,235	560	553	494	452	287	301	131	4,013	7,536	53.3%
2000	833	568	701	685	455	319	325	108	3,994	8,053	49.6%
2001	1,184	709	824	846	478	411	248	148	4,848	8,133	59.6%
2002	1,019	878	871	655	441	346	261	145	4,616	7,780	59.3%
2003	1,358	938	1,000	605	433	333	223	169	5,059	9,492	53.3%

Source: Bangladesh Economic Review, Ministry of Finance, 2005

Table 4. Model Efficiency Random Effect (Eq No. 4)

Method	m-Value	m-Prob.	d o f	R^2
One-way Random Effect	5.33	0.15	3	0.82
Two-way Random Effect	4.16	0.13	2	0.81

Table 5. Model Efficiency for Random Effect (Eq No. 5)

Method	m-Value	m-Prob.	d o f	R^2
One-way Random Effect	21.01 *	0.000	2	0.79
Two-way Random Effect	8.11 *	0.004	1	0.79

* Significant at 99% confidence level

Table 6. Model Efficiency for Fixed Effect (Eq No. 4)

Method	F-Value	F-Prob.	Numerator d o f	Denominator d o f	R^2
One-way Fixed Effect	40.1 *	0.000	7	139	0.88
Two-way Fixed Effect	31.96 *	0.000	25	121	0.89

* Significant at 99% confidence level

Table 7. Model Efficiency for Fixed Effect (Eq No. 5)

Method	F-Value	F-Prob.	Numerator d o f	Denominator d o f	R^2
One-way Fixed Effect	41.4 *	0.000	7	141	0.87
Two-way Fixed Effect	32.33 *	0.000	25	123	0.89

* Significant at 99% confidence level

Table 8. Both One-way and Two-way Random Effect Regression Output (Eq No. 4) (Dependent Variable: ln (Mij)

Model	Coefficient	t-Value	p-Value
Constant	-7.67 **	-1.85	0.07
ln (Yi)	-1.59	-1.31	0.19
ln (Yj)	1.07 *	5.52	0.0001
ln (Dij)	-2.03 *	-4.13	0.0001
ln (POPi)	7.19 *	2.76	0.007
ln (POPj)	-0.41 *	-3.12	0.002

* Significant at 99% confidence level, ** Significant at 90% confidence level

Table 9. One-way Fixed Effect Regression Output (Eq No. 4)(Dependent Variable: ln (Mij)

Model	Coefficient	t-Value	p-Value
Constant	-28.31 *	-12.8	0.0001
ln (Yi)	-1.29	-1.08	0.28
ln (Yj)	1.43 *	5.28	0.0001
ln (Dij)	0	-	-
ln (POPi)	5.65*	2.12	0.04
ln (POPj)	-0.39	-0.48	0.63

* Significant at 95% confidence level

Table 10. Two-way Fixed Effect Regression Output (Eq No. 4)(Dependent Variable: ln (Mij)

Model	Coefficient	t-Value	p-Value
Constant	-13.68 *	-2.90	0.004
ln (Yi)	0	-	-
ln (Yj)	1.41 *	5.15	0.0001
ln (Dij)	0	-	-
ln (POPi)	0	-	-
ln (POPj)	-0.54	-0.66	0.51

* Significant at 99% confidence level

Table 11. Both One-way and Two-way Random Effect Regression Output (Eq No. 5) (Dependent Variable: ln (Mij))

Model	Coefficient	t-Value	p-Value
Constant	-13.67 *	-3.39	0.0009
ln (Yi)	2.06 *	9.27	0.0001
ln (Yj)	0.78 *	5.33	0.0001
ln (Dij)	-1.57 *	-3.35	0.001

* Significant at 99% confidence level

Table 12. One-way Fixed Effect Regression Output (Eq No. 5) (Dependent Variable: ln (Mij))

Model	Coefficient	t-Value	p-Value
Constant	-31.06*	-21.07	0.0001
ln (Yi)	1.11 *	3.73	0.003
ln (Yj)	1.56 *	7.00	0.0001
ln (Dij)	0	-	-

* Significant at 99% confidence level

Table 13. Two-way Fixed Effect Regression Output (Eq No. 5) (Dependent Variable: ln (Mij))

Model	Coefficient	t-Value	p-Value
Constant	-15.39*	-3.89	0.0002
ln (Yi)	0	-	-
ln (Yj)	1.33*	5.49	0.0001
ln (Dij)	0	-	-

* Significant at 99% confidence level

Table 14. Category-Wise Yearly Imports of Bangladesh (Million US $)

Major Commodities	2001-02	2002-03	2003-04	2004-05	2005-06	2006-07
FOOD GRAINS	186	409	431	574	418	581
Milk & cream	59	61	61	86	73	83
Oil seeds	72	64	73	86	90	106
Edible oil	251	364	471	441	473	583
Pulses all sorts	88	145	120	159	164	195
Sugar	23	104	110	220	124	294
Clinker	150	144	139	170	210	240
Crude petroleum	242	267	252	350	604	524
POL	481	620	770	1252	1400	1709
Chemical	335	353	406	510	580	668
Pharmaceutical products	39	44	45	41	50	49
Fertilizer	107	109	150	332	342	357
Dyeing, tanning etc. materials	87	86	109	132	148	161
Plastics and rubber articles thereof	250	281	367	477	523	643
Raw cotton	312	393	583	666	742	858
Yarn	283	270	323	393	501	582
Textile and articles thereof	1063	1106	1295	1571	1728	1892
Staple fiber	39	41	57	75	76	97
Iron, steel and other base metals	413	455	479	680	980	985
Capital goods	2617	2735	2875	1115	1539	1929
Others	803	848	870	2823	2887	3401
Subtotal :	7913	8931	10016	12195	13684	16013
Imports of EPZ	627	727	887	952	1062	1144
Grand Total :	8540	9658	10903	13147	14746	17157
Capital Goods as % of Total Import (Except EPZ)	**33%**	**31%**	**29%**	**9%**	**11%**	**12%**

Source: Bangladesh Bank, Import Payments Data, Available at
<http://www.bangladesh-bank.org/econdata/import/categoryimp.html>

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

An Analysis on Influencing Factors of Stockholders Interests Redistribution from a Consideration View of Non-Tradable Shares Reform in Medicine Industry

Yuhe Wang
Business School, East China Normal University
Shanghai 200241, China
Add: No.500, Dongchuan Road, Minhang District
Shanghai 200241, China
E-mail: wyh_uh@hotmail.com

Abstract

As the full-circulation reform is close to an end, this paper intends to study the influencing factors of probability at certain consideration level by an empirical analysis of medicine industry and constructing a discrete choice data model, by which to research the influencing factors of stockholders interest redistribution. Results show that factors that affect the choice of consideration intervals are different and different factors prefer various probabilities of consideration levels. The higher the proportion of non-tradable stocks, the more the consideration interval is in 2.5-3.4; the higher the ROE, the lower the consideration payment is; the larger the M1, the more the consideration interval is less 2.5.

Keywords: Stockholders interest redistribution, Non-trading shares reform, Medicine industry, Consideration

1. Introduction

Non-tradable shares are special historical things appeared in the transmission of China's capital market. The co-existence of tradable shares and non-tradable shares brings about many problems in China's capital market. Majority shareholders and minority shareholders are competing for interests instead of pursuing common interests. The pricing mechanism in capital market is distorted so that the stock price fails to reflect its true value. In 1997, scholars have noticed this problem and put forward some suggestions. The sales of state shares have adopted the share allotment, share allocation, and issuing method. However, all these measures do not settle the problem of non-tradable shares. Even in 2001, a bear market appeared and lasted for four years. In 2005, China Securities Regulatory Committee issued the "Notice for Relevant Problems in Pilot Listed Companies' Non-Tradable Shares Reform" that took interests of small and medium shareholders into consideration and advanced the consideration concept for the first time in the scheme.

Consideration refers to the payments or compensations paid by non-tradable shareholders to tradable shareholders in order to make non-tradable shares circulate in market. Consideration payment is the core of non-tradable shares reform. Although the non-tradable shares reform is over, the factors that affect the allocation of shareholders interests are still there and at work. Non-tradable shares reform is a great Chinese way that embodies shareholders interest redistribution. Therefore, it is still meaningful to study the consideration in non-tradable shares reform.

2. Review of literature

How to establish the consideration level has already aroused a wide discussion in theoretical field and drawn some useful conclusions. In general, these conclusions are derived from two types of study angles. The first type is the angle of theoretical deduction. With the precondition of a series of assumptions and requirements, put forward calculation formulas and theoretical frameworks for potential consideration schemes in non-tradable shares reform, which need to be carried out by listed companies later. This kind of studies is mostly done by listed companies' entrusted industry before or in the non-tradable shares reform.

As far as principles of non-tradable shares reform are concerned, some have already been accepted widely in academic field, such as protecting interests of minority shareholders and preventing against lose of state assets (Weidong Zhai, 2006; Xiaoqiu Wu, 2006). Concerning consideration payment, Xiaoqiu Wu thinks it should be various. In his opinion, there is an average consideration rate (30% or so) in market and all companies' considerations should fluctuate around this rate. Furthermore, non-tradable shares could be taken as a kind of shares issuance. Then, non-tradable shares, shares allotment, and new issuance can serve as references one another (2006). Jinhui Qiu makes a comparative analysis on the consideration payments used by listed companies in China (2006). Jianqing Chen advances that the consideration needs a common standard and puts forward formulas respectively for theoretical maximum, minimum, balance standards (2006). In *Securities Guide*, a special report concludes the non-changeable theoretical value pricing, the super-normal P/E ratio pricing, and the analogical pricing. Weidong Zhai sets up a pricing model under the assumptions of full efficiency of market pricing, investors' same pricing predictions for circulation, and sufficient market capitals (2006). Changsong Jin analyzes consideration's influencing factors and gets a theoretical mode by means of correlation degree analysis according to the principle of same ownership for same share (Apr. 2006). Ying Zhou takes full circulation reform as a game process and constructs a game model based on two parties' effect functions and game rules, establishing optimal payment rate by calculating maximum effect (Jun. 2006).

As the non-tradable shares reform is deepening, more and more companies achieve the reform. In theoretical fields, some research results based on statistical analysis of large samples appear which contribute a lot to some final conclusions. Lihong Xiao, a graduate from North China Electric University, makes a regression analysis on listed companies that have already pay for considerations and concludes that the consideration is determined by various factors and it is hard to explain it by certain unified explanatory variables. Furthermore, different listed companies have different standards for evaluating consideration rate because of their different market values (2007). Xiaoli Li and Jianping Yang advance that the non-tradable shares reform process has characters of diversified consideration ways, various consideration bases, unbalanced consideration levels, and frequent adjustments (2006). Xinhong Chen makes an empirical analysis on ST listed companies' consideration payment ways and points out that ST listed companies can pay for considerations by multiple institutional innovative ways, such as imputing well performing assets (2007). Jianshen Zhao makes a multiple regression analysis on listed companies' non-tradable shares reform consideration, finds out four practical decisive factors, and concludes that bilateral consideration payment is a win-win strategy (2007). Pengchao Wu performs an empirical analysis on 330 listed companies' non-tradable shares reform and finds that the market game mechanism adopted by non-tradable shares reform can guarantee the rationality of consideration scheme and interests of two types of shareholders to certain degree (2006). Xuesong Qian and Wei Zou analyze the consideration issue by constructing a signal transmission mode. They find that the consideration is not only a compensate mechanism for tradable shareholders but a channel that transmits companies' private information to investors.

As the non-tradable shares reform is close to an end, this paper, from the second angle, analyzes the factors that affect listed companies' consideration payments in medicine industry by means of an empirical analysis method based on the discrete data variable model, with the hope of drawing useful conclusions.

3. Theoretical assumptions and selection of explanatory variables

In this paper, select main explanatory variables from three angles: market index, listed companies' situation, and macro economy.

(1) Share capital structure: be replaced with proportion of non-tradable shares

Relevant articles show that the larger the proportion of non-tradable shares, the bigger the changes of supply-and-demand relation after circulation, and then the higher the consideration payments. The two have a positive correlation. That is an important variable of listed companies' situation.

(2) Relevant financial index: In general, listed companies will release some important financial indexes of recent three years in their non-tradable shares reform instructions. Besides, quite a few of companies takes these financial indexes as explanatory variables in establishing their consideration payments. Therefore, select two most important financial indexes. In this paper, data of the two indexes is the average during three years before the non-tradable shares reform.

ROE/profit margin on net assets: A stock's intrinsic value depends on the discount value of its future cash flow. ROE/profit margin on net assets is an important index used to evaluate companies' former business performances and also a financial index used to predict the stock tendency.

Assets-liabilities ratio: Financial indexes affect listed companies' performances in capital market and their market values. Therefore, here include this index into the analysis mode.

(3) Stock's rate of return and stock index's rate of return

In theory, one stock's performance in market and the whole market performances relate with the distribution of shareholders' interests. The higher the rate of return, the stronger the bargaining ability of majority shareholders should be. And the consideration level will be lower. This paper selects the average rate of return in ten exchange days before establishing the consideration.

(4) M1: Money supply is an important macro economic index influencing the operation of capital market. The non-tradable shares reform happens in a background of global over supply of money. Therefore, the impacts of money supply on consideration payment are not ignorable.

4. Model selection

4.1 Selection of variables and data

A Consideration payment level (delivery rate) K

According to the formula, convert the payment levels under different payment ways into proportions of share allocation for the sake of convenient comparison (formula is in appendix).

Considerations have different payment levels. Relevant articles have also pointed out the existence of consideration payment levels' clustering effect. Therefore, group data according to certain standard and get four different cluster intervals. Because of different values of K, the variable Y is different.

Y=1 (K<=2.5)
=2 (2.5<K<=3.0)
=3 (3.0<K<=3.4)
=4 (K>=3.4)

4.2 Modeling method

Considering the fact that more than 100 listed companies have different situations, the OLS regression may lead to the heteroscedasticity issue that will affect the exactness of estimation. Therefore, this paper adopts the multi-variable discrete choice model to study the probability of listed company i selecting certain consideration interval and the influencing factors.

In other words, the explained variable y_i has S discrete choices (S>2).

Namely, $y_i = S$

Suppose the sample company selects S, $S \in A = \{1, 2, \ldots, S\}$; i=1,…,n.

Also, introduce a virtual continuous random variable U_{is}, define:

$y_i = S$, suppose $U_{is} = \max_{k \in A} \{U_{ik}\}$.

Meanwhile, $U_{is} = V_{is} + \varepsilon_{is}$,

Here V_{is} is the explanatory variable of the model. ε_{it} is the residual item response to different statistic distribution.

The explanatory variable V_{is} includes: (1) The variable X_i that has nothing to do with the value of explained variable; (2) The variable Z_{is} that has a correlation with the value of explained variable. Then, V_{is} can be represented as:

$$V_{is} = x_i' \beta_s + z_{is}' \gamma.$$

Here, the coefficient β_s is correlated to the choice of explained variable, but γ not.

Define p_{is} is the probability of sample i selecting s. Namely:

$$p_{is} = P(y_i = s) = P(U_{is} = \max_{k \in A}\{U_{ik}\}) = P(V_{is} + \varepsilon_{is} > V_{ik} + \varepsilon_{ik}, k \neq s).$$

Then the logarithm likelihood function is:

$$l(\theta) = \log L(\theta) = \sum_{i=1}^{n} \sum_{s=1}^{S} Y_{is} \log p_{is}(\theta),$$

Thereof, suppose $y_i = s$, $Y_{is} = 1$; Under other conditions, $Y_{is} = 0$.

For a multi-variable logit model, then:

$$p_{is} = \frac{\exp(V_{is})}{\sum_{k=1}^{S} \exp(V_{ik})} = \frac{1}{1 + \sum_{k=1, k \neq s}^{S} \exp(V_{ik} - V_{is})}.$$

Finally: $\log(p_{is} / p_{ik}) = V_{is} - V_{ik} = x_i'(\beta_s - \beta_k) + (z_{is} - z_{ik})'\gamma.$

4.3 Model's parameter estimation and test result

In this model, estimated parameters are $\hat{p}_{is}$ and explanatory variable parameter $\hat{\beta}_s$ and $\hat{\gamma}$. If there are constants in the model, the maximum likelihood value estimation and test are similar to normal F test. The referenced model in test includes just constants. Namely:

$V_{is} = \alpha_s, \ \ s = 2,\ldots,S.$

Suppose there are K parametric variables in explanatory variable interval x_i' and C parametric variables in z_{is}'. K(S-1) + C parametric variables need to be estimated in the model. And the logarithm likelihood function is $l(\hat{\theta})$. The reference model has S-1 parametric variables. Its logarithm likelihood function is $l(\hat{\alpha})$. Then, statistic $2l(\hat{\theta}) - 2l(\hat{\alpha})$ distributes in form of $\chi^2[(K-1)(S-1)+C]$.

In addition, we can use logit model to analyze the impacts of explanatory variables on possibilities p_s of explained variables. The partial derivative results are:

$$\frac{\partial p_s}{\partial x_m} = \hat{p}_s(\hat{\beta}_{sm} - \sum_{k=1}^{S} \hat{\beta}_{km}\hat{p}_k),$$

$$\frac{\partial p_s}{\partial z_{sm}} = \hat{\gamma}_m \hat{p}_s(1-\hat{p}_s),$$

$$\frac{\partial p_s}{\partial z_{rm}} = \hat{\gamma}_m \hat{p}_s \hat{p}_r, \quad r \neq s.$$

The quasi elasticity coefficient is $\frac{\partial p_s}{\partial \log x_m} = \frac{\partial p_s}{\partial x_m}\overline{x_m},$. Thereof, $\overline{x_m}$ is the sample average of explanatory variable x_m.

Suppose x_m rises by q percent, p_s will increase by $q \times \partial p_s / \partial \log x_m$ percent.

The elasticity coefficient is: $\frac{\partial p_s}{\partial x_m}\frac{x_m}{p_s}$.

In calculating the t statistic of series of parametric variables, we take $\hat{p}_s$ as a non-random variable. It proves that the partial derivatives, quasi elasticity coefficients, and elasticity coefficients have completely same t distributing statistic.

5. The discrete choice model's parameter estimation and test result

In this model, we select 119 samples.

In medicine industry, there are 132 listed companies. Two of them do not perform the non-tradable shares reform and five of them adopt the warrant method. Because warrant has different delivery conversion formulas that can lead different results, the five listed companies are excluded from samples in this paper. Five listed companies adopt the afterward conditional super-addition consideration way that needs certain triggering conditions. This way is few and hard to make quantization. Therefore, for the sake of convenient calculation, these life companies are also deleted from samples in this paper.

Then, 120 listed companies are right for samples in this paper. However, one company can not provide with complete data. So, the final samples are 119 listed companies.

The model estimation follows an order from the common to the special (table 1).

Parameter estimation and test result is in table 2.

From the test results in the table above, the model passes the F test.

Firstly, according to the test results, considerations' choice probabilities distribute uniformly. In the medicine industry, the convergence effect, namely centering on 3.5, mentioned in former articles does not appear.

Secondly, factors have different impacts on different consideration. In the state 4, impacts of constants are most prominent. Proportion of non-tradable shares has significant impacts on all considerations. As the consideration rises, the impacts of ROE/profit margin on net assets on probability of certain consideration increases significantly. The money supply has significant negative impacts on all conditions. In conclusion, the more the money supply, the higher the nominal value of financial assets, the stronger the bargaining ability of majority shareholders. Then the majority shareholders will be more capable of decreasing consideration payments.

Thirdly, the prominent factors do not include indexes of market performance. It means performances of one stock and market do not have impacts on distribution of shareholders' interests.

Next, we will further analyze the influencing factors' partial derivatives, elasticity coefficients, and quasi elasticity coefficients on probabilities of all considerations (table 3).

From the table above, we can conclude:

(1) If the ROE/profit margin on net assets rises by 1%, the probability of consideration less 2.5 will rise by 0.00102%, the probability of consideration between 2.5 and 3.0 will rise by 0.000239%, the probability of consideration between 3.0 and 3.4 will drop by 0.00036%, and the probability of consideration larger than 3.4 will drop by 0.00058%.

(2) If the proportion of non-tradable shares rises by 1%, the probability of consideration less 2.5 will drop by 0.19855%, the probability of consideration between 2.5 and 3.0 will rise by 0.22246%, the probability of consideration between 3.0 and 3..4 will rise by 0.45654%, and the probability of consideration larger than 3.4 will drop by 0.00410%.

(3) If the money supply rises by 1%, the probability of consideration less 2.5 will rise by 1.6119%, the probability of consideration between 2.5 and 3.0 will drop by 0.22588%, the probability of consideration between 3.0 and 3..4 will drop by 0.31107%, and the probability of consideration larger than 3.4 will drop by 1.0749%.

6. Results analysis

This paper analyzes the factors that affect the probabilities of considerations by constructing a discrete choice data Logit model based on 118 listed companies' data in medicine industry. The model study results are as follow:

Firstly, for different considerations, the main influencing factors are different. Especially, the proportion of non-tradable shares, namely the share ownership structure, has significant impacts on each consideration interval.

Secondly, in the aspect of market factors, the rate of return in stock market and the rate of return of one stock do not have significant effects. In theory, the two rates are supposed to affect the bargaining ability of majority shareholders to a great degree. But the result is contrary. It proves that there is no significant correlation between choice of consideration in non-tradable shares reform and one stock's market performance. Companies' situations and macro variables are the significantly influencing factors that affect the distribution of shareholders' interests.

Thirdly, impacts of macro factors are not ignorable. Some impacts of M1 on probability are over 1%. It means that macro economic factors are also important in determining the consideration. As M1 rises, the consideration begins to drop gradually. In this case, overage of money supply forces money to pursue for lesser assets. Under this background, shares, as a kind of financial assets, become relatively scarce, what will greatly enhance the majority shareholders' bargaining ability.

Fourthly, in the aspect of companies' situation, higher ROE/profit margin on net assets will make listed companies tend to choose lower consideration. The impacts of assets-liabilities ratio are not significant.

ROE/profit margin on net assets is an important index to evaluate listed companies' market values. The higher it is, the better the listed company' performances are. Then its majority shareholders possess stronger bargaining ability.

In theory, the assets-liabilities ratio can affect listed companies' market values. However, because the correlation between assets-liabilities ratio and market value is not simply positive or negative and every listed company is different in certain aspect and each has different optimal leverage rate, the positive impacts and the negative impacts may cause a zero impact among amounts of samples.

Fifthly, the rise of the proportion of non-tradable shares will make consideration center to an interval from 2.5 to 3.4. Taking the state 1 as the basic state, then the higher the proportion of non-tradable shares, the higher the probability of low consideration. It is in accord with the normal logics and the empirical tests and theoretical analyses in most articles. The smaller the proportion of non-tradable shares, the fewer the supply for market after the non-tradable shares reform, the smaller the capital pressure, the lesser the drop of price, and the lesser the consideration payment.

7. Conclusion

By analyzing the sample data of 119 listed medicine companies that listed in the A board in Shanghai and Shenzhen Stock Markets, this paper finds that:

(1) The money supply, the proportion of non-tradable shares, and the ROE/profit margin on net assets have most significant impacts on the choice of consideration. And different factors have different impacts on the probability of consideration choice. Each choice results from integrated effects of all factors.

(2) One stock's market performance and whole capital market's performance do not have significant impacts on certain listed company's consideration. Their impacts mainly focus on the company's situation and macro environment.

(3) Macro economic factors have certain effects on the choice of consideration, whether the market participators realize it or not, and the effects are objective and should not be neglected.

We can further draw these conclusions:

(1) The proportion of non-tradable shares is an important influencing factor that affects the distribution of shareholders' interests. Although its influences disappear on the surface, as the non-tradable shares reform is closed, it is evolved into the quantity and proportion of newly tradable shares. Therefore, we can still take this factor as the basis for protecting interests of small and medium shareholders.

(2) Listed company's condition is still an important influencing factor. Better conditions and higher ROE/profit margin on net assets mean small and medium shareholders can obtain relatively direct interests. In re-distributing interests, such as consideration payment, the trend to be lower is rational.

(3) Macro factors, especially money supply, have great impacts on shareholders' behaviors. According to studies above, tight monetary policy is still appropriate. At present, an appropriate tight monetary policy can help to create a favorable macro environment for China's capital market.

References

Chen, Jianqing. (2006). Studies on the reform of integrated consideration standard. *Journal of Xi'an University of Finance and Economics*. No. 4.

Chen, Xinhong. (2007). Empirical analysis and research on ST stock's non-trading shares reform. *Financial Observe*. No.1.

Jin, Changsong & Li, Fan. (2006). The model of the full circulation. *Journal of Harbin University of Science and Technology*. No.4.

Li, Xiaoli & Yang, Jianping. (2006). Analysis on issue of consideration in pilot stage of equity division reform. J*ournal of Guangxi University of Finance and Economics*. No.4.

Qian, Xuesong & Zou, Wei. Choice of consideration means for non-tradable shares reform ------ theoretical mode and empirical research. *Nanjing Business Review*. No.12.

Qiu, Jinhui. (2006). A comparative analysis of listed company's consideration payment in China. *Corporate Finance*.

Wu, Pengchao. (2006). Theoretical and empirical analysis on the influencing factors of consideration payment. *Economic Review*. No.8.

Wu, Xiaoqiu. (2006). Theoretical problems in non-trading shares reform ------ on new changes in China's capital market in full circulation. *Financial & Trade Economics*. No.2.

Qu, Weidong. (2006). A study on the consideration in non-trading shares reform. *Research on Economics and Management*.

Zhao, Jianshen & Zhong, Jianxin. (2007). A win-win reform ------ An analysis on factors of consideration level on non-trading shares reform. *Productivity Research*. No.16.

Zhou, Ying & Wang, Siliang. (2006). Research on game issue of complete liquidity consideration. *Journal of Dalian University of Technology (social sciences)*. No.6.

Zhu, Huacheng & Ye, Shaizheng. (?). Evaluation of 42 pilot companies' consideration schemes. *Securities Guide*.

Table 1.

Explained variable Y				
Coefficient	Standard deviation	t-statistic	t-statistic's P	
Constant(S2)	3.24262	4.147	0.782	0.436
NON(S2)	0.0520244	0.02305	2.26	0.026
ROE(S2)	-0.00965126	0.01497	-0.645	0.520
M1(S2)	-5.79455e-005	3.618e-005	-1.60	0.112
Constant(S3)	2.40020	4.441	0.540	0.590
NON(S3)	0.0717504	0.02599	2.76	0.007
ROE(S3)	-0.0213216	0.01482	-1.44	0.153
M1(S3)	-6.28142e-005	3.869e-005	-1.62	0.107
Constant(S4)	8.10068	5.179	1.56	0.121
NON(S4)	0.0379071	0.02444	1.55	0.124
ROE(S4)	-0.0283987	0.01474	-1.93	0.057
M1(S4)	-9.71538e-005	4.605e-005	-2.11	0.037

Table 2.

Logarithm likelihood value	-152.963987	Number of states	4
Sample observe value	119	Number of estimated parameters	12
Zero line logarithm likelihood value	-163.9155	F test Chi^2(9)	21.903 [0.0092]**
AIC	329.927973	AIC/T	2.77250398
Y sample average	1.35294	Sample variance	1.23678
Newton estimation method (esp1=0.0001; esp2=0.005): strong convergence			

State	Frequency	Probability	estimation	Logarithm likelihood value
State 1	35	0.29412	0.29412	-37.79
State 2	32	0.26891	0.26891	-40.55
State 3	27	0.22689	0.22689	-37.34
State 4	25	0.21008	0.21008	-37.29
Total	119	1.00000	1.00000	-153.0

Notice: In this table, state 1 means consideration is less 2.5. State 2 means consideration is between 2.5 and 3.0. State 3 means consideration is between 3.0 and 3.4. State 4 means consideration is larger than 3.4.

Table 3.

State choice probability	
State 1	0.29412
State 2	0.26891
State 3	0.22689
State 4	0.21008

Partial derivative

	Sample average	State 1	State2	State 3	State 4
Constant	1.0000	-0.91717	0.033411	-0.16295	1.0467
NON	60.033	-0.011245	0.0037086	0.0076048	-6.8446e-005
ROE	0.88276	0.0039409	0.0010078	-0.0017975	-0.0031512
M1	1.0907e+005	1.4778e-005	-2.0709e-006	-2.8520e-006	-9.8549e-006

Quasi elasticity coefficient

	State 1	State 2	State 3	State 4
Constant	-0.91717	0.033411	-0.16295	1.0467
NON	-0.67507	0.22264	0.45654	-0.0041090
ROE	0.0034789	0.00088966	-0.0015868	-0.0027818
M1	1.6119	-0.22588	-0.31107	-1.0749

Elasticity coefficient

	State 1	State 2	State 3	State 4
Constant	-0.26976	0.0089846	-0.036971	0.21990
NON	-0.19855	0.059869	0.10358	-0.00086324
ROE	0.0010232	0.00023924	-0.00036003	-0.00058440
M1	0.47407	-0.060740	-0.070579	-0.22582

t-statistic	State 1	State 2	State 3	State 4
Constant	-1.2807	0.045264	-0.23391	1.3417
NON	-2.8271	0.95123	1.9217	-0.019388
ROE	1.4369	0.44685	-1.0588	-2.0225
M1	2.3470	-0.32253	-0.47161	-1.4195

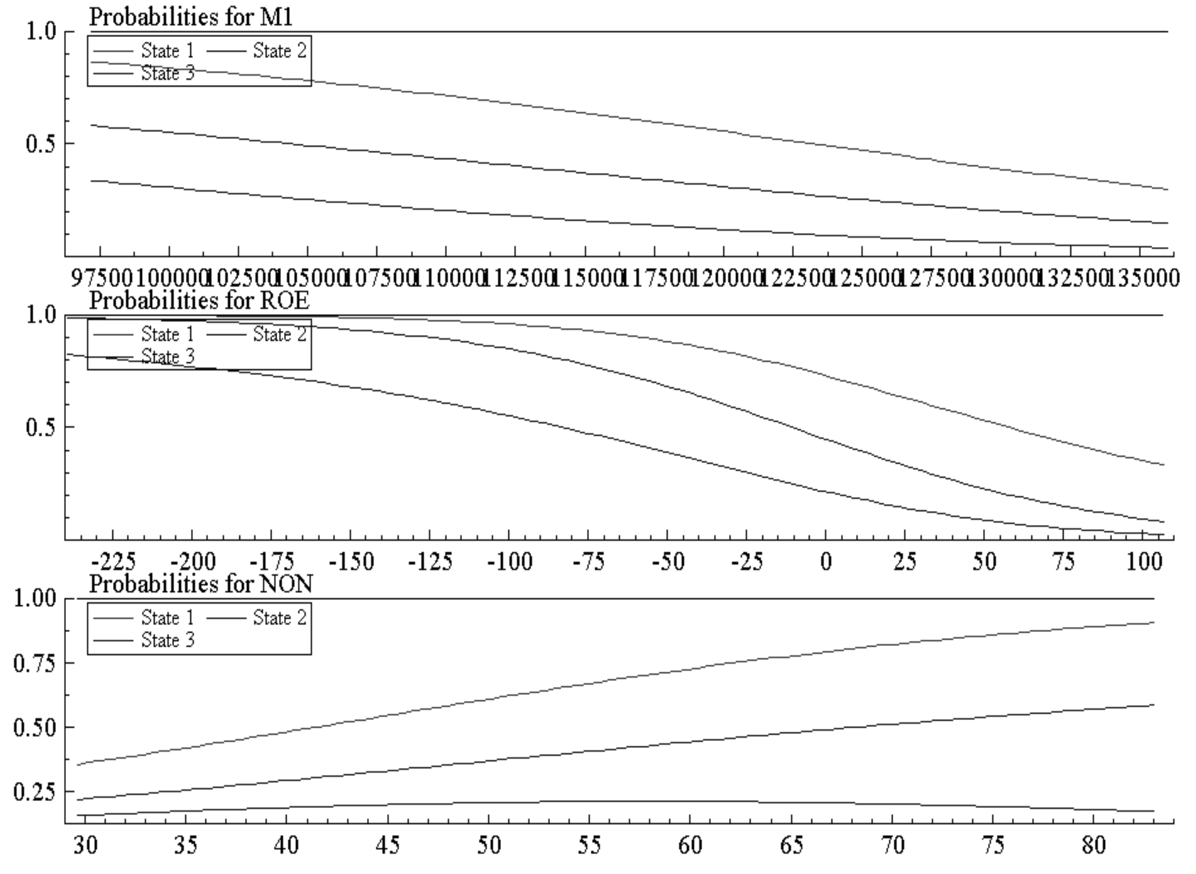
Probabilities for M1
1.0
0.5
State 1
State 2
State 3
97500
100000
102500
105000
107500
110000
112500
115000
117500
120000
122500
125000
127500
130000
132500
135000
Probabilities for ROE
1.0
0.5
State 1
State 2
State 3
-225
-200
-175
-150
-125
-100
-75
-50
-25
0
25
50
75
100
Probabilities for NON
1.00
0.75
0.50
0.25
State 1
State 2
State 3
30
35
40
45
50
55
60
65
70
75
80

Figure 1.

Appendix: Samples of listed companies

Conversion formula:

000004STShenzhen CAU Technology	000028Accord Pharmaceutical	000078Neptunus Bioengineering	000153Fengyuan Pharmaceutical	000411Int'l Group
000415Huitong Group	000416Huaxin Industry	000423Dong-E E-Jiao Group	000518Sihuan Bioengineering	000538Yunnan Baiyao
000545Jilin Pharmaceutical	000566Hainan Haiyao	000590Ziguang Guhan	000591Tongjunge	000597Northeast General Pharmaceutical
000603 STWeida	000605 STSihuan	000607Huali Pharmaceutical	000623Jilin Aodong	000626Ruyi Group
000627Tianmao Group	000650Renhe Pharmaceutical	000628 STGaoxin	000705Zhejiang Zhenyuan Pharmaceutical	000739Puluo Kangyu Natural Medicine
000765Xinhua Pharmaceutical Chemical	000766Tonghua Golden-Horse Pharmaceutical	000788Southwest Synthetic Pharmaceutical	000790Hoist Group	000796Baoshang Group
000809Zhonghui Pharmaceutical	000819Yueyang Xingchang	000915Shanda Wit Science and Technology	000919Jinling Pharmaceutical	000963Huadong Medicine
000979ST Keyuan	000989 Jiuzhitang	000990Chengzhi	002001NHU Company	002004Huapont Pharmaceutical
002007Hualan Biological Engineering	002016Well Science and Technology	002020Jingxin Pharmaceutical	002022Kehua Bio-Engineering	002030Daan Gene
002038Shuanglu Pharmaceutical	600055Wandong Medical Equipment	600062Double-Crane Pharmaceutical	600079Renfu Science and Technology	600080ST Jinhua
600085Tongrentang	600095Haerbin High-Tech	600129Taiji Group	600133East Lake Hi- tech	600146Dayuan Chemical
600161Tiantan Biological Products	600196Fuxing Pharmacy	600200Jiangsu Wuzhong	600201Jinyu Group	600211Tibet Pharmaceutical
600216Zhejiang Medicine	600222Tailong Pharmaceutical	600223ST Wanjie	600226Shenghua Biok Biology	600252Zhongheng Group
600253Tianfang Corporation	600267Haizheng Pharmaceutical	600276Hengrui Medicine	600285Lingrui Pharmaceutical	600297Meiluo Pharmaceutical
600329 ST Zhongxin	600332Guangzhou Pharmaceutical	600351Yabao Pharmaceutical	600380Health Investment Holdings	600385 ST Jintai
600420Modern Pharmaceutical	600421Guoyao Technology	600422Kunming Pharmaceutical	600436Pientzhuang Pharmaceutical	600466 ST Dikang
600479Qianjin Pharmaceutical	600488Tianyao Pharmaceutical	600490Synica Corporation	600511Guoyao Group	600518Kangmei Pharmaceutical
600530Jiaoda Onlly	600535Tasly Group	600556 ST Beisheng	600557Kangyuan Pharmaceutical	600568 ST Qianyao
600572Conba Pharmaceutical	600587Xinhua Medical Instrument Factory	600594Yibei Pharmaceutical	600607Shanghai Industrial Pharmaceutical Investment	600613Wingsung Data Technology
600624Fudan Forward	600645ST Wangchunhua	600666Southwest Pharmeceutical	600668Jianfeng Group	600706 ST Changxin
600713Nanjing Medical	600730China High-Tech	600750Jiangzhong Pharmaceutical	600771Dongsheng Science & Technology	600773 ST Yalong
600778Friendship Group	600779Swell Fun	600789Lukang Pharmaceutical	600797Zheda Innovation Technology	600812North China Pharmaceutical
600829Sanchine Pharmaceutical	600833No.1 Pharmacy	600842Zhongxi Pharmaceutical	600851Haixin Group	600866Star Lake Bioscience
600867Tonghua Dongbao Pharmaceutical	600869S&P Pharmaceutical	600976Wuhan Jianmin Pharmaceutical	600993Mayinglong Pharmaceutical	

K: Allotment proportion $N1$: Proportion of non-tradable shares $N2$: Proportion of tradable shares

1. The converted rising shares taking tradable shares as the cardinal number (expanding shares)

$K=bN1/(N1+N2+bN2)$

Thereof, b refers to the proportion of expanded shares.

2. Shrinking shares

$K=N1(1-a)/(aN1+N2)$

Thereof, a refers to the proportion of shrunk shares.

3. Cash issue

$K=c/p$

Thereof, c means cash and p means the stock price in the cash issue day.

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Board Structure and Corporate Performance in Malaysia

Zubaidah Zainal Abidin
Institute of Graduate Studies, Universiti Teknologi MARA
Block 8, Intekma Resort & Convention, 40450 Shah Alam, Selangor, Malaysia
Tel: 60-3-5522-5435 E-mail: drzubaidah@salam.uitm.edu.my

Nurmala Mustaffa Kamal
Faculty of Accountancy, Universiti Teknologi MARA
Km 12 Jalan Muar, 85200 Segamat, Johor, Malaysia
Tel: 60-7-935-2000 E-mail: nurma181@johor.utim.edu.my

Kamaruzaman Jusoff (Corresponding author)
Yale University, Earth Observation Centre, Environmental Science Centre
21 Sachem St, New Haven, CT 06511, USA
Tel: 203-432-1384 E-mail: jusoff.kamaruzaman@yale.edu

Abstract

This study examines the association between board structure and corporate performance, where performance is defined as the value added (VA) efficiency of the firm's physical and intellectual resources rather than the more commonly used financial terms or profitability ratios. It is argued that the inclusion of intellectual elements into the measurement provides a long-term measurement of corporate performance. The VA efficiency of the firm's total resources is calculated using the Value Added Intellectual Coefficient (VAIC) methodology developed by an Austrian, Ante Pulic. The four board characteristics that are of interest in this study are board composition, directors' ownership, CEO duality and board size. Based on a randomly selected sample of 75 companies listed on Bursa Malaysia, it is found that board composition and board size have a positive impact on firm performance, while the effects of directors' ownership and CEO duality on the VA efficiency of firm's total resources are not established. The outcome of the study shows that the emphasis on the importance of outside directors on the board by The Malaysian Code on Corporate Governance and by the requirements of Bursa Malaysia is deemed pertinent to the long-term corporate performance.

Keywords: Intellectual capital, Board structure, Corporate governance, Value-Added efficiency

1. Introduction

The board of directors have been largely criticised for the decline in shareholders' wealth and corporate failure. They have been in the spotlight for the fraud cases that had resulted in the failure of major corporations, for example, Enron, WorldCom and Global Crossing. Some of the reasons stated for these corporate failures are the lack of vigilant oversight functions by the board of directors, the board relinquishing control to corporate managers who pursue their own self-interests and the board being remiss in its accountability to stakeholders. As a result, various corporate governance reforms have specifically emphasised on appropriate changes to be made to the board of directors in terms of its composition, structure and ownership configuration. The Malaysian Code on Corporate Governance (MCCG) of 2000 proposes that the board should include a balance of executive directors and non-executive directors in order to ensure that the board's decision making is not dominated by a certain party. The best practices of the code also recommend that the responsibilities of the chairman and the chief operating officer should not be held by the same person to ensure that there is a balance of power and authority. The appointments to the board should be made by a nomination committee and the directors should undergo an orientation and education programme (training). Other mechanisms of corporate governance include audit committees and a sound internal control structure.

The underlying objective for these regulations and recommendations is to create an effective board that can perform its stewardship responsibilities and protect the interests of the shareholders. This implies that certain board characteristics may lead to better corporate performance. In addition, corporate performance needs to be measured using a long-term indicator, instead of the more commonly used short-term financial ratios. This is to determine the impact of the board structure on the sustainable performance of the company. Hence, the measurement of corporate performance needs to include the firm's total (physical and intellectual) resources. The board characteristics that are of particular interest in this study are board composition, directors' share ownership, Chief Executive Officer (CEO) duality and board size. Hence, the impact of each of these characteristics on corporate performance needs to be determined to substantiate the desirable board structure which would produce favourable outcome on the value-added efficiency of the resources of Malaysian companies. Intellectual capital is of particular interest in this study due to the increasing importance of this resource to companies, especially in this era of knowledge economy and due to the increasing prevalence of knowledge-intensive industries. According to Baron (2003), about 75% of market value relies on intangible assets, most notably human capital, which is one of the components of intellectual capital. Yet there is very little account taken of the contribution people make to the success of the organisation. The impact of intellectual capital on markets is huge as it is concerned with value creation for the long-term development of capabilities and competencies. Baron (2003) also argued that in advanced economies, intellectual capital is the only distinctive asset which cannot be imitated easily (especially human capital which includes skills, talent, motivation, and know how of people). According to Bontis et al. (2000), the estimates in 1999 show that 37.5% of the workforce in Malaysia is in the service sector, which is a knowledge-intensive sector. These demonstrate that intellectual capital and its management is gaining importance in the country. Due to the rise in the number of knowledge-intensive organisations and the increasing awareness of the importance of intellectual capital to firms, it is doubtful that traditional measures are adequate to represent the true picture of corporate performance. A more comprehensive measure is needed to incorporate intellectual capital resources of a firm.

Corporate performance can be measured using long-term market performance measures and other performance measures that are non market-oriented measures or short-term measures. Some examples of these measures include market value added (MVA), economic value added (EVA), cash flow growth, earnings per share (EPS) growth, asset growth, dividend growth, and sales growth (Coles et al., 2001, Abdullah, S.N., 2004). Dehaene et al. (2001) used return on equity (ROE) and return on assets (ROA) as proxies for corporate performance in Belgian companies and Chen et al. (2005) utilised the market-to-book ratio in their research on firms in Hong Kong. Judge et al. (2003) used a series of indicators including financial profitability, customer satisfaction, product/service quality, capacity utilisation and process improvements to assess firm performance. In their article, Roos and Roos (1997) defined intellectual capital as the "…sum of the 'hidden' assets of the company not fully captured on the balance sheet", and this includes the tacit knowledge of the members of the organisation, and what is left in the company when they leave. Intellectual capital is argued to be the most important source of sustainable competitive advantage and it comprises both human capital and structural capital. Human capital represents the knowledge, skills, motivation and capability of the individual employee to provide solutions to the customers. On the other hand, structural capital represents the organisational capability to meet market requirements. A significant part of structural capital is customer capital, which basically is the relationship developed with the customers, suppliers, network partners and investors of the company. Another part of structural capital is organisational capital, and this includes business processes, strategic processes, productions processes and business development capital (Roos and Roos, 1997).

Realising the importance of intellectual capital to a firm's performance, there have been numerous attempts and approaches to intellectual capital reporting. According to Mouritsen et al. (2001), attempts made by companies to report on intellectual capital include the preparation of intellectual capital statements that combine numbering, visualisation and narration to account for organisational value creation. The emphasis here is on creating value, and Cooper (2000), in his paper, likened intellectual capital to the conversion of knowledge into something valuable. In other words, intellectual capital is defined as intellectual material that has been formalised and leveraged to produce higher valued asset. Value creation is presented as an effect of the connections between human, structural and customer capital. One of the approaches to "valuing" suggests that the object of "valuing" is to create (more) value and to generate value via the transformation or improvement of corporate routines and practices. So, 'value added' in this study is defined as the wealth created (or contributed) by the firm through the utilisation of its key productive resources (Ho & Williams, 2003).

In their study, Barsky and Marchant (2000) stated that due to the nature of knowledge-based organisations and their dependence on intellectual capital, the way they view the value creation process has also changed. Traditional financial measures such as return on capital and earnings per share (measures that constitute most corporate performance management systems) tell investors and management little about the true performance of the company. Since accounting practices have failed to keep pace with the growing importance of intellectual capital, many of these assets remain unaccounted for and unmonitored. Without tools to capture and measure intellectual capital, many firms wind up mismanaging their intellectual assets or, worse, destroying knowledge value simply because managers misunderstand the nature of the company's resources.

Shareholders (principals) hire managers (agents) to make decisions that are in the best interest of the shareholders. The separation of ownership from control implies that the principal (shareholders) cannot exercise full control over managerial actions. Opportunistic behaviour is assumed in agency theory, and there exists information asymmetry between the principals and agents (managers). Information asymmetry occurs where management have the competitive advantage of information within the company over that of the owners. These lead to two important conflicts between management and shareholders. One of the conflicts is in terms of differing objectives. In his paper, Grinyer (1995) argued that the main objective of the shareholders is to maximise personal wealth. Meanwhile, managers may have a wider range of economic and psychological needs such as power, reputation and higher salaries. This means some decisions of managers are motivated by self-interest, which reduces the welfare of the principal. In other words, given their ability to run the company with little check on their activities by shareholders, managers may be tempted to place less emphasis on maximising shareholder returns (profitability, share price and dividend payouts) and more on expanding the asset base, increasing turnover at the expense of profitability and paying themselves higher salaries. Another source of conflict is differing time horizons, where the agent may be eager to take actions which have relatively short-run pay-offs in order to demonstrate success, whereas shareholder interests may be better served by longer-term actions (Evans & Weir, 1995).

Board composition refers to the number of independent non-executive directors on the board relative to the total number of directors. An independent non-executive director is defined as independent directors who have no affiliation with the firm except for their directorship (Clifford and Evans, 1997). There is an apparent presumption that boards with significant outside directors will make different and perhaps better decisions than boards dominated by insiders. The Malaysian Code on Corporate Governance (2000) recommends, as a best practice, that there needs to be balance on the board of directors with at least one third of the board members should be independent directors. This is to ensure the effectiveness of the independent directors in maintaining the objectivity in board decisions. The argument for the need of independent non-executive directors on the board substantiated from the agency theory which states that due to the separation between ownership and control, managers (given the opportunity) would tend to pursue their own goals at the expense of the shareholders (Jensen and Meckling, 1976). Hence, by having independent non-executive directors on the board, these directors would help to monitor and control the opportunistic behaviour of management, and assist in evaluating the management more objectively. Furthermore, Brickley and James (1987) argued that outside directors also contribute to reduce management consumption of perquisites. In the absence of such monitoring by outside directors, managers might have the incentive to manage earnings in order to project better performance results and hence increase their compensation. Empirically, studies on the association between independent non-executive directors and firm performance have shown mixed results. In their study among Belgian companies, Dehaene et al. (2001) found a significant positive relationship between the number of external directors and return on equity, which lends support to the notion that outside directors provide superior benefits to the firm as a result of their independence from firm management and this is taken into account by investors in making investment decisions. This is also supported by Dahya and McConnell (2003) who found evidence in the UK that investors appear to view appointments of outside CEOs as good news, and this is reflected in the announcement period stock returns. On top of that, in their study, Lee et al. (1999) found that the appointment of a financial outside director to the board of a public corporation is associated with positive abnormal returns among medium-size firms. These firms, which may have limited access to financial markets and less financial expertise, benefit substantially from these appointments.

On the contrary, there were also studies that found negative association between independent non-executive directors and firm performance. One such study is by Agrawal and Knoeber (1996), who discovered a significant negative relationship between board outsider and firm performance. This is also supported by the findings of Bhagat and Black (1999) who established that firms with majority outside directors perform worse than other firms. These studies show that independent non-executive directors do not necessarily have positive impact on firm performance, implying that in these cases perhaps the independent non-executive directors do not play their roles effectively. Apart from their role in firm performance, independent non-executive directors may also have an effect on the comprehensiveness of financial disclosures by firms. A study by Chen and Jaggi (2000) discovered that there is a significant positive relationship between the ratio of independent non-executive directors on corporate board and the comprehensiveness of financial disclosures. This implies that the inclusion of such directors on boards improves the firm's compliance with the disclosure requirements, which in turn enhances the comprehensiveness and quality of disclosures. So, independent non-executive directors are useful for monitoring board activities and improving the transparency of corporate boards.

A potentially important factor that may reduce manager–shareholder conflicts is stock ownership by board members (both executive and non-executive). To the extent that board members own part of the firm, they develop shareholder-like interests and are less likely to engage in behaviour that is detrimental to shareholders. In other words, managerial shareholdings help align the interests of shareholders and managers since as the company's performance increases, the managers benefit via their equity interests in the company (Jensen and Meckling, 1976). Therefore, managerial ownership is argued to be inversely related to agency conflicts between managers and shareholders, and to be positively related to corporate performance. However, Morck et al. (1988) argued that higher levels of managerial equity ownership may decrease financial performance since managers with significant ownership stakes may gain such power that they neglect or become less considerate of the interests of other shareholders. This is because they are in the position where they have considerable voting rights and they are also the ones who make the judgement on how to run the company. This may lead them to make decisions that confer benefits to themselves to the detriment of other stakeholders. One of the key monitoring mechanisms advocated by the agency perspective is the separation of the roles of CEO from chairperson. If the two roles are not separated, this means that the CEO also chairs the group of people in charge of monitoring and evaluating the CEO's performance, and hence duality exists. This situation also gives rise to possible conflict of interest and may impair the independence of the monitoring group. This is because in such situation, the ability of the CEO/Chairperson to exercise independent self-evaluation is questionable (Rechner and Dalton, 1989). Fosberg and Nelson (1999) discovered that firms that switch to the dual leadership structure (separated roles between the CEO and the chairman) to control agency problems experienced a significant improvement in performance which is measured by the operating income before depreciation, interest and taxes to total assets ratio. Dehaene et al. (2001) found evidence that where the functions of chairman and chief executive are combined, the return on assets is significantly higher than otherwise, which suggests a positive relationship between duality and firm performance. They argued that when the chairman is also active as the CEO in the daily activities of the firm, he will try to invest as much as possible to increase the size of the firm or to boost his personal status.

On the contrary, Rechner and Dalton (1989) found no significant difference between shareholders returns of companies with CEO duality and those that separate the two roles. The study consists of companies from the Fortune 500 group, and the data was collected from the year 1978 to 1983. They concluded that there is little justification to infer that it is an unprofitable move for a company to have CEO duality. There are also studies that examine the relationship between CEO duality and corporate voluntary disclosures. Voluntary disclosure is seen as an act of accountability and transparency on the firm's behalf and poor disclosure has also been blamed for the recent financial crisis in the Asian region. So, more demands have been made of the board to disclose more corporate information. Gul and Leung (2004) found that in Hong Kong, when the roles of the CEO and the chairman are combined, the levels of voluntary corporate disclosures are lower. This also supports the notion for the two roles to be separated. However, they added that the negative relationship is weaker for firms with higher proportion of independent non-executive directors. This suggests that the expertise of independent non-executive directors moderates the negative CEO duality and corporate disclosure relationship. The MCCG does not specify a desirable board size; instead it proposes that every board examines its size while keeping in mind the impact of the number upon its effectiveness. The findings of previous studies have shown mixed results with regard to the relationship between this board characteristic and performance. A study involving a sample of large U.S. industrial corporations between 1984 and 1991 by Yermack (1996) presents evidence consistent with theories that small boards of directors are more effective, as he found an inverse association between board size and firm value. In that study, the effect of board size on two variables, namely investors' valuation of the company and profitability ratios, are tested. It is found that when the board consists of between four to ten members, the investors' valuation of the company and the profitability ratios decrease steadily. When the board size is more than ten, there is no significant relationship between the board size and investors' valuation, but the profitability ratios decrease less rapidly. The result of his study is robust to numerous controls for company size, industry membership, inside stock ownership, growth opportunities, and alternative corporate governance structures.

The main purpose of this study is to examine the relationship between board structure and the VA efficiency of physical and intellectual resources within the Malaysian business context. This study is a replication of the study conducted by Ho and Williams (2003) who investigated the link between corporate board features and corporate performance for a sample of publicly traded firms in South Africa, Sweden and the UK. In contrast to most prior literature, performance in their study is defined as the value added (VA) efficiency of the firm's resources rather than the more commonly used financial terms or profitability ratios. The VA efficiency of the firm's total resources is calculated using the Value Added Intellectual Coefficient (VAIC) methodology developed by Ante Pulic (1998).

2. Methods

2.1 Theoretical framework

Based on the extensive literature, four board characteristics (board composition, directors' ownership, CEO duality and board size) have been identified as possibly having an impact on corporate performance and these characteristics are set as the independent variables in the framework.

2.2 Control Variables

Drawing from a review of related corporate disclosure researches (for example, Vafeas and Theodorou (1998)), five control factors (profitability, leverage, dividend yield, R&D sensitivity and firm size) are included in the theoretical model designed for this study. These factors have been known to have an impact on corporate performance, and hence need to be controlled in the study. According to Wahab et al. (2004), the inclusion of the control factors also reduces the risk of model misspecification due to missing variables. The inclusion of the control factors in the theoretical model is depicted in **Figure 1**:

The dependent variable is the VA efficiency of total resources, which is used as a measure of corporate performance. The relationship between each of the independent variable and corporate performance is hypothesised as follows:

H_1: There is a positive relationship between the percentage of independent non-executive directors and the VA efficiency of firm's total resources.

H_2: There is a positive relationship between directors' ownership and the VA efficiency of firm's total resources.

H_3: There is a negative relationship between CEO duality and the VA efficiency of firm's total resources.

H_4: There is a negative relationship between board size and the VA efficiency of firm's total resources.

2.3 sampling, Instrumentation and Variable Measurement

2.3.1 Sample

All companies listed on the main board of Bursa Malaysia constitute the population of this study. However, firms belonging to the financial services industry and regulated utility companies are excluded from the population. This is due to the special regulatory environment in which they operate. This follows from the argument that regulation masks the efficiency differences across firms, potentially rendering governance mechanisms less important (Vafeas & Theorodou, 1998, Singh & Davidson, 2003).

From this set of population, 75 companies have been randomly selected using the random case selection function of the SPSS software. These 75 companies are selected without discriminating between the different industries included in the study. Besides that, the convenient sampling technique is also applied in this research, where the availability of the annual reports of the chosen companies on the Bursa Malaysia website also plays a determining role in the inclusion of the company in the final list. In other words, companies that have been chosen by the random sampling function but do not have annual reports readily available on the Bursa Malaysia website are eliminated from the sample list and the random sampling technique is repeated to replace these companies.

The sample size of 75 is about 13% of the population and is deemed sufficient for the purpose of the statistical analyses that are planned to be performed to study the relationship between board structure and VA efficiency of total resources of the firms. It should also be noted that the small sample size may not render the results of the study to be generalisable.

2.3.2 Content analysis

This study is based on content analysis of the annual reports of the sample companies in order to obtain data to measure the dependent and independent variables. According to Guthrie et al. (2004), content analysis is one of the more widely used research method applied in investigating the frequency and type of intellectual capital reporting. A technique for gathering data, content analysis involves codifying qualitative and quantified information into pre-defined categories in order to derive patterns in the presentation and reporting of information (Guthrie and Petty, 2000).

The data required for the purpose of this study is collected from 2003 fiscal year annual reports of the chosen publicly traded firms listed on Bursa Malaysia. The data collection technique used is mainly content analysis of these annual reports and database searches using the Datastream facilities. In the case of unavailability of data on the database, manual search on the annual reports of the companies are conducted. The manual search also applies to the information required from other parts of the annual reports, such as from the Directors' Profile and the Statement on Corporate Governance. In this study, content analysis involves reading the annual reports of each company in the sample and coding the information contained therein.

2.3.3 Variable measurement

According to the VAIC concept, human capital is not part of input (costs) but a resource equal to capital employed. Hence, firms should strive to achieve a maximum result (increase the efficiency of its resources in adding value) by utilising its resources as well as possible. Also known as the Austrian Approach, the VAIC methodology is used in this study to measure the value added efficiency of the companies' physical and intellectual capital. The methodology is considered a universal indicator and it shows the abilities of a company in value creation and representing a measure for business efficiency in a knowledge-based economy (Pulic, 1998). It also enables the collection of intellectual capital leverage to key success processes and is easy to calculate using information already accounted and reported in annual reports (thus minimises the costs to preparers and stakeholders).

Formally, VAIC is a composite sum of three indicators formally termed: (1) Capital Employed Efficiency (CEE) – indicator of VA efficiency of capital employed; (2) Human Capital Efficiency (HCE) – indicator of VA efficiency of human capital; and (3) Structural Capital Efficiency (SCE) – indicator of VA efficiency of structural capital. This is represented by Formula 3.1:

$$\mathbf{VAIC = CEE + HCE + SCE} \tag{3.1}$$

Where VAIC = value added intellectual coefficient

CEE = capital employed efficiency coefficient

HCE = human capital efficiency coefficient

SCE = structural capital efficiency coefficient

It is important to note here that even though the term VAIC stands for value added intellectual coefficient, it actually represents the value-added efficiency of the firm's total resources, and not just the intellectual resources. It is used as an alternate measure of firm performance in this study, as opposed to the more commonly used profitability and other financial ratios in previous studies.

The CEE, HCE and SCE are indicators of the efficiency of each of these elements in creating (or adding) new value. For example, if HCE is 2.8, this means that each RM1 invested in human capital has added value of RM1.80 to the firm. CEE represents the efficiency of VA of physical resources, while the total of HCE and SCE represents the efficiency of VA of intellectual capital A higher VAIC coefficient indicates a better management of physical (CE) and intellectual (HC and SC) resources.

The first step of calculating the three indicators is to calculate the firm's total VA:

$$\mathbf{VA = I + DP + D + T + M + R} \qquad \mathbf{(3.2)}$$

Where:

I = Interest expense

DP = Depreciation expense

D = Dividends

T = Corporate taxes

M = Equity of minority shareholders in net income of subsidiaries

R = Retained profits

All the above information is obtained from the financial statements of the companies. After obtaining the VA, the three indicators in the VAIC methodology (CEE, HCE and SCE) are calculated using the following formulas:

$$\mathbf{CEE = VA/CE} \qquad \mathbf{(3.3)}$$

$$\mathbf{HCE = VA/HC} \qquad \mathbf{(3.4)}$$

$$\mathbf{SCE = SC/VA} \qquad \mathbf{(3.5)}$$

Each of the above indicators and its formulas are described in Table 1:

The regression results of the VAIC against the independent and control factors will form the evidence supporting or rejecting hypotheses $\mathbf{H_1}$, $\mathbf{H_2}$, $\mathbf{H_3}$ and $\mathbf{H_4}$.

For each of the independent variables, the variable name and measurement technique is shown in Table 2 below. The proxy measures for the control factors are described in Table 3.

2.4 Statistical analyses

For the purpose of empirical analysis, this study uses descriptive analysis, Pearson correlation analysis and linear multiple regression as the underlying statistical tests. A descriptive analysis of the data obtained is conducted to obtain sample characteristics. The Pearson correlation analysis is executed to check for multicollinearity problem among the variables.

The multiple regression analysis is performed on the dependent variable, VAIC, to test the relationship between the independent variables (board structure features) with VA efficiency of firms' total resources. The regression model utilised to test the relationship between the board characteristics and VAIC is as follows:

$$\text{VAIC} = ß_0 + ß_1\text{OutDir} + ß_2\text{InsOwn} + ß_3\text{Duality} + ß_4\text{BrdSize} + \text{Control Variables} + \varepsilon_i$$

Where:

$ß_0$ = Intercept coefficient

$ß_i$ = Coefficient for each of the independent variables

Control Variables = Represents the control factors included in the regression analysis, which are profitability (ROA), leverage (Leverage), dividend yield (DivYield), R&D sensitivity (R&D) and firm size (FrmSize).

ε_i = Error term

3. Results and Discussion

3.1 Board characteristics and ownership

Consistent with expectation, the mean for directors' ownership suggests that directors of publicly listed firms in Malaysia generally have sizeable ownership stakes in the company compared to their counterparts in Western developed economies such as the Sweden and the United Kingdom (see, Ho and Williams, 2003). This is perhaps due to the higher number of family-owned and managed companies in Malaysia. The descriptive statistics for the independent variables indicate that the average number of directors on the board in the selected companies is about 8 persons. On average, the VA efficiency coefficient for the total resources (VAIC) of the firms in this study is 5.020, and the mean VA efficiency coefficients for each intellectual and physical resource are 4.829 and 0.191 respectively. The noticeable difference between the mean and median of the VA efficiency of physical resources indicates the presence of extremely low values for that variable that brings down the mean figure to 0.191 even though 50% of the sample obtained a coefficient of more than 0.311. This is also evident from the high standard deviation for the variable, which is 16.99. The average VAIC of the firms in this study is quite high compared to the findings by Ho and Williams (2003) in their study. They found that UK firms have the highest average VAIC of 5.317, followed by South African firms with 4.270 (Table 4).

In Table 5, it can be seen that 70.7% of the companies comply with the recommendations of the MCCG by separating the roles of the chairman and the CEO of the company. Relative to the study of Swedish, British and South African firms by Ho and Williams (2003), the level of duality of 29.3% of the sample data in this study is fairly low. However, this is higher than a previous finding in the Malaysian setting by Abdul Rahman and Mohd Haniffa (2002). They found that only 11.8% of their sample practises CEO duality, based on the pooled data over a period of five years from 1996 to 2000.

3.2 Regression analysis

A Pearson correlation analysis is performed on the independent variables to check for the degree of multicollinearity among the variables. The results are shown in Table 6, and it can be seen that directors' ownership is significantly correlated to dividend yield (at p = 5%), while profitability is significantly correlated to leverage and dividend yield (both at 1% level). There is also a significantly positive association between dividend yield and R&D sensitivity (at p = 1%) and firm size (at p = 5%). The results also indicate a positive correlation between R&D sensitivity and firm size, at the 5% significance level. The highlighted correlations in the box represent the correlation coefficients between the independent variables and the control factors.

Even though there are significant correlations among some of the independent variables and also between some of the independent variables and the control factors, none of the coefficients exceeds 0.8, which is used as an indicator of serious multicollinearity (Gujarati, 1992). Hence, it may be concluded that multicollinearity is not a serious problem in this case.

By setting the VAIC coefficient as the dependant variable, and the board characteristics and control factors as independent variables, a regression analysis is performed. This aims to address the first research objective of testing the relationship between board characteristics and the VA efficiency of firms' total resources. The results of this analysis are shown in Table 7.

From the output of the analysis in Table 7, the analysis of variance (ANOVA) test returns a significant p-value of 0.000 which means there is sufficient evidence to infer that at least one of the explanatory variables is linearly related to VAIC, and the model seems to have some validity. The regression results which indicate that many of the independent and control variables are significant also supports the prior conclusion that there is no indication of the existence of serious multicollinearity in these models (Gujarati, 1992).

The R^2 value, which indicates the explanatory power of the independent variables, is 0.825. This means that 82.5% of the variation in the VA efficiency coefficient of firms' total resources is explained by the variation in the independent and control variables. According to Keller and Warrack (2003), the R^2 value does not have a critical value that enables a conclusion to be drawn. However, in general, the higher the R^2 value, the better the model fits the data. In this study, the R^2 value is quite high since only about 20% of the variation in the dependent variable is unexplained by the model, denoting a strong relationship between the explanatory variables and the VAIC.

3.3 The Impact of Board Composition on VAIC

Table 7 shows that the coefficient for percentage of independent non-executive directors on the board (t = 2.92) is significant at the 1% level. In addition, the coefficient is positive at 1.31. This means that for each additional increase in the percentage of independent non-executive directors to total board size, the VAIC increases on average by 0.131%, holding other explanatory variables constant. Hence, it can be inferred that hypothesis $\mathbf{H_1}$ is supported. In other words, as the percentage of independent non-executive directors on the board increases, the VA efficiency of the firms' total resources also increases. This is consistent with Ho and Williams' (2003) findings with respect to Swedish sample firms in their study. This finding also supports the notion that independent non-executive directors contribute to corporate performance as a whole.

3.4 The Impact of Directors' Ownership on VAIC

Table 7 depicts that the coefficient for directors' ownership is insignificant even at the 10% level. This reveals that there is insufficient evidence to infer that there is a linear relationship between directors' share ownership in the company and the VA efficiency of the company's total resources. Hence, hypothesis $\mathbf{H_2}$ is rejected. Nevertheless, it is interesting to observe that the coefficient is negative, which is contrary to the expectation in the theoretical model. In their study, Ho and Williams (2003) found that the coefficient for directors' ownership is significantly positive for their sample of Swedish firms, but not significant for South African and British samples. The insignificant coefficient in this study may mean that directors in Malaysia are not motivated by equity interests in the company. This is perhaps because they are more highly rewarded in the form of perquisites and allowances.

3.5 The Impact of CEO Duality on VAIC

It is important to note here that the CEO duality variable is a dummy variable, since it constitutes the value of either 0 or 1. Table 7 demonstrates that the coefficient for CEO duality ($t = 0.156$) is insignificant even at the 10% level. The insignificant t value for this variable means that there is no significant difference in the VA efficiency of total resources between firms that separate the roles of the CEO and the chairman and those that practice CEO duality. However, an interesting point to note here is that the coefficient is positive while the theoretical model predicts a negative relationship between this variable and the dependent variable (VAIC). Thus, hypothesis H_3 is rejected and it can be concluded that there is no evidence to support the notion that there is a relationship between CEO duality and VAIC.

The findings contradict the results of a prior research on firms in Malaysia by Abdul Rahman and Mohd Haniffa (2002). In their study, CEO duality is found to reduce the effectiveness of the board, resulting in poor company performance. However, their study uses accounting measures (ROA and ROE) as proxies for corporate performance, which may account for the difference in the outcomes. Furthermore, as argued by Judge et al. (2003), even with the existence of a formal law that separates the roles of CEO and the chairperson, informal activities may still undermine the law and firm performance. So, even though most of the firms in this study separate the two roles, it may just be a case of form over substance, and hence the company fails to reap the actual benefits of such practice.

3.6 The Impact of Board Size on VAIC

The results of the regression analysis in Table 7 show that the coefficient for total number of directors (board size) is significant at the 10% level. However, the coefficient is positive, suggesting that there is a moderately significant positive relationship between board size and VAIC. This is contrary to the theoretical model and the stated hypothesis, which predicts a negative relationship between board size and company performance. Hence $\mathbf{H_4}$ is rejected. This finding indicates that a larger board size contributes more towards firm performance as a whole. This suggests that a larger board size means that there are more ideas and skills that can be shared among board members. It presents evidence against the argument by Eisenberg et al. (1998) who suggested that a large board size is more vulnerable to being overpowered by the CEO.

It also proposes that the board of directors in Malaysian firms perform more effectively in a larger group. This is opposite to the findings by Ho and Williams (2003), where they concluded that board size is statistically insignificant against VAIC in their regression for all three countries (South Africa, Sweden and the UK). This implies that board size plays a more imperative role in Malaysia compared to these countries.

3.7 Control Factors

The regression results also show that for the control factors included in the analysis, four of them (ROA, Leverage, R&D Sensitivity and Firm Size) are significant at the 1% level. The coefficient for dividend yield is found to be insignificant, even at the 10% level. As expected, all relationships between the control factors and VA efficiency coefficient are positive, except for R&D Sensitivity which is negatively related to the dependent variable. This is perhaps due to the lack of awareness of the importance of R&D disclosure in the financial statements, and such disclosure may viewed negatively by the users.

One of the limitations of this study is its small sample size, which consists of only 75 companies. Most studies involving corporate governance structure and firm performance had used all non-financial companies listed on the board in its sample (see, for example, Abdul Rahman and Mohd Haniffa, 2002). Ho and Williams (2003) used about 84 to 108 companies for each of the countries included in their study, and Judge et al.'s (2003) study comprises of 113 firms. Due to the numerous variables and items that need to be collected, the choice of 75 companies in this study is more manageable, and is still valid for statistical analysis. However, an extension of the study to include all the companies listed on Bursa Malaysia would further enhance the generalisability of the findings

4. Conclusion

The purpose of this research is to examine the importance of one of corporate governance aspects, namely board structure. Compared to previous literature on board structure-performance relationship, this study uses the VAIC methodology to measure corporate performance, which includes the physical and intellectual resource bases of the firm. In general, the results of this study provide evidence that a higher proportion of independent non-executive directors on the board has a positive impact on firm performance based on the VAIC measurement. Meanwhile, at a higher significance level (10%), the board size is found to have a positive impact on the VA efficiency of firms' total resources. However, the effects of directors' ownership in the company and CEO duality on the VA efficiency of firms' total resources are not established. The above findings are also similar for the relationships between each of the board characteristics and the VA efficiency of the intellectual resources of the firms.

The results imply that the requests for a minimum number (one-third of the board) of independent non-executive directors on the board by the Bursa Malaysia Listing Requirements and the MCCG are deemed very important. This is because independent non-executive directors possess a diverse background, attributes, characteristics and expertise, which may improve board processes and decision-making, and consequently firm performance. Independent non-executive directors also play a vital role in the long-term performance of the company, as they contribute significantly to the performance of intellectual resources of the firm. There is also evidence to suggest that a large board size performs effectively and there seems to be no communication and coordination problem among the board members. This is contrary to most US studies which found that a small board size is more effective and performs better. This is perhaps due to the differences in the culture and nature of the firms, as this particular study is conducted in a developing Asian country. Regardless of the board composition and structure adopted by the company, it is important that they make adequate disclosure on this matter in the annual reports so that users can decide for themselves.

It would also be more meaningful to perform a comparative analysis between Malaysia and other countries. This comparative analysis could serve to gauge Malaysia's VAIC performance and also its intellectual capital performance against countries with similar or opposite settings. A comparative analysis could be performed between Malaysia and another developing nation, or with a developed nation. Other corporate governance variables may also be included in the model, such as cross-directorship and family ownership.

References

Abdullah, S.N. (2004). Board Composition, CEO Duality and Performance among Malaysian Listed Companies, *Corporate Governance*, 4, 4, pp 47-61.

Abdul Rahman, R. and Abdul Hamid, N. (2003). The Role of Board Characteristics as a Response to Agency Problem among Acquiring Firms in Malaysia, presented at the Asian Accounting Association Conference, Seoul, Korea, 20-21 October.

Abdul Rahman, R. and Haniffa, R. (2003). Internal Control Mechanisms and Performance of Malaysian Listed Firms, presented at the ANZAM 2003 Conference, Perth, Australia, 2-4 December.

Abdul Rahman, R. and Mohd Haniffa, R. (2002). CEO Duality and Corporate Performance: Evidence in Malaysia, presented at the International Conference on Corporate Governance: Trends and Challenges in the Millennium, Kuala Lumpur, Malaysia, 29-30 October.

Agrawal, A. and Knoeber, C.R. (1996). Firm Performance and Mechanisms to Control Agency Problems Between Manager and Shareholders, *Journal of Financial and Quantitative Analysis*, 31, pp. 377-389.

Baron, A. (2003). Measuring Human Capital: The Key to Organizational Success?, *Development and Learning in Organizations*, 17, 3, pp 7-9.

Barsky, N.P. and Marchant, G. (2000). The Most Valuable Resource Measuring and Managing Intellectual Capital, *Strategic Finance*, 81, 8.

Bhagat, S. and Black, B. (1999). The Uncertain Relationship between Board Composition and Firm Performance, *Business Lawyer*, 54, pp. 921-963.

Bontis, N., Chua, W.C.K. and Richardson, S. (2000). Intellectual Capital and Business Performance in Malaysian Industries, *Journal of Intellectual Capital*, 1, 1, pp 85-100.

Bontis, N., Dragonetti, N.C., Jacobsen, K. and Roos, G. (1999). The Knowledge Toolbox: A Review of the Tools Available to Measure and Manage Intangible Resources, *European Management Journal*, 17, 4, pp 391-402.

Brennan, N. (2001). Reporting Intellectual Capital in Annual Reports: Evidence from Ireland, *Accounting, Auditing & Accountability Journal*, 14, 4, pp 423-436.

Brickley, J.A., Coles, J.L. and Jarrell, G. (1997). Leadership Structure: Separating the CEO and Chairman of the Board, *Journal of Corporate Finance*, 3, pp 189-220.

Brickley, J.A. and James, C.M. (1987). The Takeover Market, Corporate Board Composition and Ownership Structure: The Case of Banking, *Journal of Law and Economics*, 30, pp. 161-180.

Byrd, J.W. and Hickman, K.A. (1992). Do Outside Directors Monitor Managers? Evidence from Tender Offer Bids, *Journal of Financial Economics*, 32, pp 195-221.

Chen, C.J.P. and Jaggi, B. (2000). Association between Independent Non-Executive Directors, Family Control and Financial Disclosures in Hong Kong, *Journal of Accounting and Public Policy*, 19, 4-5, pp. 285-310.

Chen, Z., Cheung, Y.L., Stouraitis, A. and Wong, A.W.S. (2005). Ownership Concentration, Firm Performance, and Dividend Policy in Hong Kong, *Pacific-Basin Finance Journal*, Article in Press – Corrected Proof, available online 23 March.

Clifford, P. and Evans, R. (1997). Non-Executive Directors: A Question of Independence, *Research and Theory Papers*, 5, 4.

Coles, J.W., McWilliams, V.B. and Sen, N. (2001). An Examination of the Relationship of Governance Mechanisms to Performance, 27, 1, pp 23-50.

Cooper, J. (2000). Turning Knowledge into Value (Against a Backdrop of Tangibles, the Intangible Reality of Intellectual Capital Flows Throughout the Realm of Biotechnology Firms across Canada), *CMA Management*, 74, 8, p22-25.

Cutting, B. and Kouzmin, A. (2000). The Emerging Patterns of Power in Corporate Governance - Back to the Future in Improving Corporate Decision Making, *Journal of Managerial Psychology*, 15, 5, pp. 477-507.

Dahya, J. and McConnell, J.J. (2003). Outside Directors and Corporate Board Decisions, *Journal of Corporate Finance*, Article in Press, Corrected Proof.

Dehaene, A., De Vuyst, V. and Ooghe, H. (2001). Corporate Performance and Board Structure in Belgian Companies, *Long Range Planning*, 34, 3, pp 383-398.

Demb, A. and Neubauer, F.F. (1992). *The Corporate Board*, Oxford University Press, New York.

Eisenberg, T., Sundgren, S. and Wells, M. (1998). Larger Board Size and Decreasing Firm Value in Small Firms, *Journal of Financial Economics*, 48, pp 35–54.

Evans, J. and Weir, C. (1995). Decision Processes, Monitoring, Incentives and Large Firm Performance in the UK, *Management Decision*, 33, 6, pp 32-38.

Fosberg, R.H. and Nelson, M.R. (1999). Leadership Structure and Firm Performance, *International Review of Financial Analysis*, 8, 1 pp 83-96.

Ghosh, A. (1998). Does Accounting Based Performance Really Improve Following Corporate Acquisitions?, Working Paper, Zickling School of Business, Baruch College (CUNY), New York, USA.

Goodstein, J., Gautum, K. and Boeker, W. (1994). The Effect of Board Size and Diversity on Strategic Change, *Strategic Management Journal*, 15, pp. 241-250.

Grace, M., Ireland, A. and Dunstan, K. (1995). Board Composition, Non-Executive Directors' Characteristics and Corporate Financial Performance, *Asia-Pacific Journal of Accounting*, pp. 121-137.

Grinyer, J.R. (1995). A New Concept for Amortising Goodwill, University of Dundee.

Gujarati, D. (1992). Essentials of Econometrics, Singapore: McGraw-Hill.

Gul, F.A. and Leung, S. (2004). Board Leadership, Outside Directors' Expertise and Voluntary Corporate Disclosures, *Journal of Accounting and Public Policy*, 23, 5, pp. 351-379.

Guthrie, J. and Petty, R. (2000). Intellectual Capital: Australian Annual Reporting Practices, *Journal of Intellectual Capital*, 1, 3, pp 241-251.

Guthrie, J., Petty, R., Yongvanich, K. and Ricceri, F. (2004). Using Content Analysis as a Research Method to Inquire into Intellectual Capital Reporting, *Journal of Intellectual Capital*, 5, 2, pp 282-293.

Hampel Report, The (1998). www.legal500.com, 15/10/2004.

Herremans, I.M. and Isaac, R.G. (2004). The Intellectual Capital Realization Process (ICRP): An Application of the Resource-based View of the Firm, *Journal of Managerial Issues*, 16, 2, pp 217-231.

Ho, C.A. and Williams, S.M. (2003). International Comparative Analysis of the Association between Board Structure and the efficiency of Value Added by a Firm from its Physical Capital and Intellectual Capital Resources, *The International Journal of Accounting*, 38, 4, pp 465-491.

Huang, C.C. (2004). Intellectual Capital Reporting, *Accountants Today*, August, pp 22-24.

Ingley, C.B. and van der Walt, N.T. (2003). Board Composition: Building Better Boards, *Corporate Governance*, 3, 4, pp 5-17.

Jensen, M.C. and Meckling, W.H. (1976). Theory of the Firm: Managerial Behaviour, Agency Costs and Ownership Structure, *Journal of Financial Economics*, 3, pp. 305-360.

Judge, W.Q., Naoumova, I. and Koutzevoi, N. (2003). Corporate Governance and Firm Performance in Russia: An Empirical Study, *Journal of World Business*, 38, 4, pp 385-396.

Keller, G. and Warrack, B. (2003). Statistics for Management and Economics, United States: Thomson Brooks/Cole.

Klapper, L.F. and Love, I. (2004). Corporate Governance, Investor Protection, and Performance in Emerging Markets, *Journal of Corporate Finance*, 10, 5, pp 703-728.

Kuala Lumpur Stock Exchange Listing Requirements, (1994).

Laing, D. and Weir, C.M. (1999). Governance Structures, Size and Corporate Performance in UK Firms, *Management Decision*, 37, 5, pp 457-464.

Lee, Y.S., Rosenstein, S. and Wyatt, J.G. (1999). The Value of Financial Outside Directors on Corporate Boards, *International Review of Economics and Finance*, 8, 4, pp. 421-431.

Malaysian Code of Corporate Governance, The (2000).

Mohamed Abdullah, M.R. (2003). Functions of an Independent Director, in *Independent Directors: Perceptions, Roles and Responsibilities*, Malaysian Institute of Corporate Governance, Percetakan Sempana Score Sdn Bhd, Kuala Lumpur.

Morck, R., Shleifer, A. and Vishny, R.W. (1988). Management Ownership and Market Valuation: An Empirical Analysis, *Journal of Financial Economics*, 20, pp 293-315.

Mouritsen, J., Larsen, H.T. and Bukh, P.N. (2001). Valuing the Future: Intellectual Capital Supplements at Skandia, *Accounting, Auditing & Accountability Journal*, 14, 4, pp 399-422.

Pass, C., Robinson, A. and Ward, D. (2000). Performance Criteria of Corporate Option and Long-Term Incentive Plans: A Survey of 150 UK Companies 1994-1998, *Management Decision*, 38, 2, pp 130-137.

Patton, A. and Baker, J.C. (1987). Why Won't Directors Rock the Boat?, *Harvard Business Review*, 65, pp. 10-18.

Pearce, J.H. and Zahra, S.A. (1992). Board Composition form a Strategic Contingency Perspective, *Journal of Management Studies*, 29, pp. 411-438.

Peng, M.W., Buck, T. and Filatotchev, I. (2003). Do Outside Directors and New Managers Help Improve Firm Performance? An Exploratory Study in Russia Privatization, *Journal of World Business*, 38, 4, pp. 348-360.

Pettigrew, A.M. and McNulty, T. (1995). Power and Influence in and around the Boardroom, *Human Relations*, 48, pp. 845-873.

Pulic, A. (1998). Measuring the Performance of Intellectual Potential in Knowledge Economy, http://www.vaic-on.net, 2 August 2004.

Rechner, P.L. and Dalton, D.R. (1989). The Impact of CEO as Board Chairperson on Corporate Performance: Evidence vs. Rhetoric, *The Academy of Management Executives*, 111, 2, pp 141-143.

Randoy, T. and Goel, S. (2003). Ownership Structure, Founder Leadership, and Performance in Norwegian SMEs: Implications for Financing Entrepreneurial Opportunities, *Journal of Business Venturing*, 18, 5, pp 619-637.

Roos, G. and Roos, J. (1997). Measuring your Company's Intellectual Performance, *Long Range Planning*, 30, 3, pp 413-426.

Ryan Jr., H. E.., and Wiggins, R. A. (2004). Who is in Whose Pocket? Director Compensation, Board Independence, and Barriers to Effective Monitoring, *Journal of Financial Economics*, pp 497–524.

Sarbanes-Oxley Act, The (2002).

Short, H. (1996). Non-Executive Directors, Corporate Governance and the Cadbury Report: A Review of the Issues and Evidence, *Research Papers*, 4, 2.

Singh, M. and Davidson, W.N. (2003). Agency Costs, Ownership Structure and Corporate Governance Mechanisms, *Journal of Banking & Finance*, 27, 5, pp 793-816.

Srinivasan, N. (2003). Enhancing the Contributions of Independent Directors in *Independent Directors: Perceptions, Roles and Responsibilities*, Malaysian Institute of Corporate Governance, Percetakan Sempana Score Sdn Bhd. Kuala Lumpur.

Vafeas, N. Theodorou, E. (1998). The Relationship between Board Structure and Performance in the UK, *British Accounting Review*, 30, pp 383-407.

Wahab, E.A.A., How, J.C.Y. and Verhoeven, P. (2004). Institutional Investors and Firm Performance in Malaysia, Working Paper, Curtin Business School, Curtin University of Technology, Bentley, Australia.

Weisbach, M.S. (1988). Outside Directors and CEO Turnover, *Journal of Financial Economics*, 20, pp. 431-460.

Williams, S.M. (2000). Relationship Between Board Structure and a Firm's Intellectual Capital Performance in an Emerging Economy, Working Paper, University of Calgary, Alberta, Canada.

Yermack, D. (1996). Higher Market Valuation of Companies with a Small Board of Directors, *Journal of Financial Economics*, 40, pp 185-211.

Table 1. Definition and Calculation of the Three Indicators of the VAIC Methodology

Indicator	Definition	Formula
CEE (Capital Employed Efficiency)	Represents the efficiency of VA of physical resources	$\frac{VA}{CE}$ Where: CE = Capital employed = Book value of net assets
HCE (Human Capital Efficiency)	Represents the efficiency of VA of intellectual (human) resources	$\frac{VA}{HC}$ Where: HC = Total salary and wages of the firm. This represents the salary and benefit expenses to the employees and officers of the firm.
SCE (Structural Capital Efficiency)	Represents the efficiency of VA of intellectual (structural) resources	$\frac{SC}{VA}$ Where: SC = VA – HC

Table 2. Variable Name, Description, Type and Measurement for Independent Variables

Variable Name	Description	Type	Measurement
OutDir	Percentage of independent non-executive directors on the board	Ratio	$\frac{\text{No. of outside directors X 100\%}}{\text{Total No. of Director}}$
InsOwn	Percentage of outstanding shares owned by directors	Ratio	$\frac{\text{No. of Ordinary Shares Owned X 100\%}}{\text{Total No. of Ordinary Shares}}$
Duality	CEO Duality	Nominal	(1=Yes, 0=No)
BrdSize	Board size	Ratio	Number of directors on the board

Table 3. Control Factors, Type and Their Proxy Measures

Control Factor	Type	Measurement
Profitability (ROA)	Ratio	Profitability is measured by using the return on assets ratio, which is the net income plus after-tax interest expense, and divided by total assets: $\frac{\text{Net income + Interest expense } (1 - \text{Tax rate*})}{\text{Total assets}}$ * The company tax rate in 2003 was 28%
Leverage (Leverage)	Ratio	Leverage is measured by dividing the total debt with the total shareholders' equity as reported in 2003 annual report: $\frac{\text{Total Debt}}{\text{Total Shareholders Equity}}$
Dividend Yield (DivYield)	Ratio	Measured by the cash dividends paid as a percentage of total shareholder equity: $\frac{\text{Cash Dividends Paid}}{\text{Total Shareholder Equity}}$ x 100%
R&D Sensitivity (R&D)	Nominal	Firm is defined as R&D sensitive if a separate disclosure on the amount of R&D expense is made in its annual report: (1=R&D sensitive, 0=Otherwise)
Firm Size (FrmSize)	Ratio	Natural log of annual sales, i.e. ln (annual sales). Similar to most board structure-firm performance literature (for example, Brickley et al., 1997, Vafeas and Theodorou, 1998), the annual sales figures is transformed using the logarithmic transformation to address non-normality concerns with residuals.

Table 4. Samples by Efficiency Indicators, Percentage of Independent non-executive directors, Directors' Share Ownership and Total Number of Directors

	N	Mean	Median	Std. Deviation	Minimum	Maximum
Efficiency Indicators:						
VA efficiency of total resources (VAIC)	75	5.019576	5.178239	8.7481525	-53.5755	19.9552
VA efficiency of intellectual resources (HCE + SCE)	75	4.828975	4.912069	7.9582582	-46.3405	19.2838
VA efficiency of physical resources (CEE)	75	0.190601	0.311429	0.9086316	-7.235	0.8939
Independent Variables:						
Percentage of independent non-executive directors	75	38.5853	37.5	11.24894	10	87.5
Director's ownership	75	11.6925	4.03	16.98664	0	65.21
Total number of directors	75	7.84	8	1.661	5	13

Table 5. Samples by CEO Duality

		Frequency	**Percent**	**Valid Percent**
Valid	Separate roles between chairman and CEO	53	70.7	70.7
	Duality exists	22	29.3	29.3
	Total	**75**	**100.0**	**100.0**

Table 6. Pearson Correlation Analysis Results (n = 75)

	Percentage of independent non-executive directors	**Directors' ownership**	**CEO duality**	**Board size**	**Profitability**	**Leverage**	**Dividend yield**	**R&D sensitivity**	**Firm size**
Percentage of independent non-executive directors	1								
Directors' ownership	-.148	1							
CEO duality	.055	.095	1						
Board size	-.209	-.112	-.133	1					
Profitability	**-.147**	**.201**	**.018**	**.070**	1				
Leverage	**.134**	**.069**	**.111**	**.073**	.503(*)	1			
Dividend Yield	**.030**	**-.236(**)**	**.054**	**-.016**	.341(*)	.042	1		
R&D Sensitivity	**.088**	**-.169**	**.157**	**.112**	.108	.019	.463(*)	1	
Firm size	**.170**	**-.107**	**.067**	**.077**	.176	.003	.232(**)	.279(**)	1

* Correlation is significant at the 0.01 level (2-tailed).

** Correlation is significant at the 0.05 level (2-tailed).

Table 7. Regression of Independent Variables and Control Factors on VA Efficiency of Firms' Total Resources (VAIC)

Model Summary				
R	**R Square**	**Adjusted R Square**	**Std. Error of the Estimate**	**Durbin-Watson**
.908	.825	.801	3.9015317	1.834

ANOVA					
	Sum of Squares	**df**	**Mean Square**	**F**	**Sig.**
Regression	4673.806	9	519.312	34.116	.000**
Residual	989.427	65	15.222		
Total	5663.233	74			

Coefficients					
	Unstandardised Coefficients		**Standardised Coefficients**	**t**	**Sig.**
	B	**Std. Error**	**Beta**		
(Constant)	-30.720	7.738		-3.970	.000**
Percentage of independent non-executive directors	.131	.045	.169	2.920	.005**
Director's ownership	-.047	.030	-.091	-1.570	.121
CEO = Chairman	.161	1.033	.008	.156	.877
Total number of directors	.516	.293	.098	1.763	.083*
Profitability	.403	.056	.518	7.178	.000**
Leverage	1.027	.144	.461	7.128	.000**
Dividend Yield	.034	.168	.013	.201	.842
R&D Sensitivity	-5.138	1.727	-.183	-2.974	.004**
Firm size (Sales)	1.390	.401	.196	3.464	.001**

* Significant at 10% level

** Significant at 1% level

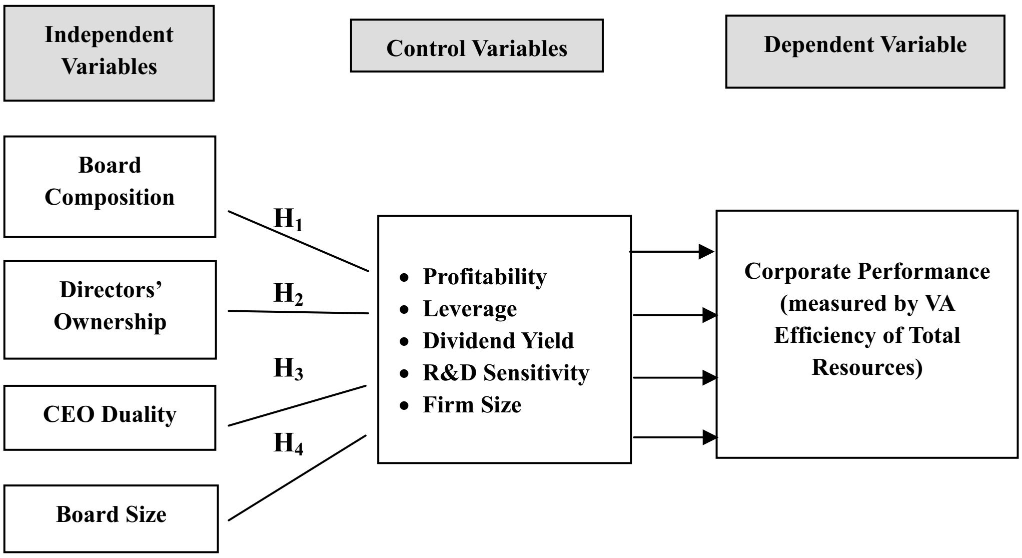
Independent Variables
Control Variables
Dependent Variable
Board Composition
Directors' Ownership
CEO Duality
Board Size
H_1
H_2
H_3
H_4
Profitability
Leverage
Dividend Yield
R&D Sensitivity
Firm Size
Corporate Performance (measured by VA Efficiency of Total Resources)

Figure 1. The Framework for the Relationship between Board Features and the Efficiency of VA by Firm's Total Physical and Intellectual Resources

International Journal of Economics and Finance
Vol. 1, No. 1
February 2009
www.ccsenet.org/journal.html

Audit Independence and Expectation Gap: Empirical Evidences from Iran

Mahdi Salehi
Accounting and Management Department
Zanjan University
D.N 1 Nagilo Alley, Hidaj City, Zanjan Province, Zanjan, Iran
Tel. 00989121425323 E-mail:mahdi_salehi54@yahoo.com

Ali Mansoury
Accounting and Management Department
Zanjan University

Zhila Azary
Research Staff, Ministry of Agriculture
Zanjan Branch, Iran
E-mail: Zhl_azary1@yahoo.com

Abstract

This paper discusses about auditor independence and audit expectation gap. An independent auditor is essential because of the separation of ownership from the management; the independent factor is the foundation of the public accounting profession and upon its maintenance depend the profession's strength and its stature. Independence is fundamental to the reliability of auditors' reports. The causes for reducing independence are economic dependence of the auditor on the client; audit market competition; the provision of non-audit services (NAS); the regulatory framework. Due to these causes auditors are not able to produce a fair report. Therefore "auditor independence is a key element of the audit expectation gap". If auditors are independent it reduces the expectation gap.

Keywords: Audit, Auditor independence, Iran, Audit expectation gap

1. Introduction

The concept of auditor independence changed during the late 19th and early 20th centuries. Because there was a large shift from capital coming from some sources to capital deriving primarily from domestic sources, noted economic sectors means articulated this change that large corporations were based on the separation of ownership from management, which they also emphasized the growing importance and role of accounting and auditing. The auditors' primary duty was to serve the needs of proprietary interest, which comprises of shareholders and general public. The creation of SEC laid emphasis on standards for financial reporting and auditing. The concept of auditor independence shifted in favor of objectivity and neutrality in reporting.

The audit of accounts in the corporate sector by an independent auditor is obligatory by statute, which defines his duties, rights and powers. It is essential because of the separation of ownership from the management in the corporate sector as the former needs someone who can keep a professional watch on the latter and to whom they can trust for the reliability of accounts as the preparation of financial statement is the prerogative of the management. The auditor has not much to suggest on the form and adequacy of financial statement and independent auditor is responsible for his report. Independence is fundamental to the reliability of auditors' reports. Those reports would not be credible, and investors and creditors would have little confidence in them, if auditors were not independent in both fact and appearance, which it expressed by Council of the American Institute of Certified Public Accountants (AICPA) in a statement adopted in 1947. Auditor independence is considered the hallmark of auditing profession. Independence is viewed as the most essential factor in business sector in protecting the interest of several parties.

Independence, both historically and philosophically, is the foundation of the public accounting profession and upon its maintenance depend the profession's strength and its stature (Carey, 1970).

2. Independence

Gill and Cosseral (1996) have emphasized that independence is the cornerstone of the auditing profession. Without independence the auditor's opinion is suspect. In addition the third parties believe that without independence there is no need for external auditors. Third parties' acceptance implies that the role of the external auditors is that of an independent instrument of social control within the corporate accountability process. To maintain independence under the many pressures of practice, an auditor must be alert to any deleterious influences on his planning, investigative, or reporting independence.

The public perception of auditor independence is important because the apparent independence depends on society's perception of what could impair actual independence. The external auditors must strive to ensure that audit quality will not be compromised and that auditor performance.

2.1 Definitions of auditor independence

The concept of independence has proved difficult to define precisely (Schutze, 1994), however representative definitions are:

'The conditional probability of reporting a discovered breach' (DeAngelo, 1981);

'The ability to resist client pressure' (Knapp 1985);

'A function of character, with the attributes of integrity and trustworthiness being key' (Magill and Previts 1991);

'An absence of interest that create an unacceptable risk of bias' (AICPA 1985);

'Independence is the main means by which the auditor demonstrates that he can perform his task in an objective manner" (FEE 1995) and

'Freedom from those pressures and other factors that compromise, or can reasonably be expected to compromise, an auditor ability to make unbiased audit decisions' (ISB, 2000).

There respective definitions all reflect the importance of objectivity and integrity as the two key aspects of auditor independence (Dunmore and Falk, 2001).

2.2 Factors affecting auditor independence

The auditor's independence may be influenced by conscious inaccuracy or by unconscious in inaccuracy in the reported information. The conscious inaccuracy may arise out of several factors like,

a) Close to his client;

b) Dependency on the client for his livelihood;

c) Driven by a desire of economic and social success;

d) Close relationship with the client's executive;

e) Blood relationship or marriage relationship with his clients;

f) Acceptances of goods or services from clients directly or through his employees at confessional basis or free basis;

g) Beholden to the Board of Directors for his re-appointment; and

h) Competitive in audit market (Sucher and Maclullich, 2004).

The unconscious inaccuracy may also arise from several factors as below:

a) The auditor may rely on branch manager;

b) Rely on external confirmation while making his opinion on accounts such as confirmation from debtors, creditors, bankers etc; and

c) Rely on the management for verification and valuation of assets to a greater extent.

Normally auditors bother for their re-appointment as shareholders in the annual general meeting appoint them but in practice they are dependent on the board for their reappointment because of controlling power of the board arising out of its domination of the management by a few family blocks.

Sometimes the possibility of varying interpretations of financial accounting practices may adversely affect Auditor independence

Battie et al. (1999) identified four themes as impediments to independence. They are:

1. Economic dependence of the auditor on the client;
2. Audit market competition;
3. The provision of non-audit services (NAS); and

4. The regulatory framework.

3. Audit expectation gap

Audit is essentially entrusted with the task of reporting reality in financial statements and this reality is what the users of accounting information expect. However, the auditors may not check out the reality and this reality may fall short of user expectations. This shortfall in audit effectiveness is broadly labeled as audit expectation gap. Most of the time financial statement users consider an auditor's report to be a clean bill of health. Thus most users' expectation toward auditors is far more than what they should be. The expectation gap occurs when there are differences between what the public expects from the auditor and what the auditor actually provides. To guarantee on efficient control to the shareholders and to the general public, the auditors have to meet stringent requirements both with regard to their professional knowledge and their independence.

One of the first discussions of an expectation gap in financial statements was by Liggio (1974). He was concerned that since the late 1960 the accidentally profession had been under attack regarding the quality of its professional performance. He suggested two reasons for this: 'a greater willingness to hold others-especially professionals-accountable for perceived misconduct and the expectations gap as 'a factor of the levels of expected performance as envisioned by both the independent accountant and by the user of financial statements. The difference between these levels of expected performance is the expectation gap'.

This definition was extended a little in the Cohen Commission's (1978) terms of reference. The commission was charged, inter alia, to consider whether a gap might exist between what the public expected or needed and what auditor could and should reasonably should. They did not allow for substandard performance. It is submitted that gap which gives rise to criticism of auditors is that between what society expects from auditors and what is perceives its receives from them. It is therefore proposed that the gap more appropriately entitled "the audit expectation performance gap" may be defined as the gap between society's expectation of auditors and auditors' performance as perceived by society. Given this definition, the analysis indicates that the gap has two major components:

1. A Gap between what society expects auditors to achieve and what they can reasonably be expected to accomplish (designated the reasonableness gap).

2. A gap between what society can reasonably expect auditors to accomplish and what they are perceived to achieve (designated the performance gap).

This may be subdivided into 2:1 a gap between the auditor that can reasonably be expected of auditors and auditors existing duties as defined by the law and professional promulgations (deficient standards) and 2.2 a gap between the expected standard of performance of auditors existing duties and auditors perceived performance, as expected and perceived by society (deficient performance).

Sikka et. al., (1992) identified the independence is the main cause to reduce the expectation gap. He stated "as a first step towards reducing the expectation gap, auditing standards and hence audit objectives should be shaped by open, democratic, accountable bodies, independent of the accountancy profession and the DTI". Continuing arguments on reducing audit expectation gap, he recommends one of the steps to be taken by the profession to reduce the expectation gap is reform relating to auditor independence. Directly related to the independence issue is auditor appointment and the role of directors and senior management in that appointment.

Different underlying explanations have been offered for the continuing presence of the expectations problem. Tricker (1982) viewed the expectation gap as the result of a natural time lag in the auditing profession identifying and responding to continually evolving and expanding public expectations. Other authors argued that it was the consequence of the contradictions in a self-regulated audit system operating with minimal government intervention as evidenced in Hopwood (1990); Humphrey (1991).

Much has been written about the possibility of an audit expectations gap. Concerns over ambiguities in the role and responsibilities of auditors have led to the establishment of several government and professional investigations, which form an important part of the expectation gap literature. These include the Cohen Commission (1978); Metcalf Committee (1976); and Treadway Commission (1987); in the United States, the Cross Committee (1977); and Greenside Committee (1978); in the United Kingdom and the Adams Committee (1977) and MacDonald Commission (1988) in Canada. A common finding in all these investigations is that a gap between audit performance and expectations exists.

In the face of a growing expectation gap between auditors and the public, the accounting profession has attempted to narrow, if not eliminate, the gap and counteracts the negative consequences. Efforts to this end have included changes in accounting/auditing standards, revising the audit report and conducting official investigation as evidenced in the studies of Guy and Sullivan (1988); Hille (1993); Pound and Fensome (1993); Chenok (1994); ICAA (1994); and Koh (2000) The most relevant studies are delineated here. However, these evidences prompt a serious concern to the auditing profession as the proven existence of the expectation gap lead to severely tarnished reputation and negative perception towards the value of independent audit (Fadzly and Ahmad, 2003).

4. Empirical evidence on independence and audit expectation gap

The expectation gap stems from the differing expectations of the function of independent audit between the auditors and the public. Sweeney (1997) listed out the areas where difference in expectations mostly arises: going concern, fraud and illegal acts, independence, and duty of care.

The empirical study by Humphrey (1991) with reference to England is an extension of first studies and the main conclusions include expectations gap in auditors' role in fraud detection; the extent of auditors' responsibility to this party; the nature of balance sheet valuation; threats to auditors' independence; and auditors' ability to cope up with risk and uncertainty. He also adds "If any topic can be classified as going to the heart of the audit expectations debate, it is the issue of auditor independence." Humphrey et al. (1992) stated that the "auditor independence as a key element of the audit expectation gap".

Garcia-Benau and Humphrey (1992) investigated the expectation gap in both the UK and Spain. Auditors, finance directors and users were asked, inter alia, whether audit firms should not provide NAS to their audit clients. In both countries, the average response was close to neutral for all groups except UK auditors who expressed strong disagreement.

The Beattie et al., (1988; 1999) study examined a large set of 45 economic and regulatory factors that could impair or enhance auditor independence, using questionnaire surveys of UK audit partners, finance directors and financial journalists. At a level of 50 per cent of audit fee, the corresponding threat rankings were firth, tenth and seventeenth and at the level of 25 per cent of audit fee they dropped to fourteenth, eighteenth and twentieth ranks. NAS fees also increased economic dependence generally, and all groups ranked various measures of economic dependence among the top threat factors.

a survey taken up by Lin (2004) in China with respect to audit objectives, auditor's obligation to detect and reporting frauds and third party liability of auditors. The study evidenced the emergence of the expectation gap in China and the majority of audit independence by reducing governmental control or intervention and moving towards self-regulation of the profession. This study has a limitation in the sense that it should cast light on understanding of the institutional setting and updated development of independent audits in China and may also serve as an annotation to the recent accounting reform debates in the western world. This study investigated the views of natural shareholders regarding the role of the effects on independence due to the audit firm also providing non-audit services to their audit client. A total of 615 questionnaires were received with an overall response rate of 37.50 per cent. Shareholders were asked whether they agreed that the independent audit enabled them to rely on financial information of profits, dividend yield/payout ratios and assets/liabilities. The responses of shareholders generally confirmed that the independent audit was important in their use of financial accounting information. Similarly, the reliability factors for the audit report of the independence of auditors and audit firm reputation were tested. Both were believed to add credibility to the auditors' report. Further discussion involved shareholder perception of audit independence in three separate instances: audit firm receiving substantial consultancy fees; the auditor holding shares in the audit client and the same auditor had been retained for over seven years. Shareholder opinions revealed that auditor independence was perceived to have been impaired by the substantial consultancy fees paid, but to a lesser extent by share ownership. Long-term audit contracts were not perceived to impair auditor independence. This study also refutes the idea that auditors could maintain their independence when receiving substantial consultancy fees. Overall, the findings of this research suggest that natural shareholders place a strong reliance on regulatory matters such as the accounting standards and the corporation's law for accounting information.

Alleyne et al., (2006) investigated the appearance standard by empirically exploring both auditors' and users' perceptions of auditor independence in Barbados. Firstly, the study contributed to the existing body of knowledge in terms of providing a better understanding of the nature of auditor independence in small developing countries. Secondly, this study could inform policy makers, governments and professional accounting bodies as to how auditor independence policies and frameworks could be structured to ensure adequate regulation of the capital market. Thirdly, their study would serve to educate users and auditors about the contextual factors surrounding the role of auditor as well as the possible threats and enhancements factors affecting auditor independence. The survey instrument was divided into two sections: section one dealt with demographic factors and section two focused on 39 audit-related issues categorized under a number of generic factors. The sampling respondents identified comprised several groups such as auditors, financial directors, credit managers, investment analysts, fund managers, shareholders and government departments. The sample respondents comprised 66 auditors and 148 users. The findings of the study revealed that economic dependence of auditor on the client, the provision of non audit services, high competition, small firm size, being a sole practitioner, lengthy tenure and the size and closeness of Barbadian society were found to negatively affect perceptions of auditor independence.

5. Research methodology

The main objectives of this research are measured whether there is expectation gap between auditors and investors regarding to status of auditor independence in Iran. So far accurate answer to this question, the authors design and developed a questionnaire based on method used in Best et al (2001) in Singapore, Nazri Fadzly and Ahmad (2004) in Malaysia and Chowdhury et al (2005) in Bangladesh.

The questionnaire contain two parts namely (A) bio-data and some questions about importance of audit, importance of financial reports, importance of company profile and importance of Directors profile, which totally this section has 10 tables and (B) this section includes 9 questions regarding to the level of audit independence in Iran, which this section of questionnaire contain two parts; first contain shows on Table No.11, 9 questions which show actual level of auditor independence in Iran and second part shows on Table No.12 contains the same questions which assert investors expectation regarding to auditor independence in Iran. The measure instrument of the statements is a Five-Point Likert type scale anchored “strongly disagree” (1) and “strongly agree” (5). To test for the significant expectation gap between two parties of respondent, Mann-Whitney U test is used which is non-parametric equivalent to the t-test, considered appropriate for ordinal measurements, such as the Likert scales employed in this study. It tests whether two independent samples are from the same population.

6. Results of study

Parts (A) contain Table No. 1 for descriptive analysis which shows below:

The table shows that related information collected from 214 Iranian investors and 227 Chartered Accountants. Total 441 useable questionnaires collected from both sides.

Part (B) contains 2 tables which show status of real and expectation levels of auditor independence and perceptional level of independence in Iran.

As Table No. 1 shows, investors got more than 91% bachelor degree and less than 8% master degree, whereas more than 78% of Chartered Accountant have bachelor, more than 20% of them have master and only less than 2% have PhD in accounting qualification. As shows in the table less than 44% of investors have not any auditing information as well as knowledge; approximately 56% of them have same knowledge.

Furthermore, the table revealed status of invested in corporate which in addition of investors more than 23% of Chartered are invested in corporate sectors. However it is incredible that Chartered Accountants this much investing in corporate sectors.

Insert Table 1

Above table shows amount of reading habits in both participants, which according to this table approximately 72% of investors reading financial statements occasionally and 28% of them looking to financial statements regularly. Whereas more than 48% of auditors reading financial statements occasionally and 52% of them reading financial statements regularly and also the table presents the distribution of respondents from the viewpoint of high importance of audit of audit report and low importance, which exact 50% investors believes that the auditor report is very important and 50% of them have idea that the audit report is not important. 26.5% of auditors have this perception that audit report is not important and 73.5% of them strongly believe that audit report is very important. In addition it is shows that 35% investors have this idea that financial statements are not very important, while 17.6 auditors have same idea. 65% investors and 82.4% auditors believes that the financial statements are very important.

Analysis of statement No.1, Table No.2 about actual level of audit independence revealed that there is significant difference between investors and auditor regarding to actual level of economic dependence of the auditor on the client which 39% investors agree with this statement, while only 10% of auditors have a same idea; in this position, may be auditors want to reluctant to reveal real position of this statement.

Analysis of statement No.2 show that there is a significant gap between two parties which 36%investors believe that audit market competition impair independence, whereas only 10% of auditors agree with this idea in other words Iranian investors strongly agree that audit market competitions impair independence of external auditors.

In relation to statement No.3 results show that there is no significant gap between two parties, in other words both investors and auditors strongly believe that receiving payment for non-audit services impair independence, it my because of practicing non audit services by auditors, while there is no complete regulation with respect to that.

Insert Table 2

Analysis of statement No.4 on Table No.2 shows that there is significant gap between auditors and investors; namely 81% auditors emphasize the regulatory framework will be effected on independence of auditors but 26% investors have same perception. It could be arise from investors’ unawareness regarding auditor regulatory framework.

With regard to statement No.5 result shows that there is a significant difference between auditors and investors in relation to the corporate governance system. In other words, 79% auditors strongly agree that corporate governance system influence on independence of auditors, while 19% investors have the same believe.

Analysis statement No. 6 shows that significant gap between two parties, in other words, 8% auditors disagree that the greater use of audit committees influence to independence while, 29% investors strongly agree that the greater use of audit commits influence to auditor independence.

With regard to statement No.7 result shows that there is wide gap between two parties with regard to professional ethics guidance; 20% investors agree with this statement, while 76% auditors strongly believes that the professional ethics guidance has a direct affect to auditor independence.

Analysis statement No. 8 shows that also there is significant difference between auditors and investors in relation to receiving gifts and presentations from management; 87% auditors strongly disagree with this statement, while 37% investors strongly agree that receiving gifts and presentations from management deeply influence on auditor independence.

Analysis the last statement shows there is gap between to parties. 77%auditor strongly disagree with that the Prospects of reappointment impair audit independence, whereas 38% investors strongly agree with statement.

Table 3 shows expected level of audit independence in Iran, which the same above questions are repeated here for measuring expected level of independence from the viewpoint of auditors and investors.

According to Table No. 3 about the statement No. 1 there is a significant expectation gap between auditors and investors; 42% investors have same perception that the economic dependence of the auditor on the client, while 4% auditors agree with this statement.

Result of analysis statement No. 2 shows that there is a significant gap between two parties with regard to that the audit competition may affect to independence. 81% auditors have same perception that audit competition may independence, while, 30 investors disagree with this perception.

Analysis of the statement No.3, Table 3 shows that there is also expectation gap between two parties; 78% auditors expect that receiving payment for non-audit services will be affect on independence, while 47% investors thinking that receiving payment for non-audit service will not be affected on audit independence.

Result of statement No.4 shows that there is a significant difference between two participants. 91% auditors expecting that the regulatory framework will impair independence, whereas the investors not expecting that the regulatory framework impairs independence.

Insert Table 3

With regard to Table No. 3 about statement No.5 result shows there is expectation gap between auditors and investors in respect to the corporate governance system and independence.

87% auditors expect that corporate governance system will be affect on audit independence, while 18% investors not expecting in this manner.

Result of statement No.6 shows there is no expectation gap between two parties, in other words84% auditors and 67% investors strongly expecting that the greater use of audit committees have positive affect on independence, in this condition audit committees plying very important role to audit independence.

Result of statement No.7 shows that there is a significant gap between two parties; 84% auditors have expectation that Professional ethics guidance has positive affect on audit independence, but 13% investors expecting that the professional ethics guidance has not positive affect on independence.

With regard to result of statement No.8 found that there is no significant expectation gap between two parties with regard to receiving gifts and presentations from management; both parties strongly expecting that receiving any gifts and presentations from management impair audit independence.

Result of statement No. 9 at Table No. 3 shows that there is a expectation gap between two parties; 41% investor expecting that Prospects of reappointment for next yeas strongly impair audit independence, whereas only 8% auditors have the same expectation.

7. Conclusion

In conclusion, the maintenance of auditor independence is very important for the users of the financial statements. The more independent an auditor seems to the greater will be the confidence in his work and opinion. Independent audit is an important factor to reduce audit expectation gap, because the investor and others are expecting more from auditor. In conclusion, the maintenance of auditor independence is very important for the users of the financial statements. The more independent an auditor seems to the greater will be the confidence in his work and opinion. Independent audit is an important factor to reduce audit expectation gap, because the investor and others are expecting more from auditor.

This study found significant expectation gaps between auditors and investors on actual level of audit independence in Iran. As mentioned earlier, according to TableNo.2 except statement No.3 there are significant differences on other statements. It seems that the auditors believe the external factors (i.e. the regularity framework, and corporate governance) are more effective on depend or independent of auditors rather than eternal factors (i.e. prospects of reappointment), while investors believe the internal factors impair audit independence rather than external factors.

With regard to Table No.3 difference perceptions were found between two parties, except statements No.6 and 8 in other statements there were difference perceptions on audit independence.

For reducing such gaps at first, Iranian accounting and auditing regulators should develop scope of audit independence and show clear picture of that, second investors should aware about responsibilities as well as limitations of audit independence. In this position better communication between two parties may reduce such gaps.

References

Adams Committee. (1977). Role of the Auditor, Toronto: Canadian Institute of Chartered Accountants.

AICPA. (1978). Report of the Commission on Auditors' Responsibilities, New York: AICP.

AICPA. (1985). "A Professional Conduct", New York.

Beattie, v., Brandt, R. & Fearnley, S. (1999). "'Perceptions of Auditor Independence: UK Evidence", *Journal of International Accounting, Auditing and Taxation,* Vol.8, No.1, pp: 67-107.

Best P J, Buckby S & Tan C. (2001). "Evidence of The Audit expectation Gap in Singapore" *Managerial Auditing Journal,* Vol .16, No .3 pp: 134-144.

Carey,L. (1970) .The Rise of the Accounting Profession: To Responsibility and Authority, 1937-1969 (New York,), p. 182.

Chenok, P.B. (1994). Perception vs. Reality, *Journal of Accountancy.* pp: 47-51.

Chowdhury, R, Innes,J and Kouhy, R. (2005). "The Public Sector Audit Expectation Gap in Bangladesh", *Managerial Auditing Journal*,Vol.20, No.8,pp: 893-908.

Chowdhury,R ,Jnnes , J & Kouhy , R. (2005). "The Public Sector Audit Expectation Gap in Bangladesh", *Managerial Auditing Journal.*Vol.20,No.8,pp:893-908.

Cohen Commission. (1978). Report of the Commission on Auditor Reasonability's: Conclusions and Recommendations. New York: American Institute of Certified Public Accounts.

Cross Committee. (1977). Report of the Committee Under the Chairmanship R1. Hon the Lord Cross of Chelsea, Accountancy, 160, pp: 80-86.

DeAngelo, L.E. (1981). "Auditor Size and Audit Quality", *Journal of Accounting and Economics,* Vol.3, August, pp: 183-199.

Dunmore P.V Z & Falk, H. (2001), Joint Provision of Audit and Non-audit Services, Audit Pricing and Auditor Independence, WP.

Fadzly ,M & Ahmad,Z(2003), "The Perceived Value of Financial Statement Audit" , Proceeding of The International Conference on Quality Financial Reporting Corporate Governance, kuala Lumpur,Malaysia.

FEE :(1995) . Auditor Independence and Objectivity, Position Paper. Bruseels.

Flint, D. (1988). Philosophy and Principles of Auditing, An Introduction. Macmillan. London

Fried, D, & Schiff, A. (1981). "CPA Switches and Associated a Market Reactions", *The Accounting Review,* April, pp. 326-341.

Gill G, and Cosserat G.. (1996). Modern Auditing in Australia, Fourth Edition Wiley and Sons Australia.

Greenside Committee, (1978). "Report of the Joint Committee Appointed to Consider the Cross report and related Matters", *journal of Accountancy.* Vol. 161, pp; 124-132

Guy. D. & Suilivan, J. (1988). "The Expectations Gap in Auditing Standards", *Journal of Accountancy.*Vol.165, pp: 36-46.

Hille. (1993). "Expanding the New Zealand Audit report", *Accountants Journal,* pp: 59-61.

Hopwood, A.G.. (1990). "Ambiguity, Knowledge and Territorial Claims: Some Observations on the Doctrine of Substance over Form-S Review essay", *British Accounting Review,* Vol. 22, No.1: pp: 79-87.

Humphrey, C.G.. (1991). *Audit Expectations, in Current issues in Auditing,* London: Plul Chapma Publishing.

Humphrey, C.G., Moizer, P. &Turley, W.S. (1992). "The Audit Expectation Gap-Pusca change. Plus D'est la Meme Chose", *Critical Perspectives on Accounting,* 3: 137-161.

Independence Standard Board (ISB). (2000). Statement of Independence Concepts: A Conceptual Framework for Auditor Independence, Exposure Draft ED00-2.

Insitite of Chartered Accountants in Australia (ICAA). (1994). A research Study on Financial reporting and Auditing-Bridging the Expectation gap. Sydney.

Knapp, M. (1985). "Audit conflict: An Empirical Study of the Perceived Ability of Auditors to Resist Management Pressure", *The Accounting review,* Vol.60, No.2, April, pp: 202-211.

Koh., H,C (2000). "The Audit Expectation Gap in Singapore: An empirical Study of Company Audit Objectives", *Asian review of Accounting,* vol.8, Number 1. pp: 83-105.

Kosmala M, K. (2003). "The true and fair view construct in the context of polish transition economy: some local insights", *European Accounting Review,* vol. 12, No.3, pp.465-87.

Liggio, C.D. (1974). "The Expectation Gap: The Accountant's Legal Waterloo". *Journal of Contemporary Business,* Vol.3, No.3, pp: 27-44.

Lindsay, D. (1992). "Auditor client conflict resolution: an investigation of the perceptions of the financial community in Australia and Canada", *International Journal of Accounting,* Vol. 27.pp.342-365.

Macdonald Commission. (1988). Report on the Commission to Study the Public's expectation of Audits, Canada, Canadian Institute of Chartered Accountants.

Magill,H,T and Previts G,J. (1991). CPA Professional Responsibility : An Introduction, Southwestern Publishing Company , Cincinnati, OH.

Mautz, R.K. & Sharaf, H.A. (1961). The Philosophy of Auditing, Florida: American Accounting Association,Sarasota,F.L.

Mtcalf Committee. (1976). The Accounting Establishment, Washington D C: United States Government Printing Office.

Nazri Fadzly, M & Ahmad, Z. (2004). "Audit Expectation Gap- The Case of Malaysia", *Managerial Auditing Journal,* Vol.19, No.7, pp: 897-915.

Pound, G., and Fensome, M. (1993). "The Changing Status of AUPs", *Australian Accountant,* Pp: 35-36.

Schuetze W, P. (1994). A Mountain or molehill? Remarks by Walter P. Schuetze Chife Accountant, Securities and Exchange Commission of The United state of America to AICPA's twenty- First Annual National Conference On current SEC Developments, *Reproduced in Accounting Horizons,* Vol.8, No.1, pp: 69-75.

Shockley, R.A. (1981). "Perceptions of Auditors' Independence: An empirical Analysis", *The Accounting Review* January, pp: 785-800.

Sikka, P., Puxty, A., Wilmott, H. and Cooper, C. (1992). "The Eliminating the Expectation Gap", Canada: Chartered Association of Certified Accountants.

Sucher,P & maclullich, K. (2004). "A Construction of Auditor Independence in The Czech Republic: Local Insight", Accounting, *Auditing& Accountability Journa* . Vol.17, No.2, pp: 276-305.

Sweeney., B. (1997). "Bridging the Expectation –on Shaky Foundations", *Accountancy Ireland*, pp: 18-19.

Treadway Commission. (1987). Fraudulent Financial Reporting., New York: American Institute of Certified Public Accountants.

Tricker, R.L. (1982). Corporate Accountability and the Role of the Audit Function, London: Pitman Books.

Yost, J. A. (1995). "Auditor independence as a unique equilibrium response", *Journal of Accounting, Auditing and Finance,* Vol.10, No.1, pp: 81-103.

Table 1. Profile analysis

Statements	Investors	Auditors
Qualification:		
Graduate	195	178
Master	19	47
PhD	0	2
Audit knowledge:		
No	94	0
Yes	120	227
Invested in corporate:		
No	0	106
Yes	214	121
Reading financial statements:		
Occasionally	154	110
Regularly	60	117
Importance of audit report:		
Low	107	60
High	107	167
Importance of financial statements:		
Low	75	40
High	139	187

Table 2. Actual Level of Independence

Statements		Auditors			Investors			Z	P
		Mean	Median	Std	Mean	Median	Std		
1	Economic dependence of the auditor on the client.	2.17	2	0.85	2.82	3	1.48	-6.56	0.00
2	Audit market competition.	2.95	3	0.74	2.63	3	1.46	-2.29	0.02
3	Receiving payment for Non-audit services.	1.95	2	0.71	2.12	2	1.45	-1.55	0.12
4	The regulatory frame work.	3.70	4	0.83	2.64	3	1.25	-10.77	0.00
5	The corporate governance system.	4.29	5	1.28	2.26	2	1.34	-13.51	0.00
6	The greater use of audit committees.	2.85	3	0.87	2.62	3	1.39	-2.25	0.03
7	Professional ethics guidance.	4.34	5	1.04	2.93	3	0.73	-12.98	0.00
8	Receiving gifts and presentations from management	1.33	1	0.99	2.46	3	1.61	-7.88	0.00
9	Prospects of reappointment	2.35	2	0.98	2.74	3	1.50	-3.14	0.00

Note:* Significantly different from Auditors at $P \leq 0.01$

Table 3. Expected Level of Independence

Statements		Auditors			Investors			Z	P
		Mean	Median	Std	Mean	Median	Std		
1	Economic dependence of the auditor on the client	2.09	2	0.95	2.63	3	1.69	-2.98	0.00
2	Audit market competition	4.25	5	1.35	3.15	3	1.62	-8.35	0.00
3	Receiving payment for non-audit services	3.26	4	1.39	2.15	2	1.72	-6.92	0.00
4	The regulatory frame work	4.58	5	1.01	3.52	4	1.52	-9.55	0.00
5	The corporate governance system	4.47	5	1.14	3.32	4	1.58	-9.79	0.00
6	The greater use of audit committees	3.64	4	1.09	3.49	4	1.55	-0.54	0.59
7	Professional ethics guidance	4.52	5	0.90	3.53	3	0.93	-10.77	0.00
8	Receiving gifts and presentations from management	1.81	2	0.78	2.06	1	1.76	-0.60	0.55
9	Prospects of reappointment	2.06	2	0.94	2.60	3	1.70	-3.14	0.00

Note:* Significantly different from Auditors at P≤0.01

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Theory Survey of Stock Yield Prediction Models

Wei Shen
School of Business and Administration
North China Electric Power University
Beijing 102206, China
E-mail: shenwei_1965@126.com

Abstract

According to the difference of the modeling theories, the stock yield prediction models can be divided into two sorts. One is the traditional fluctuation rate prediction model based on statistical theory, and the other is the innovational prediction model based on theories such as NN, grey theory, support vector machines (SVM) and so on. In this article, we introduced these two models and their research actualities, and compared and analyzed the characters of these two sorts of models, and studied the problems existing in the concrete application of the stock yield prediction model in China, and put forward corresponding advices for the future development.

Keywords: Stock yield, GARCH model, SV model, NN, Grey model, SVM

1. Introduction

How to exactly describe and predict the stock yield fluctuation has been one of hotspot problems discussed by the financial domain all along. It possesses very important theoretical and practical meanings to grasp the characters and direction of stock yield fluctuation for investors to measure, avoid and manage the stock market risk. Therefore, for a long time, many scholars utilize various prediction models to empirically analyze and predict the stock yield fluctuation, and hope to find helpful revelations and useful rules. At present, from domestic and foreign relative literatures, though there are many models to predict the fluctuation of the stock yield, but according to the difference of the modeling theories, the stock yield prediction models can be divided into two sorts. One is the traditional fluctuation rate prediction model based on statistical theory, such as popular and representative models including ARCH type models and SV type models, and the other is the innovational prediction model based on theories such as NN, grey theory, support vector machines (SVM) and so on. These two sorts of model will present different characters when they are used to predict the stock yield fluctuation, but up to now, there are no relative literatures to systematically compare and analyze these two models and open out their own respectively applied conditions and situations. In this article, we compared and analyzed the characters of these two sorts of models, studied the problems existing in the concrete application of the stock yield prediction model in China, and tried to offer helpful references for the researches in this domain.

2. Stock yield prediction model

2.1 The stock yield prediction model based on statistical principle

The stock yield prediction models based on statistical principle which are applied most extensively and possess representative characters include GARCH type models and SV type models.

2.1.1 Engle's ARCH model

In 1982, Engle firstly put forward the autoregressive conditional heteroskedasticity model (ARCH model) which could better simulate the clustering and time varying properties existing in the stock fluctuation.

According to the definition put forward by Engle, the ARCH model can be expressed as

$$x_t = f(h,x) + e_t$$

$$e_t \mid y_{t-1} \sim N(0, s_t^2)$$

$$s_t^2 = a_0 + \sum_{i=1}^{p} a_i e_{t-i}^2 = a_0 + a(L)e_t^2$$

The time sequence is decided by two equations, and one is the conditional mean equation, and the other is the conditional variance equation, and h denotes the vector composed by some exogenous variables or lagged endogenous variables, x is the parameter vector of the mean equation, y_{t-1} is the information set before t term, and

$a(L)$ is the multinomial of lagged operator. In the conditional variance equation, the variance of $\{e_t\}$ presents self-correlation, i.e. the variance of error is the function of the lagged values, and e_t obeys the p-order ARCH process, which is noted as $\varepsilon_t \sim ARCH(p)$.

In 1986, Bollerslev put forward the generalized ARCH model (GARCH model). GARCH considers not only the lagged value of the disturbance item, but the lagged value of the disturbance item conditional variance. Therefore, GARCH model is a long-memory process, and it can depict the financial time sequence with long memory. In fact, GARCH process is the ARCH process with infinite orders, and we can utilize GARCH model to more conveniently describe the ARCH process with higher orders and compensate the deficiency of the ARCH model. After that, GARCH model is continually expended and perfected, and at present, the ARCH type model system including GARCH-M model, EGARCH model, A-PGARCH model, A-GARCH model, FIGARCH model and LM-GARCH model have been formed.

2.1.2 Talor's SV model

In 1986, Taylor put forward the stochastic volatility model (SV). The production of the SV model is directly correlated with the diffusion process of pricing in financial assets, and it thought not only the price of the objective assets but the fluctuation rate can be described by the Wiener Process, and SV model is the financial fluctuation model with very good application foreground.

The basic discrete SV model is

$$y_t = \varepsilon_t e^{\frac{h_t}{2}}$$

$$h_t = \alpha + \beta h_{t-1} + \eta_t$$

Where, $\{\varepsilon_t\}$ and $\{h_t\}$ are independent each other. And we generally suppose $\varepsilon_t : i.i.N(0,1), \eta_t : i.i.N(0,\sigma_\eta^2)$, and σ_η^2 is unknown. α and β are constants, and β is the durative parameter which reflects the influence of present fluctuation on future fluctuation, $|\beta| < 1$. h_t can also be extended to be a ARMA process.

Similar with the GARCH model, there are many extension forms for SV model. Kim, Shepard and Chib supposed that the ε_t didn't obey the normal distribution, but obey the t distribution, so the basic SV model can be extended as the SV-t model. Aiming at the relationship between the anticipated income and fluctuation in financial market, Kooperman and Uspensky put forward the SV-M model (SV in mean).

The parameter evaluation methods of SV model mainly include maximum likelihood method, Markov chains, Monte Carlo method and generalized matrix method, and comparing with GARCH type models, because SV model includes two stochastic processes, so SV model is theoretically better than the GARCH type model. But the parameter evaluation of SV model is relatively difficult, and in practical application, the SV model is not as extensively applied as GARCH type models.

2.2 The stock yield prediction model based on non-statistical principles

2.2.1 The grey prediction model (GM) based on grey theory

In 1982, Chinese scholar, Professor Deng Julong firstly put forward the theory of grey system which was largely concerned by the world. At present, the method of grey prediction has been extensively applied in the short term predictions of numerous industries.

The GM (1,1) model is a sort of usual grey model which is only composed by one-order differential equation containing single variable, and it is the special example of the GM (1,n) model. To establish the GM (1,1) model, we only need one material sequence $x^{(0)}$.

Suppose the original material sequence is

$x^{(0)}=[x^{(0)}(1), x^{(0)}(2), \ldots, x^{(0)}(n)]$

Use 1-AGO to produce one-order accumulated production sequence

$x^{(1)}=[x^{(1)}(1), x^{(1)}(2), \ldots, x^{(1)}(n)]$

Where, $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i)$.

Because the sequence $x^{(1)}(k)$ possesses the index increase rule, and the solution of the one-order differential equation is just the solution of the index increase form, so we can think the $x^{(1)}$ sequence fulfills the following one-order linear differential equation model.

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = u$$

Where, a is the grey endogenous control parameter of the development parameter u.

2.2.2 The NN prediction model based on bionics

The theory of ANN is a rising marginal and cross subject, and it can denote any nonlinear relationships and study, so it offers new ideas and new methods to solve many complex practical problems with uncertainty and time-varying. In the domain of prediction, Lapeds and Farber first applied NN into the prediction in 1987, which started the utilization of ANN prediction.

The learning ability of ANN can be utilized to train the nerve cell network by large numbers of samples and adjust the joint weights and thresholds, and then the assured model can be used to predict. NN can automatically study former experiences from the data samples without multifarious query and description process, and automatically approach those optimal functions which describe the rule of sample data without reference to how forms these functions have, and the function form is more complex, the property of NN is more obvious.

The basic idea of the back propagation training algorithm (BP algorithm) is to adjust and modify the joint weight and threshold to make the error minimum through the back propagation of network error, and its learning process includes forward computation and error back propagation. BP algorithm only needs one simple three-layer ANN model can realize any complex nonlinear mapping relationship from input to output.

2.2.3 The prediction model based on SVM

According to the statistical learning principle, Vapnik et al (1995, 1998) put forward the method of support vector machine learning (SVM). The method can realize the minimum experience risk and credit range through seeking minimum structured risk, and can obtain good statistical rule when the statistical samples are few. SVM algorithm is a convex optimization problem, so the local optimal solution must be the global optimal solution. SVM algorithm converts the impartible data in the original space into linear and partible high-dimensional character space, and then maps it into the original space to solve it by the core function, which can skillfully solve the problem of dimension disaster, and the complexity of the algorithm is independent of the input samples, but the selection of the core function is a difficult problem in the method of SVM.

3. Research actuality of stock yield prediction model and relative domains

3.1 Researches about the prediction model based on statistical principle

3.1.1 GARCH type models

Foreign researches have utilized the GARCH type models to implement large numbers of researches, which indicates GARCH model and its extension forms possess very good effect to describe the fluctuation of financial time sequence. French, Schwert and Stambaugh (1987) used the GARCH model to evaluate the relationship between US stock market anticipated income and fluctuation, and they found that the anticipated income was positively correlated with the predicted fluctuation. In the same year, Engle et al utilized GARCH-M model to find that the conditional variance could better explain the variance of S&P 500 index anticipated income. In 1991, Bollerslev and Engle's research also found that the risk price premium presented positive correlation with the fluctuation. The level effect of the stock price fluctuation was proved in Nelson's article (1991), Glosten, Jagannathan and Runkle's article (1993), Engle and Ng's article (1993), Fornari and Mele's article (1997) again and again. In addition, many empirical applications such as in Campbll and Hetschel (1991), Engle and Ng (1991), Pagan and Schwert (1990)'s researches also proved GARCH could offer ideal data simulation and prediction effect.

Many domestic scholars studied and applied the GARCH type models. Wei Weixian and Zhou Xiaoming (1999) utilized GARCH model to predict the fluctuation of Chinese stock market. Dinghua (1999), Zhou Zhefang and Li Zinai (2000), Tang, Qiming and Chenjian (2001) et al analyzed the ARCH effect of Chinese stock market and fit the fluctuation of the stock index. Yue Chaolong (2001), Chen Qianli (2002) and Zhou Shaofu (2002) utilized the GARCH type model to empirically study the fluctuation clustering and asymmetry of stock market income of China. Zhang Shiying and Keke (2002) systematically commented the ARCH model system. Li Shengli (2002) and Lou Yingjun (2003) discussed the level effect of Chinese stock market. Zhang Yongdong and Bi Qiuxiang (2003) empirically compared the fluctuation prediction model of Shanghai Stock Market.

3.1.2 SV type models

There are many researches about the SV model studied by foreign scholars. For example, Ghysels, Harvey and Renault (1995) detailedly summarized relative literatures about the origin, evaluation method, model expansion and fluctuation durative of the capital market stochastic fluctuation model. So, Lam, Li and Smith introduced the Markov structure conversion mechanism into the SV model, and obtained the MRS-SV model with Markov structure conversion. Kalimipalli and Susmel put forwad the different MRS-SV model to describe the level and fluctuation of short-term interest rate. Jun Yu (2002) utilized basic SV model to predict and analyze the stock market of New Zealand, and he found the basic SV model possessed good prediction ability. G.B. Durham (2007) utilized the SV-mix model to predict the S&P 500 index and the prediction effect was good.

In China, Baihuang and Zhang Shiying (2001) utilized extended SV model to empirically study the fluctuation of Shenzhen stock market, and they pointed out the extended SV model was better than SV model to describe the financial fluctuation. Yang Kelei and Mao Minglai et al (2004) used SV model to analyze the fluctuations of Shanghai index and Shenzhen index, and the result indicated that the risk of Shanghai stock market was larger than Shenzhen stock market. Wang Chunfeng (2005) compared the SV model which mean conditional distribution was the normal distribution and the SV model with conditional heavy-tailed distribution, and the empirical result indicated the SV model with conditional heavy-tailed distribution could better describe the fluctuation of Chinese stock market. Qian Haoyun (2006) adopted SV model and A-SV model to fit the daily yields of Shanghai composite index, Shenzhen composite index and Hongkong Hang Seng index, and the result showed A-SV was better than SV. Wu Qiquan et al (2006) used the SV model to study the policy effect of Chinese stock market. Mao Minglai et al (2006) comprehensively discussed the SV type models. Zhouyan and Zhang Shiying (2007) utilized the Monte Carlo simulation integral method based on Markov Chains to evaluate the SV model with continual time, and empirically studied the daily composite index of Shanghai stock market, and the result proved the validity of the model and the method.

In a word, comparing foreign and domestic scholars' research results, we can find that foreign scholars always obtain better effects because they adopt the data of US or other western countries to implement data fitting or prediction by GARCH type or SV type models, and domestic scholars usually adopt relative data of domestic stock market to implement same researches.

3.2 Researches about the prediction models based on non-statistical principles

3.2.1 Grey model

The grey prediction model put forward by Chinese scholar, Professor Deng Julong, has been applied in the short-term prediction of multiple domains such as food yield prediction and power load prediction, but in foreign literatures, there are few literatures which utilized the grey model to predict the stock fluctuation in short term. In China, many researchers have applied the grey model to the stock prediction.

Chen Haiming and Duan Jindong (2002) combined GM (1,1) model with Markov model, and established the Grey-Markov prediction model to predict Shanghai composite index, and obtained the precision of the Grey-Markov was higher than GM (1,1). Shi Jiuyu and Hu Jingpeng (2004) utilized GM (1,1) model to establish the culmination prediction model of 65 days mean running orbit for Shanghai composite index, and the prediction values were very consistent with the market values. Tan Siqian (2006) applied GM (1,1) model to implement short-term prediction for the stock price, and compared with the ARIMA model, and the result indicated the precision of the grey model was higher than the precision of ARIMA model.

3.2.2 NN prediction model

Matsuba (1991) firstly introduced the NN into the price prediction of stock market. After that, quite a few researchers utilized the NN model to predict the stock market. Hill et al (1996) compared NN with six sorts of traditional statistical prediction method, and they used 111 time sequences to predict, and the results indicated that when the short-term (monthly or season) data were adopted to predict, the NN would obviously better than traditional statistical models, but when the long-term (yearly) data were adopted, the prediction results were almost same.

There are many Chinese scholars to utilize NN to predict the stock market. Li, Minqiang and Meng Xiangze (1997) adopted the genetics algorithm based on NN to study the investment strategy of stock market. Combining with the characters of domestic stock market, Wu Wei et al (2001) utilized the better classification ability of the multiple-layer forward feedback BP network to predict the fluctuation of stock. Wu Chengdong and Wang Changtao (2002) utilized ANN NP network to predict the stock market. Liu Yongfu and Wu Haihua (2003) et al established the BPNN prediction model to predict the Shanghai composite index, and they found the convergence speed of the model was very quick, and the learning ability was strong, and the prediction precision was high, and the error rare was lower, and the method was very effective for the short-term prediction of stock index. Shang Junsong (2004), Long Jiancheng (2005) and Hujing (2007) and other scholars all utilized the NN model to study the prediction of stock market. Large numbers of empirical researches of Chinese scholars indicated it was very feasible and effective to apply ANN in the prediction of Chinese stock market.

3.2.3 SVM prediction model

H. Nakayama (2003) et al introduced the increasing learning and data abandoning method in SVM and applied these methods in the price prediction of stock. W. Huang (2005) et al used the SVM to predict the direction of the stock market. P. Pai (2005) et al combined ARIMA (auto regressive integrated moving average) model with SVM model, and put forward a sort of integrated model to predict the stock price, and the result showed the integrated model was better than the conclusion of single ARIMA or single SVM.

From relative domestic researches, Yang Yiwen and Yang Chaojun (2005) utilized SVM to make exactly multiple-step prediction for the sequence tendency of Shanghai composite index. Li, Lihui (2005) et al applied SVM in the prediction of Chinese Shanghai 180 index. Zhou Wanlong, Ma Fayao and Peng Lifang et al (2006) utilized SVM to implement short-term prediction to the stock, and Zhao Jinjing (2007)'s experimental result showed that the method had higher prediction precision than the NN and time sequence method.

As a whole, the quantity of the literature that Chinese scholars utilized innovational prediction models to predict the stock market is more than the quantity of the literature utilizing traditional prediction models. And as viewed from the prediction effect, the prediction precision of innovational prediction model would be higher than traditional statistical prediction mode. In three innovational models, the application of NN model is most extensive, and the SVM possesses the most exact prediction effect and the method which uses the grey model to implement short-term prediction is the simplest method.

4. Comparison and analysis between traditional prediction model and innovational prediction model

4.1 Different theoretical bases of modeling

The traditional stock yield fluctuation prediction model is based on the statistical analysis theory. But to dispose the stochastic process based on probability statistics requires more sample quantity and more complete and confirmatory original data. But in fact, in the practice, even if there are large numbers of sample quantity, we don't always find the rule, and even if the statistical rule can be found, but it is not always the representative one. The innovational prediction model completely breaks away the statistical theory, and it establishes the prediction model by a sort of innovational modeling thinking. For example, the grey model is based on the grey theory, and it accumulates and disposes the historical data and makes the data present the rule of index change according to the change rule of generalized energy, and then establishes the model. But ANN model is based on the ANN theory, and it establishes the NN model to predict through simulating the structure and the information processing and index function of the human brain neural system. SVM is based on the machine learning theory of statistical learning, and it makes local optimal solution to be the global optimal solution through convex optimization, which can overcome the deficiencies such as low NN convergence speed and local minimum point.

4.2 Different data requirements and processing

The prediction model based on the statistical principle requires large numbers of sample and good distribution rule, and whether for GARCH type models or SV type models, if only the sample quantity is enough large and the distribution is good, the prediction effect will be ideal. For example, the prediction effect will be better when we use GARCH model to predict the US stock index, because the development time of Chinese stock market is relatively short, and because of macro-control and stock reform, the stock index changed radically, which induced the distribution rule of data was not obvious, and there is still certain limitation to utilized this type model. Bu the innovational prediction model requires low sample quantity and distribution degree. For example, the grey model can predict the next data only by possessing 7 or 8 data. For the processing technology, the grey model will implement accumulated processing to the original data, and make the disorderly and unsystematic data present obvious index rule, and when the modeling and computation are completed, the model will decrease and deoxidize the data. The NN model adopts the data drive and black box to modeling, and it needs not experimental information, and it can pick up data character and effective predict the future through self structured adjustment under the data environment that the information resource is not complete and exact.

4.3 Different stabilities and adaptabilities of model structure

Once the prediction model based on the statistical principle, the model structure will possess strong stability, and a stable interior relationship will exist among model variables. Whether the GARCH model or the SV model, the model structures are relatively stable and simple, and they all belong to the model with single factor. But in the practice, the prediction environment is complex and changeful, and once new relationship occurs among system variables, the model will not adjust and adapt the change. The innovational prediction model is the model that one factor or multiple factors or the structure can be changed, and the computation is relatively complex, but its adaptability is better than the prediction model based on the statistical principle. For example, for the grey model, except for the basic GM (1,1) model, the grey theory can solve the high-order system through the GM (1,1) model group and can comprehensively consider the influences of multiple factors. But the NN and SVM belong to the variable structured model, and they adjust their interior structures to adopt the changes of the system variables through the learning of network to the new samples. For the nonlinear high-dimensional and high-order problems, the effect of NN and SVM will be better.

4.4 Different prediction precision and extrapolation

Comparably speaking, the prediction model based on the statistical principle has larger error and worse extrapolation, because the type of model lacks in the re-processing or re-learning process to the data samples, and the low fitting of sample induces the bad extrapolation. But the innovational prediction model possesses higher precision and strong extrapolation because the type of model possesses re-processing or re-learning process to the data. The grey model implements accumulation processing to the data, and the NN model and SVM first studies the data and then implements reasoning and optimization. So the fitting degree and the extrapolation ability of the innovational prediction model will be higher than the statistical models.

4.5 Different prediction difficulties and prediction time length

The technology of the prediction model based on the statistical principle is mature and the prediction process is relatively simple. Whether for the GARCH type model or for the SV type model, the theoretical base to establish the model is stable, and the model structure is relatively simple and the computation difficulty is relatively low. Because this type of model adopts the historical data in long time, so it can be used to predict the future in a long time. And the prediction technology of the innovational prediction model can be further improved and the prediction difficulty is large. For example, the process which utilizes the NN to predict the fluctuation of stock yield is relatively difficult because the NN needs set up the hidden layer and the weight which are directly correlated with the rationality and nicety of the prediction result. The prediction using the method of SVM comes down to the confirmation of the core function which is very difficult to be confirmed in fact. Because the innovational prediction model only requires small quantity of sample, so it is fit to implement short-term prediction for the prediction objects.

5. Problems existing in the application of the stock yield prediction model in China

5.1 The research results about the application of the stock yield prediction model in China are very limited, the research difficulty is large and the prediction precision should be further enhanced

From the index result of CJFD, in the periodical literatures from 1998 to 2007, there are four articles which utilized the GARCH type model to predict the stock market, and there are seven articles which used the SV type model to prediction the stock market, and the sum is 11. And there are five articles which used the grey model to predict the stock market, and there are 67 articles which used the NN to predict the stock market, and there are 9 articles which used the SVM to predict the stock market, and the sum is 81.

Add two above sums, the quantity of the relative articles in ten years is less than 100. The figure explains two problems at least, and one is the research in this domain in China has not be developed intensively, and relative research talents and results are very limited and the existing research is only in the stage of start. The other one is that the prediction of the stock yield fluctuation is very difficult, and the precision of the prediction hardly achieves researchers' anticipated level, so quite part of research results may not be vended because of failure prediction.

5.2 The prediction effect utilizing the innovational prediction models for Chinese stock market is better than the prediction model based on statistical principles

From the prediction results of the articles vended in China, the prediction effect of innovational model is better than the traditional statistical prediction model. The research mainly is that the prediction model based on statistical principle always very strictly requires the original data, and only under the premise that the original data distribution is good and the material is complete and the sample quantity is large, the fitting and the prediction effect will be ideal. But since the Shanghai Stock Exchange was established in 1990, there is only 18 years' history. In these 18 years, the Chinese stock market experienced drastic changes several times because of policy, supervision and stock reform, and the quantity of listed company is limited and changeful continually, so the original data are not ideal in the aspects such as distribution, sample quantity and data integrality. So in China, the prediction effect which adopts the prediction model based on statistical principle is not ideal to predict the fluctuation of stock yield. But the innovational prediction models don't require the data as the statistical models, and most of them only require small samples and short-term prediction, so at present, the prediction effect of these types of model is better than the statistical types.

5.3 Most Chinese stock yield fluctuation prediction models adopt single prediction model

From existing research literatures, the prediction models predicting the stock yield fluctuation of China all adopt single prediction model whether for the prediction model based on statistical principal or for the innovational prediction model based on non-statistical principle. The advantages utilizing single prediction model include that the model structure is relatively simple, and the influencing factors are few and the prediction difficulty is relatively low. But its disadvantages are also obvious. Single prediction model can not contain all influencing factors in one model, and any one sort of single prediction model only uses part of useful information and abandons other useful information. But the influencing factors in the stock market are numerous. Single prediction model can not contain and reflect enough information, so it can not better effectively predict the future tendency.

5.4 The technologies and methods to effectively deal with the non-quantitative factors are deficient

The factors which influence the stock yield fluctuation are very complex. Except for quantitative influencing factors such as stock price, turnover and stock index, the fluctuation is also influenced by non-quantitative factors such as policy, psychological fluctuation and international emergency. At present, most stock yield prediction models of China belong to purely quantitative models which can only compute the prediction result by using the quantitative information. But the non-quantitative information which largely influences the stock yield fluctuation can not be added into the quantitative model because of the limitation of the model and the processing technology, so the exact prediction always can not accord with the actual situation.

6. The development direction of Chinese stock yield prediction model

6.1 The innovational intelligent prediction model will be one development direction of Chinese stock yield prediction

First, the innovational prediction model can overcome many disadvantages such as incomplete data, large fluctuation and unreasonable distribution in China stock market, and it adopt the small sample data to implement short-term prediction to the stock market, and the prediction precision is relatively higher than the traditional statistical prediction model. Second, the intelligent models such as NN and SVM in the innovational models can simulate or partly simulate artificial intelligence, and implement complex nonlinear changeable structure processing to various factors influencing the stock market, and they can overcome not only the disadvantage that the single factor model can not contain sufficient information, but the disadvantage that the fixed structured model can not dispose the emergency, and they can fully reflect various sorts of information and changes influencing the stock market and increase the nicety of the prediction.

6.2 The integrated prediction model will be another development direction of Chinese stock yield prediction

Each sort of prediction model has its advantages and limitation in the acquisition of predicted information. How to exert the advantages of different prediction models and overcome their limitations and enhance the nicety of the prediction is the urgent problem which should be solved by the prediction theory and technology. The integrated prediction is to endow different weights for the prediction results of different prediction models according to certain principles, and implement weighted average, and obtain the final prediction result. This prediction method can overcome the disadvantage that the single prediction model has not sufficient information, fully exert the advantages of different prediction models, acquire information from different aspects to the large extents, and enhance the prediction level of stock yield.

6.3 The prediction model with various kinds of non-quantitative information will be one important development direction of Chinese stock yield prediction

At present, all stock yield prediction models belong to the quantitative prediction model, and the non-quantitative factors can not be added into the models, which induce the prediction losses large numbers of non-quantitative information and the nicety of the prediction will be influenced largely. How to translate various sorts of non-quantitative information influencing the stock market into quantitative information, add them into the stock yield prediction model, fully reflect the influences of non-quantitative factors such as policy, psychology and emergency and enhance the nicety of the prediction is one important development direction for the stock yield prediction model.

6.4 The continual absorptions and applications of new prediction theories and technologies will be the necessary tendency of Chinese stock yield prediction model

As viewed from the development process of the stock yield prediction model, we can see that the process is just that new prediction theories and technologies are continually introduced, absorbed, assimilated, applied and perfected. The fractal theory of GARCH model, the grey theory and technology of SV model, and the NN theory and technology of NN model all experience thus process. New prediction theories and technologies continually offer new views and more perfect and complete technologies for the prediction of the stock yield fluctuation, and they make the prediction effect of the stock yield fluctuation prediction model more go to perfection and nicety. At present, the chaos theory and the wavelet technology have been primarily applied in the prediction of stock yield fluctuation, and with the development of relative researches, new ideas will be introduced the prediction of stock yield fluctuation.

References

Engle R F. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica.* No.50. P.987-1008.

Ghysels, E.,Harvey, A.Renault. (1995). *Stochastic volatility.* In G.S. Maddala and C.R.Rao (eds.), Handbook of Statistics, vol.14:119-19. Elseier Science Publishers, Amsterdam, The Netherlands, 1995.

Hill T,O' Connor M and Remus W. (1996). Neural network models for time series forecasts. *Management science.* No.42(7). P.1082-1092.

Jun, YU. (2002). Forecasting volatility in the New Zealand stock market. *Journal of Financial Economics.* No.12. P.193-202.

Matsuba, I. (1991). Application of neural network sequential associator to long-term stock price prediction. *IJCNN'91, Singarporee.* P.1196-1202.

Pai, C.Lin. (2005). A hybrid ARIMA and support vector machines model in stock price forecasting. *Omega.* No.33(6). P.497-505.

Wuwei, Chen, Weiqiang & Liubo. (2001). Prediction of ups and downs of stock market by BP neural networks. *Journal of Dalian University of Technology.* No.41(1). p.9-12.

W. wuang, Y. Nakamori & S.Wang. (2005). Forecasting stock market movement direction with support vector machine. *Computers and Operations Research.* No.32(10). P.2513-2522.

Yang, Kelei, Mao, Minglai & Xu, Zhengguo. (2004). Comparative Research of Shanghai and Shenzhen Stock Market Based on Stochastic Volatility Model. *Journal of Tianjin University (Social Sciences).* No.10. p.334-338.

Yang, Yiwen & Yang, Chaojun. (2005). Financial Time Series Forecasting Based on Support Vector Machine. *System Engineering: Theory, Method and Application.* No.14(2). p.176-181.

Zhang, Shiying & Keke. (2002). ARCH Modeling System. *Journal of System Engineering.* No.6. p.236-245.

Zhang, Yongdong & Bi, Qiuxiang. (2003). Volatility Forecasting Models in the Shanghai Stock Market: A Comparison. *Journal of Industrial Engineering and Engineering Management.* No.17(2). p.16-19.

International Journal of Economics and Finance
Vol. 1, No. 1
February 2009
www.ccsenet.org/journal.html

Estimating Software Cost with Security Risk Potential

Nur Atiqah Sia Abdullah
Faculty of Information Technology and Quantitative Sciences
University of Technology MARA
40450 Shah Alam, Selangor Darul Ehsan, Malaysia
Tel: 60-3-5521-1175 E-mail: szeyieng@yahoo.com; atiqah@tmsk.uitm.edu.my

Rusli Abdullah, Hasan Selamat, Azmi Jaafar
Faculty of Computer Science and Information Technology
Universiti Putra Malaysia
43300 Serdang, Selangor Darul Ehsan, Malaysia
Tel: 60-3-8946-6575 E-mail: {rusli, hasan, azmi}@fsktm.upm.edu.my

Kamaruzaman Jusoff (Corresponding author)
Yale University
Yale's Centre for Earth Observation
21 Sachem St, New Haven, CT 06511, USA
Tel: 203-432-1384 E-mail: jusoff.kamaruzaman@yale.edu

Abstract

Software houses are now keen to provide secure software as requested by customers' desire with respect to security and quality of their products especially related to the software costing estimation in the software development and implementation environment. Therefore, there is a need to identify the potential security risks while estimating the application cost. In this paper, we provide a list of potential security risks throughout the system development life cycle (SDLC). This list provides useful insights for software developers and practitioners in identifying security risks so that it can be encountered as security cost in the application.

Keywords: Security factor, Software cost estimation, Software metrics

1. Introduction

Software cost estimation (SCE) is very important to estimate the cost that involved in developing software. SCE can be vary depends on different models. But most of the cost estimation models do not consider the costs that invoke security while developing software. It is because security is often an afterthought when developing software and is often bolted on late in development or even during deployment or maintenance (Shanai *et al*, 2006). Currently, secure software development has gained momentum during the past couple of years and improvements have been made. Software houses are now keen to provide secure software as requested by customers' desire with respect to security and quality of their products (Jari, 2006).

Besides, a risk discovered late in the project lifecycle becomes a fire to fight. In examining some real-world experiences, project managers faced project "fires" when risks were not identified during the planning phase. These fires often resulted in added costs or negative consequences (Frederick, 2006). But, engineering security will substantially raise software – project cost and there has been wide variation in the amount of added cost estimated by different models (Colbert et. al., 2004). Therefore, there is a need to include secure software costing while estimating the application cost. Extended the existing software cost estimation technique is essential in order to calculate the cost that includes security in application.

2. Overview of Related Works

2.1. Parametric Software Cost Estimation Models

From two decades, many studies that comparing parametric models (Kemerer, 1987; Srinivasan *et. al*., 1995; Chulani *et.*

al., 1998; Briand *et. al*., 1999) for software cost estimation have been published. There are numbers of widely used parametric models (Symons, 1991; Albrecht, 1979; Boehm, 1981; COSMIC, 2007; Putnam, 1978) are reviewed, compared and evaluated from many aspects or factors.

Kemerer (Kemerer, 1987) conducted a well-known study reporting on the relative accuracy of four software cost estimation models includes Function Points (Albrecht, 1979), SLIM (Putnam, 1978), COCOMO (Boehm, 1981), and ESTIMACS, in being able to predict the actual costs of 15 completed software development projects. The results indicated that the models require significant calibration, and the study provided insight to the factors affecting modern software development productivity. Briand et al. (1998) examined the accuracy of data-driven software cost modeling based on certain criteria and compared organization-specific with multi-organization models. The study assessed and compared the models based on criteria such as ordinary least square regression, stepwise ANOVA, Classification and Regression Trees (CART) and analogy. This study urged the need for risk analysis to be an integrated part of software cost estimation procedures (Briand *et al*., 1998).

Colbert et al (Colbert *et. al*., 2004; Colbert *et. al*., 2006) have developed a model for costing secure software-intensive systems (COSECMO), which extended from COCOMO II. COSECMO is based on behavior-analysis activities, such as analyzing industry practices with respect to security, 149 Security Targets that registered on the National Information Assurance Partnership (NIAP) website and conducting preliminary surveys of experts in software development and in security. COSECMO introduces a new cost driver (Wu *et. al*., 2006) and defines guides for setting other COCOMO drivers when costing the development of a secure system. Our current study makes the contribution of evaluating the common software cost estimations models like FPA, MkII, COCOMO II, SLIM and COSMIC-FFP from the security aspects.

2.2 IFPUG 4.1

In Albrecht Function Point Analysis (FPA) (Albrecht, 1979; Longstreet, 2008, Longstreet Consulting Inc, 2008), there are two parts in the measurement, which are Unadjusted Function Point and Adjusted Function Point. First part consists of five components, which are External Inputs (EI), External Outputs (EO), External Inquires (EQ), Internal Logical Files (ILF) and External Interface Files (EIF), and these components are evaluated by complexity weights. The second part is 14 General System Characteristics (GSCs) that measured on a six point scale. The GSCs are presented in Table 1. Based on the manual (Longstreet, 2008), the security factors are not highlighted in FPA. Existing FPA lacks in emphasizing on security costing in developing application. Besides, the issue of security is discussed by the users. The discussions include the sample scenario, user request, technical design, scenario resolution, and general discussion about user security file, access profile file, external inputs, external outputs, external enquiries and contributions to measurement (Total Metrics, 2001).

From Table 1, there are no security factors or risk assessments listed in the first level of the GSCs. However, the security factors are slackly considered in the second level of GSCs (as shown in Table 2). Therefore, any software that contains security coding or schemes, the existing FPA will not be able to estimate the cost that invoked by the built-in security coding in the software. This added cost is important because highly-secure software will increase costs based on the different models.

2.3 Software Life Cycle Model (SLIM)

Software Life Cycle Model (SLIM) is developed by Putnam (Putnam, 1978). It used validated data from over 2600 projects from industry, which stratified into nine application categories raging from microcode to business systems from IBM Canada Ltd. Laboratory. There are three inputs in SLIM, consists of software size, process productivity parameters, and management constraints. Software size is calculated by using Source Line of Code (SLOC), Function Points, Modules and Uncertainty. Process Productivity consists of methods, skills, complexity and tools. Management Constraints are maximum people, budget, schedule and required reliability.

These three inputs go through three main processes to get the minimum possible schedule, evaluate practical alternatives and optimum estimate to meet constraints. As the outputs for SLIM, there are four graphs namely Staff v.s Time, Cumulative Cost vs. Time, Probability of Success vs Time and Reliability vs. Time. This concept supports the SLIM tool for two important management indicators, which are Productivity Index and Manpower Buildup Index (Panlilio, 1994). Therefore, in Putnam's SLIM, there is no emphasis on the security aspects throughout the system development.

2.4 COSMIC-FFP

The new release COSMIC-FPP (COSMIC, 2007) has changed this method's name to COSMIC Method. COSMIC provides a standardized method of measuring a functional size of software from the functional domains commonly referred to as business application (or MIS) software and real-time software.

The COSMIC measurement method involves applying a set of models, principles, rules and processes to the Functional

User Requirements (or 'FUR') of a given piece of software. The general measurement process consists of three phases, which are the Measurement Strategy Phase, the Mapping Phase, and the Measurement Phase. During the Measurement Strategy Phase, four elements are identified, which are purpose, scope, functional users and level of granularity of the software. The Mapping Phase identifies functional processes, data groups (Entry, Exit, Write and Read data) and data attributes. The last phase is the Measurement Phase. The data movements of each functional process are identified and measured by 1 Cosmic Function Point as the size of one data movement. Then, aggregate the measurement results to get the size of the software. (COSMIC, 2007). Throughout the measurement, the security aspect is slightly mentioned as one of the examples in the Environment Constraints in Functional User Requirements (FUR). Therefore, if a software contains security coding or schemes, the existing COSMIC will not be able to estimate the cost that invoked by the built-in security coding in the software.

2.5 COCOMO II

COCOMO (COnstructive COst MOdel) which created by Barry Boehm is currently established to COCOMO II.2000 (Boehm, 1981; Boehm *et. al*., 2000). The full COCOMO II model includes three stages. Stage 1 supports estimation of prototyping or applications composition efforts. Stage 2 supports estimation in the Early Design stage of a project, when less is known about the project's cost drivers. Stage 3 supports estimation in the Post-Architecture stage of a project.

From the Effort Estimation Equation, COCOMO II (Boehm, 2000) considers the Effort Multipliers, size by using source line of code and adjusted function points, scale factors, and adaptation adjustment factors. Effort Multipliers includes product attributes, platform attributes, personnel attributes, and project attributes. These attributes are as follows:

Product Attributes

- Required software reliability (RELY)
- Database size (DATA)
- Documentation match to life-cycle needs (DOCU)
- Product complexity (CPLX)
- Required Reusability (RUSE)

Platform Attributes

- Execution time constraint (TIME)
- Main storage constraint (STOR)
- Platform volatility (PVOL)

Personnel Attributes

- Analyst capabilities (ACAP)
- Applications experience (APEX)
- Programmer capabilities (PCAP)
- Platform experience (PLEX)
- Programming language experience (LTEX)
- Personnel Continuity (PCON)

Project Attributes

- Use of software tools (TOOL)
- Multisite Development (SITE)

In additional to the 16 adjustment factors, there are two user defined factors, which their initial values are set to 1. In this version of COCOMO, the security aspects are not listed in the calculation. However, Expert COCOMO (Boehm et. al., 2000) aids in project planning by identifying, categorizing, quantifying, and prioritizing project risks. It used an automated heuristic method for conducting software project risk assessment.

2.6 Mk II FPA

Mk II Function Point Analysis (Mk II FPA) (Symons, 1991; UKSMA, 1998) is one of the most widely used SCE that managed by United Kingdom Software Metrics Association (UKSMA). It is a method for the quantitative analysis and measurement of information processing applications. It quantifies the information processing requirements specified by the user to provide a figure that expresses a size of the resulting software product. This size is suitable for the purposes of performance measurement and estimating in relation to the activity associated with the software product (UKSMA, 1998). Mk II FPA measures the functional size based on logical transactions concept. Each logical transaction consists of three components; input across an application boundary, processing involving stored data within the boundary and output back across the boundary. Besides, MK II measures the influence on the size of the application of each of 19 (or more user defined characteristics) technical characteristics on a scale of 0 to 5. Refer to Table 3 for the characteristics (UKSMA, 1998). The sum for all characteristics or Total Degrees of Influence (TDI) is used to compute Technical Complexity Adjustment (TCA). Function Point Index (FPI):

$$FPI = Wi * \sum Ni + We * \sum Ne + Wo * \sum No$$

where

$Wi = 0.58$; $We = 1.66$; $Wo = 0.26$; Ni = Input Data element Types;

Ne = Data Entity Types Referenced; No = Output Data Element Types;

Technical Complexity Adjustment (TCA):

$$TCA = (TDI * C) + 0.65$$

where

Current Industry Average Value of $C = 0.005$;

Adjusted Function Point Index (AFPI):

$$AFPI = FPI * TCA$$

From Table 3, there is a characteristic related to Security, Privacy, Auditability to measure the level of confidentiality or security. But the related scores and descriptions are as in Table 4. From Table 4, the scores for security, privacy and auditability are quite low. Besides, the details of risk assessment or security are not listed here. Therefore, when a software is embedded with security coding, the existing Mk II might under estimate the cost that invoked by the built-in security coding in the software.

3. Security standards

There are four security standards that are reviewed in this paper, which are commonly used as guidelines in considering the security cost in system and web development.

3.1 IT Security Cost Estimation Guide

This Information Technology Security Cost Estimation Guide (Department of Education, 2002) provides the security cost estimation based on Management Controls, Operational Controls and Technical Controls. For each control, there are categories of potential security risks that a developer has to estimate while calculating for the security cost. To effectively use this guidance, the user has to complete the National Institute of Standards and Technology (NIST) Self Assessment Guide for Information Technology Systems. According to this guide, security must be considered in integral part of the overall IT infrastructure in order to effectively manage risks. As risks vary by initiative and fluctuate throughout the life cycle of each system, the necessary costs for security controls will vary as well. Therefore, risks should be identified according to the life cycle phases during which they will pose the most imminent threat, as this will largely determine which fiscal year to budget for controls (Department of Education, 2004).

This guide suggested an IT Security Costing Framework. This framework comply with mandated standards of information system security, and each system, at a minimum, must meet requirements that fall within the following 17 management, operational and technical control categories. Refer to Table 5. These 17 categories align with the 17 elements of IT security identified in the NIST Self Assessment. This guide provides the calculation of costs based on government employees or contractors rates.

3.2 Common Criteria for Information Technology Security Evaluation

Common Criteria for Information Technology Security Evaluation (CC v 3.1) (CC, 2008) consists of three main parts: Part 1 – Introduction and general model; Part 2 – Security functional requirements; Part 3 – Security assurance requirements. This guide is used by COCOMO II to extend its model to COSECMO (Colbert et. al., 2006; Wu et. al., 2006). This CC philosophy is to provide assurance based upon an evaluation of the IT product or system that is to be trusted. The assurance classes include configuration management, delivery and operation, development, guidance documents, life cycle support, tests, vulnerability assessment, protection profile and security target evaluation.

Assurance family from CC v 3.1 lists the following assurance classes to be considered during the security evaluation:

- Configuration Management (ACM)
- Delivery and Operation (ADO)
- Development (ADV)
- Guidance documents (AGD)
- Life Cycle Support (ALC)
- Tests (ATE)
- Vulnerability Assessment (AVA)
- Protection Profile Evaluation (APE)
- Security Target Evaluation (AST)

For each assurance family, brief description is provided to help the evaluators to assess the product or system. The evaluation assurance level is from scale EAL1 to EAL7. For each scale, there are some assurance components to be evaluated (CC, 2008).

3.3 The Open Web Application Security Project (OWASP)

The Open Web Application Security Project (OWASP) (OWASP, 2008) is a worldwide free and open community focused on improving the security of application software. It is an open community dedicated to find and fight the causes of insecure software. It is a guide for building secure web applications and web services. OWASP has chosen Microsoft's threat risk modeling process as it works well for the unique challenges facing application security. OWASP guidelines list the following aspects to be considered:

- Threat Risk Modeling
- Handling e-Commerce Payments
- Phishing
- Secure Web Services
- Authentication
- Authorization
- Session Management
- Data Validation
- Interpreter Injection
- Canoncalization, Locale and Unicode
- Error Handling, Auditing and Logging
- File System Protection
- Buffer Overflows
- Administrative Interfaces
- Cryptography
- Configuration
- Maintenance
- Denial of Service Attacks

OWASP guidelines help to encounter the causes of the insecure software. It lists the risk factors that potentially attack an application. This security factors should be considered during the application development. Furthermore, these factors should evolve some costs in developing secure software (OWASP, 2008).

3.4 Control objectives for information and related technology (COBIT)

Control Objectives for Information and related Technology (COBIT) (ISACA, 2008) provides good practices across a domain and process framework and presents activities in a manageable and logical structure. COBIT's good practices represent the consensus of experts. These practices will help optimize IT-enabled investments. One of the objectives of COBIT is to ensure IT risks are managed appropriately. Therefore, the focus area includes risk management, which requires risk awareness. COBIT is a process-oriented and controls-based framework. There are five IT Governance focus areas, which are namely strategic alignment, value delivery, risk management, resource management and performance management. It is organized according to IT process like plan and organizes, acquire and implement, deliver and support, monitor and evaluate. During plan and organize, the user of this guide has to assess and manage IT Risks (P09). In the process of Acquire and Implement, the user has to Acquire and Maintain Technology Infrastructure (AI3). The most important process that related to security is Deliver and Support, which includes Ensure Continuous Service (DS4) and Ensure Systems Security (DS5). The aspects and considerations of COBIT are as follows:

- PO2 Define the Information Architecture
- PO3 Determine Technological Direction
- PO9 Assess and Manage IT Risks
- AI2 Acquire and Maintain Application Software
- AI3 Acquire and Maintain Technology Infrastructure
- DS4 Ensure Continuous Service
- DS5 Ensure Systems Security
- DS11 Manage Data
- DS12 Manage the Physical Environment
- ME2 Monitor and Evaluate Internal Control
- DS13 Manage Operations

4. Methodology

This methodology produces the proposed potential security factors. It consists of:

a. Review on the existing software cost estimation models such as Mk II, COSMIC-FFP, COCOMO II, SLIM and FPA, concentrating on the security aspect.

b. Review four security standards that related to IT and development like IT Security Guidelines, CC v 3.1, COBIT 4.1 and OWASP, for identifying the potential security factors that can be included in software cost estimation.

c. Integrate the lists from four security standards and categories the items into lists according to a proposed potential security factors.

Matching the proposed potential security factors in System Development Life Cycle (SDLC), which include Plan, Design, Code, Test and Deploy.

5. Results and discussions

The integrated list of proposed potential security factors is combined from four security standards and organized according to the software development life cycle as shown in Table 6. All these lists are integrated according to the steps and proposed potential security factors in Table 6. This integrated list forms a framework that helps the practitioners and developers to consider the potential security risks while planning for an application (Refer to Figure 1).

List for Security Requirements (SR)

- Specification of Management
- Abuse Cases
- Threat Risk Modeling
- Risk Analysis
- Management of IT Security
- Security Rules
- PHP Guidelines
- Functions
- IT Security Plan

List for Security Features (SF)

- User Identify
- Management
- Interpreter Injection
- Authorization
- User Attribute Definition
- Protection of Security
- Data Validation
- Authentication
- Session Management
- Cryptography
- Exchange of Sensitive Data
- Verification of Secrets
- Data Integrity

List for Functional Features (FF)

- Management Security
- Production and Input/Output Controls
- File System
- Web Services
- Incident Response Capability
- Handling E-Commerce Payments
- Configuration
- Functions Behavior
- Personnel Security

List for Attack Planning (AP)

- Phishing
- Denial of Service Attacks
- Non-bypass Ability of TSP
- Malicious Software Prevention, Detection and Correction

List for Formal Review & Sign Off (FR)

- Review of Security Controls
- Logical Access Controls
- Audit Data Generation
- Security Attribute based Access Control

List for Software Security Assurance (SSA)

- Security Testing
- Security Surveillance
- Security Monitoring

List for Final Security Review & Sign Off (FSR)

- Final Review of Security Controls
- Final Security Attribute based Access Control
- Final Logical Access Controls
- Final Audit Data Generation

List for Infrastructure Application Security Measures (ASM)

- Security Awareness
- Security Training
- Security Education

List for Software Hardening & Application Security Monitoring (SHA)

- Hardware Maintenance
- System Software Maintenance

These potential security factors are able to help the practitioners and developers to estimate the software security costs.

The proposed potential security factors are the factors that might influence the software cost estimation. As mentioned before, security is often an afterthought when developing software and is often bolted on late in development or even during deployment or maintenance. Any occurrences that caused by the security aspects will add to the existing software cost when the developers fix the problems. It is better to discover and estimates the security cost before or during the development of application. Therefore, this proposed potential security factors is a big help to the practitioners or developers while designing or developing any application.

6. Conclusions

Software houses are now keen to provide secure software as requested by customers' desire with respect to security and quality of their products (Jari, 2006). In examining some real-world experiences, project managers faced project "fires" when risks were not identified during the planning phase. These fires often resulted in added costs or negative consequences (Frederick, 2006). Engineering security will substantially raise software – project cost and there has been wide variation in the amount of added cost estimated by different models (Colbert et. al., 2006). Therefore, there is a need to include secure software costing while estimating the application cost. Based on the review of current software cost estimation models and security standards, a proposed potential security factors is suggested. This integrated list can help the software developers to encounter the built-in security factors in the application by estimating the application cost more precisely.

In the future, extending the existing software cost estimation technique is essential in order to calculate the cost that includes security in application. We also decide to visualize the integrated list by developing a prototype of software cost estimation tool and validate it through case studies in the system development companies.

References

Albrecht, A.J. (1979). Measuring application development productivity. In: SHARE/GUIDE: Proceedings of the IBM Applications Development Symposium, 79:(83-92), 1979.

Boehm, B. (1981). Software Engineering Economics. Englewood CI@, NJ Prentice-Hall, 1981.

Boehm, B., Abts, C., Brown, A.W., Chulani, S., Clark, B.K., Horowitz, E., Modachy, R., Reifer, D., Steece, B. (2000). Software Cost Estimation with COCOMO II. Prentice Hall, Inc., New Jersey. ISBN 0-13-026692-2, 2000.

Briand, L.C. El Emam K., Maxwell, K., Surmann, D., Wieczorek, I. (1998). An Assessment and Comparison of Common Software Cost Estimation Models. Technical Report, ISERN TR-98-27, International Software Engineering Research Network.

Briand, L.C., El Emam, K., Surmann, D., Wieczorek, I., Maxwell, K.D. (1999). An Assessment and Comparison of Common Software Cost Estimation Modeling Techniques. ICSE '99:(313-322), Los Angeles CA, ACM 1999.

CC. (2008). Common Criteria for Information Technology Security Evaluation, version 3.1. [Online] Available: http://www.commoncriteriaportal.org.

Chulani, S., Boehm, B., Abts, C. (1998). Software Development Cost Estimation Approaches – A Survey. University of Southern California – Computer Science Engineering, 1998.

Colbert E., Wu D., Chen Y., and Boehm B. (2004). Cost Estimation for Secure Software and Systems. University of Southern California, United States of America.

Colbert E., Wu D., Chen Y., and Boehm B. (2006). Cost Estimation for Secure Software and Systems. University of Southern California, USA.

Common Software Measurement International Consortium (COSMIC) (2007). The COSMIC Functional Size Measurement Method Version 3.0 – Measurement Manual. September 2007

Department of Education, USA. (2002). Information Technology Security Cost Estimation Guide. November 2002.

Frederick, L. (2006). The Importance of Identifying Risk during Project Planning. In Ackermann, D., Project Management In Practice, Boston University, Metropolitan College, The 2006 Project Risk and Cost Management Conference, Commonwealth Avenue.

ISACA. (2008). Control Objectives for Information and related Technology (COBIT) – Framework, Control Objectives, Management Guidelines, and Maturity Models. IT Governance Institute, USA.

Jari, R, (2006). Contracting over the Quality Aspect of Security in Software Product Markets, University of Lapland, Rovaniemi, Finland. ACM Press, Oct 2006.

Kemerer, C.F. (1987). An empirical validation of software cost estimation models. Communications of the ACM vol. 30, no. 5, 416-429, May 1987.

Longstreet Consulting Inc. (2008). [Online} Available: http://www.SoftwareMetrics.com

Longstreet, D. (2008). Function Points Training and Analysis Manual. ISBN: 0-9702439-3-6, Jan 2008.

OWASP. (2008). The Open Web Application Security Project – A Guide to Building Secure Web and Web Services. Black Hat 2nd Ed.

Panlilio N.Y. (1994). Software Estimation Using the SLIM Tool. IBM Canada Ltd Laboratory Technical Report TR 74.102.

Putnam, L.H. (1978). A general empirical solution to the macro software sizing and estimation problem. IEEE Transactions on Sofiware Engineering, ~01.4, no. 4, 345-381, July 1978.

Shanai Ardi, David Byers, Nahid Shahmehri (2006). Towards a Structured Unified Process for Software Security, ACM Press, May 2006.

Srinivasan, K., Fisher, D. (1995). Machine learning approaches to estimating software development effort. IEEE Transactions on Software Engineering, Vol. 21, no. 2, 126-137, February 1995.

Symons, C. (1991). Software Sizing and Estimating – Mark II FPA. Wiley Series in Software Engineering. John Wiley and Sons, U.K., ISBN 0-471-92985-9, 1991.

Total Metrics. (2001). [Online] Available: http://www.totalmetrics.com

United Kingdom Software Metrics Association (UKSMA). (1998). Mk II Function Point Analysis – Counting Practices Manual Version 1.3.1. September 1998.

Wu, D., and Yang, Y. (2006). Towards An Approach for Security Risk Analysis in COTS Based Development. Center for Software Engineering, University of Southern California, Los Angeles, CA. Proceeding of Software Process Workshop on Software Process Simulation.

Table 1. List of General System Characteristics (GSCs)

Characteristics	Brief Descriptions
Data communications	How many communication facilities are there to aid in the transfer or exchange of information with the application or system?
Distributed data processing	How are distributed data and processing functions handled?
Performance	Did the user require response time or throughput?
Heavily used configuration	How heavily used is the current hardware platform where the application will be executed?
Transaction rate	How frequently are transactions executed daily, weekly, monthly, etc.?
On-Line data entry	What percentage of the information is entered On-Line?
End-user efficiency	Was the application designed for end-user efficiency?
On-Line update	How many ILF’s are updated by On-Line transaction?
Complex processing	Does the application have extensive logical or mathematical processing?
Reusability	Was the application developed to meet one or many user’s needs?
Installation ease	How difficult is conversion and installation?
Operational ease	How effective and/or automated are start-up, back up, and recovery procedures?
Multiple sites	Was the application specifically designed, developed, and supported to be installed at multiple sites for multiple organizations?
Facilitate change	Was the application specifically designed, developed, and supported to facilitate change?

This list of 14 General System Characteristics (GSCs) is being measured as part of the Adjusted Function Point Analysis. The characteristics are measured on a six point scale, which is from 0 to 5.

Table 2. Security features and descriptions in the second level of GSCs

Feature	Characteristic	Descriptions / Second Level
Security	Heavily Used Configuration	Some security or timing considerations are included.
	Complex Processing	Sensitive control (for example, special audit processing) and / or application specific security processing

This table shows the characteristics and descriptions that related to security factors. There are only two GSCs that have concerned in security factors in their second level. Heavily Used Configuration has a score of 2 for its second level while Complex Processing has a score of 3.

Table 3. Technical Characteristics

Characteristics	Brief Descriptions
Data communication	Data and control information use communication facilities
Distributed Function	Application spread over two or more processors
Performance	Application response/throughout influence the application
Heavily used configuration	The equipment that will run the application is already heavily used
Transaction rates	The high rate of arrival of transactions causes problems beyond those of Performance
On-Line data entry	Terminal used for input
Design for End-user efficiency	Human factors considered in design
On-Line update	Data updated in real time
Complexity of processing	Internal complexity beyond that dealt with by entity counting conventions
Usable in Other Applications	The code is designed to be shared with or used by other applications
Installation ease	Data conversion and ease of installation were considered in design
Operations ease	Ease of operation was considered in design
Multiple sites	The application is to be used in many sites and/or many organizations
Facilitate change	Future changes to the application a consideration on design
Requirements of Other Applications	Interfaces
Security, Privacy, Audit ability	Special features of confidentiality / security
User Training Needs	Specific requirements
Direct Use by Third Parties	Degree of use/connection to the application
Documentation	-

From Table 3, there are 19 technical characteristics on a scale of 0 to 5. These characteristics measure the influence on the size of the application.

Table 4. Second level of Technical Characteristics

Score	Brief Descriptions
1	If the application has to meet personal, possibly legal, privacy requirements
1	If the application has to meet special auditability requirements
2	If the application has to meet exceptional security requirements
1	If encryption of data communications is required

Table 4 shows the second level of Technical Characteristics for Mk II. As noticed, the scores for security, privacy and auditability are ranged from 1 to 2. There are no details of risk assessment or security listed here.

Table 5. IT Security Costing Framework

Type of Control	Categories
Management	Risk Management
	Review of Security Controls
	Life Cycle
	Authorize Processing (Certification & Accreditation)
	System Security Plan
Operational	Personnel Security
	Physical and Environment Protection
	Production and Input/Output Controls
	Contingency Planning
	Hardware and System Software Maintenance
	Data Integrity
	Documentation
	Security Awareness, Training, and Education
	Incident Response Capability
Technical	Identification and Authentication
	Logical Access Controls
	Audit Trails

Table 5 shows the requirements that fall within the 17 management, operational and technical control categories, which align with the 17 elements of IT security identified in the NIST Self Assessment.

Table 6. Proposed Potential Security Factors in System Development Life Cycle

Step	Proposed Potential Security Factors
Plan (P)	Security Requirements (SR)
Design (D)	Security Features (SF)
	Functional Features (FF)
Code (C)	Attack Planning (AP)
	Formal Review & Sign Off (FR)
	Security Coding, Review & Audit (SCR)
Test (T)	Software Security Assurance (SSA)
	Final Security Review & Sign Off (FSR)
	Infrastructure Application Security Measures (ASM)
Deploy (E)	Software Hardening & Application Security Monitoring (SHA)

Table 6. shows an integrated list of proposed potential security factors, which is extracted from four security standards. These potential security factors are organized according to the software development life cycle, which consists of a cycle of Plan, Design, Code, Test and Deploy.

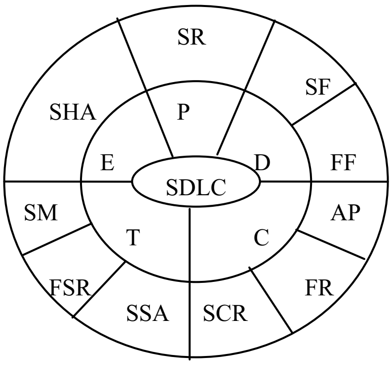

Figure 1. Proposed Potential Security Factors Framework

This framework is based on phases of SDLC and an integrated list that extracted from four security standards.

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Strengthening the Construction of the Constitutionalism and Promoting the Healthy Development of the Socialism Market Economy

Bijun Xiong
Hunan Institute of Socialism
Changsha 410001, China
Tel: 86-731-479-2700 E-mail: xbj721@163.com

Abstract

After about thirty years' reform, the socialism market economy system has been basically founded, and the function of the market in the distribution of resource is increasingly obvious. In the new century and new stage, the reform is in the process of assaulting fortified position, and we must eliminate the system obstacles which baffle the function of the socialism market system. Therefore, we have to carry out the legal basic status if the constitution, establish the legal base of the market economy, i.e. the constitutionalism, establish the democratic system of the constitutional supervision, and perfect the national power structure. And we have to confirm the rule of law of the judicial independence, perfect the judicial control, limit the right of the government, establish the public government with the democratic laws, promote the nonofficial system transformation of the constitutionalism, perfect the system of the socialism market economy, follow the rule of the market economy, establish the macro-control system according to the national situation and the constitutional order, and establish the distribution system adapting the socialism market economy through the reform of the constitutionalism. So we can strengthen the construction of the constitutionalism and promote the healthy development of the socialism market economy.

Keywords: Constitutionalism, Socialism market economy, Construction, Reform

The constitutionalism means the existence and operation of the government accord wit the constitution because its right is limited by the public opinion. American famous political commenter Thomas Paine (1737-1809) had said that "the constitutionalism is not the action of the government, but the behavior that human establish the government, and the government without the constitution is only the power of without right". And he also pointed out that "the constitution is the thing before the government, and the government is the production of the constitution" and "the function of the constitution for the freedom likes the function of the expression for the language". Therefore, the essential of the constitutionalism is that there is a set of higher law, i.e. the constitution, to limit the power of the government and its administration range on the governmental power. And we can think the limited government under the limit of law and democracy can compose the constitutionalism. Even there is the constitution without the limit of law and democracy, and there is not the real constitutionalism (Wang, 2006).

In the history, the constitutionalism had made very important contributions for the economic and social development of the western countries. In the present transformation process from the planned economic system to the market economic system, because of the faultiness of the constitutionalism, the deficiency of the market economic rule based on the constitutionalism has produced negative function (Wang, 2006). So we can adopt following measures to confirm the constitutionalism and fully exert the functions of guarantee and promotion for the economic development.

1. Putting the basic legal status of the constitution into effect and establishing the base of the legal rule of the market economy, i.e. the constitutionalism.

Up to now, China all along puts the economic development on the primacy, so the assistant construction of the constitutionalism is ignored. Because the constitution has not been effectively implemented in recent years (most economic policies, active rights and civilian obligation in present constitution can not been implemented), so people produce a sort of universal mistakes, i.e. as the essential law, the constitution has no relationship with the actual living in the society. The common legislations such as civil law, property law, company law, law of against competition by inappropriate means, securities exchange act and environmental protection law are independent of the constitution. The constitution offers a sort of meaningless reference of "mother law" for the productions of these laws, and it doesn't put the basic legal status of the constitution into effect. Because the constitution has not been as the base, so the implementation of many laws is not ideal. Chinese laws has not prevented the continual losses of stated assets, protected the successive decrease of the lands which are the national strategic resources, stopped the jobbery of the securities business, and effective restrained the counterfeit and shoddy products and coal safety accidents. Though there are many problems from the laws, but we more and more realize that only good regulations are not far enough, and they can not been implemented, even perfect laws like nothing. Nowadays, the main problems of China are not "whether there are laws?", "who participate the constitution of laws?", "what laws we should constitute?" and "what mechanisms can supervise the implementation of laws?" Theses problems can not been solved by the legal texts themselves, and they are the problems about the constitutionalism before the legal texts, and they are the preconditions for that the laws are laws, and if they can not been solved, the legal order can not be established and the economic development will get in unhealthy and un-continual dangers and confusions. So we have to put the basic legal status of the constitution into effect and establish the base of the legal rule of the market economy, i.e. the constitutionalism.

2. Establishing the democratic system supervised by the constitution and perfecting the national right structure

The national right structure includes the relationship between the central government and the local government, and the power allocation of the legislation, administration and jurisdiction. For the democracy without live constitutional base, the departments and officials of the central government and the administration discretionary powers of local governments will be infinitely expanded, and their behaviors and decisions can not be limited in the effective supervision and control of the public, and restricted by the legal systems and regulations. Under that situation, even more laws constituted by the legislatures can not restrict them, and there is not real legal democracy. So we have to establish the democratic system supervised by the constitution, supervise the constitution and the implementation of the laws and the regulations, supervise the public management of the administration departments and local government institutions, perfect the national right structure and realize real legal democracy.

3. Confirming the legal system with independent justice and perfecting the judicial control

As we know, in all governmental institutions, the judicial institution contributes more to the development of the market economy, because the market economy first means free competition, not the governmental control, but both the legislation and the administration means the control of the government in the special domain, though it is the control by the law. According to the classic liberalism theory, the market is a sort of spontaneous order produced by the economic behaviors through the voluntary trades, but sometimes the market order also needs the protection of the governmental laws, and so we have to establish a sort of constitution, i.e. the judicial constitution among the legislation, administration and justice, and it is necessary to fairly settle private confusions. Through five years' judicial reform, there are many disadvantages in the judicial institution which still could not offer firm legal guarantee and establish the perfect market order. Therefore, we must confirm the legal system of the judicial independence, perfect the legal control, strengthen the restriction of the administration behaviors and maintain the various rules of the market economy.

4. Limiting the right of the government and establishing the public government with law and democracy

As a big administration country, the characters of traditional China include not only weak justice, but strong administration. Especially after 1949, because of the influences of the all-powerful government and the development strategic selection (Wang, 2006), the power of the administration is expanded largely, and it has exceeded the limitation of any law, and it can obstacle the development of the market economy. To healthily develop the market economy, the administration right must be limited and controlled firstly. In recent years, China has gradually realized the limitation of the government, and the "limited government" has been the important principle of the administrational law. The "Administrative Permission Law" implemented in 2004 definitely embodied the principle of the limited government, i.e. if the market could be effectively adjusted by the civil decision or the market competition mechanism, the government should not set up the administrative permission (Zhang, 2004).

The economic crisis in the earl of the twentieth century indicated that the market economy had not been the myth which could sustain itself, and it needed the proper interference and protection from the government, of course, it was not the random interference of the administration, and it included the constitutions of a series of laws such as the anti-monopoly law, the securities exchange act, the social security act, the environment and consumer protection law. The effective implementation of these laws can not only protect the healthy free competition, but correct some bad social consequences in the market. So we must limit the administrative right of the government, and under the restriction of the law and democracy, the administrative right of the government only is limited to offer the protection of the national security, the maintenance of the social justice, the provision of the social public products and the compensation of the market deficiency.

5. Promoting the nonofficial system transformation of the constitutionalism and perfecting the socialism market economy system

In the market economy system, there is market failure and the government failure, and they all can be solved through the establishment of the constitutionalism, but under the special condition, the "constitutional failure" will occur. As viewed from the new system economics, the essential is that both the official system and the nonofficial system of the constitutionalism can not harmonize and integrate, and in the economic development and the social transformation process of China, the essential is the lag of the nonofficial system of the constitutionalism.

Because of many characters such as spontaneity, non-compulsion, universality, durative and lagged transformation possessed by the nonofficial system and many characters possessed by the government leading the compulsory systems, the nonofficial system transformation in the system transformation process of the marketization of China seriously lagged, and the transformation of the nonofficial system and the harmony and the integration with the official system of the constitutionalism are influenced. So we must promote the nonofficial system transformation and perfect the socialism market economic system from following aspects (Xiong, 2002). First, quicken up the gradual reform, increase the knowledge storage of modern market economic law and democracy, and drive the transformation of the nonofficial system of the constitutionalism from the exterior. Second, increase the educational investment for the ideal cultural consciousness of the modern market economy, drive the transformation of the nonofficial system of the constitutionalism from the exterior, and form the basic concept of the modern market economy. Third, accelerate the cultural basic construction of the market economy, establish new socialism market economic culture and promote the transformation of the nonofficial system of the constitutionalism. The cultural tradition decides the public attitudes, and the public attitudes directly decide the construction and perfection of the socialism market economic system. Some foreign famous experts such as Samuel Huntington and Michael Porter discussed the function between the culture and the economic development in the book of "the important function of the culture, how the value view influence the human advancement". And in Latin America, the culture is baffling the economic development, but in Japan, Korea, Singapore and some western countries, the culture has been a sort of active power to drive the economic development and the social advancement. Especially in the situation that the compulsory system transformation and traditional idea and culture are so deep in China, when the official system of the social market economy system transforms, if the nonofficial systems according with the ideal cultural bases are deficient, and if the nonofficial systems of the traditional economic system don't transform or transform slowly, the perfection and the operation of the socialism market economic system will be certainly influenced. And we are required to accelerate the cultural basic construction of the market economy, quicken up the alteration of the traditional culture, establish the socialism market economic culture combining with the traditional culture and the legal democracy of the modern market economy, and establish the bases of ideas and cultures of the constitutionalism in the process of perfecting the socialism economic system.

6. Following the market economic rules and establishing the macro control system according with the national situation according the constitutional order

In the market economy, the macro control is that the government adjusts and controls the national economic gross of by the financial and coin measure on the macro layer. The main body of the control is the central government, and the control object is the gross of the national economy. But because of the influences of the all-powerful government and the development strategic selection, the macro control of China has been the tool that the government expands the market and controls or invades the market. At present, the product overstock and production ability surplus in China are not the economic crisis of the production surplus for the economic meaning, and in the special condition with Chinese feature, the system obstacles seriously restrain the social security, medical health protection, compulsory education, middle and higher education, middle and high class houses in the nonpublic products, make them get in the deficient status, and that is not the problem of purchasing power (Liu, 2007). Therefore, we have to follow the rule of the market economy, and establish the financial and tax system with the restriction of the rules according to the constitutional order. The effect of the macro control is not simply decided by the composing or the type of the government finance or the coin policy, and it is more decided by the system environment of the macro control. For example, which constitutions decide the economic policy? What motivations do people in these constitutions have? Does the power center advocate certain policy such as the financial policy? Does the individual have enough knowledge to supervise the politicians' political activities? Does free news or free speech which makes the public the dope of the policy constitution influence the aim and the effect of the macro control? So, the precondition of the macro control of the transformation economy is to establish the financial and tax system in the restriction of the constitutionalism rules, and based on that, the macro control system according with Chinese situation can be established.

7. Establishing the distribution system according to the socialism market economy through the reform of the constitutionalism

The unfair income distribution has been the most serious problem faced by Chinese society, and it seriously influences the stability of the society and the healthy environment of the economic development. From the development experiences of various countries, the early income distribution difference in the high-speed economic development process is inevitable, but the instance like China is very few. At present, the serious inequality of the individual income in China is not the result of the marketization, and it is the result of the popular unfair trades induced by the deficiency of the system in the economic transformation process, and it is also the result of the marketization of the money-power trade. In the transformation process from the planned economy to the market economy, on the one hand, individual hierarch leaves no stone unturned to control, expand and snatch the power of the national resource distribution, and on the other hand, with the quick development of the domestic market economy, many unperfected systems offer sorts of convenient condition for translating power to coin. Under the conditions of the market trade internationalization and the financial capitalization, the money-power trade is largely released and expanded in time and space, and the social wealth flows to the people with power through sorts of money-power trade (whether it is illegal or “legal”), which induces the serious unfair income distribution. So the expanded income difference in present China is mainly induced by the power (including power monopolization, power corruption and power capitalization), i.e. the restriction of the constitutionalism and the standardized power are deficient.

Therefore, the present distribution system should be reformed and the socialism market economic distribution system should be established and perfected. First, the power offside should be limited through the reform of the constitutionalism to reduce that the power participates in the monopolization and competition. Only the power range of the government is reduced to only maintain the public order and offer public production an other public domains, the governmental right can be transparent, and the corruption can be basically eliminated (Zhaoxiao, 2007). Second, the industrial monopolization should be forbidden and the industrial free competition should be guaranteed. Third, the first distribution system should be perfected and the second distribution system should be reformed. Fourth, the distribution order should be standardized to ensure the justice of the distribution process. Fifth, many control measures such as law, economy and information should be perfected, and the macro control system of the income distribution should be established. Sixth, the low income group should be guaranteed and the high income level should be adjusted. Though above six aspects, a real fair society can be established, and the fair income distribution can be realized and the healthy and stable development of the socialism market economy can be really guaranteed.

References

Liu, Futan. (2007). Establishing the Macro Control System According with the National Situation. *Chinese Economic Report.* No.1.

Zhaoxiao. (2007). Key of the Fair Distribution Is to Standardize the Right but not to Execrate the Market. *China Development Observation.* No.1.

Wang, Xiaowei. (2006). *Constitutional Economics.* Shanghai: Lixin Accounting Press.

Xiong, Bijun. (2002). Promoting the Transformation of the Nonofficial System and Perfecting the Socialism Market Economic System. *Socialism Survey.* No.9.

Zhang, Qianfan. (2004). *Constitution, Nomocracy and Economic Development.* Beijing: Beijing University Press.

International Journal of Economics and Finance
Vol. 1, No. 1
February 2009
www.ccsenet.org/journal.html

Customers Satisfaction in Malaysian Islamic Banking

‘Ismah Osman
Faculty of Business Management, Universiti Teknologi MARA
40450, Shah Alam, Selangor, Malaysia
Tel: 60-3-554-4446 E-mail: ismah817@salam.uitm.edu.my

Husniyati Ali
Faculty of Business Management, Universiti Teknologi MARA
40450, Shah Alam, Selangor, Malaysia
Tel: 60-3-554-4446 E-mail: husni833@salam.uitm.edu.my

Anizah Zainuddin
Faculty of Business Management, Universiti Teknologi MARA
40450, Shah Alam, Selangor, Malaysia
Tel: 60-3-554-4446 E-mail: niz@salam.uitm.edu.my

Wan Edura Wan Rashid
Faculty of Office Management and Technology, Universiti Teknologi MARA
40450, Shah Alam, Selangor, Malaysia
Tel:60-3-5543-5549 E-mail: wan_edura@salam.uitm.edu.my

Kamaruzaman Jusoff (Correspondimg author)
Yale University, Yale’s Tropical Forest Institute
370 Prospect St, New Haven, CT 06511-2104, USA
Tel: 203-676-7761 E-mail: jusoff.kamaruzaman@yale.edu

Abstract

The development of the Islamic banks in Malaysia is increasingly challenging with more conventional banks participating and offering banking products and services based on Islamic principles after two decades of its establishment. Thus, it is very important for the players in the Islamic banking industry to better position themselves in order to confront and respond vigorously to the robustness of the global financial environment. This study is embarked on to identify the relative significance assigned to selection criteria in deciding a bank that has an Islamic banking system. More importantly, it would evaluate customer preference of the choices of a bank that is adapted from CARTER model. Secondly, it would investigate service quality and customer satisfaction between the full-pledged Islamic banks and the conventional banks offering Islamic banking products and services. It is hoped that the study would facilitate the management of the banks in developing their marketing strategies which is very important for the survival of the bank specifically due to intense competition arising from the foreign banks.

Keywords: Customer satisfaction, Service quality

1. Introduction

It is now a fact that consumers of all products and services tend to become more demanding due to the current marketplace that are becoming more competitive. As a result, there is a continuous increase in customer expectations and customers’ successive demands within the improvement of the quality of service (Parasuraman *et. al*., 1988). Thus, service organizations should always be able to meet the needs and demands of those sophisticated customers effectively and efficiently in order to retain them, as loyal customers tend to spend more, purchase more and act as enthusiastic advocates for the organizations (Harris & Goode, 2004). In addition, it can generate positive returns to the firm such as increased sales and lower costs as to attract new customers (Griffin, 1995) as well as more predictable profit flow (Meidan, 1996; Ostrowski, O'Brien, & Gordon, 1993), and determines price indifference, favorable word of mouth communication, customer referrals and resistance to competition (Barnes & Howlett, 1998). Eventually, one of the strategies that would enhance customer loyalty is through service quality and it has been related to success in service organizations (Wong & Sohal, 2003) including Islamic banks, which will be a significant indicator to differentiate an organization among the rest of the competitors (Parasuraman et al., 1988).

The establishment of Islamic banks has grown rigorously for the past four decades .It has spread across the whole universe and received wide acceptance by both Muslims and non-Muslims (Iqbal & Molyneux, 2005). This is because Islamic banks play their role similar to conventional banks except that they have to conform to Islamic principles and regulations (Henry & Wilson, 2004). It has been estimated that Islamic banking, which managed around US$250 billion worth of assets worldwide as of 2004, is expected to grow at the rate of 15% per annum (The Banker Magazine, 5th October 2005). Therefore, Islamic banking can no longer regarded as a business organization which is established to fulfill religious duties but what is more important, to be as competitive as possible side by side with the conventional system in alluring more customers and retain them. Inevitably, Islamic banks need to really understand the perceptions of their customers towards their business operations particularly their quality of service rendered to increase customer satisfaction and ultimately their loyalty (Dusuki & Abdullah, 2007).

Evidently, the Islamic banking and finance has contributed to the growth and sustainability of the Malaysian economy since its establishment in 1983. Since then, the strong growth of the Islamic banking in Malaysia is expected to continue due to new players who have joined the bandwagon in relation to the strong demand for Shariah-compliant financial products and services. The total deposits of Islamic banks have grown vigorously at 18.2% to account for 12.2% of banking system deposits. Assets of the Islamic banking system accounted for RM133 billion or 6.4% of total assets in the financial system compared to 5.8% in 2005 (Bank Negara Malaysia, 2006). In time, this figure is expected to increase, gearing towards competitive advantage, with the launching of more innovative products and services and the enhancement of customer centricity to cater to the needs of customers, Muslims and non-Muslims.

Against this backdrop, the present study seeks to identify the selection criteria of the customers patronizing Islamic banks in Malaysia. Specifically, this study attempts to investigate customers' perception towards quality of services based on CARTER (Othman and Owen, 2002). What makes this study different from other studies particularly like Al-Tamimi and Al-Amiri (2003), al-Zaabi (2007) or even El-Din and Shahril (2007) is that this study involves the conventional banks especially the foreign counterparts which are also offering Islamic banking products and services, for example Citibank, HSBC Bank, OCBC Bank, and also the Islamic foreign counterparts which include Kuwait Finance House. Consequently, it determines to clarify the perception of customers who patronize Islamic banks and those who patronize conventional banks in terms of the quality of service rendered and their satisfaction towards the services. Hence, the study sets out to provide answers to the following research questions:

ii) What are the factors that seems to be significant in terms of service quality based on CARTER model (Othman & Owen, 2002)

iii) To what extent do the customers in the Islamic banks are satisfied with the services provided compared to their competitors particularly from the conventional banks?

2. Methodology

The measurement of customer satisfaction in this study is based on CARTER model that has been developed by Othman and Owen (2001). However, prior to the survey, four (4) mini focus groups each consisting of five participants was used to generate a list of attributes of quality service, seen as being of importance to consumers. The finding reveals that the attributes chose by the participants close resemble to CARTER (Othman and Owen, 2001). In addition, it has been successfully scrutinized in Kuwait Finance House in terms of its compliance to Islamic guidelines, assurance, reliability, tangibles, empathy and responsiveness. The reason for adopting this model is because it emphasizes on quality which is important to customers' satisfaction and it is verified in meeting customer's requirements, specifically in terms of its importance feature and that is compliance to Islamic law. The reliability test has also been carried out and the alpha for CARTER dimensions has been high (0.79). A seven-point scale was used with anchors of very strongly disagree and very strongly agree. This scale is appropriate for the study and enabled the identification of the relative significance in selection criteria for choosing a bank which offers Islamic banking products and services.

However, it is important to note that the questions pertaining to service quality in this study does not employ the gap analysis based on Parasuraman *et. al* (1988) SERVQUAL measurement where it measures the perception of customers on service quality against their expectations to the quality of service provided by the organization. The reason being refers to Cronin and Taylor (1992) where they compared the expectation-perception gaps in opposition to perceptions alone, which they called SERVPERF and they decided that the measurement of service quality based on perception alone is adequate.

Furthermore, Othman and Owen (2001) did not measure customer service quality in Kuwait Finance House based on the expectation but they merely adopted the questions based on perceptions. Hence, we are adopting to Othman and Owen (2001) measurement which is based on processes inputs and overall satisfaction outputs.

Self-administered questionnaires is based on those six variables adapted from Othman and Owen (2001), namely, Compliance Assurance, Reliability, Tangibles, Empathy and Responsiveness. The questionnaires were distributed to the 200 adult respondents, however only 141 were used for analysis purposes. There are seventeen (17) selected commercial banks which offer Islamic banking products and services involve in this study. The selection of banks can be further developed through the nature of operation; whether they are a full-pledged Islamic banks or conventional banks having Islamic windows.

3. Results and discussion

From the data it was possible to access how Islamic Financial Institutions was viewed with respect to how satisfied customers were with the services that they have provided. Furthermore, this study discovers certain differences in terms of absolute and relative perceptions of the attributes instill in CARTER model. Results show that Reliability (mean 5.05) is perceived as the most important attributes in choosing the Islamic Financial Institutions (IFIs), followed by Assurances and Tangibles (Table 1). The least important attributes are Empathy, Responsiveness and Compliance, respectively.

In terms of the overall impression of CARTER model, it is interesting to note that Compliance is the only significant attributes that contribute to the criteria for choosing IFIs that offers Islamic product and services in the market (Table 2). However, the result from test statistics revealed that Reliability was ranked second, Tangibles was ranked third, Responsiveness was ranked fourth, Assurance was ranked fifth and finally, Empathy was ranked the least important attribute.

This study supports the research done by Othman and Owen (2001) based on the fact that compliance to Islamic principles is perceived as the most significant criteria. It is interesting to note that Malaysian in general prefer to deal with IFIs as it suits with their religious' values and principles. The study also revealed that choices of the IFIs do not much correspond with IFIs product which is "*offered free interest loan*" as the mean observed in this study is the lowest (4.61) (Table 3).

Finally, in order to compare the service quality between the full-pledged IFIs and the conventional banks offering Islamic products and services, the result has shown that the relative importance of each selection criteria as adopted from CARTER model (Othman and Owen, 2001) are differ from one bank to another depending on the type of institution (fully pledged Islamic bank or conventional bank offers Islamic window). The findings are shown in two tables, namely Table 4 shows the top five (5) rank of financial institutions offerings Islamic banking products and services, while Table 5 shows the bottom five (5) of ranking based on CARTER model.

The findings also show that Bank Islam Malaysia Berhad (BIMB) was ranked top under the compliant attribute. It is not surprising as the bank carries the word *'Islam'* as such it is expected to perform and operating fully under Islamic principles. However, as for assurance, it has been ranked five (5) which are indeed poor, and it implies to the knowledge and the courtesy of the employees and their ability to convey trust and confidence. This indicates that least attention was given by the bank as to ensure that their customers are satisfied with the services provided. As for other attributes such as Assurance, Reliability, Empathy and Responsiveness, the respondents ranked Citibank Berhad as a top provider. This indicates that Citibank Berhad has been doing very well in order to satisfy its customers. The results may also suggest that consumer have a strong satisfaction level towards the conventional bank. Perhaps, this is because customers perceive that the conventional banks, especially the international banks offer better services particularly in terms of quality. This can also be witnessed through customers' ranking of Malayan Banking Berhad. Although it is known as one of the largest commercial and Islamic bank in Malaysia, as it is also offering Islamic window, yet it has been ranked five (5) for Empathy attribute.

Table 5 stipulates that as for the Reliability, Tangibles and Responsiveness, the respondents have ranked the services rendered by BIMB as dissatisfactory. It could be inferred that the BIMB has taken little effort to improve with their services knowing that consumer will still go to their bank because of the important factor that consumer are looking at especially the Muslims, and it would be that the bank must fully adhere to the Islamic principles. As noted by Ariff (1988), if there is only one Islamic bank in a particular country, the complacency factor may set in. That is, the bank tends to behave as though it has capture a larger market especially catering to the Muslim masses since they will patronize an Islamic bank due to religiosity. So what does this tell us? Do the customers really satisfy with the services render by the full-pledged Islamic bank? Can the full-pledge Islamic banks sustain their market positioning since now we have more competitors offering innovative Islamic banking products and services?

4. Conclusion

The researchers feel that in order for the full-pledged Islamic banks to sustain their market positioning, they need to really ensure their service quality as this will entail customer satisfaction and ultimately customer loyalty. They need to remain competitive especially with their international counterparts by offering good quality products and services as well as not just only rely on their "Islamic" image and reputation to draw customers (Muslims and non-Muslims).

References

Al-Tamimi, H. A. H., & Al-Amiri, A. (2003). Analyzing service quality in the UAE Islamic Banks. Journal of Financial Services Marketing, 8(2), 119-132.

al-Zaabi, O. S. H. A. (2007, 23 April - 25 April). Measuring the perceived service quality : an empirical study of Islamic banks in the UAE. Paper presented at the IIUM International Conference of Islamic Banking and Finance, Kuala Lumpur.

Barnes, J., & Howlett, D. (1998). Predictors of equity in relationships between financial service providers and retail customers. International Journal of Bank Marketing, 16(1), 15-23.

Cronin, J. J., & Taylor, S. A. (1992). Measuring service quality : a re-examination and extension. Journal of Marketing, 56(July), 55-68.

Dusuki, A. W., & Abdullah, N. I. (2007). Why do Malaysian customers patronise Islamic banks? International Journal of Bank Marketing, 25(3), 142-160.

El-Din, S. E. I. T., & Shahril, R. A. M. (2007). Analysis of banking service quality perception : a comparative study for Islamic and conventional banks. Paper presented at the IIUM International Conference on Islamic Banking and Finance, Kuala Lumpur.

Griffin, J. (1995). Customer loyalty : how to earn it, how to keep it. San Francisco, USA: Jossey-Bass Publishers.

Harris, L. C., & Goode, M. M. H. (2004). The four levels of loyalty and the pivotal role of trust : a study of online service dynamics. Journal of Retailing, 80, 139-158.

Henry, C. M., & Wilson, R. (2004). The politics of Islamic Finance. Edinburgh: Edinburgh University Press.

Iqbal, M., & Molyneux, P. (2005). Thirty years of Islamic Banking : History, Performance and Prospect. New York, NY: Palgrave Macmillan.

Meidan, A. (1996). Marketing Financial Services. Houndmills, UK: Macmillan Press.

Ostrowski, P. L., O'Brien, T., & Gordon, G. (1993). Service quality and customer loyalty in the commercial airline industry. Journal of Travel Research, 32, 16-24.

Othman, A., & Owen, L. (2001). Adopting and measuring customer service quality (SQ) in Islamic banks : a case study in Kuwait Finance House. International Journal of Islamic Financial Services, 3(1), 1-26.

Othman, A., & Owen, L. (2002). The multidimensionality of CARTER model to measure customer service quality in Islamic banking industry : a study in Kuwait Finance House. International Journal of Islamic Financial Services, 3(1).

Parasuraman, A., Zeithaml, V. A., & Berry, L. L. (1988). SERVQUAL : a multi-item scale or measuring customer perceptions o service quality Journal of Retailing, 64(1), 12-40.

Wong, A., & Sohal, A. (2003). Service quality and customer loyalty perspectives on two levels of retail relationships. Journal of Services Marketing, 17 (5), 495-513.

Table 1. Mean for all items under CARTER Model

CARTER Model	Mean
Compliance	4.8733
Assurance	4.9726
Reliability	5.0548
Tangibles	4.9592
Empathy	4.9490
Responsiveness	4.9218

Table 2. Criteria for choosing a IFIs products and services

Attributes (CARTER)	Compliance	Assurance	Reliability	Tangibles	Empathy	Responsiveness
Chi-square	52.269	14.273	16.970	16.670	12.156	16.333
Df	16	16	16	16	16	16
Asymp. Sig	.000	.578	.388	.407	.733	.430
Rank	1	5	2	3	6	4

Kruskall Wallis Test

Table 3. Compliance Attribute

Dimension	Items	Mean	Standard deviation
compliance	The bank is operated on Islamic principles	4.94	.911
	I prefer this bank because no interest neither paid no taken	4.65	1.051
	I prefer this bank because it offers Islamic products and services	5.24	3.588
	I prefer this bank because it offers free interest loan	4.61	1.040
	I prefer this bank because it offers profit-sharing investment products	4.92	.843
	I prefer this bank because the transactions made are free from any uncertainties (gharar) or doubtful transactions	4.93	.860
	I prefer this bank because investment made does not involve any act of gambling or other acts that are against Shari'ah	4.96	.878
	I prefer this bank because the bank contributes some portion of its profits towards the payment of zakat and sadaqah	4.85	.930
	I prefer this bank because the bank performs its task by means of justice and honesty based on the teachings of the Holy Qur'an and the Sunnah	4.71	1.020

Kruskall Wallis Test

Table 4. Summary of the criterion selection of selected commercial banks.

Banks	**Compliant**	**Assurance**	**Reliable**	**Tangible**	**Empathy**	**Responsive**
BIMB	1	5				
Bank Muamalat	3	4			4	
Commerce TIJARI Bank Berhad	4					5
Kuwait Finance House (Malaysia) berhad	4		2	3		5
RHB Islamic Bank Berhad			4			
Affin Islamic Bank Berhad	2					
EONCAP Islamic Bank Berhad						
AmBank (M) Berhad		3				4
Citibank Berhad		1	1	4	1	1
HSBC Bank Malaysia Berhad	5	2	3	2	2	2
Malayan Banking Berhad					5	3
OCBC Bank(Malaysia) Berhad						
Public Bank Berhad			5	5	3	
Southern Bank Berhad				1		5
Standard Chartered Bank Malaysia Berhad						5

Kruskall Wallis Test.

Note: the lower the number the better the perception .

Table 5. Summary of the criterion selection of selected commercial banks

Banks	**Compliance**	**Assurance**	**Reliability**	**Tangible**	**Empathy**	**Responsive**
BIMB			4	4	2	4
Bank Muamalat						3
Commerce TIJARI Bank Berhad		5	2	2		
Kuwait Finance House (Malaysia) berhad		5				
RHB Islamic Bank Berhad		4				2
Affin Islamic Bank Berhad			2	2	1	
EONCAP Islamic Bank Berhad		5	2	5	4	5
AmBank (M) Berhad			3		3	
Citibank Berhad	5					
HSBC Bank Malaysia Berhad						
Malayan Banking Berhad	2	2		3		
OCBC Bank(Malaysia) Berhad						
Public Bank Berhad	1	3				1
Southern Bank Berhad	4	5	5			
Standard Chartered Bank Malaysia Berhad	4		2	2	5	

Kruskall Wallis Test.

Note: The highest the number the least customer preference.

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Asymmetric Information and Dividend Policy in Emerging Markets: Empirical Evidences from Iran

Hashem Valipor
Science and Research Branch
Azad University, Tehran, Iran
Tel: 98-917-308-6986 E-mail: h.Dvalipour@gmail.com

Vahab Rostami
Science and Research Branch
Azad University, Tehran, Iran
Tel: 98-912-241-6433 E-mail: vahab.rostami@gmail.com

Mahdi Salehi, PhD (corresponding author)
Accounting and Management Department
Zanjan University, Zanjan, Iran
Tel: 98-912-142-5323 E-mail: mahdi_salehi54@yahoo.com

Abstract

Purpose-this study investigates the effect of asymmetric information on dividend policy in listed companies in Tehran Stock Exchange.

Methodology-four main hypothesis have designed by theoretical framework, and tested on 111 listed companies in Tehran stock exchange during 2003 to 2007. The statistic analysis had done by malt- variable regression analysis.

Findings-the study findings show that there is a meaningful and reverse relationship between asymmetric information and dividend policy. Some other findings show there is a meaningful relationship between dividend policy and return on stock but there is no meaningful relationship between dividend policy with firm size and book value to market value of equity ratio.

Practical implementation-the finding of the study could be helpful for university students and users of financial information and other financial analysts in capital market.

Keywords: Asymmetric information, Dividend policy, Capital market, Signalling theory

1. Introduction

Corporate dividend policy has captured the interest of economists of this century and over the recent decades has been the subject of intensive theoretical modeling and empirical examination. Further, several factors may affect to corporate dividend policy, which one of the most important factor is asymmetric information. This paper study this issue that how asymmetric information can has effect on dividend policy by testing the relationship between asymmetric information and dividend policy. Dividend policy has been complicated problem for financial economist. Modigliani & Miller (1961) observed dividend policy has no effect on stock values in efficient markets. However, observations state that stock price can be increased when the suggestion for payment of dividend comes up and the dividend policy is important after other information. Some studies have various suggestions for impalement specific dividend policy. The dividend signal theory is one of the most famous theories between them. Upon The signaling model Bhattacharya (1979), Williams & John (1985), Rock & Miller (1985), managers know more about the real value of the firm than investors and they direct the information in the market for dividend policy. Thus, this model suggests a direct relationship between asymmetric information and dividend. Under the dividend signaling hypothesis, dividend initiations and omissions are generally considered to be important events. When a firm initiates a dividend payment, it creates a benchmark against which investors can clearly measure subsequent performance. This can be viewed in contrast to reported earnings, which are widely regarded as being subject to management is widely regarded as being subject to management. Many researches have done about signalling theory for profit dividing. Maybe this study is the first one in Iran that tests signalling model on asymmetric information and dividend policy. Private information represents information about firm value that has not been incorporated into the firm's stock price. Hasbrouck's (1991) characterization of private information as "essentially prior knowledge of public information" indicates that timing differences account for much of the distinction between private and public information. Information asymmetry in the stock market occurs when one or more investors possess private information about the firm's value while other investors are uninformed. This dichotomy of information among investors is consistent with Admati and Pfleiderer (1988), Diamond and Verrecchia (1991), Easley and O'Hara (1992), Glosten and Milgrom (1985), Kim and Verrecchia (2001), Kyle (1985), and McNichols and Trueman (1994), among others. The level of information asymmetry can be characterized by the risk of trading with a privately-informed investor. A firm's choice of disclosure quality-which it can define as the precision, timeliness, and quantity of information provided-affects this information risk by altering the distribution of public and private information among investors.

Disclosure quality is related to information risk because it affects the incentives to search for private information. Higher disclosure quality reduces the incentives to search for private information by reducing the expected benefits from obtaining private information. Diamond (1985), Hakansson (1977), and Verrecchia (1982) examine settings where public and private information are substitutes for each other. In these settings, increased public disclosures by firms generally reduce the incentives to collect costly private information. These findings suggest that higher disclosure quality leads to less private information being produced. In such a condition asymmetry of information may damage shareholders' confidence about company dividend policy, which more convincing argument in favour of dividends is the signalling hypothesis, which is associated with propositions put forward in Bhattacharya (1979), Miller and Rock (1985), John and Williams (1985), and others.

The signalling hypothesis can explain the preference for dividends over stock repurchases in spite of the tax advantage of the latter. Particularly, as suggested in Jagannathan, Stephens and Weisbach (2000), Guay and Harford (2000) and DeAngelo, DeAngelo and Skinner (2000) among others, the regular dividend signal an ongoing commitment to pay out cash. This signal is consistent with Lintner (1956) observation that managers are typically reluctant to decrease dividend levels. If changes in the levels of dividend release information to the market, then firms can reduce price volatility and influence share prices by paying dividends. However, it is only unexpected changes which have an informative value and which can thus impact prices. Therefore, the value of the signal depends on the level of information asymmetries in the market. For example, in developing countries where capital markets are typically less efficient and where information is not as reliable as in more sophisticated markets, the signalling function of dividend may be more important.

Moreover, it can be argued that information will eventually be revealed whether or not the dividend signal is sent; hence the dividend impact on prices is only temporary. It is based on the idea of information asymmetries between the different participants in the market and in particular between managers and investors. Under such conditions, the costly payment of dividend is used by managers, to signal information about the firm's prospects to the market. For example, in John and Williams' (1985) model the firm may be temporarily under-valued when investors have to meet their liquidity needs. If investors sell their holdings when the firm is undervalued, then there is a wealth transfer from old to new shareholders. However, the firm can save losses to existing shareholders by paying dividends. Although investors pay taxes on the dividends, the benefits from holding on to the undervalued firm more than offset these extra tax costs. A poor quality firm would not mimic the dividend behaviour of an undervalued firm because holding-on to over-valued shares does not increase wealth.

This question is asked to lead the study: Does Asymmetric information has effect on dividend policy? Asymmetric information is a main defect in capital market and taking a specific dividend policy is one of the most important decisions in firm. So, the above question is an important. Resent study uses the dividend prediction mistake and dispersion analysis to measure Asymmetric information between managers and investors. Several factors associated with the firm's information environment to be associated with the level of information asymmetry. The first variable we include is Size. Previous research indicates that stock prices incorporate information about large firms earlier than information about small firms. Based on the results in Atiase (1985), Bamber (1987), and Diamond and Verrecchia (1991), it may expect a negative association to size. Ayers and Freeman (2001) and Jiambalvo, Rajgopal, and Venkatachalam (2002) find evidence that current returns reflect future earnings to a greater extent when institutional ownership is higher. These findings suggest that sophisticated investors are more actively trading on private information relating to future earnings, and current prices thus reflect future earnings information to a greater extent. While Ayers and Freeman (2001) find that analyst following plays a similar role to institutional ownership, Jiambalvo, Rajgopal and Venkatachalam (2002) find that the analyst following is negatively associated with the extent that prices lead future earnings. The analysis in Easley, O'Hare and Paperman (1998) suggests that the role of analysts is more complex since the number of analysts cannot simply be used as a proxy for informed trade. They find that analyst following is positively associated with both the amount of informed and uninformed trading; with the net effect that information asymmetry is negatively associated with analyst following. The amount of private information search activities will be positively associated with the expected benefits of obtaining the information. Boot and Thakor (1993) demonstrate that the incentives for private information acquisition are increasing with a firm's debt-to-assets ratio. This result occurs because, for a given amount of private information about the value of a firm's assets, the expected profits from trading on that information in the equity market increase with the firm's leverage, ceteris paribus. Additionally, Zhang (2001) demonstrates that the endogenously-determined level of private information production increases with the volatility of earnings. When earnings volatility is higher, greater benefits arise from obtaining private information about future earnings.

2. Literature review

The paper studies on this question that how asymmetric information has effect on dividend policy in firms, by testing the relationship between asymmetric information and dividend policy. Dividend has complicated problem for financial statisticians. Modigliani & Miller (1961) proved that dividend policy is unrelated to stock value in efficient market. However Observations state stock price can be increased when the suggestion for distribution of dividend comes up. Some studies have various suggestions for handling profit dividend. The dividend signalling theory is one of the most famous theories between them. The signalling hypothesis is based on the notion of asymmetric information particularly between managers and investors. Under this assumption dividend changes are valuable in that they convey information about the firm's prospects. Indeed, Lintner (1956) observes that managers are more willing to raise rather than reduce dividend levels, and this has been widely interpreted as indicating that dividend decreases are associated with negative signals while dividend increases signal positive news. But what precisely is the nature of the information contained in dividend changes? The risk-information hypothesis claims that dividend increases signal risk reduction. Alternatively, according to the cash flow signalling hypothesis, dividend changes contain information about future cash flows. Another opinion is that dividend changes signal permanent shifts in current earnings. In any event, as noted by Allen and Michaely (1995), regardless of the precise information contained in the dividend signal, hypothesis to be valid.

The "information signaling theory", and suggests dividend reduces the asymmetric information by signaling mechanism. On the base of signaling theory Batacharia (1979), Rock and Miller (1985), Williams and John (1985), John and Lang (1971) say: managers know more about the real value of the firm than investors and they direct the information in the market by profit dividing. Thus, this model suggests a plus relationship between an asymmetric information and dividend policy.

Khang & king (2002) tested the relationship between asymmetric information and gained profit by personals during 1982 to 1995 for public corporation that SEC reported their internal exchanges. The examples showed dividend ascertained by personals profit and asymmetric information. Finding indicated firms with high dividend had lower personals profit and asymmetric information. These results found by "free cash flow theory" and "institutional monitoring theory". Generally, the results of study are: 1) dividend level is an important element for personals profit rather than the decision for dividend payment. 2) Institutional monitoring theory is not a suitable dividend policy to explain the relationship between the dividend level and asymmetric information, but free cash flow theory is the best one to explain dividend changes to assertion asymmetric information. At last there are a few evidences for this topic that dividend changes level can be a sign of asymmetric information reduction (Khang and King, 2002).

Khang & King (2006) showed amount of dividend has a meaningful and negative correlation with efficiency of personals exchanges. Information signal theory doesn't accept the result of this research.

Li & Zhao (2007) researched to figure out how asymmetric information effects on dividend policy. They studied some examples from IBES files during 1983 to 2003. The examples were included the observations of 22413 firms. Measurement was include of profit predict mistake and dispersion of deviations. Last researches showed a positive correlation between predict deviation and asymmetric information. It had been estimated if well use signaling theory for profit dividing, it will be a exact description of reality. Thus, the dividend policy of firms should have a positive and meaningful relationship with mistake and dispersion of profit prediction. The finding states that with permanent of other factors, the firms with more asymmetric information have lower probability for dividend payment, they start to distribution of dividend and increase it, but these firms divide a lower amount of profit for their investors. Likewise, findings show there is a weak and negative relation between recall stock and amount of asymmetric information.

3. Purposes and hypotheses of the study

The main purposes of the study are:

1) Research on the effect of asymmetric information on dividend policy in accepted corporations in Tehran Stock Exchange organization.

2) Ascertain the other effective elements on dividend policies in accepted corporations in Tehran Stock Exchange organization.

The below hypotheses have suggested to arriving to these purpose:

Hypothesis 1: There is a meaningful relationship between dividend policy and asymmetric information in Tehran Stock Exchange organization.

Hypothesis 2: There is a meaningful relationship between dividend policy and equity book value to market value ratio in Tehran Stock Exchange.

Hypothesis 3: There is a meaningful relationship between dividend policy and firm size in Tehran Stock Exchange organization.

Hypothesis 4: There is a meaningful relationship between dividend policy and dividend profitability in Tehran Stock Exchange organization.

3.1 Study variables structures

The relation between dividend policy and asymmetric information will be test by signaling model. Because the asymmetric information effect is quantifiable, the firms with high amount of asymmetric information must have higher dividend. If Signaling theory after controlling all other elements effecting dividend payment still has valid, we can find a positive relationship between asymmetric information and dividend policy. So, signaling theory predicts a strong and positive relationship between asymmetric information.

3.1.1 Measurement of dividend policy

To find out the role of asymmetric information in dividend policy, the study focus on annual profit payment, dividend has higher possibility to have information content. All annual dividends for common stock will be collected after change in amount of issued stock. We have used of DPS/EPS ratio to study about the effect of dividend policy that shows the percent of distributed dividend. In the other word, it shows that the firm distributed how percent of gained profit to investors.

3.1.2 Measurement of asymmetric information

The study uses the predict mistakes and real profit dispersion of firm compared with predicted profit to analyze asymmetric information. Elton, Gruber and Gultekin (1984) showed the largeness of predict mistakes are related to mistake of special factors of the firm. Their finding suggest: predict mistakes are acceptable approximation for amount of asymmetric information.

On the base of research by Li and Zhao (2007), in this study we use predicted profit standard deviation to real profit as a quantity criterion for asymmetric information measurement. The profit predicted mistake is the exact difference between the average predicted profit and real profit. Thus, we need firm that have information about real profit, predicted profit and dividend for examples.

3.1.3 The other effective variable on firm dividend policy

Li and Zhao (2007) to study on asymmetric information also entered the profitability, size and potential growth that have effect on dividend. Current research also for control other effective variables on dividend policy entered size, potential growth (BV/MV ratio) and profitability variables on its model. Forgoing variables are used in research in following manner:

3.1.3.1 Book value to Market value of equity ratio (BE/ME)

At the first the equity and common stock assertion by the information of the last balance sheet, then we calculate the market value of equity ratio by multiplies the last price of common stock in number of issued stock; after that book value of equity divided to its market value to earn this ratio.

3.1.3.2 Firm size

There are various methods to assertion the firm size. For example, someone uses total assets of the firm. Some times the sell value of the firms is used for firm’s size. Like Fama and Franch (2001) in this research, the market values of equity are utilizing as the firm size and it’s calculated by multiple the last market price of stock in number of common stock. Because the firm size is very bigger than other variables, we use a logarithm of market equity ratio to bring it near to the other variable size.

3.1.3.3 Profitability

Profitability is being calculated by dividing the common Stock net income for the year ended to "t" to book value of common stock.

3.2 Statistical tool and sample

We use the financial statements information of the accepted corporation in Tehran stock exchange organization during 2003 to 2007 to test the relationship between dividend policy and asymmetric information. The bankrupt firms are removed from the sample; also the firms with out EPS prediction for these years are removed from the sample. The data of 108 firms are used in the study for analysis.

3.3 Methodology

Because the study tries to find the relation between the asymmetric information and dividend policy variables, so Methodology include a type of correlation and uses a multivariable regression model to analyze the data. The regression model of the study is:

Dividend policy = b0 +b1 AI +b2 firm size +b3 BV/MV +b4 profitability

That AI= Asymmetric Informatio

Because for determining dividend policy was used of the percent relation of the dividend policy, therefore the model is released in following manner:

DPS/EPS = b0 +b1 AI +b2 firm size +b3 BV/MV +b4 profitability

3.3.1 Testing of hypotheses

With due attention to the above results, test of each hypothesis has been done separately.

Test of hypothesis 1

H0: β1 = 0
H1: β1 ≠ 0

H1 hypothesis explain that there is a meaningful relationship between dividend policy and asymmetric information.

As the result in table 2, H hypothesis reject on the 0.95 meaningful level (α>p-value, .05> 0.00). In the other hand, there is a meaningful relationship between dividend policy and asymmetric information. Also there is a reverse relationship between dividend policy and asymmetric coefficient with due attention to negative of asymmetric coefficient. Beta standardized coefficient is -0 .431 and shows that 0.431 of change in dividend policy deviation explain by change in asymmetric information.

Test of hypothesis 2

H0: β2 = 0
H2: β2 ≠ 0

H2 hypothesis explains there is a meaningful relationship between dividend policy and BV/MV.

As the results in table 2, H hypothesis cannot be rejected on .95 meaningful level (α > p-value, .05> 0.642). In the other hand, there is no relationship between dividend policy and BV/MV.

Test of hypothesis 3

H_0: $\beta_3 = 0$
H_3: $\beta_3 \neq 0$

H3 hypothesis explains, there is no a meaningful relationship between dividend policy and firm size.

As the results in table 2, H hypothesis cannot be rejected on 0.95 meaningful level (α < p-value, .05< 0.801). In the other hand, there is no relationship between firm size and dividend policy.

Test of hypothesis 4

H0: β4 = 0
H4: β4 ≠ 0

H4 hypothesis explains, there is a meaningful relationship between dividend policy and stock profitability.

As the results in table 5, H hypothesis rejected on 0.95 meaningful level (α >p-value, .05> 0.00). In the other hand, there is a meaningful relationship between dividend policy and stock profitability. This relationship is direct with due attention to positive amount of profitability coefficient. Standardized beta coefficient is 0.585 and shows that 0.585 of change in dividend policy deviation explained by change in stock profitability.

4. Finding of the research

The study hypotheses test has done by using multivariable regression analysis. Before the testing the study hypotheses, regression hypotheses related test has done. These hypotheses include:

a) The random variable of mistakes has normal distribution

b) The mistake variances are equal

c) The mistakes are independent

The test showed that above hypotheses are confirmed. Results of the statistic hypotheses tests are in the next table.

Insert Table 1

Table 1 shows the results of "stepwise static hypothesis test". The best regression model and independent variable that use in regression model can be determined by this test. The results show that just two variables- asymmetric information and stock profitability- are effectible factors on dividend policy that can enter to the model.

Insert Table 2

The results in table 2 explain the cause of expel the book value to market value of equity ratio and firm size variable. In BV/MV case with due attention to the "t" parameter –.466 and significant level was 0.642, the beta coefficient for this variable is not meaningful. Thus, it removed from model.

Insert Table 3

Table 3 shows the correlation after removing the variables that has increased. Durbin Watson parameter is 2086 which shows the model works better with two remaining variables.

Insert Table 4

The results of table 4 are about the test of meaningfully of all the independent variable coefficients. Results show that variable coefficients are meaningful on level of confidence .99, before and after removing variables.

Insert Table 5

Table5 shows the final results of regression model. The above table shows two models of the final model. With due attention to the above results, the final regression equation is: Dividend policy =.620 - .000075 AI +.119 profitability.

5. Conclusion

The study is about the effect of asymmetric information on dividend policy. Analysis based on signaling model. This model explains managers know more about the real value of the firm than investors and they direct the information in the market by profit dividing. Book equity to market equity ratio (BE/ME) variable, firm size variable and stock profitability variable are the control variables. Results of study explain there is a reverse and meaningful relationship between dividend policy and asymmetric information. It mean, increasing the asymmetric information reduce the dividend between investors. Also the contrary of this relationship is right. The results are compatible with signaling model. The results of last studies by Batcheria (1979), Rock, Miller, William and John (1985), John and Long (1991), Khang and King (2002) explain there is a meaningful and positive relationship between dividend policy and asymmetric information, and results of this study has not conformity with them. But it's conformity with study of Lee and Jaho (2007). With due attention to the relationship between dividend policy and asymmetric information, there is a probability that managers and subordinates effect on the policies to increases their profit. Other finding explain that in control variables that have a probability of effect on dividend policies just stock profitability has affect on dividend policies. The firm size variable and BE/ME variable have no affect on profit dividing for common stock holders.

6. Suggestions

With due attention to results, asymmetric information is very important for financial analysts and investors; because there is a probability that managers know more about the real value of the firm than investors and direct the information in the market by profit dividing. If it be a real issue, it could be harmful for investors out of firm.

7. Agenda for future research

Because of the results of this study, and the researches of other countries have shown that there is a relationship between dividend policies and asymmetric information, so we can search on the relationship between asymmetric information in firm and managers and subordinates profit in stock markets in next researches.

References

Admati, A. R. & P. Pfleiderer P. (1988). Selling and Trading on Information in Financial Markets, *American Economic Review*, 78, 2, 96-103 196.

Allen, F., Bernardo, A. & Welch, I. (2000). A Theory of Dividends Based on Tax Clientele, *Journal of Finance* 55, 2499–2536.

Atiase, R. K. (1985). Pre-disclosure Information, Firm Capitalization, and Security Price Behavior around Earnings Announcements, *Journal of Accounting Research*, 23, 1, 21-36.

Ayers, B. C. & Freeman, R. N. (2001), *Evidence That Price Leads of Earnings Increase with Analyst Following and*

Institutional Ownership, University of Georgia Working Paper.

Bamber, L. S. (1987). Unexpected Earnings, Firm Size, and Trading Volume around Quarterly Earnings Announcements, *The Accounting Review*, 62, 3, 510-532.

Bhattacharya, S. (1979). Imperfect Information, Dividend Policy, and "the Bird in the Hand" Fallacy, *Bell Journal of Economics,* Vol. 10, 259-270.

Boot, A. and Thakor, A. (1993). Security Design, *Journal of Finance*, Vol. 48, No. 4, 1349-1378.

Diamond, D. W. & Verrecchia, R. E. (1991). Disclosure, Liquidity, and the Cost of Capital, *Journal of Finance,* Vol. 46, No. 4, 1325-1360.

Diamond, D. W. (1985). Optimal Release of Information by Firms, *Journal of Finance*, Vol. 40, No. 4, 828- 862.

Easley, D. & O'Hara, M. (1992). Time and the Process of Security Price Adjustment, *Journal of Finance,* Vol. 47, No. 2, P. 29.

Easterbrook, Frank H. (1984). Two Agency-cost Explanations of Dividends, *American Economic Review,* Vol. 74, 650-659.

Elton, E.J., Gruber, M.G. & Gultekin, M.N. (1984). Professional Expectations: Accuracy and Diagnosis of Errors, *Journal of Financial and Quantitative Analysis*, Vol. 19, 351-363.

Fama, E.F. & French, K.R. (2001). Disappearing dividends: changing firm characteristics or lower propensity to pay? *Journal of Financial Economics*, Vol.60, No. 1, pp.3-43.

Finnerty, Joseph. (1976). Insiders and market efficiency, *Journal of Finance*, Vol. 31, pp. 1141-1148.Gompers, Paul A., and Andrew Metrick, 1998, Institutional investors and equity prices, working paper 6723, National Bureau of Economic Research.

Glosten, L. & Milgrom, P. (1985). Bid, Ask, and Transaction Prices in a Specialist Market with Heterogeneously Informed Agents, *Journal of Financial Economics*, Vol.14, pp. 71-100.

Hakansson, N. (1977). Interim Disclosure and Public Forecasts: An Economic Analysis and Framework for Choice, *Accounting Review*, Vol. 52, 396-416.

Hasbrouck, J. (1991). Measuring the Information Content of Stock Trades, *Journal of Finance*, Vol. 46, No. 1, p. 29.

Jaffe, Jeffrey. (1974). Special information and insider trading, *Journal of Business*, Vol. 47, pp. 410-428.

Jeng, Leslie, Andrew Metrick, & Joseph Zeckhauser, (1999). The profits to insider trading: a performance-evaluation perspective, working paper 6913, National Bureau of Research.

Jensen, Michael. (1986). Agency costs of free cash flows, corporate finance, and takeovers, *American Economic Review,* Vol. 76,pp. 323-329.

Jiambalvo, J., Rajgopal S. & Venkatachalam, M. (2002). Institutional Ownership and the Extent to Which Stock Prices Reflect Future Earnings, *Contemporary Accounting Research*, Vol.19, No. 1, pp. 117-145.

John, K. & Williams, J. (1985). Dividends, Dilution, and Taxes: A Signaling Equilibrium, *Journal of Finance,* Vol. 40, pp. 1053-1070.

John, Kose, & Larry Lang. (1991. Insider trading around dividend announcements: Theory and Evidence, *Journal of Finance*, Vol. 46, pp. 1361-1389.

Khang, K. & King, T. H (2006), Does Dividend Policy Relate to Cross-Sectional Variation in Information Asymmetry? Evidence from Returns to Insider Trades, *Financial Management,* Vol. 35, pp. 71-94.

Khang, Kenneth & Tao-Hsien Dolly King. (2002). Is Dividend Policy Relate to Information Asymmetry? Evidence from Returns to Insider Trades, *Financial Management*, Vol. 35, pp. 71-94.

Kim, O. & Verrecchia, R. E. (2001). The Relation among Disclosure, Returns, and Trading Volume Information, *Accounting Review*, Vol. 76, No. 4, pp. 633-654.

Kyle, A., (1985). Continuous Auctions and Insider Trading, *Econometrica,* Vol. 53, pp. 1315-1335.

Lakonishok, Josef & Inmoo Lee, (2001). Are insider trades informative? *Review of Financial Studies*, Vol. 14, pp. 79-111.

Li, Kai & Xinlei Zhao, (2007). Asymmetric Information and Dividend Policy, *Journal of Financial Management,* forthcoming, available at: http://ssrn.com.

McNichols, M.& Trueman, B. (1994). Public Disclosure, Private Information Collection, and Short-Term Trading, *Journal of Accounting and Economics*, Vol. 17, No. 1,2, pp. 69-94.

Miller, M. & Modigliani, F. (1961). Dividend Policy, Growth and the Valuation of Shares, *Journal of Business,* Vol. 34, pp. 411-433.

Miller, M. & Rock, K. (1985). Dividend Policy under Asymmetric Information, *Journal of Finance*, Vol. 40, pp. 1031-1051.

Seyhun, Nejat. (1986). Insiders' profits, costs of trading, and market efficiency, *Journal of Financial Economics,* Vol. 47, pp. 410-428.

Verrecchia, R. E. (1982). The Use of Mathematical Models in Financial Accounting/Discussions, *Journal of Accounting Research*, Vol. 20, p.55.

Zhang, G. (2001), Private Information Production, Public Disclosure, and the Cost of Capital Theory and Implications, *Contemporary Accounting Research*, Vol. 18, No. 2, pp. 363-384.

Table 1.

Variables Entered/Removed[a]

Model	Variables Entered	Variables Removed	Method
1	X4_PROF	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).
2	X1_AI	.	Stepwise (Criteria: Probability-of-F-to-enter <= .050, Probability-of-F-to-remove >= .100).

a. Dependent Variable: Y_PD

Table 2.

Excluded Variables[c]

Model		Beta In	t	Sig.	Partial Correlation	Collinearity Statistics
						Tolerance
1	X1_AI	-.431[a]	-7.393	.000	-.332	.542
	X2_BV_MV	-.024[a]	-.466	.642	-.022	.807
	X3_FS	-.012[a]	-.253	.801	-.012	.986
2	X2_BV_MV	.075[b]	1.508	.132	.072	.751
	X3_FS	-.062[b]	-1.419	.157	-.067	.963

a. Predictors in the Model: (Constant), X4_PROF

b. Predictors in the Model: (Constant), X4_PROF, X1_AI

c. Dependent Variable: Y_PD

Table 3.

Model Summary[c]

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate	Durbin-Watson
1	.293[a]	.086	.084	.27544	
2	.432[b]	.187	.183	.26011	2.086

a. Predictors: (Constant), X4_PROF

b. Predictors: (Constant), X4_PROF, X1_AI

c. Dependent Variable: Y_PD

Table 4.

ANOVA[c]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	3.146	1	3.146	41.468	.000[a]
	Residual	33.534	442	.076		
	Total	36.680	443			
2	Regression	6.844	2	3.422	50.577	.000[b]
	Residual	29.837	441	.068		
	Total	36.680	443			

a. Predictors: (Constant), X4_PROF

b. Predictors: (Constant), X4_PROF, X1_AI

c. Dependent Variable: Y_PD

Table 5.

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.655	.017		39.331	.000
	X4_PROF	.059	.009	.293	6.440	.000
2	(Constant)	.620	.016		37.798	.000
	X4_PROF	.119	.012	.585	10.023	.000
	X1_AI	-7.50E-05	.000	-.431	-7.393	.000

a. Dependent Variable: Y_PD

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Empirical Forecasting Practices of a British University

Rozainun Abdul Aziz
Accounting Research Institute and Faculty of Accountancy
Universiti Teknologi MARA, 40450,
Shah Alam, Selangor,
Malaysia
Tel: 60-3-5544-4747 E-mail: rozainun@salam.uitm.edu.my

D.F. Percy
Salford Business School University of Salford
Greater Manchester M5 4WT,
United Kingdom
Tel: 44-161-295-4710 E-mail: d.f.percy@salford.ac.uk

Kamaruzaman Jusoff (Corresponding author)
Yale University
Centre for Earth Observation-Environmental Science Centre
21 Sachem St, New Haven, CT 06511, USA
Tel: 203-676-7761 E-mail: jusoff.kamaruzaman@yale.edu

The research is financed by University Teknologi MARA and University of Salford (Sponsoring information)

Abstract

This article is based on a single case study aimed at examining behavioral issues of forecasting, in particular the role and practice of forecasting in a British university settings. Key variables were identified in establishing associations between the variables that provide suitable criteria for the purpose of this study. Data collection was based on questionnaires distributed to people involved and interviews which were held with prominent staff of the University. Fisher-exact tests were performed to identify significant associations between variables. Results indicated the various levels of perceptions and practices of forecasting produced by the people involved at the University. The study implies that useful insights can be gathered through forecasting from a different perspective of the non-profit making service industry.

Keywords: Empirical, British university, Forecasting, Questionnaire survey, Interviews

1. Introduction

Forecasting relates to estimating the future in an unknown situation and a *forecast* is the estimate of an unknown value (Armstrong, 2001). Our context relates to predicting actual values of quantitative results in the future to meet organization's objectives. The forecasting archive tends to concentrate on presenting forecasting methods and on strengthening those methods. Indeed, the literature relating to forecasting procedures hardly covers any studies of universities. Jones *et al* (1997) suggested that it is now more imperative to turn towards examining other issues of forecasting besides the forecasting methods.

Diamantopoulos (2001) commented that since few studies have actually looked at the role of forecasting (Hughes, 2001), more emphasis should be placed on linkages with organizational characteristics, which supports this study. Gardner (1991) encouraged studies into how forecasting is accomplished in organizations. Dawes *et al* (1994) pointed out that model implementation to promote better understanding of the practical forecasting process, in a university in our case, will lead to a 'more careful crafting of the tool'. Though budgets are allocated, it is imperative that the limited budgets are utilized with care. In supporting this, universities via their faculties and schools must forecast their future incomes and expenditures appropriately. Wright and Mechling (2002) indicated in their study that forecasting is important for service organizations. It is a well-known fact that forecasting is a vital element of planning activities for any type of organization (Wheelwright et. al., 1998). Therefore, the purpose of this paper is to highlight the role of forecasting in a British university, especially in the UK which is mainly funded by the Higher Education Funding Council for England (HEFCE).

2. Method and materials

A single case study approach, as recommended by Yin (1994), is employed which offers a rich description of current forecasting practice in developing a greater understanding of the discipline in practice (Moon and Fitzgerald, 1996). The techniques of an exploratory case study (Winklhofer & Diamantopoulos, 1996; Winklhofer et al., 1996; Miles & Huberman, 1994) are adopted owing to the fact that studies relating to forecasting in a university setting are under-explored. Whilst the university plans for its future activities, one of the key tasks involved is forecasting. It is necessary for all faculties, schools, departments and other units to prepare forecasts to allocate budgets to ensure smooth running of the activities concerned. The role of forecasting at the university can be viewed in many different ways depending on the purpose of forecasting. This paper demonstrates the formulation of models to support such combinations and results.

Respondents were interviewed at the University of Salford and they vary in terms of designation and size of the units which they are attached to, with employees ranging between under 50 to more than 200 and with annual income from under £5 million to about £10 million. In general, the university makes forecasts just like any other established firms in industry, and this is as a vital exercise in the organization. Formal methods that are suitable for the university environment are adopted but tend to be less sophisticated than those which other business firms use. The law of parsimony or Occam's Razor was applied in that the model includes only required and important variables and does not include all reasonable predictor variables automatically. Differential equations were introduced to render a complete modeling aspect.

3. Results and discussion

3.1 The forecasting parameters and process

At the university, one of the main reasons for forecasting is for budget allocation. Another reason is as a guidance of what to expect in carrying out activities in the future. The forecasts are used as a check and guide against the activities that will be carried out. Resources involved are negotiated and discussed at various levels to achieve the units' objectives. One more reason for preparing forecasts is for users who are the senior management, heads of school, faculty accountants and administrative officers.

The main forecasts are student numbers, human resources, income and expenditure and budget balancing forecasts. Forecasts are prepared on time before implementation with various levels of frequency, depending on the importance of forecasts to be prepared, the urgency of preparing and reviewing based on a time schedule for meetings and feedback between various levels within the units and the University as a whole. Formal methods are used but only to suit the units' needs in line with fulfilling the University's overall objectives. Discussions and meetings are held to impart feedback in the course of reviewing and revising the forecasts throughout a given period before finalizing them. A simple presentation of the unit's plans for its activities and resource requirements is given in statements and reports such as receipts and payments accounts, and income and expenditure accounts. These statements are brought together into one main plan for the university.

There are two main distinctions in the people involved, namely preparers and users. It can be seen that the preparers need input from users in order to produce the forecasts. Looking at the responses from the questionnaires and interviews, respondents are not too sure whether they are either preparers or users or both. In general, there are mixed feelings amongst the forecast preparers and users towards the forecasting function as a whole. However, they try to put their best efforts into forecasting and produce results even though constrained by various resources.

3.2 Modeling the university forecasting scenario

The number of students is large when tuition fees are low since many people can afford low fees. On the other hand, when tuition fees are high student numbers will be compromised. There are many possible functions that might be suitable for modelling the association between tuition fees and numbers of students, two of which are illustrated in Figure 1 subject to arbitrary scaling on both axes. The mechanism of this relationship is explained using a differential equation as developed below.

Let y = 'number of students enrolling'

x= 'tuition fee payable per student'

If *x* is small, then an increase in *x* will reduce *y* substantially, so $\frac{dy}{dx}$ is large (though finite) and negative. Conversely, when *x* is large, then a change in *x* will hardly affect *y*, so $\frac{dy}{dx}$ is small and negative, such that $\frac{dy}{dx} \to 0$ as $x \to \infty$. Moreover, it is reasonable to suppose that $\frac{dy}{dx}$ is a concave function of *x* for simplicity, which implies that

$$\frac{d^3 y}{dx^3} < 0 \tag{1}$$

for all $x > 0$ in this context.

Restricting likely functions to just two parameters for reasons of parsimony, there are two obvious candidates for functional representations of the first derivative (rate of change of enrolment with respect to fee payable):

$$\text{(i)} \quad \frac{dy}{dx} = \frac{-a}{1+bx} \qquad \text{for a, b} > 0 \tag{2}$$

$$\text{(ii)} \quad \frac{dy}{dx} = -ae^{-bx} \qquad \text{for a, b} > 0 \tag{3}$$

(so that *dy/dx* = -*a* when *x* = 0 for both (i) and (ii)). The second of these corresponds to f(*x*) in Figure 1.

4. Conclusions

A clearer perspective of empirical evidence cultured with mathematics and functional relationships was able to be developed where the outcomes of improved teamwork and decision making were established. These differential equations are applicable in many aspects of associating variables and changes in behavior patterns for related variables in forecasting.

References

Armstrong, J. (2001). Principles of Forecasting: A Handbook for Researchers and Practitioners, Kluwer Academic Publishers.

Dawes, R., R. Fildes, M. Lawrence & K. Ord. (1994). The Past and Future of Forecasting Research. International Journal of Forecasting (10): 151-159.

Diamantopolous, A. (2001).Research in Forecasting. International Journal of Forecasting 18(3): 479-480

Gardner, S. (1991). The State of the Institute, International Institute of Forecasters Newsletters, 4: 2-5.

Hughes, M. C. (2001). Forecasting Practice: Organisational Issues. Journal of Operational Research Society 52: 143-149.

Jones, V. S., S.I. Bretschneider, & W.L. Gorr (1997). Organisational Pressures on Forecast Evaluation: Managerial, Political and Procedural Influences. Journal of Forecasting 16: 241-254.

Miles, M. B. & A. M. Huberman (1994). Qualitative Data Analysis, Sage Publications, U.S.A.

Moon, P. & L. Fitzgerald (1996). Delivering Goods at TNT: The Role of the Performance Measurement System. Management Accounting Research 7: 431-457.

Wheelwright, S.C. & D.G. Clarke (1976), Corporate Forecasting: Promise and Reality, Harvard Business Review, November-December: 40-198

Winklhofer, H. & A. Diamantopoulos (1996). "First Insights into Export Sales Forecasting Practice: A Qualitative Study. International Marketing Review 13(4): 52-81.

Winklhofer, H., A. Diamantopoulos & S.F. Witt (1996). Forecasting Practice: a Review of the Empirical Literature and Agenda for Future Research, International Journal of Forecasting, 12: 193-221.

Wright, C.M. & G. Mechling (2002), The Importance of Operations Management Problems in Service Organisations, OMEGA – The International Journal of Management Science, 30: 77-87.

Yin, R.K. (1994), Case Study Research: Design and Methods, Sage Publications, Thousand Oaks, third edition, California, USA.

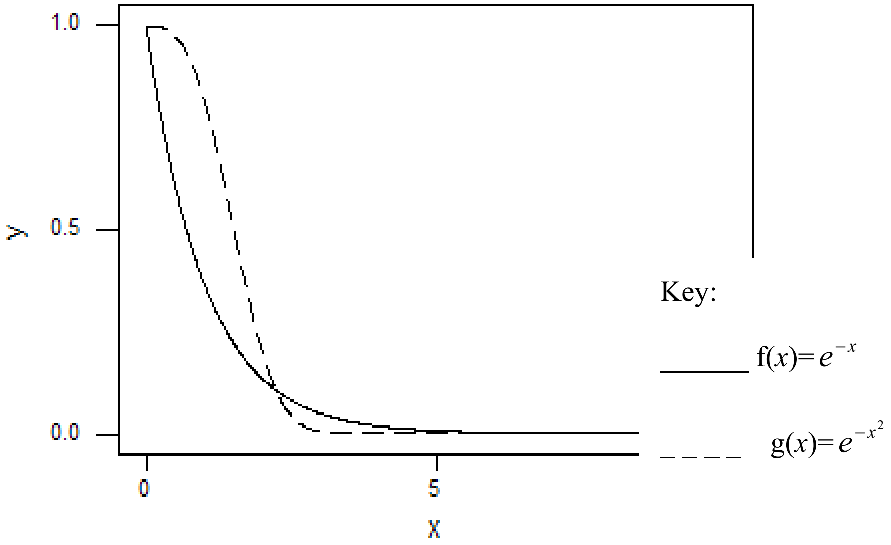

Figure 1**.** Possible models for dependency of y on x.

International Journal of Economics and Finance
www.ccsenet.org/journal.html
Vol. 1, No. 1
February 2009

Tacit Knowledge Sharing and SMEs' Organizational Performance

Rohana Ngah
Faculty of Business Management
Universiti Teknologi MARA
ohor Campus, KM 12, Jalan Muar, 85200 Segamat, Johor DT, Malaysia.
Tel: 60-16-206-5765 E-mail: rohanangah@johor.uitm.edu.my

Kamaruzaman Jusoff (Corresponding author)
Yale University
School of Forestry & Environmental Studies
205 Prospect St, New Haven, CT 06511-2104, USA
Tel: 203-676-7761 E-mail: jusoff.kamaruzaman@yale.edu

Abstract

Tacit knowledge (TK) exists in every corner of SME, in its structure and relationship especially in its people. Even though knowledge management will benefit SME as they are rich of tacit knowledge but lacking in expertise, financial capital and infrastructure, knowledge management (KM) is almost impossible for Small and Medium Enterprise (SME) currently. Since SME have efficient and informal communication network, knowledge sharing is essential especially in an organization with scarce resources. Therefore, in the wake of knowledge-based economy, tacit knowledge sharing is the best tool for SME in enhancing competence and organizational performance which suit its needs and background. Tacit knowledge sharing is still at infancy especially in SME.

Keywords: Tacit knowledge sharing, Organizational performance, Small and medium enterprises

1. Introduction

More than 99% of total business establishments in Malaysia are from SME (SME Annual Report 2006). SMEs have different structure and they develop different pattern of behavior especially in adapting to the environment (Hashim, 2007). Even though, SMEs operate in almost all industries, they are different in scope and importance (Hashim, 2007). This difference varies from industry to industry and also from country to country. In Malaysia, SMEs are largely found in four broad economic segments, namely, manufacturing, services, basic materials producers and Agriculture (Hashim, 2007). Even though SME are considered a backbone of industrial development in Malaysia (Ndubisi & Salleh, 2006), SME contribution to the economy is still relatively small (SME Annual Report 2006). As Malaysian government has made development of SMEs a high priority area, full supports is given to SME to enhance their performance in Ninth Malaysia Plan and Third Industrial Master plan which offers a platform for financial capital, training, information, policies and strategies. In 2004, National SME Development Council, which consist of 10 members, has been set up to further develop and sustain the growth of the SMEs sector in the national economy (Hashim, 2007). The efforts are to improving the infrastructure for developing SMEs, building the capacity and capability of SMEs and enhancing SMEs access to financing.

Recently, researchers and practitioners alike have been increasingly advocating knowledge sharing as essential process of the knowledge economy. Knowledge sharing would lead to knowledge accumulation, knowledge dissemination and knowledge acquisition. While there are debates on the importance between explicit and tacit knowledge, sharing this knowledge is ultimately very crucial in the organization. However, the last few decades, there has been an increasing interest in the tacit dimension of knowledge (Pathirage et al., 2007) but yet still considered to be relatively unexplored and not fully understood (Zack, 1999) compared to work on explicit knowledge (Leonard & Sensiper, 1998, Holtshouse, 1998, Pathirage et al., 2007). Nonaka & Takeuchi (1998) argue that tacit knowledge is the one that create competitive advantage for the companies. The nature of knowledge in SMEs is almost all in tacit nature (Cohen & Kaimnekais, 2007). Nonaka & Takeuchi (1995) suggest that the flow from tacit to explicit knowledge would benefit the organization. According to Wong & Radcliffe (2000), tacit knowledge in SMEs is difficult to turn into explicit therefore sharing this tacit knowledge is the best to help improving SMEs' performance particularly when tacit knowledge sharing would takes place during socialization and in an informal ways where SME have efficient and informal communication network (Egbu et al, 2005). The cost of knowledge management either IT or human resources are costly (Choi & Lee, 2003; Du et al, 2005). Therefore most of small and medium-sized firms cannot afford to have or adopt knowledge management system (McAdam & McCreedy, 1999; Kuan and Aspinwall, 2004). However, knowledge sharing cost is bearable for SMEs (Du et al, 2003). Tacit knowledge exists profusely in SMEs and tacit knowledge sharing is prevalent in SMEs through face-to-face meeting and brainstorming. As SMEs is superior in internal networking and efficient in communication, knowledge sharing is very rampant (Desouza & Awazu, 2006). Knowledge is the means for SMEs to overcome poor business environment and change the complex business environment to be manageable (Kim, 2002). The purpose of this paper is to highlights the gap of tacit knowledge sharing in Small and Medium Entreprises.

2. Nature of tacit knowledge

Knowledge lies in human minds and exists only if there is a human mind to the knowing (Widen-Wulff & Suomi, 2007). There are three dimensions of knowledge: width, depth and tacitness (Nooteboom, 1993). Knowledge can be created by intentional and resource-consuming efforts (Du et al, 2007). The neglect of the tacit knowledge based on people and ideas has undoubtedly reduced the corporate market place's capability for true innovation and sustainable competitiveness (Gamble and Blackwell, 2001). Knowledge is a resource or stock in the organization (Coulson-Thomas, 2003) that exist along the organization. Knowledge is valuable, rare, not easy to substitute and difficult to imitate, making it a competitive advantage for the organization. A firm cannot have the same knowledge as the others' because different firms have different human and knowledge resources (Kim, 2002). No organization can have a total same knowledge. Tacit knowledge is one of strength that organization has that is more difficult to transfer or copy (Nooteboom, 1993). In the resource-based theory, Hisrich et al (2008) further highlight that in order for firms to create its unique resources which are rare, valuable and non-immitigable; it has to exploit its internal knowledge. Thorpe et al (2005) stress that performance in SME is related to the ability of entrepreneurs and managers in integrate their knowledge of markets such as customers, suppliers, and stakeholders. Most of literature on knowledge in SMEs adopts Barney's RBT (1991).

Knowledge is classified into two types as tacit and explicit by Polanyi (1966, p.135 – 146). Explicit knowledge is the type of knowledge that can be easily documented and shaped (Choi and Lee, 2003). It can be created, written down, transferred and followed among the organizational units verbally or through computer programs, patents, diagrams and information technologies (Keskin, 2005; Choi & Lee, 2003). Tacit knowledge is what embedded in the mind (Choi & Lee, 2003), can be expressed through ability applications; is transferred in form of learning by doing and learning by watching. Based on Polanyi (1966), all knowledge has tacit dimensions. It can be completely tacit, semiconscious or unconscious knowledge held in people's heads and bodies (Leonard and Sensiper 1998) Tacit knowledge can be classified into two dimensions: technical and cognitive (Pathirage et al, 2007). Technical encompasses information and expertise in relation to "know-how" while cognitive consists of mental models, beliefs and values. Tacit knowing embodied in physical skills resides in the body's muscles, nerves and reflexes and is learned through practice. And tacit knowledge also embodied in cognitive skills (Leonard & Sensiper 1998). While explicit ready to be explored, tacit knowledge is difficult to be extracted without the consent of the knowledge owner. Tacit knowledge and explicit knowledge complete each other and they are important components of knowledge management approaches in organizations (Beijerse, 1999). Keskin (2005) found that tacit knowledge, explicit knowledge and performance are closely related especially when the external environment is hostile. The great virtue of tacit knowledge is that it is automatic, requiring little or no time or thought (Stewart, 2000). Stewart (2000) holistic emphasize that tacit knowledge tends to be local as well as stubborn because it not found in manuals, books, databases or files. It is oral. It is created and shared around water cooler or over coffee break. Tacit knowledge spreads when people meet and tell stories. As tacit knowledge remains hidden, unspoken and tacit, this knowledge can either knowledge embodied in people and social networks or knowledge embedded in the processes and products that people create (Horvath, 2007). Effective KM requires a symbiosis between explicit and tacit knowledge in line with technology and human resource processes (Choi and Lee, 2003). Almost all studies of knowledge management emphasize on knowledge flow.

Knowledge must be continuously circle and flow in the organization. As long there is a stock of knowledge, during any period of time, there is a flow of knowledge (Stewart, 2000). Making knowledge available to others and capturing a new knowledge as well has been described by Nonaka (1991) as spiral of knowledge. Nonaka & Takeuchi (1995) examine the concept in terms of a knowledge spiral encompassing four basic patterns of interaction between tacit and explicit knowledge: socialization, externalization, combination and internalization. There is no specific definition of this spiral knowledge but the main idea of the spiral is the sharing of knowledge of an individual and share it with others and eventually acquire a new knowledge which is simply knowledge sharing. According to Nonaka & Takeuchi (1995), tacit knowledge need to be convert into explicit in order to make it useful. And the process is through socialization and informal. However, Leonard & Sensiper (1998) pointed out that knowledge not necessarily needs to be explicit in utilizing it. Knowledge can remained tacit but collective tacit knowledge to lead to creativity and innovation (Leonard & Sensiper 1998).

3. Tacit knowledge sharing & organization performance

Not much literature review or company includes knowledge sharing as part of its key components as knowledge sharing is considered as difficult to measure (Christensen, 2007). But the bottom line is knowledge sharing is critical to a firm's success (Davenport & Prusak, 1998). The major problems of knowledge sharing are to convince, coerce, direct or otherwise get people within organization to share their information (Gupta et al, 2000). For organization, knowledge sharing is capturing, organizing, reusing and transferring experience-based knowledge that reside within the organization and making that knowledge available to others in the business. The interesting characteristics of knowledge is that its value grows when shared (Bhirud et al, 2005). The question is whether people are unaware of unable to articulate their tacit knowledge? Tacit knowledge sharing can take place at anytime, anywhere in the organization. Wah et al (2005) believe that an individual will only involve in knowledge sharing if such conditions exist, namely opportunities to do so, communication modality, expectation of the benefits of members accrue, expectation of the cost of not sharing knowledge, context compatibility for those who shared, motivation is crucial precondition for knowledge sharing, personal compatibility and liking and opportunism (associated with transaction cost analysis). Riege (2005) suggested that three important elements of knowledge sharing are individual, structure and technology would help organization in encouraging knowledge sharing in the organization. As tacit knowledge is very personal and very valuable, organization should concentrate on utilizing and capitalizing its tacit knowledge sharing. It is argued that the most effective means to transfer valuable tacit knowledge is actually not to codify it but rather to transfer it through an implicit mode (Schenkel & Teigland, 2008).

Tacit knowledge can be either knowledge embodied or knowledge embedded (Horvath, 2007). Embodied knowledge resides in the people minds while knowledge embedded is shown in products, processes or documents. The value of tacit knowledge is only known through its outcomes - innovation and organizational performance. Gold et al (2001) emphasize that knowledge infrastructures such as technology, structure and culture along with knowledge acquisition, conversion, application and protection are essential organizational capabilities for higher organizational performance. They believe that these key infrastructures would enable maximization of social capital. Technology dimensions are part of effective knowledge management include business intelligence, collaboration, distributed learning, knowledge discovery, knowledge mapping, opportunity generation as well as security (Gold et al., 2005) while structure is important to optimize knowledge sharing process within the firm. And the most significant hurdle of knowledge management or knowledge sharing in particular is organizational culture. Shaping culture is the central of firm's ability to manage its knowledge. Husted et al (2005) reveal that extrinsic motivators such as reward (monetary incentives) are related to knowledge exploitation while intrinsic motivators such as self-fulfilling task are related to knowledge exploration. Pathirage et al (2007) highlighted studies by Grant (1996) and Zander and Kogut (1996) that tacit knowledge can be integrated externally through relational networks those span organizational boundaries that are paramount for superior performance. Knowledge sharing leads to higher organizational performance (Du et al, 2007; Widen-Wulff & Suomi, 2003, 2007, Darroch & McNaughton, 2002) especially when knowledge sharing capabilities is combined with organizational resources (Widen-Wulff and Suomi, 2007). Choi & Lee (2002) in Keskin (2005) indicated that applying tacit and explicit oriented strategies is imperative for firm performance by large sized firms in western countries, but scant on firms' performance of SMEs in developing countries. As small businesses may lack performance measurement frameworks, these businesses should start with simple performance measures in their performance measurement framework, as suggested by Keegan et al (1989; 1994) in Ali (2003). Small business should strive for simplicity and keep their performance measurement system focused and simple (Ali, 2003). Ali (2003) had proposed a performance measurement framework that suit SMEs as developed by Neely et al (Ali, 2003) on balanced-scorecard. The expression "small firm" has no single definition, mainly because of the wide diversity of businesses. In the literature, some of the most widely used criteria to delineate a "small business" include size, number of employees, sales volume, asset size and type of customers (Ali, 2003). The size of the businesses and the increasingly competitive market force small and medium enterprises to consider more cost-effective processes than large enterprise must consider. As well, small businesses have to deal with unique operational and limited financial and economic resources (Ali, 2003). The strength of small business lies in greater motivation, better survey of the entirety of a project, tacit knowledge in unique skills, more informal communication along shorter lines, less bureaucracy, greater proximity to the market and to own production (Nooteboom, 1994). In short, in order for SME to perform better in market, they need supportive culture, active knowledge sharing activity and innovation process.

Every organization deal with many issues and SMEs are not exempted. Based on the problems faces by SMEs attempting to sustain their market share and to increase growth, they require strategic and operational planning, management of human resource, decision making related to financial, technical and marketing issues; and performance measurement (Ali, 2003).

4. Tacit knowledge sharing in SMEs

The strength of SMEs lies in motivation, good network, tacit knowledge in unique skills , shorter informal communication, less bureaucracy, greater proximity to market and internally which is important to be innovative (Nooteboom, 1993). Tacit knowledge sharing is ubiquitous in informal and without bureaucracy (Egbu et al, 2005).

Informal face-to-face of social interaction is the most effective technique used in the sharing of knowledge in SME whereby Egbu (2005) highlights knowledge in SME is in tacit nature. Individuals who set up SMEs do so because they have knowledge in key areas of competencies and think they can compete using such knowledge (Desouza & Awazu, 2006). In small business, knowledge tend to be shallow (no functional specialists) and tacit. Shallow can be modified by supplementing from external sources while narrow knowledge leads to tacit knowledge on how to run a business (Nooteboom, 1993). According to Davenport and Prusak (2000), network plays an important role in developing knowledge and innovative thinking especially during informal meetings. Most of consulting and service firms have organized their informal networks into formal networks in ensuring that all knowledge is acquired necessarily. Tacit knowledge helps SME to be different from its competitors particularly on their resources as based on resource-based theory (Barney, 1991) which should be valuable, rare, limited and non-imitable. The more tacit knowledge is shared, the harder is the imitation (Leonard & Sensiper 1998). When tacit knowledge is shared and combined to become collective tacit knowledge, it would then lead to creativity and innovation (Leonard & Sensiper 1998). The concept of tacit knowledge is very important in the context of innovation and its diffusion (Nooteboom 1993), explain that tacit knowledge, hard to capture in the term "know-how". Nooteboom (1993) refers tacit knowledge – "know-how" consists partly of technical skills – the kind of informal (Nonaka, 1998), of the small business as craftsmanship. Innovation is a distinction drawn in communication (Pohlmann, 2005) which is the root of knowledge sharing. Tacit knowledge sharing is essential in SME with scarce resources (Egbu et al, 2005).

5. Conclusion

Moving towards knowledge-based economy, SMEs should capitalize its internal knowledge which residing within its employees - tacit knowledge. Sharing this tacit knowledge would help SME to be creative and innovative thus enhancing its performance. The ability of SMEs to innovate and improve continuously has been proved to be related to the employees' skills and knowledge (Nonaka, 1998). Focusing on the internal knowledge – tacit knowledge – and share it in the organization as collective tacit knowledge would lead SME for a better performance especially in producing new products and services. With the impact of globalization, SME has to find their own resources that suit to their environment and their capabilities. Utilizing their very own resource in their people, equip with entrepreneurial orientation of the owners/managers of SME and supports from government, Malaysian SMEs could achieve higher performance thus contribute to the economy.

References

Ali, Imran. (2003). *A performance measurement framework for a small and medium enterprise*. Dissertation PhD. University of Alberta, Canada.

Beijerse, R. P. u. (2000). "Knowledge management in small and medium-sized companies: knowledge management for entrepreneurs." *Journal of Knowledge Management* 4(2): 162-179.

Bhirud, S., Rodrigues, L., & Desai, P. (2005). Knowledge sharing practices in KM: A case study in Indian software subsidiary. *Journal of Knowledge Management Practices.* 2 (2): 83-90

Hisrich, D.R. & M. P.Peters, (2008). *Entrepreneurship*. Singapore, McGraw-Hill.

Darroch, J. & Mc Naughton, R. (2002). Examining the link between knowledge management practices and types of innovation. *Journal of Intellectual Capital*. 3(3):210-222.

Davenport, T. H. & L. Prusak (2000). *Working Knowledge: How organizations manage what they know*. Boston, Harvard Business School Press.

Davenport, T.H & Prusak, L. (1998). Working *Knowledge: How Organizations Manage What They Know*. Cambridge , MA: Harvard Business School Press.

Desouza, K.C & Awazu, Y. (2006). Knowledge Management at SMEs: five peculiarities. *Journal of Knowledge Management*. . 10(1):32-43.

Du, R, Ai, S., Ren, Y. (2007). Relationship between knowledge sharing and performance: A survey in Xian, China. *Expert Systems with Applications*. .32:38-46.

Egbu .O. C., Hari, S.,and Renukappa, H. S. (2005). " Knowledge management for sustainable competitivenss in small and medium surveying practices".*Structural Survey*. 23(1):7-21.

Drucker, F. P. (1986). *Innovation and Entrepreneurship*. New York, Harper and Row.

Gamble, P.R. and Blackwell, J. (2001). *Knowledge Management: A state of the art guide.* Kogen Page. London, UK.

Gold H. A., Malhotra, A., & Segars, H. A. (2001). Knowledge management: An Organizational Capabilities Perspective. *Journal of Management Information Systems.* 18(1):185-214

Grant R.M. (1996). Toward a knowledge-based theory of the firm. *Strategic Management Journal*. 17(10):100-122.

Gupta, B., Iyer, L., & Aronson, J.E. (2000). Knowledge management: practices and challenges. *Industrial Management and Data Systems.* 100(1):17-21.

Hashim, M. K. (2007). *SMEs in Malaysia: A Brief Handbook*. Kuala Lumpur, August Publishing Sdn Bhd.

Horvath, J.A. (2007). Working with Tacit Knowledge. IBM Institute For Knowledge Management

Husted, K., Michailova, S., & Minbaeva, D. (2005). *Knowledge sharing and organizational performance: the role of extrinsic and intrinsic motives*. 8th International Human Resource Management Conference, Cairns, Australia, Copenhagen Business School.

Keskin, H. (2005). The relationships between explicit and tacit oriented KM strategy and Firm Performance. *Journal of American Academy of Business*, Cambridge Hollywood 7 (1), pp 169-176

Kuan Y.W & E. Aspinwall. (2004). "Characterizing knowledge management in the small business environment". *Journal of Knowledge Management*. 8(3) pp 44-61.

Lee, H. & Choi, B. (2003). Knowledge Management Enablers, Processes, and Organizational Performance: An Integrative View and Empirical Examination. *Journal of Management Information Systems*. . 20 (1):179-228

Leonard, D. & S. Sensiper (1998). The role of tacit knowledge in group innovation.. *California Management Review* 40(3): 112-132.

Man, T. W. Y., Lau, T., & Chan, K.F. (2002). The competitiveness of small and medium enterprises: A conceptualization with focus on entrepreneurial competencies. *Journal of Business Venturing* 17: 123-142.

McAdam, R and McCreedy, T. (2004). Barriers to innovation within small firms in a peripheral location. *International Journal of Entrepreneurial Behavior and Research*. 10(3):206-221.

Nonaka, I. & Takeuchi, H. (1995). The knowledge-creating company: How Japanese Companies Create the Dynamics of Innovation. Oxford University Press. New York, NY.

Nonaka, I. (1998). *The knowledge-creating company*. Boston, Harvard Business School Press.

Nooteboom, B. (1999). Innovation, learning and industrial organization. *Cambridge Journal of Economics*. . 23(2):127-150.

Pathirage, C. P., D. G. Amaratunga, and Haigh, R. (2007). Tacit knowledge and organizational performance. *Journal of Knowledge Management* 11(1): 115-126.

Pohlmann, M., Gebhardt, H. and Etzkowitz, H**.** (2005). The Development of Innovation Systems and the Art of Innovation Management – Strategy, Control and the Culture of Innovation.. *Technology Analysis and Strategic Management*. . 17(1):1-7.

Polanyi, M. (1966). The Tacit Dimension. Routledge and Kegan Paul Ltd, London.

Riege, A. (2005). Three-dozen knowledge-sharing barriers managers must consider. Journal *of Knowledge Management*. 9(3):18-35

Schencel, A. & Teigland, R. (2008). Improved organizational performance through communities of practice. *Journal of Knowledge Management*. 12 (1):106-118.

SME Annual Report 2006.

Stewart, T. A. (2000). *Intellectual Capital: The new wealth of organizations*. London, Nicholas Brealey Publishing

Wah, C.Y., Menkhoff, T., Loh, B., & Evers, H. (2005). Theorizing, Measuring and Predicting Knowledge Sharing Behavior in Organizations – A Social Capital Approach. *Proceedings of the 38th Hawaii International Conference on System Sciences.* Hawaii.

Widen-Wulff, G. & Suomi, R. (2003). Building a Knowledge Sharing Company – Evidence From the Finnish Insurance Industry. *Proceedings of the 36th Hawaii International Conference on System Sciences.*

Widen-Wulff, G. & Suomi, R. (2007). Utilization of Information Resources for Business Success: The Knowledge Sharing Model. *Information Resources Management Journal.* 20(1):46-67

Wong, W.L.P & Racliffe, D.F. (2000). The Tacit Nature of Design Knowledge. Technology Analysis and Strategic Management. 12(4): 493- 512

Yang, J. (2007). The impact of knowledge sharing on organizational learning and effectiveness. *Journal of Knowledge Management*. 11(2): 83-90.

Zander, U. & B. Kogut (1995). Knowledge and the speed of the transfer and imitation of organizational capabilities: An empirical test. *Organization Science* 6(1): 76-92.

Zack, M. (1999). Developing a knowledge strategy. *California Management Review*. 41(3): 125-145.s.

International Journal of Economics and Finance
Vol. 1, No. 1
February 2009
www.ccsenet.org/journal.html

Analysis on the Service Trade between China and ASEAN

Chunmei Yang
College of Economics
Tianjin Polytechnic University
Tianjin 300384, China
E-mail: jane_ycm@126.com

Abstract

With the economic globalization and regional economic integration, today's world witnesses the new round of industrial restructure with service and service trade playing an increasing role in the national economy and the center of international industry shifting from manufacturing to service. This article studies the trade structure, complementary advantages of both China and ASEAN, then put forward the countermeasures for improving our service trade structure and enforcing the competitiveness of service industrial system on the basis of mutual benefit.

Keywords: ASEAN, Service trade, Cooperation

Since the Framework Agreement on Comprehensive Economic Cooperation between China and ASEAN (Association of Southeast Asian Nations) signed by China and ASEAN in November 2002, in which it is decided that the establishment of China-ASEAN Free Trade Area in 2012, the negotiation of service trade has been going on several years. The Service Trade Agreement was finally signed on January 14, 2007 and took effect on July 1, 2007. According to the Agreement, on the basis of the commitment made to the WTO, China will make new commitment in 26 sectors under construction, environment preservation, transportation, sports and business. On the other hand, ASEAN makes commitment to open its markets in finance, telecommunications, education, tourism, construction, and medicines, including the further opening up of these markets and permission of solely-funded and equity joint ventures and other cooperation for other party, relaxing in the limit of stake proportion in establishing companies.

In general, there are differences between two parties in each other's development basis, industrial structure, and developing stage. These differences create the complementary trade in service

1. The volume and structure of China's Service Trade

China's service trade has been developing rapidly. The total volume of service's import and export in 2007 is 25 times the size of that in 1990. The average annual increase rate is up by 18.4%, which is two times over that of the world. While the international competitiveness in our service trade is till lower than the developed countries and even some developing countries. From 1990 to 2007, our service's export volume as a share in our total export is keeping lower than 10%, which is only half of the world's average level. The total service's import and export volume in 2007 is 250.9 million dollars, up by 30.88% sharply than previous year, but it is only account for 4% of the world's service import and export. Meanwhile our trade in service has been being in deficit for a long time. In 2007, China's services exports grew faster than imports. Deficit in services trade shrank noticeably to 7.7 billion, down by 14.6%over the previous year, and 20.3% lower than the largest deficit in 2004, the year with the largest trade deficit in services. All detail data is provided in Table 1.

Insert Table 1 Here

Insert Table 2 Here

As showed in Table 2, the overall competitiveness of our service trade is weak. Its TC index is keeping negative, which indicates China's service trade is of comparative disadvantage and not being improved for a long time. During the same period, the TC index of EU and the USA is 0.8 and 0.12 respectively. As for sector, China's service import and export is unbalance. The TC index of only both travel and other commercial service is keeping in positive, with certain competitiveness. While the transportation is being in negative because of its turning to capital and technological-intensive service manner, in which China is weak comparatively. The TC index of construction has been getting better in recent five years as a result of rich labor resource and lower labor cost. For most part of time, the index of communication, computer and information is in positive by taking advantage of national monopoly. As for the index of high value-added service sectors, such as insurance, finance, royalties and license fees, consulting, film, audiovisual, and so on, is negative, which shows that the competitiveness of China's capital and technological-intensive service trade is comparatively lower even though it is improving unceasingly.

2. Comparison of Advantage in Service Trade between China and ASEAN

China's service trade with ASEAN has been developing rapidly of late years. ASEAN has become the key export market of China's investment, contract engineering, labor service collaboration. Meanwhile the investment from ASEAN in marine and air transportation, financial service, construction and engineering service is also an important composed part of China's service import. The actual economic benefit gained from the service trade has become the important basis and powerful motive force for the development of relations between China and ASEAN.

Insert Table 3 Here

The above comparison in table 3 indicates that China and the 10 nations of ASEAN have their own advantages in terms of sector structure, which result in the great complementary between two parties. The Chinese businesses should attach great importance to the ASEAN market and choices offered by the opening up of service trade market at both sides and catch the business opportunities.

3. Countermeasures for enhancing cooperation in service trade between China and ASEAN

To sum up, there is competition in the service sector basis, resources, development of service trade and outsourcing between China and ASEAN, but their level of economic development and sector structure are multistoried and ladder shaped, which create the diversity and complementary in aspect of both parties' comparative advantages. So the bilateral cooperation in service trade should endeavor to enhance collaboration in their complementary fields, to achieve mutual –development and win-win in their competitive fields.

3.1 To keep on strengthening the competitiveness in traditional service sector

The distinctive natural and cultural sceneries of both China and ASEAN serve as plentiful travel resources and are also the basis for forming and developing the competitiveness of each other's travel service. As travel accounts for big share in both China and ASEAN's service trade export, and as developing countries, both have had gained their own competitiveness in this traditional travel service at the same time, there should be more competition than cooperation in travel service field between China and ASEAN. Under the pressure from ASEAN after the further opening of travel market, China should be devoted to develop the distinguish and diversity travel services, such as green travel, cultural travel, submarine travel, ecological travel, internet travel and other creative travel projects. The knowledge-based economy is changing the traditional competitive manner from resources and passenger flow direction to technology and information. The competition in future travel service will focus on creating travel resources and particular travel services according to the passengers' desire, interest, economic situation and required time instead of selling certain scenery point or travel rout only. Meanwhile, we should enlarge the opening of travel and use foreign capital to adjust the investment structure of our service. For example, we can transform the service projects of our old and famous national hotel by using foreign capital, outsource human resources training of our travel agents, the information system's developing and managing, landscape hotel and other low-value-added links in order to advance the management level of China's travel agents.

3.2 To consolidate the cooperation in transportation and finance

In the field of transportation, the half of the import and export between China and ASEAN is carried by marine transportation. The volume of transportation trade from China to ASEAN has increased by 22.7% within recent five year, and the Mekong Rive transportation is further developed at the same time. The increasing transportation of cargo and passengers between China and ASEAN build the solid basis for developing the transportation service of both side. So different from travel, there are more cooperation than competition in transportation service between China and ASEAN. China should reinforce the construction of logistic infrastructural facilities in order to speed um the development of modern transportation and logistic, meanwhile, take advantage of the huge population of China-ASEAN Free Trade Area to develop the scale economy.

Finance is usually regarded as "the first driving force" to the development of regional economy. For the purpose of larger scope, wider field and higher lever cooperation, both sides should enhance their collaboration in the field of finance in order to make finance play its function of configuration guidance and market regulation. For middle and long term, the cooperation between China and ASEAN should focus on financial supervision and currency in order to avoid the fluctuation and financial crises caused by hot money, and the illegal activities, such as arbitrage of exchange, money laundering, swindle, and so on, which will damage the financial system and solid economy greatly.

3.3 To improve the level of high-technology service

Technology service is the core of world's service trade. Computer and information services are of high increasing potential as high-value-added sectors. Except Singapore, the competitiveness in computer and information service of China and other ASEAN members is almost the same and need to be improved further. Both as being developing countries should prioritize the development of these sectors. While the volume and risks of investment in these sectors are huge, so it is necessary for the government to provide guidance and subsidies in their R&D and production capability. On the other hand, facing the entering of transnational corporations from Singapore, who has been of international competitiveness in computer and information services, we should guide its FDI efficiently to those high-value-added sectors, meanwhile we should develop our software outsourcing in order to build solid basis for us to develop software and export services related computer and information in overseas market.

3.4 To enhance bilateral investment and enlarge service trade

The signing of the Service Trade Agreement by China and ASEAN is inevitably bringing out more bilateral investment. China's enterprises' investment in ASEAN market, which owns 500 million consumers, not only benefit for enlarging the present advantages of manufacturing industries but also promote the service export related to the manufacturing industries. Because of the difference of service trade development among the members of ASEAN, their process of opening the service trade and acceptability are not the same at all. We should choose pointedly the investment project according to each country's different characteristic. For example, Singapore is of comparative advantage in the knowledge-intensive service, capital, management, attracting investment and facility in language. Then we can invite investment from Singapore and put into our high-technology service sectors. On the other hand, we can invest in the service sectors of ASEAN market in order to promote the economic development of both sides.

In conclusion, there is great potential cooperation in service trade between China and ASEAN. Only if both parties engage in enlarging collaboration, making resources conformity sufficiently, building up systems and coordination policies well, can they survive from the new round of industrial restructure with service and advance the service industry successfully.

References

Shi, Zhiyu. (2007). Analysis on the Trade Deficit between China and ASEAN, Economic Problems Explore, (2).http://www.mofcom.gov.cn.

Xu, Ningning. (2007). Review on the Fifteen Years' Economic and Business Cooperation between China and ASEAN, Chinese Forum, (3).

Yang, Chunmei. Associate professor of College of Economics, Tianjin Polytechnic University; Research field: international trade.

Table 1. China's Service Imports and Exports from 1990 to 2007

USD 100Million

Year	China's import volume	As a share in world export (%)	World export volume	China's import volume	As a share in world import (%)	World import volume	China's import and export volume	As a share in world import and export (%)	World import and export volume
1990	57	0.7	7805	41	0.5	8206	98	0.6	16011
1991	69	0.8	8244	39	0.5	8510	108	0.6	16754
1992	91	1.0	9238	92	1.0	9471	183	1.0	18709
1993	110	1.2	9413	116	1.2	9596	226	1.2	19009
1994	164	1.6	10332	158	1.5	10438	322	1.6	20770
1995	184	1.6	11849	246	2.0	12015	430	1.8	23864
1996	206	1.6	12710	224	1.8	12697	430	1.7	25407
1997	245	1.9	13203	277	2.1	13056	522	2.0	26259
1998	239	1.8	13503	265	2.0	13350	504	1.9	26853
1999	262	1.9	14056	310	2.2	13883	572	2.0	27939
2000	301	2.0	14922	359	2.4	14796	660	2.2	29718
2001	329	2.2	14945	390	2.6	14941	719	2.4	29886
2002	394	2.5	16014	461	209	15793	855	2.7	31807
2003	464	2.5	18340	549	3.0	18023	1013	2.8	36363
2004	621	2.8	21795	716	3.4	21328	1337	3.1	43123
2005	739	3.1	24147	832	3.5	23613	1571	3.3	47760
2006	914	3.4	27108	1003	3.8	26196	1917	3.6	53304
2007	1216	3.7	32572	1293	4.2	30591	2509	4.0	63163

Source: Derived from WTO International Trade Statistics Database, China Balance of Payment

Table 2. TC Index for China's Service Trade by sector from 1997 through 2006

Year Sector	1997	1998	1999	2000	2001	2002	2003	2004	2005	**2006**
Transportation	-0.54	-0.57	-0.53	-0.48	-0.42	-0.41	-0.4	-0.34	-0.29	**-0.24**
Travel	0.2	0.16	0.13	0.11	0.12	0.14	0.07	0.15	0.15	**0.17**
Communications	-0.03	0.6	0.51	0.7	-0.17	0.08	0.2	-0.03	-0.1	**-0.02**
Construction	-0.34	-0.31	-0.22	-0.25	-0.01	0.13	0.04	0.05	0.23	**0.15**
Insurance	-0.71	-0.64	-0.81	0.92	-0.85	-0.88	-0.87	-0.88	-0.85	**-0.88**
Financial service	-0.85	-0.72	-0.2	-0.11	0.12	-0.28	-0.21	-0.2	-0.04	**-0.72**
Computer and Information	-0.47	-0.43	0.09	0.15	0.14	-0.3	0.03	0.13	0.06	**0.26**
Royalties and license fees	-0.82	-0.74	-0.83	-0.88	-0.89	-0.92	-0.94	-0.9	-0.94	**-0.94**
Consulting	-0.15	-0.19	-0.3	-0.29	-0.26	-0.34	-0.22	-0.2	-0.07	**-0.03**
Advertising, publicity	-0.01	-0.11	0.004	0.05	0.04	-0.03	0.03	0.1	0.2	**0.2**
Film, audiovisual	-0.63	-0.44	-0.67	-0.34	-0.21	-0.53	-0.35	-0.62	-0.06	**0.08**
Other. commercial services	0.19	0.07	0.02	0.07	0.12	0.28	0.4	0.31	0.28	**0.27**
Total	**-0.07**	**-0.05**	**-0.09**	**-0.08**	**-0.08**	**-0.08**	**-0.08**	**-0.07**	**-0.05**	**-0.05**

Source: Derived from Balance of Payment of China, 1997-2006

Table 3. Comparison of Service Trade Resources between China and ASEAN

Country	**Sectors with advantage and Potentiality**	**Sectors needed to be developed**
China	Construction, marine transportation, travel, computer and information	Financial service, insurance, consulting
Brunei	Travel and related service, financial service, cooperative exploitation of oil and natural gas	Commercial service, transportation
Cambodia	Travel and relating service, construction and related engineering	Commercial service, telecommunication service, environment and public facility
Indonesia	Transportation, communication, post and cable service, consulting	Financial service, insurance, travel
Laos	Electricity, travel and related service	Transportation, communication
Malaysia	Travel and related service, financial service	Commercial service
Myanmar	Energy exploitation, construction, mining	Energy and human resources exploitation, travel, transportation communication
the Philippines	Information and related service, paging hub, commercial purchasing service	Travel, banking and security
Singapore	Air transportation, financial service, hotel, exhibition service	Gambling, construction
Thailand	Travel, environment and financial service	Construction and related services
Vietnam	Labor service	Education, commercial service, technological service, financial service

Editorial Board

Cindy Xu	Canadian Center of Science and Education, Canada
Evelyn Devadason	University of Malaya, Malaysia
Guneratne Wickremasinghe	Victoria University, Australia
Hussain A. Bekhet	University Tenaga Nasional, Malaysia
Hussin Abdullah	Universiti Utara Malaysia, Malaysia
Mahdi Salehi	Zanjan University, Iran
Mohd Nasir bin Mohd Yatim	Universiti Tenaga Nasional, Malaysia
Mohd Rosli Mohamad	University of Malaya, Malaysia
N.M.P. Verma	Babasaheb Bhimrao Ambedkar University, India
Turan Subasat	University of Bath, UK
Ying Jin	Qingdao Technological University, China

A journal archived in Library and Archives Canada
A journal indexed in Canadiana (The National Bibliography)
A journal indexed in AMICUS
A peer-reviewed journal in economics, accounting and finance

International Journal of Economics and Finance

Semiannual

Publisher	Canadian Center of Science and Education
Address	4915 Bathurst St.Unit #209-309, Toronto, ON. M2R 1X9
Telephone	1-416-208-4027
Fax	1-416-208-4028
E-mail	ijef@ccsenet.org
Website	www.ccsenet.org
Printer	William Printing Inc.
Price	CAD.$ 20.00

ISSN1916-971x

9 771916 971005 01