공학박사 학위논문

Empirical Research on Common Preferred Stock Spread Index and its Application on Investment Strategy

우선주식과 보통주식 수익율 스프레드 지수를 이용한 금융투자 실증 연구

2019년 8월

서울대학교 대학원

산업공학과 금융리스크공학 전공

이 창 주

Abstract

Empirical Research on Common Preferred Stock Spread Index and its Application on Investment Strategy

Lee, Changju

Department of Industrial Engineering

The Graduate School

Seoul National University

After various financial crises, the importance of financial market analysis for financial risk management has been emphasized. To minimize the risk of losing money from unforeseen financial crises, it became critical to develop a market index that could both evaluate the current financial market and explain the future market return. In addition, it became imperative that the investment strategy could continue to make profit during the financial crises. Many researchers have tried to evaluate the financial market and explain the future market return with various risk index such as the VIX (CBOE volatility index)

index, and TED index (the spread between three-month LIBOR interest rate and three-month US treasury bill interest rate) and valuation ratios such as Price to earning ratio, Price to book ratio, CAPE(Cyclically Adjusted Price Earning Ratio) and price to operational earning ratio. Previous studies attempted to explain future market returns better by adapting existing indexes and ratios. However, in this dissertation, we introduce a new market index called; Common Preferred Spread index (CPS index), which was empirically tested to confirm that it could not only evaluate current stock market condition but also has explanatory power for the future market return. First, this dissertation explains the CPS-index using the spread return between common and preferred stock pairs and shows that CPS-index has explanatory power for long term market return. We observed that common stocks are more sensitive to the market condition than preferred stocks so the CPS-index tends to oscillate according to market conditions. The future realized market return increases when CPS-index is low, and vice versa. We present that there is an inverse relationship between CPS-index and the future realized return of S&P500 index. By comparing the fitting validity of statistical models including correlation analysis and linear regression between the future realized return and each of CPS-index, VIX index, TED index, CAPE ratio and S&P500 index, we confirm the superior power of CPS-index to explain the future realized return. We developed a trading strategy based on CPS-index and assessed how to enhance the predictive power of CPS-index through stepwise regression, Granger causality test and neural network method. Second, we conduct an empirical analysis on CPS-index in comparison with currently existing valuation ratios such as the price to earning ratio, price to book ratio, and price to operational earning ratio. The multivariate regression method is applied to test whether adding CPS-index as an independent variable significantly increases the explanatory power of regression for the future market return. According to the test results of every multivariate regression model, the CPS-index as an independent variable has the most market predictability power among other benchmark independent variables of regression. In addition, we also discovered the optimal parameters to use the CPS-index. Lastly, a new Pairs Trading strategy is proposed, using common stocks and preferred stocks. Unlike the traditional method of Pairs Trading, which is based on two different stocks that were moving together in the past, a common stock and its preferred stock as a pair are used in Pair Trading. This new method could reduce the risk of losing money from the traditional method of Pairs Trading. We explain through the test results of every portfolio that this new method of Pairs has the highest Sharpe Ratio.

Keywords: Preferred Stocks, Common Stocks, Pairs Trading, Predictive Index, Market valuation Index, Spread Return, Common preferred spread index

Student Number: 2015-30238

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Research motivation and purpose

After experiencing the 2008 financial crisis, developing a new investment strategy and index that could prevent such incidents or warn the investors became an important issue to solve in the financial engineering area. Therefore, many researchers began been working on developing a new index and investment strategies; however, it required too much financial background knowledge and computing resources for ordinary people. This research presents new investment strategies and index that could be easily operated and understood by average investors by analyzing the relationship between common stocks and preferred stocks.

Among various investment strategies, Pairs Trading was the most intriguing strategy because it could yield profit in any financial situation. Pairs Trading strategy is as follows: First, find two stocks that move together historically. Second, when the spread between those two stocks are widened, short the overpriced one and buy the underpriced one. Last, if those two stocks move similar to history, prices will converge, and the investors will make a profit. Since Pairs Trading only makes profit from the spread, investors can always make profit as long as the pair converges at some point in the future. Therefore, even during aforementioned financial crisis, portfolios that used Pairs Strategy empirically tested to yield profits. However, there are areas of weaknesses in Pairs Trading. First, finding a pair of stocks that move together requires complicated calculation and computation power because investors have to find this pattern and movement from more than 4000 stocks. Second, in Pairs Trading, it is essential to identify that the two stocks have moved together in the past, but that they will continue to move together in the future. If the pairs move in different directions in the future, the investors would lose a lot of money. This takes a considerable amount of gambling in the probability that this pair would stay in synchronized movement in the future. To address this leap of faith, we suggest operating Pairs Trading with a common stock and its preferred stocks because for a pair of common stocks and its preferred stocks, if they moved together in the past, they will undoubtedly move together in the future. By using this strategy, investors could save computation time, cut down on risk, and make profit more safely.

After finding the possibility of using preferred stock to develop existing investment strategy, we wanted to build a market index that could evaluate the market condition with preferred stocks. Instead of using two similar common stocks, we introduce a novel index obtained from the spread between the average price of common stocks and that of their corresponding preferred stocks so they can readily be used for our strategy for Pairs Trading. First, we select common stocks accompanying preferred stocks and make return indexes for common and preferred stocks separately, by averaging each. Second, we make a return spread index called 'CPS-index (Common Preferred Spread Index)' by subtracting the preferred stock index from the common stock index.

We claim that common stocks are overvalued when CPS-index is large and undervalued when CPS-index is small. The idea underlying this claim is as follows. Based on our observation, the common stocks are more sensitive to market conditions than that of their preferred stocks, so the prices of common stocks are likely to be overvalued in the bull market condition and undervalued in the bear market condition. Empirically, we show that the common stock index is more correlated with market index than the preferred stock index. For this reason, the spread between returns of common stocks and their preferred stocks can indicate the future realized return in the stock market. Even though there are 250 preferred stocks to pair with common stocks in the United States stock market, this index can represent stock market condition because those 250 common stocks having preferred stocks are mostly large company stocks that have large market capital. In addition, we confirm that the CPS-index using the return spread between common and their preferred stocks have better prediction power for the future realized market returns.

Throughout this dissertation, we focus on empirical analysis on this CPS-index and introduce ways to develop it. Using preferred stocks, we empirically showed that Pairs Trading with preferred stocks can make profit with less risk, and using this CPS-index, investors can evaluate market condition and could predict the market in order to invest in a safer condition.

1.2 Theoretical Background

There has been a lot of research on developing market index that can predict future market returns. Many researchers focused on proving the predictive power of valuation ratios such as dividend to price ratio, book to market ratio, earning to price ratio and payout yield (Boudoukh, Michaely, Richardson, & Roberts, 2007; Campbell, 1987; Campbell & Shiller, 1988; Eugene F Fama & French, 1988; Kothari & Shanken, 1997). In addition, there are many researches on relation between business cycle variables and expected rate of return (Campbell, 1987; Eugene F. Fama & French, 1989). It is difficult to decide whether each index actually has predictive power. Ang and Bekaert (Ang, Bekaert, & Wei, 2007) showed that dividend yield has no predictive power in the long horizon. Boudoukh and Michaely (Boudoukh et al., 2007) discovered that net payout yield and net equity issuance have the predictive power and that long horizon tests are not more powerful than short horizon tests. Li (Li, Ng, & Swaminathan, 2013) showed that aggregate implied cost of capital has predictive power in long horizon. Chacko et al. (Chacko, Das, & Fan, 2016) introduced index based measure of illiquidity that may have market predictive power. Ioannis and Sermpinis (Ioannis & Sermpinis, 2016) used VIX index to model and trade the US implied volatility indices. Badshah et al. (Badshah, Frijns, Knif, & Tourani-Rad, 2016) used the VIX index to find a relationship with intraday return-volatility relation. Kanas (Kanas, 2012) used VIX index to model the risk return relation for the S&P 100. Chen et al. (Chen, Bao, & Zhou, 2016) used analysis of candlestick pattern to predict stock market. For neural network predictive power related work, Kristjanpoller and Minutolo (Kristjanpoller & Minutolo, 2015) used artificial neural network and GARCH model to predict the gold price volatility, and Chen used the neural method to predict the emerging financial market. There are many studies using network to make the predictive model. Heiberger (Heiberger, 2018) used stock network, and machine learning techniques to predict economic growth while Caraiani (Caraiani, 2017) tried to find significant predictive power property using financial network.

On the other hand, many research papers on the topic of market valuation have focused on the financial bubble. The following researchers commonly emphasized that financial bubble is extra price added to fundamental value; Garber (M.Garber, 1990) defined financial bubble as part of the price movement that cannot be explained by fundamentals. Kindleberger and Aliber (Kindleberger & Aliber, 2005) defined bubble as an upward price movement over an extended range that then implodes. Brunnermeier (Brunnermeier, 2009) argued that bubbles are typically associated with dramatic asset price increases followed by a collapse. Yangru (Yangru, 1997) defined bubbles as the difference between the fundamental value and the market price allowing. In contrast, there are many researchers who define and improve the financial bubble detections. Shiller, LeRoy and Porter used variance bound test for pricing the equity (LeRoy & Porter, 1981; Shiller, 1981). In addition, Blanchard, Watson and Tirole developed this variance bound test to detect financial bubble (Blanchard & Watson, 1982; Tirole, 1982) . Although this variance bound test to detect financial bubble is criticized by other researchers, they are the first to try detecting financial bubble. Diba and Grossman tried to explain the theoretical properties of bubbles (Diba & Grossman, 1987). Recently, Phillips used the unit root behavior of key fundamental financial variables to detect the bubbles in the 2008 subprime mortgage (Phillips, Wu, & YU, 2011). Song (Song, Ko, & Chang, 2018) proposed the non-linear marginal expected shortfall models and their minimum spanning trees for analyzing the systemic risk of the US financial system during the financial crisis.

Unlike the previous studies, we focused on the fundamental theory that people would lose money if they invest when the stock market in comparably overvalued compared to the past, and people would gain money if they invest when the stock market is undervalued compared to the past. In this study, we otnained the idea of evaluating market condition from Pairs Trading and preferred stocks. Gatev et al. (Gatev, Goetzmann, & Rouwenhorst, 2006) showed that Pairs Trading is actually worked in the financial market. Pairs Trading is a market neutral trading strategy that uses two stocks that have been moved similarly; after that, when two stocks are diverging buy undervalued stocks and sell overvalued one and wait until the mispricing will correct itself in the future (Gatev et al., 2006). Instead of using two similar stocks, we used common stocks and its preferred stocks to find a spread and using this spread, we built 'CPS-index' which indicates the spread between common stocks and preferred stocks in market.

For Pairs Trading parts, researchers concentrated on how to make profit with Pairs Trading with more safe investment strategy. They empirically tested the mean excess return of portfolios and try to find out the best strategy empirically. Nath examines the implementation of a simple Pairs Trading strategy with automatic extreme risk control using the entire universe of securities in the highly liquid secondary market for U.S. government debt. Nath also examines the excess rate of return with different diverge strategy. He showed that using data from the repo and money market, estimates are also made of the distribution of absolute returns after accounting for financing and transaction costs. (Nath, 2004) Perlin examines the excess return of Pairs Trading in Brazilian stock market. We used the minimum distance method to find pairs with different learning time periods. He used different Pairs Trading strategy that were not limited to the two standard deviation divergece strategy. The main conclusion of this simulation is that pairs-trading strategy was a profitable and market-neutral strategy in the Brazilian market. Such profitability was consistent over a region of the strategy's parameters. (Perlin, 2009). Hong and Susmel researched pairs-trading strategies for 64 Asian shares listed in their local markets and listed in the U.S. as ADRs. Given that all pairs are cointegrated, they are logical choice for pairs-trading. They find that pairs-trading in this market delivers significant profits. The results are robust to different profit measures and different holding periods (Hong & Susmel, 2003) .

1.3 Research Overview

The rest of this dissertation is organized as follows. In Chapter 2, the CPS index (Common Preferred Spread index) is introduced with its various empirical analysis. This empirical analysis contains the relationship between CPS index and future realized return. VIX (CBOE volatility index) index, TED index (the spread between three-month LIBOR interest rate and three-month US treasury bill interest rate), CAPE (Cyclically Adjusted Price Earning ratio) and S&P500 index are used as bench mark index to compare; we used Univariate regression, Stepwise regression, Granger Causality test, and Neural network to explain CPS-index's market predictive power compared to bench mark indexes. In Chapter 3, the practical use of CPS-index is introduced. Comparing with representative market valuation ratios, such as price to earnings ratios, price to book ratios, and OPER (price to operation earnings ratio), we tested CPS-index as the market predictors. By combining with existing market valuation ratios, we tested whether CPS can enhance the market predictive power. Using parameter tuning, optimal parameter for the CPS index was suggested. In Chapter 4, there is empirical analysis on Pairs Trading with common stocks and these preferred stocks. It tested various pairs structuring methods and operation strategies to find out optimal strategies for different time periods. Lastly, the summary and contributions of this dissertation are reviewed in Chapter 5 with limitations and possible future works to be conducted.

Chapter 2

Development of Common Preferred Spread Index

2.1 Introduction

The issue of developing new index that can explain future stock market return is of great interest to researcher and practitioner in finance. There has been a lot of research on predicting future market return based on the usefulness of valuation ratios such as dividend-to-price ratio, book-to-market ratio, earnings-to-price ratio and payout yield (Boudoukh et al., 2007; Campbell, 1987; Campbell & Shiller, 1988; Eugene F Fama & French, 1988; Kothari & Shanken, 1997). Even though the presence of an index having market predictive power is yet debatable, the construction of an index indicating either overvalued or undervalued market condition is still interesting area to research on.

In this paper, we introduce a novel index obtained from the spread between the average price of common stocks and that of their corresponding preferred stocks. Common stocks and their preferred stocks can be used for Pairs Trading. Pairs Trading is a market neutral trading strategy that uses two stocks moving similarly. In this strategy, traders buy undervalued stocks and sell overvalued ones when two stocks are diverging and wait until the mispricing will correct itself in the future (Gatev et al., 2006). Instead of using two similar common stocks in Pairs Trading, we focus on the pairs of common stock and its accompanied preferred stock and develop an index constructed from the return spread between them. First, we select common stocks accompanying preferred stocks and make return indexes for common and preferred stocks separately by averaging each of them. Second, we make return spread index called 'CPS-index (Common Preferred Spread Index)' by subtracting preferred stock index from common stock index.

We claim that stock market is overvalued when CPS-index is large and undervalued when CPS-index is small. The idea underlying this claim is as follows. Based on our observation, the common stocks are more sensitive to market condition than their preferred stocks so that the prices of common stocks are likely to be overvalued in the bull market condition and undervalued in the bear market condition. Empirically, we present that the common stock index is more correlated with market index than the preferred stock index. For this reason, the spread between returns of common stocks and their preferred stocks can indicate the future realized return of stock market. As common stocks are sensitive to market, common stock index itself can play a role of market predictor. However, we confirm that the CPS-index using the return spread between common and their preferred stocks have better prediction power for the future realized market return.

In stock market, only limited companies have preferred stocks, and this could lead to bias results that can not represent stock market. However, Companies which have preferred stocks are mostly top 500 market capital in US stock market and market index such as DOW 30, KOSPI 200, KOSDAQ 150, FTSE 100, NIKKEI 225 contain a smaller number of companies in their index and yet represents the market well.

Throughout this paper, we prove our claim that when the spread between common stocks and those preferred stocks are large, it indicates the stock market is overvalued and when the spread is small, the stock market is undervalued by presenting CPS-index has highly correlated with long term future realized return of market and using CPS index with existing predictive indexes can enhance the market predictability. Because, the best way to present when the market is overvalued or undervalued is to present its correlation with long term future realized return. In addition, we develop a trading strategy when to take long or short positions in the market according to the movement of CPS-index. The details are as follows.

First, we build CPS-index using the spread between common stock index and their preferred stock index. Second, we show that CPS-index is highly correlated with the future realized stock market return for various time horizons from one month to fortyeight months. We confirm that they have a negative relationship in that the future realized return is high (low) when CPS-index is low (high). Third, we evaluate the explanation power of CPS-index about the future realized market return in comparison with other benchmark competitors such as VIX (CBOE volatility index) index, TED index (the spread between three-month LIBOR interest rate and three-month US treasury bill interest rate), CAPE(Cyclically adjusted price ratio) and market return itself, S&P500. To assess the explanation power, we compute correlations and conduct univariate regression analysis between the S&P500 future realized return and each of CPS-index and other bench mark indexes. Fourth, we provide an investment strategy using CPS-index about when to take a long or short position in the stock market, and present historical test result of our strategy. Fifth, we use multivariate regression, variance inflation factor, C_p, stepwise regression and F-test to present adding CPS-index to existing market indexes to improve the market predictability. Sixth, using stepwise regression, we find out useful time lag variables for CPS-index that helps to explain long term market return. In addition, using Granger causality test, we show that CPS-index affects the market future realized return. Lastly, using neural network, we predict future realized return of market and confirm that CPS-index has better prediction power than other benchmark indexes such as VIX, TED, CAPE and S&P500.

2.2 Related Literature

Before we construct the CPS-index, Emanuel helped us to understand the value of preferred stocks. He suggested the theoretical model for valuing preferred stocks (Emanuel, 1983). Even though there is not enough research on relationship between commons stock and its preferred stocks, Linn and Pinegar (Linn & Pinegar, 1988) find interesting results on the effect of issuing preferred stock on common and preferred stockholder wealth and Stickel (Stickel, 1986) shows that the different effects on common stocks and preferred stocks price after the changes of preferred stock ratings. As above mentioned, there were some papers about explaining corresponding common stock price with preferred stock; however, there is no paper relating spread between common stocks and their preferred stocks to explain future market return, and evaluating the stock market condition. Therefore, we focused on studying papers that is related to predicting market return and evaluating the stock market condition.

There has been a lot of research on return prediction. Many researchers were focused on proving the predictive power of valuation ratios such as dividend to price ratio, book to market ratio, earning to price ratio and payout yield (Boudoukh et al., 2007; Campbell, 1987;

Campbell & Shiller, 1988; Eugene F Fama & French, 1988; Kothari & Shanken, 1997). In addition, there are many researches on relation between business cycle variables and future realized return (Campbell, 1987; Eugene F. Fama & French, 1989). It is hard to decide whether each index has a predictive power. Ang and Bekaert (Ang et al., 2007) showed that dividend yield has no predictive power in the long horizon. Boudoukh and Michaely (Boudoukh et al., 2007) found that net payout yield and net equity issuance have the predictive power and long horizon tests are not more powerful than short horizon tests. Li (Li et al., 2013) showed aggregate implied cost of capital has predictive power in the long horizon. Chacko et al. (Chacko et al., 2016) introduced index based measure of illiquidity that have market predictive power. Ioannis and Sermpinis (Ioannis & Sermpinis, 2016) used VIX index to model and trade the US implied volatility indexes. Badshah et al. (Badshah et al., 2016) used VIX index to find a relationship with intraday return-volatility relation. Kanas (Kanas, 2012) used VIX index to model the risk return relation for the S&P 100. Chen et al. (Chen et al., 2016) used analysis of candlestick pattern to predict the stock market. Kristjanpoller and Minutolo (Kristjanpoller & Minutolo, 2015) used artificial neural network and GARCH model to predict gold price volatility. Chen. (Chen et al., 2016) used the neural method to predict the emerging financial market. There are many studies using the network to make predictive models. Heiberger (Heiberger, 2018) used stock network, and machine learning techniques to predict economic growth. Caraiani (Caraiani, 2017) tries to find significant predictive power property using the financial network.

On the other hand, there has been a lot of research papers on related evaluating stock market condition with financial bubble; especially, they related overvalued stock market condition with financial bubble. Researchers commonly emphasized that the financial bubble is an extra price added to the fundamental value. Garber (M.Garber, 1990) defined financial bubble as the part of the price movement that cannot be explained by fundamentals. Kindleberger and Aliber (Kindleberger & Aliber, 2005) defined bubble as an upward price movement over an extended range that then implodes. Brunnermeier (Brunnermeier, 2009) argued that bubbles are typically associated with dramatic asset price increase followed by a collapse. Yangru (Yangru, 1997) defined bubbles as the difference between the fundamental value and the market price allowing. Oscar and Moritz studied asset price bubble using bubbles in housing and equity markets in 17 countries over the past 140 years (Jorda, Schularick, & Taylar, 2015). There are many researchers who defined and improved the financial bubble detections. LeRoy and Porter (Diba & Grossman, 1987) used the variance bound test for equity pricing. In addition, Blanchard and Watson (Blanchard & Watson, 1982) and Tirole

(Tirole, 1982) developed this variance bound test to detect financial bubbles. Although this variance bound test to detect financial bubble is criticized by other researchers, they are the first people to try detecting financial bubbles. Diba and Grossman (Diba & Grossman, 1987) tried to explain the theoretical properties of the bubbles. Phillips (Phillips et al., 2011) used the unit root behavior of key fundamental financial variables to detect bubbles in the 2008 subprime mortgage. There are some researchers relate stock market bubble with monetary Policy. They successfully showed how monetary policy affects stock market and financial bubble (Gali & Gambetti, 2015) (Caraiani & Adrian, 2018). Recently, Song et al. (Song et al., 2018) proposed the non-linear marginal expected shortfall models and their minimum spanning trees for analyzing the systemic risk of the US financial system during the financial crisis. Recently, Greenwood, Shleifer, and You (Greenwood, Shleifer, & You, 2017) showed that sharp price increase of an industry portfolio does not predict unusually low return forward and this sharp price increase the predictability of probability of market crash. There were various methods to evaluating the stock market condition with financial bubble; however, there are few researches on how to evaluate stock market using the pairs of common and preferred stocks.

2.3 Method

2.3.1. Spread between Common Stocks and their Preferred Stocks

Let c_t^j and p_t^j be the observed prices of common stock j and its accompanied preferred stock at time t, respectively; both stocks should be available at time t. c_t^j only includes the common stocks that has corresponding preferred stock at time t. We calculate the cumulative daily log returns of c_t^j and p_t^j starting from t_0 as follows.

$$r_{c_t}^j = \ln\left(c_t^j / c_{t_0}^j\right) \quad (2.1)$$

$$r_{p_t}^j = \ln\left(p_t^j / p_{t_0}^j\right) \quad (2.2)$$

We collected the prices of 246 stocks having both common and preferred categories from January 1st 2000 to June 30th 2016, which amounts to 4304 days, and averaged them. Note that t_0 indicates the January 1st 2000 for the stocks available on January 1st 2000, otherwise t_0 is the date when preferred stock is issued. We denote the averages of $r_{c_t}^j$ and $r_{p_t}^j$ as C_t and P_t, respectively.

$$C_t = \frac{\sum_{j=1}^{A_t} r_{c_t}^j}{A_t} \quad t = 1,2,\dots\dots 4304 \qquad (2.3)$$

$$P_t = \frac{\sum_{j=1}^{A_t} r_{p_t}^j}{A_t} \quad t = 1,2,\dots\dots 4304 \qquad (2.4)$$

$$CPS_t = C_t - P_t \quad t = 1,2,\dots\dots 4304 \qquad (2.5)$$

where A_t is the number of common and preferred stock pairs available at time t. We define CPS_t as the spread index between C_t and P_t. C_t is average cumulative daily log return of common stocks that have corresponding preferred stocks at time t and P_t is average cumulative daily log return of corresponding preferred stocks. Note that C_t is distinct from S&P500 index because C_t is constructed from the common stocks that have preferred stocks at time t. Using correlation analysis, we observe that C_t is more sensitive to stock market condition than P_t.

2.3.2. CPS Index and Future Realized Market Return.

Let $r_{t,t+20\cdot i}$ be the i month ahead realized market return of S&P500 index, assuming that investors take long position of S&P500 index at time t and clear them at time $t+20\cdot i$,

$$r_{t,t+20\cdot i} = \frac{S\&P500_{t+20\cdot i} - S\&P500_t}{S\&P500_t} \qquad i = 1,2,\dots 48$$

$$\mathrm{t} = 1,2,\dots\dots 4304 - 20\cdot i \qquad (2.6)$$

We assume there is 20 trading days in each month. The correlation between CPS_t and $r_{t,t+20\cdot i}$ be defined as follows.

$$corr_i = \frac{\langle CPS_t, r_{t,t+20\cdot i}\rangle - \overline{CPS}_t \times \bar{r}_{t,t+20\cdot i}}{\mathrm{s}(CPS_t)\times \mathrm{s}(r_{t,t+20\cdot i})} \qquad \mathrm{i} = 1,2,\dots\dots 48 \qquad (2.7)$$

where $\langle CPS_t, r_{t,t+20\cdot i}\rangle = \frac{1}{N}\sum_{t=1}^{N} CPS_t \cdot r_{t,t+20\cdot i}$,

$$\overline{CPS}_t = \frac{1}{N}\sum_{t=1}^{N} CPS_t,\ \ \bar{r}_{t,t+20\cdot i} = \frac{1}{N}\sum_{t=1}^{N} r_{t,t+20\cdot i},$$

$$\mathrm{s}(CPS_t) = \sqrt{\frac{1}{N}\sum_{t=1}^{N} {CPS_t}^2 - (\overline{CPS}_t)^2},$$

$s(r_{t,t+i\cdot 20}) = \sqrt{\frac{1}{N}\sum_{t=1}^{N} r_{t,t+20\cdot i}{}^2 - \left(\bar{r}_{t,t+20\cdot i}\right)^2}$ and N indicates the number of trading days amounting to 4304 days.

For an example, $corr_{12}$ indicates the correlation between CPS index and 12-month ahead realized return. Using $corr_i$, we show that there is a relationship between CPS_t and $r_{t,t+i\cdot 20}$. To compare CPS_t with other benchmark indexes, correlations between $r_{t,t+20\cdot i}$ and each of VIX_t (VIX index), TED_t (TED spread), $S\&P500_t$ (S&P500 index), $CAPE_t$ (Cyclically adjusted price ratio), C_t and P_t are computed. Note that VIX is CBOE volatility index to measure market sensitivity, and TED is the spread between LIBOR interest and T-Bill interest to measure credit risk.

We regress $r_{t,t+20\cdot i}$ ($i = 1,2, \ldots 48$) on CPS_t. For the comparison analysis, we also conduct univariate regression analysis using VIX_t, $TED_t, CAPE_t$, $S\&P500_t$, C_t and P_t as follows.

$$r_{t,t+i\cdot 20} = \beta_0 + \beta_1 x_t + \varepsilon_t \qquad (2.8)$$

where x_t is one of $CPS_t, VIX_t, TED_t, CAPE_t, S\&P500_t, C_t,$ and P_t.

We define the regression performance measure for $\hat{r}_{t,t+20\cdot i}$, which is the estimate of $r_{t,t+20\cdot i}$ in equation (2.8) , $\sqrt{MSE_i}/\hat{\sigma}_i$, where

$MSE_i = \frac{1}{N}\sum_{t=1}^{N}(r_{t,t+20\cdot i} - \hat{r}_{t,t+20\cdot i})^2$ and $\hat{\sigma}_i{}^2 = \frac{1}{N-1}\sum_{t=1}^{N}(r_{t,t+20\cdot i} - \bar{r}_{t,t+20\cdot i})^2$. We use a normalized measure $\sqrt{MSE_i}/\hat{\sigma}_i$ along with MSE_i because MSE_i becomes larger as *i* increases. Also we use the adjusted regression coefficient, $R_{adj}^2 = 1 - \frac{(1-R^2)(N-1)}{N-k-1}$, where *k* is the number of explanatory variables, to confirm the validity of regression.

We claim that $r_{t,t+20\cdot i}$ becomes higher when CPS_t is lower and vice versa. According to our claim, investors could make profit if they take long (short) position of stocks when CPS_t is low (high). Specifically, we find a relationship between CPS_t and $r_{t,t+20\cdot i}$ when CPS_t is higher upper threshold or lower than lower threshold.

We define the average future realized returns for high and low CPS_t, respectively, as follows.

$$r_{i,avg}^h = \frac{1}{|T^h|}\sum_{t\in T^h} r_{t,t+20\cdot i}$$

$$\text{where } T^h = \{t: CPS_t > \text{upper threshold}\} \qquad (2.9)$$

$$r_{i,avg}^l = \frac{1}{|T^l|}\sum_{t\in T^l} r_{t,t+20\cdot i}$$

$$\text{where } T^l = \{t: CPS_t < \text{lower threshold}\} \qquad (2.10)$$

In this paper, we set upper and lower thresholds to be 0.1 and -

0.2, respectively based on our empirical knowledge. This implies that we set the upper threshold as the condition where common stocks are 10% overvalued compared to their preferred stocks, and set the lower threshold as the condition where common stocks are 20% undervalued compared to their preferred stocks. We denote T^h and T^l as the sets of days when CPS_t is over the upper threshold and below the lower threshold, respectively. Note that $|T^h|$ and $|T^l|$ are the numbers of elements belong to T^h and T^l, respectively.

We claim that stock market is overvalued when CPS_t is higher than upper threshold and undervalued when CPS_t is lower than lower threshold. If our claim is correct, $r_{i,avg}^h$ becomes positive value and $r_{i,avg}^l$ becomes negative one. To check the validity of our claim in the consideration of risk measure, we use Sharpe ratios for $r_{i,avg}^h$ and $r_{i,avg}^l$ as $(r_{i,avg}^h - r_{i,f})/\sigma_i^h$ and $(r_{i,avg}^l - r_{i,f})/\sigma_i^l$, respectively.

Note that $\sigma_i^h = \sqrt{\frac{1}{|T^h|}\sum_{t\in T^h} r_{t,t+20\cdot i}{}^2 - \left(r_{i,avg}^h\right)^2}$, $\sigma_i^l = \sqrt{\frac{1}{|T^l|}\sum_{t\in T^l} r_{t,t+20\cdot i}{}^2 - \left(r_{i,avg}^h\right)^2}$, and $r_{i,f}$ is the three months US treasury bill rate at the i month ahead time.

2.3.3. Multivariate Regression Analysis

In this section, we used multivariate regression to present how the regression explains the future realized return by adding CPS index to existing index. We normalized each index with z-score normalization to compare the coefficients of each variable and analyze the p-value to decide which index is statistically significant in multivariate regression. The empirical results are present in section 2.5.3.

By comparing the adjusted r-square value of each regression model, we present that adding CPS-index as independent variable is significantly enhances the market predictability compared to adding other existing ratios.

$$r_{t,t+i\cdot 20} = \beta_0 + \beta_1 CPS_t + \beta_2 VIX_t + \beta_3 TED_t + \beta_4 CAPE_t + \beta_5 S\&P500_t + \varepsilon_t \quad (2.11)$$

In addition, detecting the presence of multicollinearity, we calculate the variance inflation factor for each index. The variance inflation factor is the ratio of variance in a model with multiple terms, divided by the variance of a model with one term alone. It quantifies the severity of multicollinearity in an ordinary least squares regression analysis. It provides an index that measures how much the variance of an estimated regression coefficient is increased because of collinearity.

To find out the best subset model for different $r_{t,t+i\cdot 20}$, we used Mallow's C_p and stepwise regression. Mallow's C_p is a technique for model selection in regression. The C_p statistics is defined as a criteria to assess fits when models with different numbers of parameters are being compared.

$$C_p = \frac{SSE_p}{MSE(X_1,\ldots,X_{P-1})} - (n - 2p) \qquad (2.12)$$

Where SSE_p is the error sum of squares for the fitted subset regression model with p parameters (i.e., with $p-1\,X$ variables). Therefore $MSE(X_1, \ldots, X_{P-1})$ is the mean square error for the regression model with every p parameters. If the model is correct, then C_p will tend to be close to or smaller than p.

In addition, using F-test, we present that adding CPS_t variables is statistically significant on multivariate regression in equation (2.11) for each of $r_{t,t+20\cdot i}$ $(i = 1,2, \ldots 48)$. In equation (2.11), F-test is assessed by null hypotheses of $H_0\colon \beta_1 = 0$. In addition, we compare the r-square of multivariate regression that includes CPS index to r-square of multivariate regression that excludes CPS index.

2.3.4. Stepwise Regression and Granger Causality Test

To improve market return predictability, we add past values of CPS_t as explanatory variables. For each of $r_{t,t+20\cdot i}$ $(i = 1,2, \dots 48)$, we use CPS_t values of one day to five days, two weeks, one month, two months, three months, six months, nine months and twelve months past as well as current one to predict $r_{t,t+20\cdot i}$ as follows.

$$r_{t,t+20\cdot i} = \text{constant} + \sum_{j\in S} \beta_j CPS_{t-j} + \varepsilon_t \qquad (2.13)$$

where $S = \{0, 1,2,3,4,5,10,20,40,60,120,180,240\}$ and ε_t is an error. We select statistically significant variables using stepwise regression approach.

For Granger causality test, we show that CPS_t variables affects the *i*-month ahead realized return.

$$r_{t,t+20\cdot i} = \text{constant} + \sum_{j=1}^{\tau} \alpha_j r_{t-j,t-j+20\cdot i} + \sum_{j=1}^{\tau} \beta_j CPS_{t-j} + \varepsilon_t \qquad (2.14)$$

In equation (2.14), the Granger Causality test is assessed by the F-test with null hypotheses of $H_0: \beta_j = 0$ for j=1,2,...., τ, where τ is the maximum time lag.

2.3.5. Neural Network and Prediction

We use neural network method to predict $r_{t,t+20\cdot i}$, $i = 3,4,\ldots,36$. Explanatory variables for each model are current index and previous ones which dates back to one day to five days, two weeks, one month, two months, three months, six months, and twelve months. For neural network model, we use first eighty five percent of data minus $20 \cdot i$ days for training and last fifteen percent of data for testing.

We need to assign a long time period including both pre-financial crisis and post-financial crisis periods. Ten hidden layers are used for our neural network model, and the average of one thousand predictions are obtained. We compare MSE and $\sqrt{\text{MSE}}/\hat{\sigma}$ of the prediction using CPS_t with those using VIX_t, TED_t, and $S\&P500_t$. In addition, we make a confusion matrix and evaluate how accurate CPS_t predicts the sign of $r_{t,t+20\cdot i}$.

2.4 Data

Thomson Reuters Datastream (www.financial.thomsonreuters.com) provides time series data of common stocks' closed prices and their preferred stocks' closed prices from January 1st 2000 to June 30th 2016 in NYSE (New York stock exchange). In addition, we use index time series data such as S&P500, TED spread, VIX from January 1st 2000 to June 30th 2016. TED spread is difference between the 3-month U.S. Libor rate and the 3-month U.S. Treasury bill rate from Federal Reserve Economics Data (FRED) database. VIX is Chicago CBOE volatility index from Chicago Board Options Exchange(CBOE). There are 246 pairs of common and preferred stocks, and 4304 trading dates.

2.5 Empirical Results

Table 2.1 presents the Pearson correlations between any pair of indexes $CPS_t, VIX_t, TED_t, CAPE_t, S\&P500_t, C_t,$ and P_t. According to Table 1, C_t is more sensitive to $S\&P500_t$ than P_t, and CPS_t is more correlated with $C_t,$ than P_t. This result supports that common stocks are more sensitive to market condition than their preferred stocks.

Table 2.1 : Pearson Correlations for each index.

	CPS_t	VIX_t	TED_t	$CAPE_t$	$S\&P500_t$	C_t	P_t
CPS_t	1.0000	-0.4567	0.4468	0.3518	0.3834	0.7848	0.2085
VIX_t		1.0000	0.1441	-0.3030	-0.4922	-0.6928	-0.5939
TED_t			1.0000	0.0934	0.0069	0.1500	-0.2148
$CAPE_t$				1.0000	0.2469	0.5419	0.5057
$S\&P500_t$					1.0000	0.3409	0.1156
C_t						1.0000	0.7685
P_t							1.0000

Fig. 2.1 shows CPS_t, $S\&P500_t, C_t,$ and P_t from January 1st 2000 to June 30th 2016 and dashed line of financial crises. There are 4304 available trading dates ($t = 1, \ldots, 4304$), in CPS_t and $S\&P500_t$ index graphs. According to Fig. 2.1, we can infer that $S\&P500_t$ is heading downward when the value of CPS_t is above of 0.1 which is the upper

threshold, and is heading upwards when the value of CPS_t is below of -0.2 which is the lower threshold. The dashed line in Fig.2.1 indicates the subprime mortgage crisis in 2008.

We empirically test our claim using data from New York Stock Exchange. We build CPS_t, $t = 1,2, \ldots \ldots 4304$, from January 1st 2000 to June 30th 2016, and observe that CPS_t is relatively high before the market crashes and relatively low before the market is recovered (See Fig. 2.1).

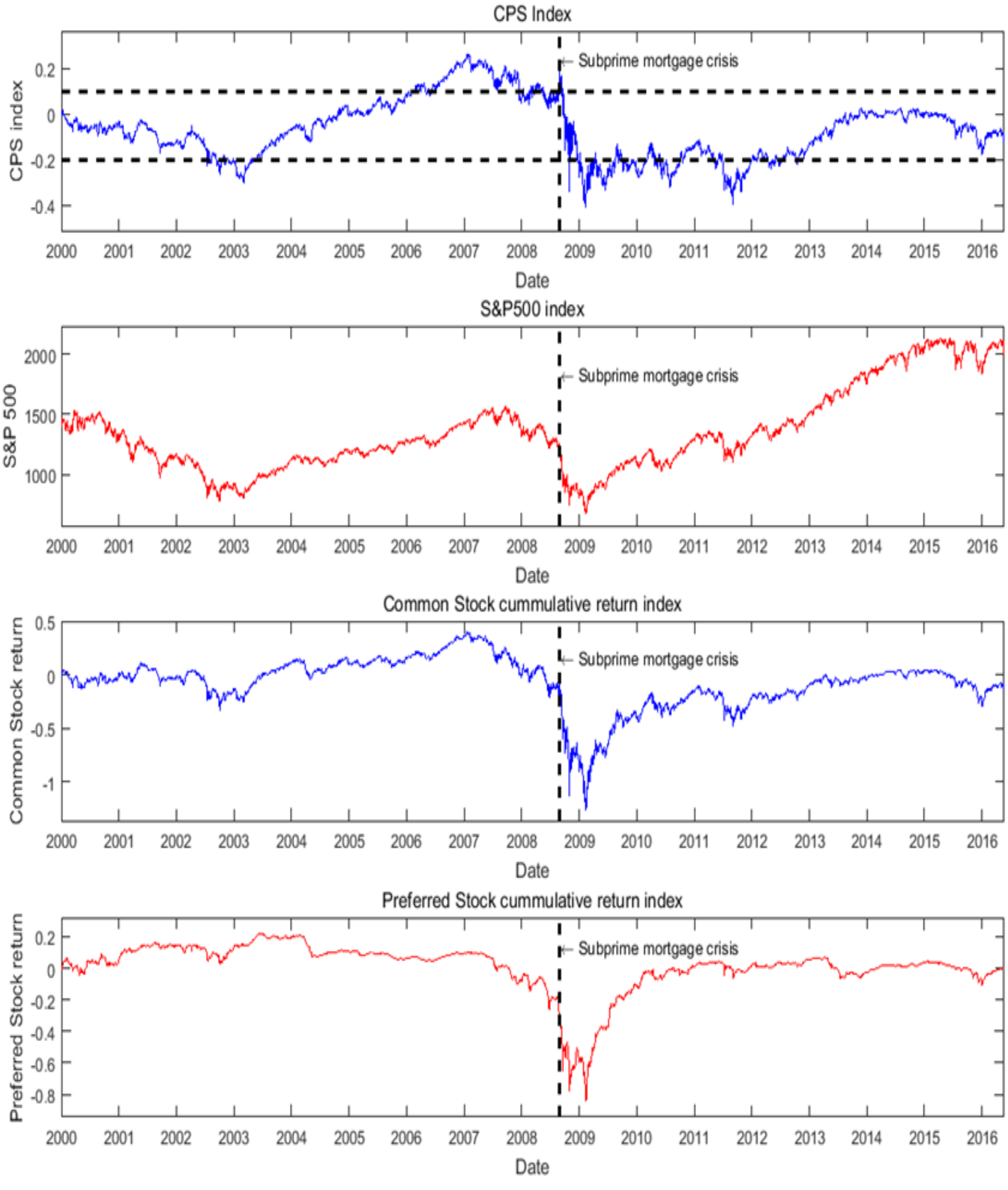

Fig. 2.1 Plots of $CPS_t, S\&P500_t, C_t,$ and P_t .

2.5.1. Correlation and Univariate Regression Results

To express the clearer relationship between CPS_t and $r_{t,t+20\cdot i}$, we draw scatter plots of $r_{t,t+20\cdot i}$ (i=12,24,36,48) against CPS_t in Fig. 2.2 In Fig.2, we also draw linear regression lines and the information about this linear regression is in Table 2.2. As the number of trading days related to $r_{t,t+20\cdot i}$ is 4304$-20 \cdot i$ days, there are 4064, 3824, 3584, and 3344 points in the scatter plots of $r_{t,t+20\cdot i}$ (i=12,24,36,48), which are one through four years ahead realized return, against CPS_t. The values of slope (β), T-stats, and R-square of each regression line are shown in the Table 2.2. We examine the univariate forecasting power of $CPS_t, VIX_t, TED_t, CAPE_t, S\&P500_t, C_t,$ and P_t for $r_{t,t+20\cdot i}$ in equation (2.8) of section 2.3.2. We claim that the high value of CPS_t predicts low value of $r_{t,t+20\cdot i}$ as high CPS_t indicates the large possibility of stock market being overvalued.

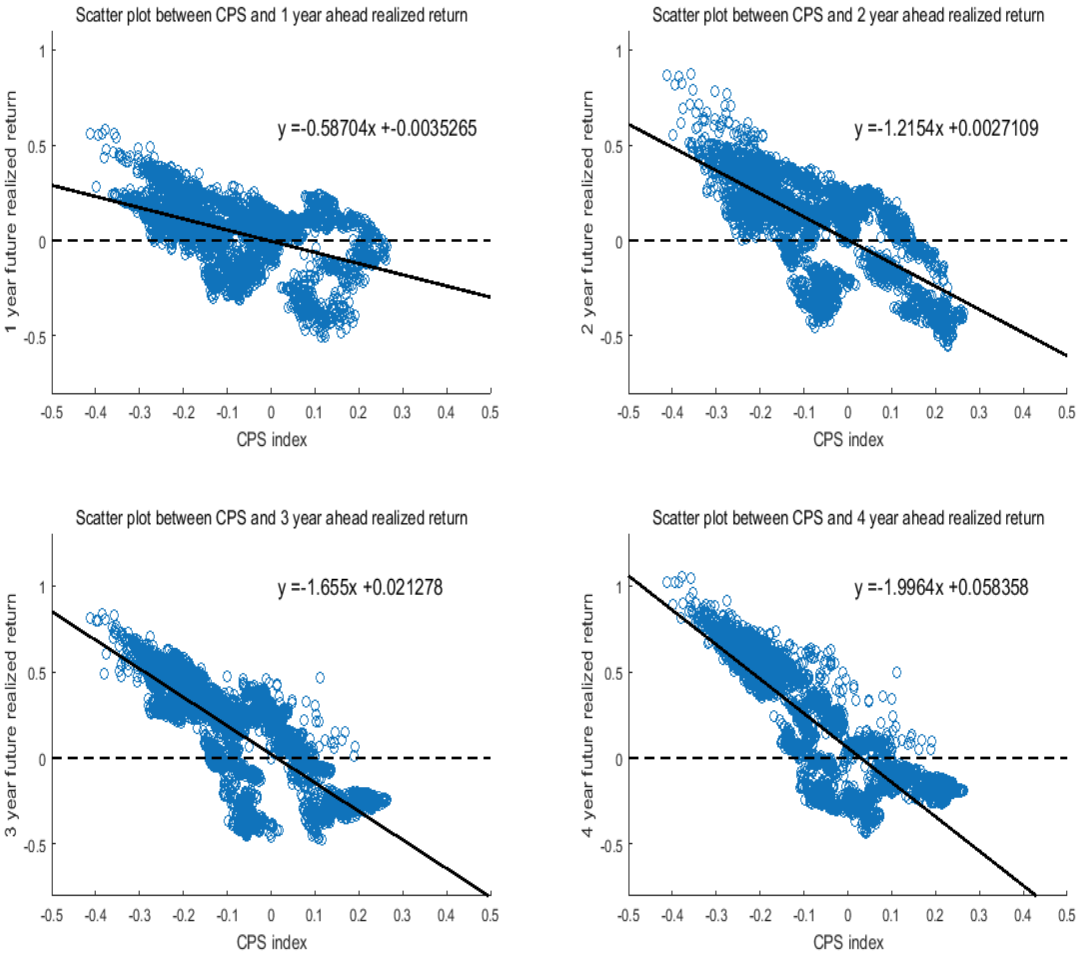

Fig. 2.2 Scatter plot of CPS and 1, 2, 3, 4 year ahead realized return

Table 2.2

Univariate regression for the CPS, VIX, TED, CAPE, S&P500, C_t and P_t.. This table summarizes single regression results in equation(2.8). The dependent variable in these regression is i month ahead realized return, and the independent variables are CPS, VIX, TED, CAPE, S&P500, C_t and P_t. Note that i is the forecasting horizon in months and β is the slope of linear regression. $\sqrt{\mathbf{MSE}}/\hat{\boldsymbol{\sigma}}$ and adj. R^2 are obtained from the linear regression.

i	$\boldsymbol{\beta}$	T-Stat	$\sqrt{\mathbf{MSE}}/\hat{\boldsymbol{\sigma}}$	adj. $\boldsymbol{R^2}$
Panel A: CPS-index				
1	-0.0645	-12.278	0.9824	0.0349
6	-0.3147	-25.899	0.9266	0.1414
12	-0.5870	-33.919	0.8985	0.1927
18	-0.8600	-42.895	0.8488	0.2795
24	-1.2154	-56.178	0.7817	0.3889
30	-1.4592	-63.917	0.7598	0.4227
36	-1.6550	-74.434	0.7029	0.5059
42	-1.8554	-84.658	0.6280	0.6056
48	-1.9964	-86.948	0.5847	0.6581
Panel B: VIX-index				
1	0.0001	1.841	0.9992	0.0015
6	0.0009	4.602	0.9971	0.0058
12	0.0023	8.003	1.0115	-0.0232
18	0.0028	7.667	1.0188	-0.0379
24	0.0052	12.301	1.0296	-0.0602
30	0.0086	18.486	1.0412	-0.0841
36	0.0088	18.387	1.0466	-0.0954
42	0.0127	25.424	0.9830	0.0337
48	0.0154	30.064	0.9438	0.1092
Panel C: TED-index				

1	-0.0175	-10.675	0.9866	0.0267
6	-0.1296	-36.644	0.8696	0.2438
12	-0.1962	-37.277	0.8800	0.2255
18	-0.2399	-37.338	0.8838	0.2188
24	-0.2654	-33.915	0.9212	0.1514
30	-0.2661	-29.097	0.9793	0.0410
36	-0.2804	-28.307	0.9822	0.0352
42	-0.3036	-27.718	0.9512	0.0952
48	-0.2958	-23.617	0.9505	0.0965
Panel D: S&P500-index				
1	-0.000006	-3.348	0.9986	0.0027
6	-0.000037	-9.725	0.9932	0.0136
12	-0.000075	-15.461	1.0069	-0.0139
18	-0.000136	-19.033	1.0022	-0.0045
24	-0.000221	-24.186	1.0068	-0.0137
30	-0.000309	-29.047	1.0295	-0.0599
36	-0.000393	-32.930	1.0240	-0.0485
42	-0.000553	-39.560	0.9716	0.0560
48	-0.000766	-44.753	0.9227	0.1487
Panel E: CAPE-index				
1	-0.00001	-7.617	0.9934	0.0131
6	-0.00005	-24.598	0.9348	0.1262
12	-0.00013	-41.132	0.8403	0.2939
18	-0.00022	-51.092	0.7757	0.3982
24	-0.00037	-58.156	0.7285	0.4693
30	-0.00054	-64.124	0.6884	0.5261
36	-0.00072	-62.210	0.6934	0.5192
42	-0.00096	-58.194	0.7111	0.4943
48	-0.00116	-54.900	0.7252	0.4740

Panel F: Common Stocks (C_t)				
1	-0.024845	-8.410	0.9920	0.0160
6	-0.152571	-22.097	0.9464	0.1043
12	-0.308802	-31.522	0.8965	0.1963
18	-0.439649	-37.993	0.8557	0.2678
24	-0.611876	-47.463	0.7933	0.3707
30	-0.720008	-51.260	0.7649	0.4150
36	-0.772181	-52.294	0.7531	0.4328
42	-0.899873	-61.259	0.6928	0.5200
48	-1.003671	-67.945	0.6481	0.5799
Panel G: Preferred Stocks (P_t)				
1	-0.004988	-1.093	0.9920	0.0000
6	-0.099640	-9.004	0.9464	0.0188
12	-0.240889	-14.773	0.8965	0.0508
18	-0.323004	-16.133	0.8557	0.0617
24	-0.433334	-18.232	0.7933	0.0798
30	-0.486050	-18.154	0.7649	0.0815
36	-0.454962	-15.760	0.7531	0.0646
42	-0.586623	-19.086	0.6928	0.0949
48	-0.720544	-22.230	0.6481	0.1286

According to Table 2.2, as i increases, the relationship between CPS_t and $r_{t,t+20\cdot i}$, becomes stronger. In addition, according to β (slope of the regression line) for each regression line indicates that the correlation between CPS_t and $r_{t,t+20\cdot i}$ is negative. This supports our claim that when CPS_t is high (low), stock market is overvalued (undervalued). Panels A-G of Table 2.2 present univariate regression results for CPS_t, VIX_t, TED_t, $CAPE_t$, $S\&P500_t, C_t$, and P_t, respectively, based on daily data from January 2000 to June 2016. In the regression analysis, $CPS_t, VIX_t, TED_t, S\&P500_t, C_t$ have negative slope coefficients, while VIX_t has positive slope coefficient. All coefficients are statistically significant at all horizons with 5% significant level in two-sided test and the corresponding p-values are close to zero. The adj R^2 of regression using CPS_t is much larger than that using each of VIX_t, TED_t, and $S\&P500_t$. In the prediction performance, other indices are inferior to CPS_t for the cases of more than 12 months ahead, implying that CPS_t has the better predictability for $r_{t,t+20\cdot i}, i \geq 12$.

To investigate the validity of CPS_t as an explanatory variable, the correlation between CPS_t and $r_{t,t+20\cdot i}$ $(i = 1,2,\ldots 48)$ is compared with those between $r_{t,t+20\cdot i}$ and each of $VIX_t, TED_t, CAPE_t, S\&P500_t, C_t$, and P_t in Fig. 3.

As shown in Figure 2.3, $CPS_t, TED_t, CAPE_t, S\&P500_t, C_t$, and P_t have a negative correlation with $r_{t,t+20\cdot i}$ while VIX_t has a positive correlation with $r_{t,t+20\cdot i}$. Correlation of CPS_t and $r_{t,t+20\cdot i}$ becomes stronger as i increases. Among $CPS_t, VIX_t, TED_t, CAPE_t, S\&P500_t, C_t$, and P_t, CPS_t is most correlated with $r_{t,t+20\cdot i}$

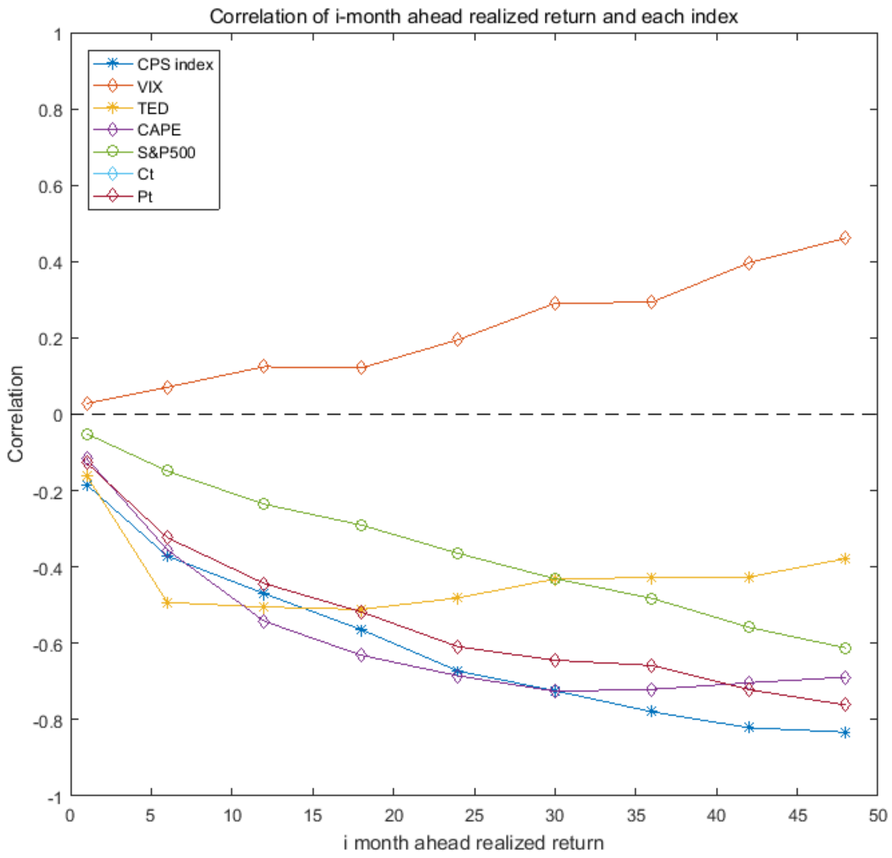

Fig 2.3 Correlation of CPS and i-month ahead realized return

2.5.2. Investment Strategy with CPS Index

In this section, we investigate what happens to future realized return when CPS_t gets higher than upper threshold in equation (2.9) or lower than lower threshold in equation (2.10). According to our claim we discussed in section 2.3.3, selected high (low) CPS_t is followed by a low (high) $r_{t,t+20 \cdot i}$.

As shown in equations (2.9) and (2.10) we select the dates when CPS_t is higher than the upper threshold of 0.1 or lower than the lower threshold of -0.2. Two threshold values of 0.1 and -0,2 are set based of our empirical tests and there is much room for the revision of threshold values in more scientific way.

In Table 2.3, we summarize $r_{i,avg}^{h}$ and $r_{i,avg}^{l}$ value for different time horizon ($i = 1,6,12,18,24,36,42,48$). Note that $r_{i,avg}^{h}$ is the average *i*-month ahead realized return when CPS_t in higher than 0.1 and $r_{i,avg}^{l}$ is the average *i*-month ahead realized return when CPS_t is lower than -0.2.

According to Table 2.3, $r_{i,avg}^{h}$ are negative values in all horizons while $r_{i,avg}^{l}$ are positive values in all horizons. It is evident that investors are likely to lose money if they take long positions when CPS_t is higher than 0.1 while investors are likely to earn money if they take long position when CPS_t is lower than -0.2. The lower level of CPS_t implies undervalued stock market condition so that there is much possibility of market improvement. In Table 3, $r_{i,avg}^{h}$ decreases until *i* is thirty months then increases afterwards. From this, we can estimate the cycle time for overvalued stock market to be undervalued is thirty months. We observe that $r_{i,avg}^{l}$ is continuously getting bigger as *i* increases.

In Table 2.4, we compute Sharpe Ratio for the $r_{i,avg}^{h}$ and $r_{i,avg}^{l}$ for different time horizon to consider volatility of future realized return. We use US three months treasury bill rate as the risk free rate at *i*-month ahead time. According to Table.4, the Sharpe Ratio for $r_{i,avg}^{h}$ is the smallest at the time horizon of forty two months ($i = 42$), and Sharpe Ratio for $r_{i,avg}^{l}$ is the largest at the time horizon of forty eight months ($i = 48$).

In Table 2.5 we set the different upper bound threshold and compute $r_{i,avg}^{h}$ and its possible loss for different time horizon. ut is upper threshold, n is the number of days when CPS index is higher than upper threshold, and $p_{i,sum}^{h}$ is the sum of profit if investor invest \$1 every time CPS index is higher than upper threshold. According to Table 2.5, $r_{i,avg}^{h}$ is lowest when the upper threshold is 0.2 and investing 24 month, and $p_{i,sum}^{h}$ is lowest when the upper threshold is 0.05 and investing 36 month. This is great opportunity to take short position and make profit. If investor investing \$100 in short position every time CPS is index is higher than 0.05 and investing 36 months, he or she could make $14,838 as a profit.

In Table 2.6 we set the different lower bound threshold and compute $r_{i,avg}^{l}$ and its possible profit for different time horizon. ul is lower threshold, n is the number of days when CPS index is lower than lower threshold, and $p_{i,sum}^{l}$ is the sum of profit if investor invest $1 every time CPS index is lower than lower threshold. According to Table 2.6, $r_{i,avg}^{l}$ is highest when the lower threshold is -0.25 and investing 48 months, and $p_{i,sum}^{h}$ is highest when the lower threshold is -0.05 and investing 48 month. This is great opportunity to take long position and make profit. If investor investing $100 in long position every time CPS is index is lower than -0.05 and investing 48 months, he or she could make $82,047 as a profit.

Table 2.3

Average future realized return of $r^h_{i,avg}$ and $r^l_{i,avg}$. in equations (2.9) and (2.10), respectively.

i	$r^h_{i,avg}$	$r^l_{i,avg}$
1	-0.0037	0.0264
6	-0.0213	0.1242
12	-0.0702	0.2071
18	-0.1526	0.2467
24	-0.2220	0.3079
30	-0.2588	0.3443
36	-0.2414	0.4178
42	-0.1825	0.5077
48	-0.1361	0.6134

Table 2.4

Sharpe Ratio for the $r^h_{i,avg}$ and $r^l_{i,avg}$ for different time horizon.

i	$r^h_{i,avg} - r_{i,f}$	$\sigma^h_{i,}$	Sharpe ratio	$r^l_{i,avg} - r_{i,f}$	$\sigma^l_{i,}$	Sharpe ratio
1	-0.0058	0.0475	-0.1219	0.0243	0.0567	0.4292
6	-0.0337	0.1237	-0.2723	0.1118	0.1041	1.0744
12	-0.0952	0.1905	-0.4997	0.1821	0.1186	1.5354
18	-0.1903	0.2205	-0.8631	0.2089	0.0924	2.2613
24	-0.2726	0.1905	-1.4313	0.2573	0.1596	1.6119
30	-0.3225	0.1343	-2.4006	0.2806	0.1368	2.0511
36	-0.3183	0.1045	-3.0448	0.3409	0.1096	3.1099
42	-0.2728	0.0844	-3.2314	0.4175	0.1130	3.6959
48	-0.2399	0.0745	-3.2192	0.5096	0.1142	4.4631

Table 2.5

Return and loss for the different upper bound for different time horizon.

ut	0.05		0.10		0.15		0.20		0.25	
n	762		524		300		159		15	
i	$r^h_{i,avg}$	$p^h_{i,sum}$	$r^h_{i,avg}$	$p^h_{i,sum}$	$r^h_{i,avg}$	$p^h_{i,sum}$	$r^h_{i,avg}$	$p^h_{i,sum}$	$r^h_{i,avg}$	$p^h_{i,sum}$
1	-0.01	-5.57	0.00	-1.95	0.00	-1.07	0.01	0.85	-0.04	-0.54
6	-0.04	-33.20	-0.02	-11.14	0.00	-0.77	0.03	4.36	0.03	0.38
12	-0.08	-61.20	-0.07	-36.78	-0.07	-21.21	-0.04	-6.04	-0.05	-0.78
18	-0.11	-83.14	-0.15	-79.97	-0.21	-62.83	-0.16	-25.14	-0.11	-1.72
24	-0.15	-113.57	-0.22	-116.34	-0.32	-96.72	-0.39	-62.53	**-0.39**	-5.88
30	-0.18	-140.91	-0.26	-135.60	-0.33	-98.61	-0.36	-56.85	-0.37	-5.48
36	-0.19	**-148.38**	-0.24	-126.50	-0.25	-76.23	-0.25	-39.50	-0.24	-3.67
42	-0.18	-134.40	-0.18	-95.62	-0.20	-60.49	-0.23	-36.44	-0.19	-2.85
48	-0.13	-96.41	-0.14	-71.30	-0.17	-50.61	-0.18	-27.89	-0.19	-2.85

Table 2.6

Return and loss for the different lower bound for different time horizon.

lt	0.05		0.10		0.15		0.20		0.25	
n	2042		1578		1219		771		316	
i	$r^l_{i,avg}$	$p^l_{i,sum}$	$r^l_{i,avg}$	$p^l_{i,sum}$	$r^l_{i,avg}$	$p^l_{i,sum}$	$r^l_{i,avg}$	$p^l_{i,sum}$	$r^l_{i,avg}$	$p^l_{i,sum}$
1	0.005	9.80	0.01	15.24	0.01	16.08	0.02	19.22	0.04	12.20
6	0.03	59.85	0.06	90.65	0.08	98.24	0.11	86.43	0.17	53.74
12	0.06	130.50	0.11	173.96	0.16	191.93	0.19	148.39	0.23	72.71
18	0.11	214.61	0.17	262.46	0.22	269.04	0.25	193.16	0.28	89.94
24	0.16	324.49	0.24	374.68	0.29	358.48	0.33	251.89	0.39	124.01
30	0.21	430.36	0.30	480.03	0.36	436.63	0.38	295.03	0.44	139.71
36	0.27	555.06	0.37	580.67	0.43	526.23	0.46	355.56	0.53	167.69
42	0.34	702.20	0.44	695.60	0.50	612.31	0.54	415.28	0.63	198.52
48	0.40	**820.47**	0.51	800.49	0.58	701.04	0.62	475.04	**0.71**	223.24

2.5.3. Multivariate Regression Analysis Results

In this section, we use multivariate regression and F- Test that we discuss in section 2.3.3 to find out significant CPS_t and its relationship with $r_{t,t+20\cdot i}$. In addition, using Variance inflation factor, we show that CPS_t does not have collinearity with any other indexes.

Table 2.7 present multivariate regression results using CPS_t variable and existing variables. The p-value for CPS index is all zero for the regression of $r_{t,t+20\cdot i}$, $i = 1,6,12,18,24,30,36,42,48$; we can proclaim that CPS_t variables is statistically significant for all different time lags.

Table 2.7

Multivariate regression for the CPS, VIX, TED, CAPE and S&P500. This table summarizes multivariate regression results in equation.(2.11). The dependent variable in these regression is *i* month ahead realized return, and the independent variables are CPS, VIX, TED, CAPE and S&P500. Note that *i* is the forecasting horizon in months. *Coeff.* and *p-value* are obtained from the multivariate regression.

i		*Intercept*	CPS_t	VIX_t	TED_t	$CAPE_t$	$S\&P500_t$	adj. R^2
1	*Coeff.*	0.0028	-0.0066	-0.0022	-0.0038	-0.0033	0.0000	0.0457
	p-value	**0.0000**	**0.0000**	0.0183	**0.0000**	**0.0000**	0.9944	
6	*Coeff.*	0.0167	-0.0065	0.0002	-0.0495	-0.0318	-0.0046	0.3447
	p-value	**0.0000**	**0.0014**	0.9318	**0.0000**	**0.0000**	**0.0060**	
12	*Coeff.*	0.0356	-0.0214	-0.0077	-0.0667	-0.0771	-0.0095	0.5154
	p-value	**0.0000**	**0.0000**	**0.0020**	**0.0000**	**0.0000**	**0.0000**	
18	*Coeff.*	0.0602	-0.0591	-0.0328	-0.0630	-0.1106	-0.0124	0.6490
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	
24	*Coeff.*	0.0886	-0.1124	-0.0423	-0.0467	-0.1362	-0.0103	0.7501
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0001**	
30	*Coeff.*	0.1164	-0.1439	-0.0259	-0.0330	-0.1579	-0.0002	0.8030
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.9291	
36	*Coeff.*	0.1459	-0.1877	-0.0423	-0.0144	-0.1610	-0.0010	0.8562
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.6993	
42	*Coeff.*	0.1781	-0.2062	-0.0018	-0.0249	-0.1604	0.0155	0.8746
	p-value	**0.0000**	**0.0000**	0.5049	**0.0000**	**0.0000**	**0.0000**	
48	*Coeff.*	0.2047	-0.2372	0.0098	-0.0103	-0.1657	0.0241	0.8777
	p-value	**0.0000**	**0.0000**	**0.0008**	**0.0004**	**0.0000**	**0.0000**	

Note: *Coeff.* represents the beta coefficient of each variable.

Bold in *p-value* indicates the 1% significance.

Table 2.8

Variance Inflation Factor for each index in equation.(2.11).

Variable	$(VIF)_{variable}$
CPS	2.0577
VIX	1.7921
TED	1.5471
CAPE	1.1813
S&P500	1.3863

Table 2.8 present the Variance inflation factor for the each variables in multivariate regression in equation (2.11). According to the results, all the variance inflation factors are less than 5; therefore, it present there is no multicollinearity problem in this multivariate regression model.

Table 2.9

C_p and stepwise regression results for the multivariate regression in equation (2.11) .

i	C_p	Selected beta by C_p	Selected beta by stepwise regression	adj. R^2
1	4.00	$\beta_1, \beta_2, \beta_3, \beta_4$	$\beta_1, \beta_2, \beta_3, \beta_4$	0.0460
6	4.01	$\beta_1, \beta_3, \beta_4, \beta_5$	$\beta_1, \beta_3, \beta_4, \beta_5$	0.3448
12	6	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	0.5154
18	6	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	0.6490
24	6	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	0.7501
30	4.01	$\beta_1, \beta_2, \beta_3, \beta_4$	$\beta_1, \beta_2, \beta_3, \beta_4$	0.8031
36	4.15	$\beta_1, \beta_2, \beta_3, \beta_4$	$\beta_1, \beta_2, \beta_3, \beta_4$	0.8563
42	4.44	$\beta_1, \beta_3, \beta_4, \beta_5$	$\beta_1, \beta_3, \beta_4, \beta_5$	0.8746
48	6	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	$\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$	0.8777

Table 2.9 present the C_p and stepwise regression results for the multivariate regression in equation (2.11). In table 2.9, we present the selected subset when C_p is less than p; otherwise we use entire variables for the multivariate regressions. For the regression of $r_{t,t+20\cdot i}$, $i = 1,6,12,18,24,30,36,42,48,$ we select variables, which are statistically significant. Among the variables, mostly selected β in equation (2.11) are $\beta_1, \beta_3, \beta_4$, which correspond to CPS_t, TED_t and $CAPE_t$. In addition, we present stepwise regression results in Table 2.7. According to stepwise regression results, mostly selected β in equation (2.11) are also $\beta_1, \beta_3, \beta_4$, which correspond to CPS_t, TED_t and $CAPE_t$.

Table 2.10

F-test for multivariate regression adding CPS-index to the model.

i	**F-statistics**	**P-value**
1	150.741	0.000
6	50.124	0.000
12	133.090	0.000
18	1839.968	0.000
24	3155.954	0.000
30	4085.442	0.000
36	5540.457	0.000
42	7167.019	0.000
48	7560.027	0.000

Table 2.10 presents the F-test results of adding CPS index to multivariate regression is statistically significant. According to the results, p-value for the F-test is all zero in all different time lags. Therefore, we can proclaim that adding CPS-index to currently existing indexes is statistically significant.

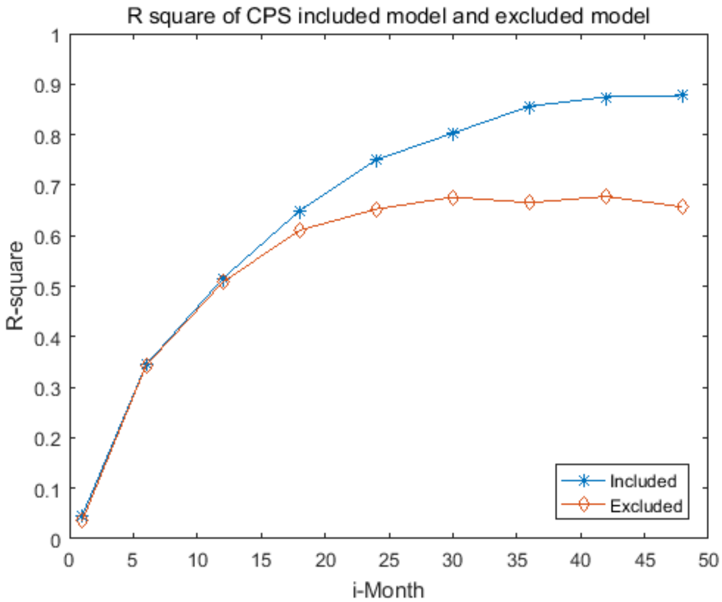

Fig.2.4 R-square of CPS included model (CPS, VIX, TED, CAPE and S&P500 as independent variables) and excluded model (VIX, TED, CAPE and S&P500 as independent variables)

Figure 2.4 presents the R-square of multivariate regression that CPS index is included and CPS index is excluded. Included model have CPS, VIX, TED, CAPE and S&P500 as independent variables; excluded model have VIX, TED, CAPE and S&P500 as independent variables.

According to the figure, multivariate regression that include CPS index as independent variable has higher R-square value compare to the multivariate regression that exclude CPS index as independent variable when the i-month is bigger than 18 months in dependent variable $r_{t,t+20\cdot i}$.

2.5.4. Stepwise Regression and Granger Causality Test Results

In this section, we use stepwise regression and Granger Causality Test that we discuss in section 2.3.4 to find out significant time lag variables for CPS_t and its relationship with $r_{t,t+20\cdot i}$. In addition, using Granger Causality test, we show that CPS_t affects $r_{t,t+20\cdot i}$ in different time horizon of i-months. Table 2.11 present stepwise regression results for the selection of CPS_t variables, based on daily data from January 2000 to June 2016. The adj.R^2 value goes as high as 0.75 in the case of $r_{t,t+960}$. For the regression of $r_{t,t+20\cdot i}$, $i = 1,6,12,18,24,30,36,42,48$, we select CPS_t variables, which are statistically significant. For an example in Table 2.9 β_0, β_{60}, β_{180}, β_{240}, which correspond to CPS_t, CPS_{t-60}, CPS_{t-180}, CPS_{t-240}, are statistically significant to predict twelve months ahead realized return, $r_{t,t+240}$. Among time lag variables, mostly selected β in equation (2.13) is β_0, which corresponds to CPS_t.

Table 2.12 shows the Granger Causality test result for CPS_t

variables in equation (2.14). As F-statistics are bigger than critical values in all time horizon, CPS_t variables affect $r_{t,t+20\cdot i}$. The significance level of this test is 0.05, and we set τ to be 1 and 10.

Table 2.11

Step-wise regression for CPS index. This table summarizes step-wise regression results in equation (2.11). The dependent variable in these regression is i month future realized return, and the independent variables are CPS-index. Selected $\boldsymbol{\beta}$s are statistically significant and determined using step-wise regression. $\sqrt{\mathbf{MSE}}/\hat{\boldsymbol{\sigma}}$ and adj. R^2 are computed for the step-wise regression.

i	Selected beta	$\sqrt{\mathbf{MSE}}/\hat{\boldsymbol{\sigma}}$	adj. $\boldsymbol{R^2}$
1	β_0, β_{60}, β_{180}, β_{240}	0.9772	0.0450
6	β_0, β_{120}, β_{180}, β_{240}	0.9153	0.1623
12	β_0, β_{60}, β_{180}, β_{240}	0.8486	0.2799
18	β_0, β_{180}, β_{240}	0.7918	0.3731
24	β_0, β_{20}, β_{180}, β_{240}	0.6965	0.5149
30	β_0, β_3, β_{60}, β_{120}, β_{240}	0.6054	0.6335
36	β_0, β_2, β_{120}, β_{240}	0.5309	0.7181
42	β_0, β_3, β_{60}, β_{120}, β_{240}	0.4994	0.7506
48	β_0, β_2, β_{180}	0.4972	0.7528

Table 2.12

Granger Causality test of CPS-index for the i month ahead realized return τ is the maximum time lag in equation (2.12)

i	F-statistics	Critical Value
$\tau = 1$		
1	33.999	3.844
6	29.703	3.844
12	38.693	3.844
18	35.881	3.844
24	39.945	3.844
30	29.123	3.844
36	39.708	3.844
42	43.926	3.844
48	40.932	3.844
$\tau = 10$		
1	588.506	2.607
6	719.435	2.607
12	791.519	2.607
18	525.085	2.216
24	457.803	2.216
30	473.231	2.217
36	575.777	2.374
42	499.825	2.217
48	617.137	2.375

2.5.5. Neural Network Prediction Results

We use ten hidden layers to reduce over fitting risk and computational time. Neural network performs better than linear regression for the prediction of $r_{t,t+20\cdot i}$.

Table 2.13 presents neural network results for the training set of $CPS_t, VIX_t, TED_t, CAPE_t$ and $S\&P500_t$, respectively, using daily data from January 2000 to June 2016. For each prediction of $r_{t,t+20\cdot i}$, we repeat prediction a thousand times and average the resulting MSEs. According to our results, in long term time horizon, the MSE and $\sqrt{\text{MSE}}/\hat{\sigma}$ of CPS_t is much smaller than those of VIX_t, TED_t, $CAPE_t$ and $S\&P500_t$. In the prediction of $r_{t,t+720}$ using CPS_t, $\sqrt{\text{MSE}}/\hat{\sigma}$ decrease to 0.4639 while the minimum values of $\sqrt{\text{MSE}}/\hat{\sigma}$ for VIX_t , TED_t , $CAPE_t$, and $S\&P500_t$ are 0.6371, 0.5771, 0.5073 and 1.986, respectively. Note that $\sqrt{\text{MSE}}/\hat{\sigma}$ for CPS_t becomes smaller than those for VIX_t, $TED_t, CAPE_t$ and $S\&P500_t$ in the prediction of $r_{t,t+20\cdot i}, i \geq 30$

Table 2.14 presents the confusion matrix for the prediction accuracy of Neural Network model using CPS_t in terms of the of prediction direction (sign). Note that $\hat{r}_{t,t+i\cdot 20} > 0$ ($\hat{r}_{t,t+i\cdot 20} < 0$) indicates prediction is positive (negative) and $r_{t,t+i\cdot 20} > 0$ ($r_{t,t+i\cdot 20} < 0$) represents actual return is positive (negative). Performance of CPS_t

with different time horizons is shown in table 2.14. Figure 2.5 shows that CPS_t predicts the sign of $r_{t,t+20\cdot i}$ accurately when i is bigger than twelve months ($i \geq 12$).

Table 2.13

Neural network for the CPS, VIX, TED, CAPE and S&P500 index. The dependent variable is i month ahead realized return, and the explanatory variables are CPS, VIX, TED, CAPE and S&P500 separately. MSE and $\sqrt{\text{MSE}}/\hat{\sigma}$ is computed from the test set of neural network.

i	**MSE**				
	CPS-index	**VIX-index**	**TED-index**	**CAPE-index**	**S&P500-index**
3	0.0056	0.0027	0.0022	0.0024	0.0369
6	0.0105	0.0050	0.0041	0.0034	0.0768
9	0.0168	0.0087	0.0074	0.0056	0.1495
12	0.0196	0.0136	0.0118	0.0079	0.3172
15	0.0198	0.0231	0.0176	0.0113	0.6298
18	0.0240	0.0292	0.0228	0.0137	0.6731
21	0.0268	0.0301	0.0308	0.0161	0.6096
24	0.0241	0.0310	0.0450	0.0173	0.5463
27	0.0233	0.0333	0.0457	0.0189	0.4051
30	0.0199	0.0369	0.0388	0.0210	0.3434
33	0.0172	0.0425	0.0339	0.0192	0.2747
36	0.0153	0.0411	0.0313	0.0183	0.2808
	$\sqrt{\textbf{MSE}}/\hat{\boldsymbol{\sigma}}$				
	CPS-index	**VIX-index**	**TED-index**	**CAPE-index**	**S&P500-index**
3	0.9717	0.6808	0.6132	0.6330	2.5039
6	0.9192	0.6371	0.5771	0.5195	2.4853
9	0.9214	0.6616	0.6099	0.5308	2.7496
12	0.8633	0.7192	0.6712	0.5470	3.4749
15	0.7802	0.8411	0.7348	0.5881	4.3957
18	0.7848	0.8652	0.7647	0.5922	4.1557
21	0.7613	0.8059	0.816	0.5905	3.6293
24	0.6736	0.7646	0.9215	0.5710	3.2095
27	0.628	0.7502	0.8791	0.5654	2.6163
30	0.5631	0.7663	0.7854	0.5779	2.3369
33	0.5129	0.8049	0.7191	0.5415	2.0474
36	0.4639	0.7601	0.6632	0.5073	1.986

Table 2.14.

Confusion matrix for the sign of prediction periods in Neural Netwrok with CPS-index

i=3	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$	i=6	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$	i=9	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$
$\hat{r}_{t,+20\cdot i} > 0$	119	4	$\hat{r}_{t,+20\cdot i} > 0$	210	3	$\hat{r}_{t,+20\cdot i} > 0$	327	16
$\hat{r}_{t,+20\cdot i} < 0$	303	176	$\hat{r}_{t,+20\cdot i} < 0$	207	173	$\hat{r}_{t,+20\cdot i} < 0$	93	148
i=12	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$	i=15	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$	i=18	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$
$\hat{r}_{t,+20\cdot i} > 0$	396	87	$\hat{r}_{t,+20\cdot i} > 0$	443	116	$\hat{r}_{t,+20\cdot i} > 0$	500	57
$\hat{r}_{t,+20\cdot i} < 0$	23	69	$\hat{r}_{t,+20\cdot i} < 0$	3	4	$\hat{r}_{t,+20\cdot i} < 0$	0	0
i=21	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$	i=24	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$	i=27	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$
$\hat{r}_{t,+20\cdot i} > 0$	507	41	$\hat{r}_{t,+20\cdot i} > 0$	535	4	$\hat{r}_{t,+20\cdot i} > 0$	529	1
$\hat{r}_{t,+20\cdot i} < 0$	0	0	$\hat{r}_{t,+20\cdot i} < 0$	0	0	$\hat{r}_{t,+20\cdot i} < 0$	0	0
i=30	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$	i=33	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$	i=36	$r_{t,t+20\cdot i} > 0$	$r_{t,t+20\cdot i} < 0$
$\hat{r}_{t,+20\cdot i} > 0$	521	0	$\hat{r}_{t,+20\cdot i} > 0$	512	0	$\hat{r}_{t,+20\cdot i} > 0$	503	0
$\hat{r}_{t,+20\cdot i} < 0$	0	0	$\hat{r}_{t,+20\cdot i} < 0$	0	0	$\hat{r}_{t,+20\cdot i} < 0$	0	0

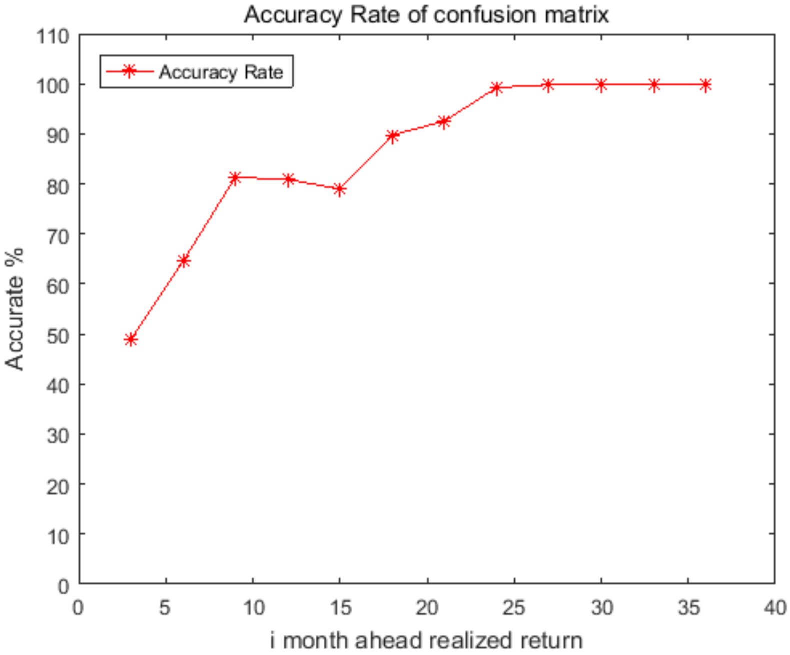

Fig.2.5 Accuracy rate of confusion matrix of neural network with CPS-index

2.6. Conclusion

This paper has shown that the return spread between common stocks and their preferred stocks could be used to assess the stock market condition and predict the future market return. We reach several conclusions as follows. First, using pairs of common stocks and preferred stocks, we built CPS-index that measures the spread between the cumulative returns of common stocks and their preferred stocks in market. We investigated whether this CPS-index determines either of the overvaluation or the undervaluation of common stocks against corresponding preferred stocks. Second, we showed that there is strong negative relationship between CPS-index and future realized return of stock market, while the relationship between the future realized return and any of VIX, TED, S&P500 is not so significant; among the existing indexes, only CAPE index has strong negative relationship. Third, the correlation, MSE and adjusted R squared value of the regression between CPS-index and future realized return show that CPS-index has a significant power to explain the future realized market return in 21 months or up to 48 months ahead of time. Furthermore, we showed that the average future realized return increases when CPS-index is below the lower threshold of -0.2, and decreases when CPS-index is above the upper threshold of 0.1. In addition, we estimated the cycle time for overvalued stock market to be undervalued is thirty months. This implies that if CPS-index is higher than 0.1, the investor should take a short position and wait for thirty months to maximize the profit. Forth, using multivariate regression and variance inflation factor, we present that adding CPS-index to currently existing indexes is statistically significant to enhance the market predictability and there is no multic ollinearity problem using CPS index and other existing indexes. In addition, using C_p and stepwise regression to present mostly selected β in equation (2.11) are $\beta_1, \beta_3, \beta_4$, which correspond to CPS_t, TED_t and $CAPE_t$. Fifth, we confirmed that past and current CPS-index affects future realized market return using stepwise regression and Granger Causality test. Lastly, we applied neural network to predict the realized market return and found that CPS-index provides better prediction results in any time horizon of more than twenty seven months.

There are some limits in our research on CPS-index. If we analyze longer period of data than sixteen years from 2000 to 2016 used in this paper, our results might be more robust. As we applied CPS-index to the US market data only, we need to investigate whether CPS-index has the equivalent prediction power in other global markets such as Europe and China. Also the appropriate upper and lower threshold values of CPS-index need to be studied in more scientific way to maximize the profit of our investment strategy. A portfolio consisting of pairs of common stock and its preferred stock can be developed using CPS-index and verified in comparison with other benchmark portfolios.

Chapter 3

Empirical Analysis of Common Preferred Spread Index

3.1 Introduction

In both academic researchers and financial practitioners have sought to develop variables that highly related to stock market returns. They have looked at various metrics as future predictors of share price performance. There has been a lot of research on predicting future market return based on the usefulness of valuation ratios such as dividend-to-price ratio, book-to-market ratio, earnings-to-price ratio and payout yield (Boudoukh et al., 2007; Campbell, 1987; Campbell & Shiller, 1988; Eugene F Fama & French, 1988; Kothari & Shanken, 1997). According to previous research, size, interest rate, beta, price to book ratio (PBR), price to earning ratio (PER), price to operation earning ratio(OPER) and dividend yield to be good indicators among others. Intuitively, PER, OPER, and the PBR could be viewed as a multiple that the market attaches to earnings, operational earnings and book value respectively. Thus, it is possible to make profitable trading strategies based on earnings, operational earning and book value. It is effective with existing variables; however, these variables could enhance market predictive power using with new developed index.

In this chapter, we empirically present how CPS-index works with other existing valuation ratios. In previous chapter we compared the market predictability with VIX (CBOE volatility index) index, TED index (the spread between three-month LIBOR interest rate and three-month US treasury bill interest rate) because when we first build CPS-index we thought as risk measure index. Therefore, we compared with existing risk indexes that have market predict power. However, after we build CPS-index we found that it has market predict power and we realized it has potential ability to explain future realized return. Therefore, in this chapter, we focused on how CPS-index works with other existing valuation ratios that known as market predict power such as PER (Price to earning ratio), PBR(price to book ratio), price to operation earning ratio(OPER) to explain future realized return. In addition, we developed the CPS-index by finding parameter for the CPS-index.

First, we evaluate the explanation power of CPS-index about the future realized market return in comparison with other benchmark competitors such as PER (Price to earning ratio), PBR(price to book ratio), price to operation earning ratio(OPER) and market return itself, S&P500. To assess the explanation power, we compute correlations and conduct univariate regression analysis between the S&P500 future realized return and each of CPS-index and other bench mark indexes. Second, we used multivariate regression to present how explanatory power of each regression improved by adding CPS-index as independent variable. We present that adding CPS-index as independent variable to each regression significantly increase the explanatory power to predict future realized return. Lastly, we used moving window method to find optimal parameter for the CPS-index. The previous CPS-index is highly depends on the starting day of the CPS-index. If the spread between common stocks and their preferred stock were large at starting day of the CPS-index, the entire CPS index would have comparably low value; on the other hand if the spread between common stocks and their preferred stock were small at starting day of the CPS-index, the entire CPS index would have comparably high value. Therefore, using moving window method we try to find the optimal parameter for the CPS index. In addition, using Granger causality test, we show that developed CPS-index affects the market future realized return.

3.2 Method

3.2.1. Empirical Analysis of Spread between Common Stocks and their Preferred Stocks

We used CPS method from the previous Chapter. Let c_t^j and p_t^j be the observed prices of common stock j and its accompanied preferred stock at time t, respectively; both stocks should be available at time t. c_t^j only includes the common stocks that has corresponding preferred stock at time t. We calculate the cumulative daily log returns of c_t^j and p_t^j starting from t_0 as follows.

$$r_{c_t}^j = \ln(c_t^j / c_{t_0}^j) \quad (3.1)$$

$$r_{p_t}^j = \ln(p_t^j / p_{t_0}^j) \quad (3.2)$$

We collected the prices of 246 stocks having both common and preferred categories from January 1st 2000 to June 30th 2016, which amounts to 4304 days, and averaged them. Note that t_0 indicates the January 1st 2000 for the stocks available on January 1st 2000, otherwise t_0 is the date when preferred stock is issued. We denote the averages of

$r_{c_t}^j$ and $r_{p_t}^j$ as C_t and P_t, respectively.

$$C_t = \frac{\sum_{j=1}^{A_t} r_{c_t}^j}{A_t} \qquad \text{t} = 1,2, \ldots \ldots 4304 \qquad (3.3)$$

$$P_t = \frac{\sum_{j=1}^{A_t} r_{p_t}^j}{A_t} \qquad \text{t} = 1,2, \ldots \ldots 4304 \qquad (3.4)$$

$$CPS_t = C_t - P_t \qquad \text{t} = 1,2, \ldots \ldots 4304 \qquad (3.5)$$

where A_t is the number of common and preferred stock pairs available at time t. We define CPS_t as the spread index between C_t and P_t. C_t is average cumulative daily log return of common stocks that have corresponding preferred stocks at time t and P_t is average cumulative daily log return of corresponding preferred stocks. Note that C_t is distinct from S&P500 index because C_t is constructed from the common stocks that have preferred stocks at time t. Using correlation analysis, we observe that C_t is more sensitive to stock market condition than P_t.

3.2.2. Empirical Analysis of CPS Index and Future Realized Market Return.

Let $r_{t,t+20\cdot i}$ be the i month ahead realized market return of S&P500 index, assuming that investors take long position of S&P500 index at time t and clear them at time $t + 20 \cdot i$,

$$r_{t,t+20\cdot i} = \frac{S\&P500_{t+20\cdot i} - S\&P500_t}{S\&P500_t}$$

$$i = 1,2, \dots 48 \qquad \mathrm{t} = 1,2, \dots\dots 4304 - 20 \cdot i \qquad (3.6)$$

We assume there is 20 trading days in each month. The correlation between CPS_t and $r_{t,t+20\cdot i}$ be defined as follows.

$$corr_i = \frac{\langle CPS_t, r_{t,t+20\cdot i}\rangle - \overline{CPS}_t \times \bar{r}_{t,t+20\cdot i}}{\mathrm{s}(CPS_t) \times \mathrm{s}(r_{t,t+20\cdot i})} \qquad \mathrm{i} = 1,2, \dots\dots 48 \qquad (3.7)$$

where $\langle CPS_t, r_{t,t+20\cdot i}\rangle = \frac{1}{N}\sum_{t=1}^{N} CPS_t \cdot r_{t,t+20\cdot i}$, $\overline{CPS}_t = \frac{1}{N}\sum_{t=1}^{N} CPS_t$, $\bar{r}_{t,t+20\cdot i} = \frac{1}{N}\sum_{t=1}^{N} r_{t,t+20\cdot i}$, $\mathrm{s}(CPS_t) = \sqrt{\frac{1}{N}\sum_{t=1}^{N} {CPS_t}^2 - (\overline{CPS}_t)^2}$, $\mathrm{s}(r_{t,t+i\cdot 20}) = \sqrt{\frac{1}{N}\sum_{t=1}^{N} {r_{t,t+20\cdot i}}^2 - \left(\bar{r}_{t,t+20\cdot i}\right)^2}$ and N indicates the number of trading days amounting to 4304 days.

For an example, $corr_{12}$ indicates the correlation between CPS index and 12-month ahead realized return. Using $corr_i$, we show that there is a relationship between CPS_t and $r_{t,t+i\cdot 20}$. To compare CPS_t with other benchmark indexes, correlations between $r_{t,t+20\cdot i}$ and each of PBR_t (Price to book ratio), PER_t (Price to earning ratio), $OPER_t$ (price to operation earning ratio), and $S\&P500_t$ (S&P500 index) computed. Note that PBR is price to book ratio which to measure market conditions we used quarterly book value and daily S&P500 index to compute daily price to book ratio for the S&P500; PER is price to earnings ratio which to decide whether market conditions is overvalued or under we used quarterly book value and daily S&P500 index to compute daily price to book ratio for the S&P500. OPER is price to operational earnings ratio which to decide whether market conditions is overvalued or under we used quarterly book value and daily S&P500 index to compute daily price to book ratio for the S&P500.

We regress $r_{t,t+20\cdot i}$ $(i = 1,2,\dots 72)$ on CPS_t. For the comparison analysis, we also conduct univariate regression analysis using CPS_t PBR_t PER_t, $OPER_t$, and $S\&P500_t$

$$r_{t,t+i\cdot 20} = \beta_0 + \beta_1 x_t + \varepsilon_t \qquad (3.8)$$

where x_t is one of CPS_t, PER_t, PBR_t, $OPER_t$, and $S\&P500_t$

We define the regression performance measure for $\hat{r}_{t,t+20\cdot i}$, which is the estimate of $r_{t,t+20\cdot i}$ in equation (3.8) , $\sqrt{MSE_i}/\hat{\sigma}_i$, where $MSE_i = \frac{1}{N}\sum_{t=1}^{N}(r_{t,t+20\cdot i} - \hat{r}_{t,t+20\cdot i})^2$ and $\hat{\sigma}_i{}^2 = \frac{1}{N-1}\sum_{t=1}^{N}(r_{t,t+20\cdot i} - \bar{r}_{t,t+20\cdot i})^2$. We use a normalized measure $\sqrt{MSE_i}/\hat{\sigma}_i$ along with MSE_i because MSE_i becomes larger as i increases. Also we use the adjusted regression coefficient, $R_{adj}^2 = 1 - \frac{(1-R^2)(N-1)}{N-k-1}$, where k is the number of explanatory variables, to confirm the validity of regression.

3.2.3. Empirical Analysis of Multivariate Regression, Variance Inflation Factor and F-test

In this section, we used multivariate regression to present how the regression explain the future realized return by adding CPS index to existing ratio and index. By comparing the adjusted r-square value of each regression model, we present that adding CPS-index as independent variable is significantly enhance the market predictability compared to adding other existing ratios. In addition, detecting the presence of multicollinearity, we calculate the variance inflation factor for each index. The variance inflation factor is the ratio of variance in a model with multiple terms, divided by the variance of a model with one term alone. It quantifies the severity of multicollinearity in an ordinary least squares regression analysis. It provides an index that measures how much the variance of an estimated regression coefficient is increased because of collinearity.

$$r_{t,t+i\cdot 20} = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \varepsilon_t \qquad (3.9)$$

In equation (3.9), the $x_{1,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$, and $S\&P500_t$ and $x_{2,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$ that is not used in $x_{1,t}$. Therefore, we have ten different regressions to compare those market predictability.

$$r_{t,t+i\cdot 20} = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \beta_3 x_{3,t} + \varepsilon_t \qquad (3.10)$$

In equation (3.10), the $x_{1,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$, and $S\&P500_t$ and $x_{2,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$ that is not used in $x_{1,t}$ and $x_{3,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$that is not used in $x_{1,t}$ and $x_{2,t}$. Therefore, we have ten different regressions to compare those market predictability.

$$r_{t,t+i\cdot 20} = \beta_0 + \beta_1 x_{1,t} + \beta_2 x_{2,t} + \beta_3 x_{3,t} + \beta_4 x_{4,t} + \varepsilon_t \qquad (3.11)$$

In equation (3.11), the $x_{1,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$, and $S\&P500_t$; $x_{2,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$that is not used in $x_{1,t}$; $x_{3,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$that is not used in $x_{1,t}$ and $x_{2,t}$; $x_{4,t}$ is one of CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$that is not used in $x_{1,t}$, $x_{2,t}$ and $x_{3,t}$. Therefore, we have four different regressions to compare those market predictability.

For every multivariate regression, we compute the r-square value and compare the market predictability. The main purpose of this multivariate regression is to find out adding CPS index as independent variable enhance the market predictability of the multivariate regression.

In this section, we used multivariate regression to present how the regression explains the future realized return by adding CPS index to existing index. We normalized each index with z-score normalization to compare the coefficients of each variable and analyze the p-value to decide which index is statistically significant in multivariate regression. The empirical results are present in section 3.5.3.

By comparing the adjusted r-square value of each regression model, we present that adding CPS-index as independent variable is significantly enhances the market predictability compared to adding other existing ratios.

$$r_{t,t+i\cdot 20} = \beta_0 + \beta_1 CPS_t + \beta_2 PER_t + \beta_3 PBR_t + \beta_4 OPER_t + \beta_5 S\&P500_t + \varepsilon_t \quad (3.12)$$

In addition, using F-test, we present that adding CPS_t variables is statistically significant on multivariate regression in equation (3.12) for each of $r_{t,t+20\cdot i}$ $(i = 1,2,\ldots 48)$. In equation (3.12), F-test is assessed by null hypotheses of $H_0\colon \beta_1 = 0$. In addition, we compare the r-square of multivariate regression that includes CPS index to r-square of multivariate regression that excludes CPS index.

3.2.4. Optimal Starting Point for CPS index.

In this section, we developed CPS index by finding the optimal parameter. The weakness of the CPS index is that depends on the t_0 which is January 1st 2000 in this paper, could change the CPS value. If the spread between common stocks and their preferred stock were large at t_0, the CPS index would have comparably low value; on the other hand if the spread between common stocks and their preferred stock were small at t_0, the CPS index would have comparably high value. Therefore, using moving window method we try to find the optimal parameter for the CPS index.

$$r_{c_t}^{j,m} = \ln\left(c_t^j / c_{t-m*240}^j\right) \qquad t = m \cdot 240, \dots \dots 4304 \quad (3.13)$$

$$r_{p_t}^{j,m} = \ln\left(p_t^j / p_{t-m*240}^j\right) \qquad t = m \cdot 240, \dots \dots 4304 \quad (3.14)$$

$$C_{t,m} = \frac{\sum_{j=1}^{A_t} r_{c_t}^{j,m}}{A_t} \qquad t = m \cdot 240, \dots \dots 4304 \qquad (3.15)$$

$$P_{t,m} = \frac{\sum_{j=1}^{A_t} r_{p_t}^{j,m}}{A_t} \qquad t = m \cdot 240, \dots \dots 4304 \qquad (3.16)$$

$$CPS_{t,m} = C_{t,m} - P_{t,m} \quad \mathrm{t} = \mathrm{m} \cdot 240, \dots \dots 4304 \qquad (3.17)$$

$$r_{t,t+i\cdot 20} = \beta_0 + \beta_1 x_t + \varepsilon_t \qquad (3.18)$$

The equation (3.13), (3.14), (3.15), (3.16), and (3.17) is very similar to equation from previous section except in previous section, CPS indicates the difference between cumulative log return of common stocks that have corresponding preferred stocks and those preferred stocks from January 1st 2000; however, in those equation m-year CPS index, $CPS_{t,m}$, indicates the difference between cumulative log return of common stocks that have corresponding preferred stocks and those preferred stocks from m-years ago. By finding optimal m-year, we can build robust CPS-index that is not depends on t_0. In equation (3.18) we used $CPS_{t,m}$ as independent variable and $r_{t,t+i\cdot 20}$ as dependent variables to compare the adjusted R-square for each regression with different m-years.

After finding the right parameter for the CPS-index, we used Granger Causality test to present its market predictability. For Granger causality test, we show that $CPS_{t,m}$ variables affects the *i*-month ahead realized return.

$$r_{t,t+20\cdot i} = \mathrm{Constant} + \sum_{j=1}^{\tau} \alpha_j r_{t-j,t-j+20\cdot i} + \sum_{j=1}^{\tau} \beta_j CPS_{t-j,m} + \varepsilon_t$$

(3.19)

In equation (3.19), the Granger Causality test is assessed by the F-test with null hypotheses of $H_0: \beta_j = 0$ for j=1,2,...., τ, where τ is the maximum time lag.

3.3 Data

Thomson Reuters Datastream (www.financial.thomsonreuters.com) provides time series data of common stocks' closed prices and their preferred stocks' closed prices from January 1st 2000 to June 30th 2016 in NYSE (New York stock exchange). In addition, we use index time series data such as S&P500, price to book ratio, price to earning ratio and price to operational earning ratio from January 1st 2000 to June 30th 2016.

3.4 Empirical Results

Table 3.1 presents the Pearson correlations between any pair of indexes CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$. According to Table 3.1, Price to Book Ratio and Price to Operational Earing Ratio are highly correlated. CPS-index is not correlated with other variables.

Table 3.1

Pearson Correlations for each index.

	CPS_t	PER_t	PBR_t	$S\&P500_t$	$OPER_t$
CPS_t	1	-0.3338	0.3215	0.3834	-0.0796
PER_t		1	-0.0880	-0.3464	0.4919
PBR_t			1	0.1371	0.6885
$S\&P500_t$				1	-0.1011
$OPER_t$					1

3.4.1. Empirical Results of Correlation and Univariate Regression

The values of slope (β), $\sqrt{\text{MSE}}/\hat{\sigma}$ and R-square of each regression line are shown in the Table 2. We examine the univariate forecasting power of CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$ for $r_{t,t+20\cdot i}$ in equation (3.8). We claim that the high value of CPS_t predicts low value of $r_{t,t+20\cdot i}$ as high CPS_t indicates the large possibility of stock market being overvalued.

According to Table 3.2, as i increases, the relationship between CPS_t and $r_{t,t+20\cdot i}$, becomes stronger until the i is 48 months. In addition, according to β (slope of the regression line) for each regression line indicates that the correlation between CPS_t and $r_{t,t+20\cdot i}$ is negative. This supports our claim that when CPS_t is high (low), stock market is overvalued (undervalued). Panels A-E of Table 3.2 present univariate regression results for CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$, respectively, based on daily data from January 2000 to June 2016. In the regression analysis, CPS_t, PBR_t, $OPER_t$ and $S\&P500_t$ have negative slope coefficients, while PER_t has positive slope coefficient. All coefficients are statistically significant at all horizons with 5% significant level in two-sided test and the corresponding p-values are close to zero. The adj R^2 of regression using CPS_t is larger than that using each of

PER_t, PBR_t, $OPER_t$ and $S\&P500_t$ when i is bigger than 30 months. In the prediction performance, other indices are inferior to CPS_t for the cases of more than 30 months ahead, implying that CPS_t has the better predictability for $r_{t,t+20\cdot i}, i \geq 30$.

To investigate the validity of CPS_t as an explanatory variable, the correlation between CPS_t and $r_{t,t+20\cdot i}$ $(i = 1,2, \ldots 72)$ is compared with those between $r_{t,t+20\cdot i}$ and each of CPS_t, PBR_t, $OPER_t$ and $S\&P500_t$ in Fig. 3.1.

As shown in Table 3.2, CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$ have a negative correlation with $r_{t,t+20\cdot i}$ while PER_t has a positive correlation with $r_{t,t+20\cdot i}$. Correlation of CPS_t and $r_{t,t+20\cdot i}$ becomes stronger as *i* increases. Among CPS_t, PER_t, PBR_t, $OPER_t$ and $S\&P500_t$, CPS_t is most correlated with $r_{t,t+20\cdot i}$

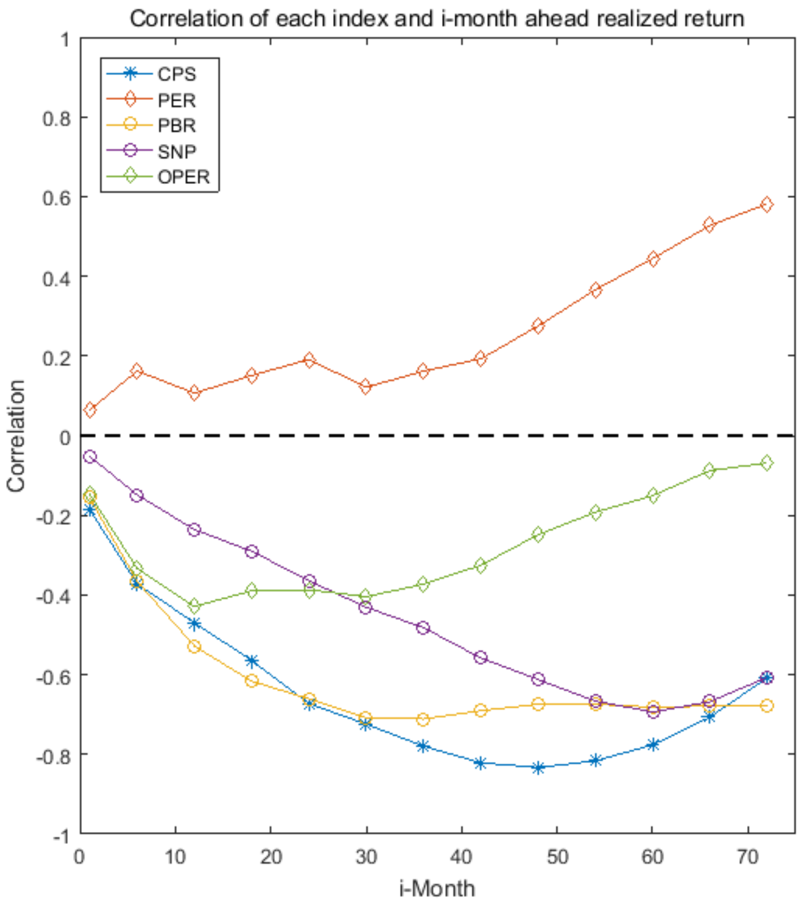

Fig. 3.1 Correlation of Each index and i-month ahead realized return

Table 3.2: Univariate regression for the CPS, PER, PBR, OPER and S&P500 index. This table summarizes single regression results in equation (3.8). The dependent variable in these regression is i month ahead realized return, and the independent variables are CPS, PER, PBR, OPER and S&P500 index. Note that i is the forecasting horizon in months and β is the slope of linear regression. $\sqrt{\mathbf{MSE}}/\hat{\boldsymbol{\sigma}}$ and adj. R^2 are obtained from the linear regression.

i	$\boldsymbol{\beta}$	$\sqrt{\mathbf{MSE}}/\hat{\boldsymbol{\sigma}}$	adj. $\boldsymbol{R^2}$
Panel A: CPS-index			
1	-0.0647	0.9824	0.0349
6	-0.3188	0.9266	0.1414
12	-0.5939	0.8985	0.1927
18	-0.8589	0.8488	0.2795
24	-1.1967	0.7817	0.3889
30	-1.4262	0.7598	0.4227
36	-1.6093	0.7029	0.5059
42	-1.8170	0.6280	0.6056
48	-1.9981	0.5847	0.6581
54	-2.1697	0.5834	0.6596
60	-2.2390	0.6221	0.6129
66	-2.1687	0.6830	0.5336
72	-1.9489	0.7625	0.4186
Panel B: PER			
1	0.00014	0.9981	0.0038
6	0.00085	0.9869	0.0261
12	0.00082	0.9944	0.0111
18	0.00142	0.9887	0.0226
24	0.00213	0.9817	0.0362
30	0.00150	0.9928	0.0144
36	0.00211	0.9869	0.0260
42	0.00268	0.9814	0.0368

48	0.00405	0.9616	0.0753
54	0.00584	0.9307	0.1338
60	0.00755	0.8960	0.1972
66	0.00914	0.8497	0.2780
72	0.01006	0.8136	0.3380

Panel C: PBR			
1	-0.01083	0.9880	0.0238
6	-0.06177	0.9307	0.1337
12	-0.13152	0.8494	0.2785
18	-0.18692	0.7875	0.3799
24	-0.23642	0.7511	0.4359
30	-0.28124	0.7066	0.5007
36	-0.29691	0.7033	0.5054
42	-0.30785	0.7233	0.4769
48	-0.32044	0.7390	0.4538
54	-0.34965	0.7388	0.4541
60	-0.38027	0.7318	0.4644
66	-0.39330	0.7346	0.4604
72	-0.39416	0.7370	0.4568

Panel D: S&P500-index			
1	-0.000006	0.9988	0.0024
6	-0.000037	0.9890	0.0219
12	-0.000075	0.9719	0.0553
18	-0.000136	0.9571	0.0839
24	-0.000221	0.9314	0.1325
30	-0.000309	0.9026	0.1854
36	-0.000393	0.8763	0.2322
42	-0.000553	0.8300	0.3111
48	-0.000766	0.7909	0.3745
54	-0.001069	0.7456	0.4440

60	-0.001304	0.7207	0.4806
66	-0.001367	0.7443	0.4460
72	-0.001262	0.7952	0.3676

Panel E: OPER			
1	-0.00672	0.9894	0.0211
6	-0.03749	0.9428	0.1111
12	-0.07213	0.9036	0.1836
18	-0.08095	0.9211	0.1516
24	-0.09634	0.9223	0.1494
30	-0.11344	0.9149	0.1630
36	-0.11177	0.9279	0.1389
42	-0.10541	0.9457	0.1057
48	-0.08648	0.9687	0.0616
54	-0.07320	0.9815	0.0366
60	-0.06203	0.9887	0.0224
66	-0.03718	0.9963	0.0074
72	-0.02953	0.9978	0.0044

3.4.2. Investment Strategy Using CPS Index with Other Variables

In this section, we investigate what happens to explaining future realized return when CPS_t is added as an independent variable to the current existing valuation ratios. According to our claim we discussed in section 3.3.3, Adding CPS-index as an independent variable to multivariate regression should increase the adjusted R-square value. In addition, using Variance inflation factor, we show that CPS_t does not have collinearity with any other indexes.

Table 3.3

Variance Inflation Factor for each index in equation.(3.12).

Variable	$(VIF)_{variable}$
$\boldsymbol{CPS_t}$	3.4444
$\boldsymbol{PER_t}$	2.4015
$\boldsymbol{PBR_t}$	4.5272
$\boldsymbol{S\&P500_t}$	3.8511
$\boldsymbol{OPER_t}$	4.6317

Table 3.3 presents the Variance inflation factor for the each variables in multivariate regression in equation (3.12). According to the results, all the variance inflation factors are less than 5; therefore, it present there is no multicollinearity problem in this multivariate

regression model.

In Table 3.4, we summarize Adjusted R-square for multivariate regression with two independent variables using the CPS, PER, PBR, OPER and S&P500 index as independent variables and the dependent variable is i month ahead realized return in equation (3.9). We bold the highest adjusted r-square value for each multivariate regression with different i-month ahead realized return. According to our results, adding CPS-index as independent variable significantly increase the adjusted r-square. Multivariate regression using combination of CPS-index and Price to book ratio as independent variables and Multivariate regression using combination of CPS-index and OPER as independent variables explain the future realized return well throughout all time horizon. According to the results, in multivariate regression with two independent variables model that having CPS-index as one of independent variables explains i-month future ahead return better than those of multivariate regression with two independent variables model which do not having CPS-index as independent variable.

In Table 3.5, we summarize Adjusted R-square for multivariate regression with three independent variables using the CPS, PER, PBR, OPER and S&P500 index as independent variables and the dependent variable is i month ahead realized return in equation (3.10). We bold the highest adjusted r-square value for each multivariate regression with different i-month ahead realized return. According to our results, adding CPS-index as independent variable significantly increase the adjusted r-square. Multivariate regression using combination of CPS-index, price to earning ratio and OPER as independent variables explains the i-month ahead realized return the best from until 24 month ahead realized return and Multivariate regression using combination of CPS-index, Price to Book Ratio and Price to Earning Ratio as independent variables explain the long time horizon ahead realized return well . According to the results, in multivariate regression with three independent variables model that having CPS-index as one of independent variables explains i-month ahead realized return better than those of multivariate regression with three independent variables model which do not having CPS-index as independent variable.

In Table 3.6, we summarize the adjusted r-square for multivariate regression with four independent variables using the CPS, PER, PBR, OPER and S&P500 index as independent variables and the dependent variable is i month ahead realized return in equation (3.11). We bold the highest adjusted r-square value for each multivariate regression with different i-month ahead realized return. According to our results, adding CPS-index as independent variable significantly increases the adjusted r-square. According to the results, in multivariate regression with four independent variables model that having CPS-index as one of independent variables explains i-month future ahead return better than those of multivariate regression with four independent variables model which do not having CPS-index as independent variable.

In figure 3.2, the adjusted R-square of each univariate regression and each multivariate regression is presented. According to the figure 3.2, regressions that have CPS-index as independent variable have higher adjusted R-square value compared to those not.

In Table 3.7 presents multivariate regression results using CPS_t variables and existing variables. The p-value for CPS index is all zero for the regression of $r_{t,t+20\cdot i}$, $i = 1,6,12,18,24,30,36,42,48,54,60,66,72$; we can proclaim that CPS_t variables is statistically significant for all different time lags.

Table 3.4 : Adjusted R-square for multivariate regression using the CPS, PER, PBR, OPER and S&P500 index as independent variables and the dependent variable is *i* month ahead realized return in equation (3.9).

i	CPS PER	CPS PBR	CPS S&P500	CPS OPER	PER PBR	PER S&P500	PER OPER	PBR S&P500	PBR OPER	S&P500 OPER
1	0.0336	0.0438	0.0341	**0.0596**	0.0261	0.0045	0.0451	0.0245	0.0265	0.0254
6	0.1394	0.2056	0.1379	**0.2702**	0.1505	0.0353	0.2533	0.1428	0.1460	0.1495
12	0.2233	0.3789	0.2210	**0.4385**	0.2819	0.0555	0.3191	0.3016	0.2861	0.2773
18	0.3195	**0.5299**	0.3187	0.5078	0.3892	0.0853	0.3098	0.4161	0.3823	0.2850
24	0.4531	**0.6728**	0.4540	0.6455	0.4536	0.1343	0.3443	0.4974	0.4452	0.3456
30	0.5392	**0.7749**	0.5307	0.7340	0.5040	0.1903	0.2997	0.5813	0.5147	0.4261
36	0.6159	**0.8391**	0.6157	0.7893	0.5145	0.2356	0.2969	0.5979	0.5320	0.4473
42	0.6797	**0.8703**	0.6901	0.8188	0.4919	0.3156	0.2681	0.5998	0.5193	0.4806
48	0.6933	**0.8766**	0.7130	0.7960	0.4927	0.3743	0.2645	0.5999	0.5355	0.4756
54	0.6725	**0.8662**	0.6980	0.7542	0.5302	0.4504	0.3054	0.6279	0.5753	0.5023
60	0.6256	**0.8363**	0.6468	0.6900	0.5794	0.5062	0.3580	0.6446	0.6191	0.5129
66	0.5537	**0.7814**	0.5452	0.5931	0.6215	0.5160	0.4088	0.6089	0.6555	0.4580
72	0.4789	**0.6840**	0.4143	0.4505	0.6529	0.4960	0.4661	0.5446	0.6568	0.3728

Table 3.5: Adjusted R-square for multivariate regression using the CPS, PER, PBR, OPER and S&P500 index as independent variables and the dependent variable is i month ahead realized return in equation (3.10).

i	CPS PER PBR	CPS PER S&P500	CPS PER OPER	CPS PBR S&P500	CPS PBR OPER	CPS S&P500 OPER	PER PBR S&P500	PER PBR OPER	PER S&P500 OPER	PBR S&P500 OPER
1	0.0436	0.0339	0.0683	0.0441	0.0604	0.0594	0.0260	0.0466	0.0451	0.0285
6	0.2075	0.1396	0.3355	0.2056	0.2706	0.2711	0.1532	0.2628	0.2585	0.1659
12	0.3809	0.2249	0.4815	0.3791	0.4429	0.4466	0.3014	0.3274	0.3558	0.3261
18	0.5307	0.3211	0.5505	0.5302	0.5474	0.5178	0.4167	0.3906	0.3670	0.4169
24	0.6732	0.4563	0.6923	0.6736	0.6891	0.6600	0.4986	0.4536	0.4337	0.4974
30	0.7875	0.5550	0.7493	0.7769	0.7893	0.7592	0.5859	0.5174	0.4584	0.5815
36	0.8467	0.6336	0.8082	0.8404	0.8482	0.8162	0.5995	0.5342	0.4794	0.5978
42	0.8755	0.7064	0.8357	0.8723	0.8752	0.8507	0.6011	0.5225	0.5020	0.6000
48	0.8766	0.7157	0.8259	0.8776	0.8765	0.8231	0.6008	0.5367	0.5070	0.6080
54	0.8703	0.6987	0.8135	0.8675	0.8694	0.7847	0.6396	0.5753	0.5549	0.6494
60	0.8519	0.6577	0.7871	0.8367	0.8461	0.7212	0.6774	0.6227	0.5947	0.6851
66	0.8128	0.5856	0.7327	0.7840	0.8010	0.6108	0.6828	0.6662	0.5893	0.6903
72	0.7491	0.5122	0.6568	0.7044	0.7245	0.4609	0.6732	0.6832	0.5734	0.6640

Table 3.6: Adjusted R-square for multivariate regression using the CPS, PER, PBR, OPER and S&P500 index as independent variables and the dependent variable is *i* month ahead realized return in equation (3.11).

i	CPS PER PBR S&P500	CPS PER PBR OPER	CPS PER S&P500 OPER	CPS PBR S&P500 OPER	PER PBR S&P500 OPER
1	0.0440	0.0816	0.0691	0.0602	0.0468
6	0.2079	0.3928	0.3360	0.2718	0.2706
12	0.3818	0.4885	0.4834	0.4490	0.3585
18	0.5315	0.5580	0.5532	0.5507	0.4218
24	0.6747	0.7007	0.6967	0.6940	0.5017
30	0.7951	0.7904	0.7659	0.7977	0.5879
36	0.8517	0.8487	0.8244	0.8544	0.6025
42	0.8813	0.8761	0.8563	0.8821	0.6051
48	0.8780	0.8766	0.8386	0.8778	0.6099
54	0.8706	0.8705	0.8233	0.8697	0.6493
60	0.8519	0.8520	0.7933	0.8461	0.6881
66	0.8182	0.8133	0.7331	0.8062	0.7004
72	0.7706	0.7507	0.6572	0.7468	0.6905

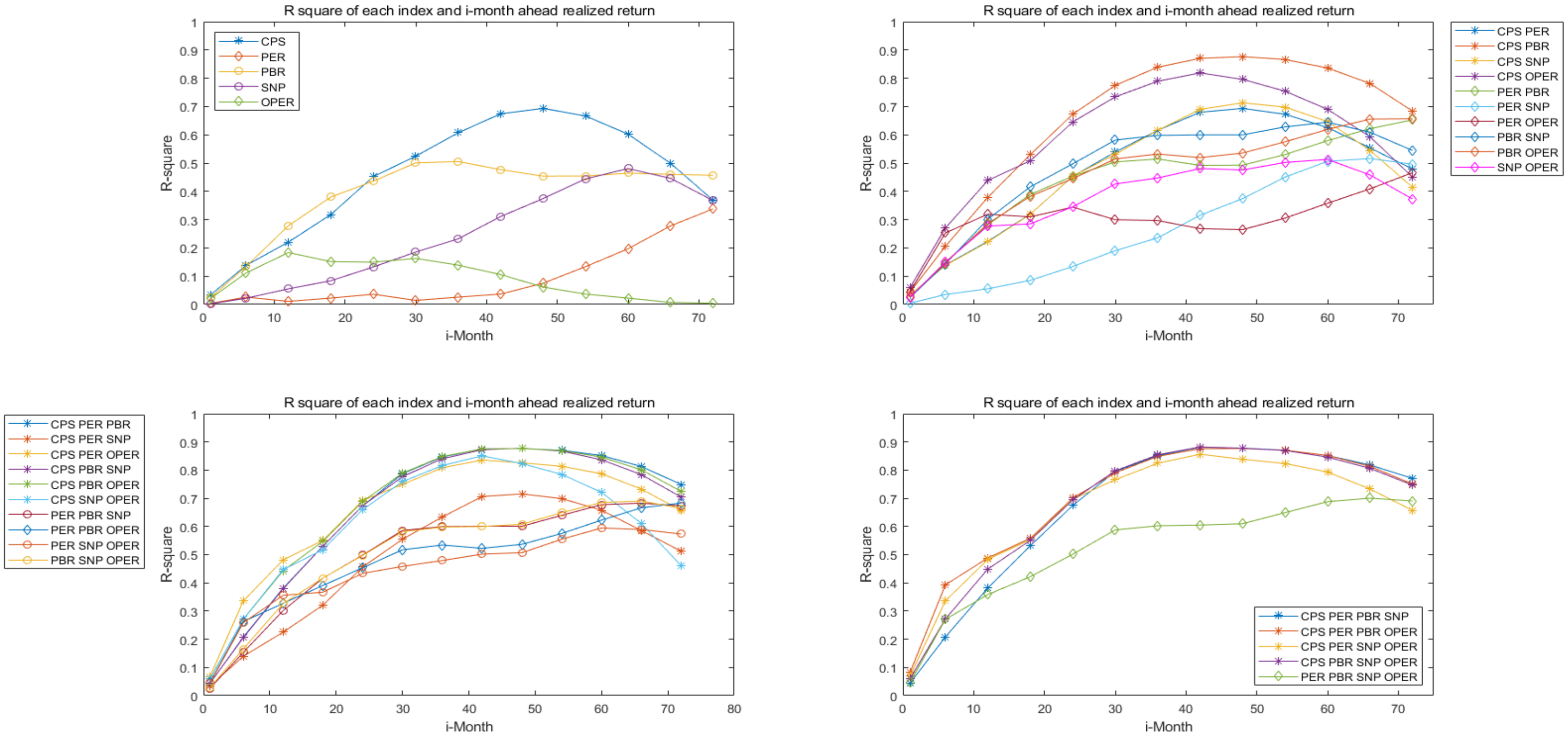

Fig.3.2 R-square of each univariate regressions and multivariate regressions

Table 3.7

Multivariate regression for the CPS, PER, PBR, OPER and S&P500. This table summarizes multivariate regression results in equation.(3.12). The dependent variable in these regression is *i* month ahead realized return, and the independent variables are CPS, PER, PBR, OPER and S&P500. Note that *i* is the forecasting horizon in months. *Coeff.* and *p-value* are obtained from the multivariate regression.

i		*Intercept*	CPS_t	PER_t	PBR_t	$S\&P500_t$	$OPER_t$	adj. R^2
1	*Coeff.*	0.0028	-0.0107	0.0111	0.0112	0.0019	-0.0206	0.083
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.0111	**0.0000**	
6	*Coeff.*	0.0167	-0.0486	0.0629	0.0563	0.0029	-0.1109	0.393
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.0588	**0.0060**	
12	*Coeff.*	0.0356	-0.0760	0.0554	0.0310	-0.0098	-0.1286	0.491
	p-value	**0.0000**	**0.0000**	**0.0020**	**0.0000**	**0.0000**	**0.0000**	
18	*Coeff.*	0.0602	-0.0960	0.0324	-0.0363	-0.0111	-0.0815	0.560
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	
24	*Coeff.*	0.0886	-0.1393	0.0397	-0.0448	-0.0171	-0.0989	0.704
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0001**	
30	*Coeff.*	0.1164	-0.1608	-0.0207	-0.1103	-0.0337	-0.0446	0.800
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	
36	*Coeff.*	0.1459	-0.1873	-0.0153	-0.1147	-0.0315	-0.0432	0.855
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	
42	*Coeff.*	0.1781	-0.2143	-0.0185	-0.1183	-0.0373	-0.0345	0.883
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	
48	*Coeff.*	0.2047	-0.2317	-0.0081	-0.1537	-0.0182	0.0003	0.878
	p-value	**0.0000**	**0.0000**	0.0154	**0.0000**	**0.0000**	0.9442	
54	*Coeff.*	0.2282	-0.2416	0.0196	-0.1832	-0.0085	0.0093	0.871
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.0182	0.0927	
60	*Coeff.*	0.2485	-0.2409	0.0499	-0.2174	0.0067	0.0143	0.852

	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	0.1390	0.0251	
66	*Coeff.*	0.2603	-0.2385	0.0751	-0.2642	0.0561	0.0290	0.819
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	
72	*Coeff.*	0.2717	-0.2281	0.1068	-0.3099	0.1247	0.0417	0.773
	p-value	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	**0.0000**	

Note: *Coeff.* represents the beta coefficient of each variable.

Bold in *p-value* indicates the 1% significance.

Table 3.8

F-test for multivariate regression adding CPS-index to the model.

i	**F-statistics**	**P-value**
1	169.03	0.000
6	845.47	0.000
12	1056.59	0.000
18	1237.61	0.000
24	2607.43	0.000
30	3913.51	0.000
36	6257.59	0.000
42	8251.05	0.000
48	7336.30	0.000
54	5512.13	0.000
60	3437.12	0.000
66	1956.50	0.000
72	1031.90	0.000

Table 3.8 presents the F-test results of adding CPS index to multivariate regression is statistically significant. According to the results, p-value for the F-test is all zero in all different time lags. Therefore, we can proclaim that adding CPS-index to currently existing indexes is statistically significant.

3.4.3. Parameter Tuning and Granger Causality Test

In this section, we use different m year periods to find optimal parameter for the CPS-index. As mentioned previously, the original CPS-index indicates the difference between cumulative log return of common stocks that have corresponding preferred stocks and cumulative log return of corresponded preferred stock with base time t_0. In this reason, depends on t_0, the overall value for the CPS-index could be large or small. Therefore, through the empirical analysis, we try to find the optimal past base time. In other words, we try to build an optimal m-year CPS-index which indicates that difference between cumulative log return of common stocks that have corresponding preferred stocks and cumulative log return of corresponded preferred stocks with base time m-years ago.

In Figure 3.3, we presented the adjusted R-square of univariate regression using m-year CPS-index as independents variable and i-month ahead realized return as dependent variable. According to the results, when m is 7 years and 8 years, the regression model has highest adjusted R-square value.

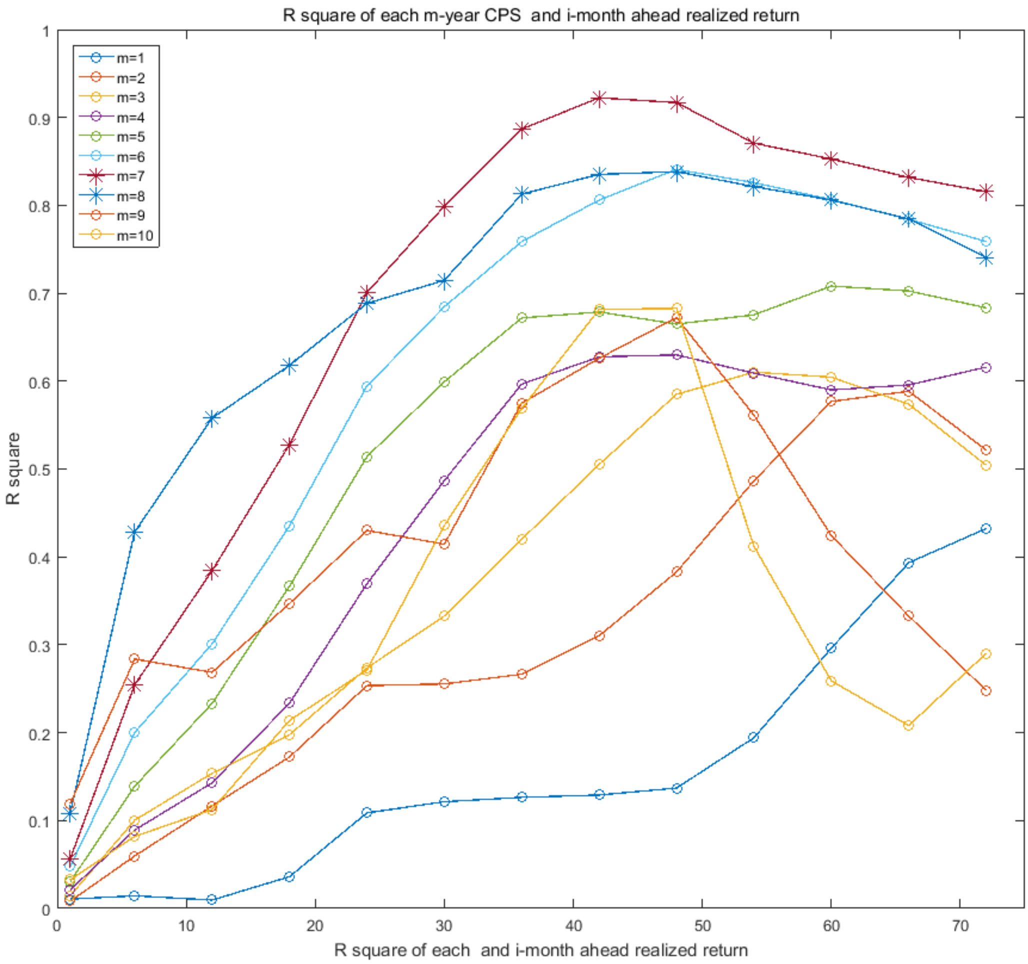

Fig.3.3 Adjusted R-square of each univariate regressions and multivariate regression

Table 3.9

Adjusted R-square for univariate regression using the m-year CPS index as independent variables and *i* month ahead realized return as dependent variable is in equation (18). For each different i and m value we get different regressions and its adjusted R-square.

	m=1	2	3	4	5	6	7	8	9	10
i=1	0.0106	0.0082	0.0144	0.0213	0.0305	0.0478	0.0564	0.1083	**0.1190**	0.0326
6	0.0144	0.0597	0.1004	0.0891	0.1389	0.1999	0.2544	**0.4283**	0.2838	0.0819
12	0.0096	0.1165	0.1538	0.1430	0.2333	0.3011	0.3843	**0.5580**	0.2686	0.1125
18	0.0361	0.1723	0.1977	0.2335	0.3666	0.4353	0.5277	**0.6180**	0.3469	0.2139
24	0.1087	0.2535	0.2736	0.3693	0.5140	0.5943	**0.7007**	0.6885	0.4302	0.2712
30	0.1214	0.2555	0.3328	0.4868	0.5991	0.6846	**0.7991**	0.7147	0.4144	0.4357
36	0.1264	0.2665	0.4200	0.5966	0.6718	0.7588	**0.8868**	0.8128	0.5748	0.5688
42	0.1290	0.3102	0.5051	0.6276	0.6787	0.8061	**0.9224**	0.8353	0.6260	0.6816
48	0.1370	0.3828	0.5850	0.6298	0.6649	0.8410	**0.9169**	0.8383	0.6723	0.6828
54	0.1942	0.4866	0.6102	0.6091	0.6755	0.8258	**0.8708**	0.8213	0.5613	0.4120
60	0.2961	0.5768	0.6045	0.5900	0.7082	0.8066	**0.8527**	0.8060	0.4242	0.2589
66	0.3934	0.5885	0.5738	0.5956	0.7026	0.7843	**0.8318**	0.7845	0.3326	0.2081
72	0.4323	0.5216	0.5042	0.6160	0.6834	0.7588	**0.8156**	0.7408	0.2482	0.2902

Table 3.10

Adjusted R-square for univariate regression using the 7-year CPS, PER, PBR, OPER and S&P500 index as independent variables and *i* month ahead realized return as dependent variable is in equation (3.18).

m=7	CPS	PER	PBR	OPER	SNP
i=1	**0.0564**	0.0141	0.0187	0.0001	0.0020
6	**0.2544**	0.0733	0.1237	0.0028	0.0201
12	**0.3843**	0.0451	0.2514	0.0045	0.0299
18	**0.5277**	0.0603	0.3820	0.0004	0.0287
24	**0.7007**	0.0680	0.4829	0.0074	0.0448
30	**0.7991**	0.0250	0.5646	0.0273	0.0670
36	**0.8868**	0.0352	0.6276	0.0259	0.0909
42	**0.9224**	0.0360	0.7005	0.0181	0.1365
48	**0.9169**	0.0652	0.7493	0.0005	0.1934
54	**0.8708**	0.1118	0.8114	0.0014	0.3214
60	0.8527	0.1510	**0.8613**	0.0053	0.4901
66	0.8318	0.1847	**0.8772**	0.0091	0.6468
72	0.8156	0.2191	**0.8733**	0.0118	0.7482

Table 3.11

Adjusted R-square for univariate regression using the 8-year CPS, PER, PBR, OPER and S&P500 index as independent variables and *i* month ahead realized return as dependent variable is in equation (3.18).

m=8	CPS	PER	PBR	OPER	SNP
i=1	0.1083	0.0158	**0.0256**	0.0002	0.0015
6	**0.4283**	0.0745	0.1440	0.0029	0.0194
12	**0.5580**	0.0401	0.2373	0.0069	0.0313
18	**0.6180**	0.0517	0.3007	0.0010	0.0394
24	**0.6885**	0.0618	0.3679	0.0308	0.0788
30	**0.7147**	0.0111	0.4553	0.0818	0.1185
36	**0.8128**	0.0200	0.5016	0.0690	0.1281
42	**0.8353**	0.0193	0.5548	0.0488	0.1603
48	**0.8383**	0.0486	0.6086	0.0052	0.1920
54	**0.8213**	0.1005	0.6884	0.0003	0.2850
60	**0.8060**	0.1464	0.7653	0.0048	0.4376
66	0.7845	0.1834	**0.7868**	0.0085	0.5980
72	0.7408	0.2193	**0.7750**	0.0076	0.7090

Table 3.12

Granger Causality test of 7-year CPS-index for the *i* month ahead realized return τ is the maximum time lag in equation (3.19)

i	F-statistics	Critical Value
$\tau = 1$		
1	18.763	3.858
6	31.447	3.858
12	29.320	3.858
18	53.775	3.858
24	83.228	3.858
30	20.415	3.858
36	60.556	3.858
42	58.144	3.858
48	102.211	3.858
$\tau = 10$		
1	102.839	2.621
6	151.376	2.621
12	134.237	2.621
18	161.569	2.231
24	218.943	2.621
30	198.347	2.621
36	145.280	2.621
42	271.550	2.621
48	259.875	2.621

In table 3.9, Adjusted R-square for univariate regression using the m-year CPS index as independent variables and i month ahead realized return as dependent variable is in equation (3.18). For each different i and m value we get different regressions and its adjusted R-square. According to the results, explaining the less than 18 month ahead realized return, using 8 year CPS-index as an independent variable has the highest adjusted R-square value; when the regression explain more than 18 month using 7 year CPS-index as an independent variable has the highest adjusted R-square value. Therefore, we tested those 7 year CPS-index and 8-year CPS index with existing ratios in table 7 and table 8 in following.

Table 3.10 presents the Adjusted R-square for univariate regression using the 7-year CPS, PER, PBR, OPER and S&P500 index as independent variables and *i* month ahead realized return as dependent variable is in equation (3.18) section 3.3.4. We bolded the highest adjusted R-square value for explaining each i-month ahead realized return. According to the results, 7-year CPS-index explains the future return until 54 month ahead realized return, and then price to book ratio explains the future return after 54 month ahead realized return. From this result, we can conclude that using 7-year CPS-index to explain future market return is valid.

Table 3.11 presents the Adjusted R-square for univariate regression using the 8-year CPS, PER, PBR, OPER and S&P500 index as independent variables and i month ahead realized return as dependent variable is in equation (3.18) section 3.2.4. We bolded the highest adjusted R-square value for explaining each i-month ahead realized return. According to the results, 8-year CPS-index explains the future return until 60 month ahead realized return, and then price to book ratio explains the future return after 60 month ahead realized return. From this result, we can conclude that using 8-year CPS-index to explain future market return is valid.

Table 3.12 shows the Granger Causality test result for 7 year CPS index, $CPS_{t,7}$ variable in equation (3.19). As F-statistics are bigger than critical values in all time horizon, $CPS_{t,7}$ variable affect $r_{t,t+20\cdot i}$. The significance level of this test is 0.05, and we set τ to be 1 and 10.

3.5. Conclusion

This chapter has shown that the cumulative log return spread between common stocks and their preferred stocks could be used to assess the stock market condition and predict the future market return better using currently existing valuation ratios such as price to earning ratio, price to book ratio and price to operational earnings ratio. We reach several conclusions as follows. First, using univariate regression, we presented that CPS-index has better explanatory power of explaining market future returns compared to currently existing valuation ratios such as price to earning ratio, price to book ratio and price to operational earnings ratio especially predicting more than 30 months ahead. Since CPS-index is not built by any of earning or book value, this result gave us the possibility to enhance the market predictability by using it together. Second, we used multivariate regression to present adding CPS-index to currently existing valuation ratio enhanced the market predictability. We used multivariate regression models using from two independent variables to four independent variables. In every multivariate regression model that have the most market predictability power having CPS-index as independent variable. According to results, we can proclaim that using CPS-index we can explain long term market return better. Lastly, we developed CPS index by finding right parameter for the CPS-index. We presented that CPS-index that build with the difference between cumulative log return of common stocks that have corresponding preferred stock and those preferred stocks from seven years explains the market future return better than any of other parameter. We showed that using seven year CPS-index and eight year CPS-index have significantly high explanatory power. By comparing with currently existing predictive ratios, we showed that both seven year CPS-index and eight year CPS-index have better market predictability power than currently existed ratios. In addition, we confirmed that past and current seven year CPS-index affects future realized market return using Granger Causality test.

There are some limits in our research on this chapter. The research conducted on price to earnings ratios and price to book used more than eighty years of data periods, therefore if we analyze longer period of data than sixteen years from 2000 to 2016 used in this paper, our results might be more robust. As we applied CPS-index to the US market data only, we need to investigate whether CPS-index has the equivalent prediction power in other global markets such as Europe and China. For the finding optimal starting point for the CPS index, we need to have concrete rule to determine the starting point. Eliminating the periods when stock price rise or fall rapidly, could help the accuracy of this research, however, the period for the rapid rise or rapid fall were too limited to be significant. Therefore, we will build more robust model when we receive longer period of data for the preferred stocks.

Chapter 4

Empirical Analysis of Pairs Trading Using Preferred Stocks.

4.1 Introduction

Both financial practitioner and academic profession has long been interested in quantitative methods of speculation. One of most commonly used short term strategy is called “Pairs Trading.” The strategy has at least a 20-year history on Wall Street and is among the proprietary “statistical arbitrage” tools currently used by hedge funds as well as investment banks. This Pairs Trading strategy is surprisingly simple. First, find two stocks that moves together historically. Second, when the spread between those two stocks are widen, short the overpriced one and buy the underpriced one. Last, if those two stocks moves similar to history, prices will converge and the investors will profit. It is hard to believe that such a simple strategy, based solely on historical price dynamics and simple contrarian principles, could make profit in complicated stock market (Gatev et al., 2006).

Among various investment strategies, Pairs Trading was the most intriguing strategy because it can make profit in any financial situation. Pairs Trading strategy is as follows. First, find two stocks that moves together historically. Second, when the spread between those two stocks are widen, short the overpriced one and buy the underpriced one. Last, if those two stocks moves similar to history, prices will converge and the investors will profit. Since the Pairs Trading only makes profit from the spread, investors can always make profit as long as the pair converge in the future. Therefore, empirically tested, during the financial crisis, portfolios that used Pairs Trading strategy make profit during that periods. However, there are weaknesses existed in Pairs Trading. First finding pairs that moves together requires complicated calculation and computation power because investors have to find two stocks that moves together from more than 4000 stocks. Seconds, in Pairs Trading, it is essential to find two stocks that move together in the past and will move together in the future because if they move in wrong direction in the future, the investors would lose a lot of money. However, in current Pairs Trading strategy, investors only consider the two stock that move together in the past and we believe this pair could move different direction in the future. Therefore, we suggest operate Pairs Trading with common stock and its preferred stocks because for common stocks and its preferred stocks if they move together in the past they will move together in the future. By using this strategy investors could save computation time and make profit more safely. We presented that using common stock and its preferred stock for Pairs Trading strategy, investors can make more profit considering its risk. On the other hand, it is important to figure out which Pairs Trading method to use in different situation. The optimal Pairs Trading strategy could be different when the stock market is in pre-crisis, financial-crisis, and post-crisis. Therefore, in this chapter, we examine the risk and return characteristics Pairs Trading with common stocks and preferred stocks compare to existed Pairs Trading methods, and find an optimal Pairs Trading strategy for different financial periods. For the analysis we sued daily data over the period 2000 through May 2015.

4.2 Background and Literature Review.

According to Gatev et al. (2006), in the mid-1980s, the Wall Street quant Nunzio Tartaglia assembled a team of physicists, mathematicians, and computer scientists to uncover arbitrage opportunities in the equities markets (Gatev et al., 2006). Tartaglia's group of former academics used sophisticated statistical methods to develop high-tech trading programs, executable through automated trading systems, which took the intuition and trader's ''skill'' out of arbitrage and replaced it with disciplined, consistent filter rules. Among other things, Tartaglia's programs identified pairs of securities whose prices tended to move together. They traded these pairs with great success in 1987 a year when the group reportedly made a $50 million profit for the firm. Although the Morgan Stanley group disbanded in 1989 after a couple of bad years of performance, Pairs Trading has since become an increasingly popular ''market neutral'' investment strategy used by individual and institutional traders as well as hedge funds.

Previous works on Pairs Trading are fallows. Jegadeesh and Titman's finding that contrarian profits are in part due to over-reaction to company-specific information shocks rather than price reactions to common factors (Jegadeesh & Titman, 1993). According to Lo and

Mackinlay, trading strategy using mean reverting trading is highly due to the stock market overreaction. He proclaims that when assets are temporarily associated with one or more financial instruments, when they move, there is an undercurrent or overreaction. The risk of various trading techniques using a similar paradigm was analyzed in previous studies. Gatev analysis the Pairs Trading method (Gatev et al., 2006).

The research on Pairs Trading has largely focused on the study of finding a suitable pair in a large number of assets, and there is an empirical study on whether non-ideal profit occurs when fair trading is performed in various markets. A study on finding a suitable pair has been divided into two types according to the calculation method as follows. First one is minimum distance method. It is a way to normalize the price of an asset over a period of time and to pair the two assets with the minimum difference in normalized prices. Calculation is also simple and intuitive, so it is widely used. The other one is stochastic residual method. Elliott suggested this method by proclaim the difference of two stocks will mean reverting in the future (Elliott, Hoek, & Malcolm, 2005).

Nath examines the implementation of a simple Pairs Trading strategy with automatic extreme risk control using the entire universe of securities in the highly liquid secondary market for U.S. government debt. Nath also examines the excess rate of return with different diverge strategy. He showed that Using data from the repo and money market, estimates are also made of the distribution of absolute returns after accounting for financing and transaction costs. (Nath, 2004)

Perlin examines the excess return of Pairs Trading in Brazilian stock market. We used minimum distance method to find pairs with different learning time periods. He used different Pairs Trading strategy that not limited to 2 standard deviation diverge strategy. The main conclusion of this simulation is that pairs-trading strategy was a profitable and market-neutral strategy at the Brazilian market. Such profitability was consistent over a region of the strategy's parameters. (Perlin, 2009).

Hong and Susmel researched pairs-trading strategies for 64 Asian shares listed in their local markets and listed in the U.S. as ADRs. Given that all pairs are cointegrated, they are logical choice for pairs-trading. They find that pairs-trading in this market delivers significant profits. The results are robust to different profit measures and different holding periods (Hong & Susmel, 2003) .

In this study, we suggest the optimal Pairs Trading method for different financial market situation. To make this suggestion, we interpreted practitioner description of Pairs Trading as straightforwardly as possible. As we mentioned previously, our rules for Pairs Trading is first “find stocks that move together,” and second “take a long-short position when they diverge and unwind on convergence” We put positions on at a one-standard deviation diverge strategy and two-standard deviation diverge strategy to compare the excess rate of return.

4.3 Methodology

In this chapter, there are two stage for our Pairs Trading strategy. First, we find pairs that moves together for 12-month period. Second, we trade them in the next 6-month period. Both 12-month and 6 months are chosen arbitrarily and have remained our horizons since the beginning of the study.

4.3.1 Pairs Formation

There are three methods for the pairs formation; we wanted to answer the previous question "How to define two stocks that moves together?" 1) pairs with common stock and its preferred stock that co-integrate each other. 2) Using correlation method to find stocks that moves together. 3) Using co-integration method to find stocks that moves together.

For pairs with common stock and its preferred stock method, we tested all the pairs that forms with common stocks and its preferred stock. We construct a cumulative total returns index for each stock over the formation period. For each time periods, we test the co-integration test and out of pairs that co-integrated each other, we selected top twenty pairs that minimizes the sum of squared deviations between the two normalized price series. Pairs are thus formed by exhaustive matching in normalized daily “price” space, where price includes reinvested dividends.

For finding pairs using correlation method, we tested all stocks from S&P 500. Before doing this process, we screen out all stocks from the S&P500 daily files that have one or more days with no trade. This serves to identify relatively liquid stocks as well as to facilitate pairs formation. Next we construct a cumulative total returns index for each stock over the formation period. We then choose the top twenty pairs that mostly correlated each other.

For finding pairs using co-integration method, we tested all stocks from S&P 500. Before doing this process, we screen out all stocks from the S&P500 daily files that have one or more days with no trade. This serves to identify relatively liquid stocks as well as to facilitate pairs formation. Next we construct a cumulative total returns index for each stock over the formation period. For each time periods, we test the co-integration test and out of pairs that co-integrated each other, we selected top twenty pairs that minimizes the sum of squared deviations between the two normalized price series. We then choose the top twenty pairs that mostly co-integrated each other.

4.3.2 Trading Strategy and Periods

We set the training period for 12-months and operating period for 6-month. During the training period, we use correlation and co-integration method to find pairs that moves together. Once we find the right pairs that moves together, we operate Pairs Trading strategy for 6-months. As mentioned previously, we select trading rules based on the proposition that we open a long–short position when the pair prices have diverged by a certain amount and close the position when the prices have reverted. If the pair does not revert until the end of operating periods, we clear out the long and short position with price at the end of operating periods. Following practice, we base our rules for opening and closing positions on a standard deviation metric. When the gap between pair is widen, we take short position the winner and take long position the loser, and when the gap is closed, we make a profit. We trade according to the rule that opens a position in a pair on the day following the day on which the prices of the stocks in the pair diverge by one or two historical standard deviations. In Pairs Trading, it is important to set up the optimal trading strategy according to financial market situation. Therefore, we compare the monthly excess return on Pairs Trading strategy with opens a position in a pair which the prices of the stocks in the pair diver by one historical standard deviations and Pairs

Trading strategy with opens a position in a pair which the prices of the stocks in the pair diver by two historical standard deviations. One standard deviation strategy have more frequent trading opportunities but with limited excess returns; two standard deviation strategy have less frequent trading opportunities but have more excess returns.

To find out which Pairs Trading strategy would perform best in different financial situation, we analysis the excess return of Pairs Trading in four different periods: Pre-crisis (from 2005-01-03 to 2007-07-31), Subprime-crisis (from 2007-08-01 to 2009-12-07), European-crisis (from 2009-12-08 to 2012-04-27), and Post-crisis (2012-04-28 to 2015-05-30). For each different periods, we compare the monthly excess return of each Pairs Trading strategy. In addition, we used Sharpe ratio to compare the monthly excess return with its risk.

4.3.3 Excess Return Computation

We used an excess return computation method from Gatev (2006). Calculating excess earnings in a portfolio of pairs is not an issue because pairs can be opened and closed at various points during a six-month trading period. Pairs that open and converge during the trading interval have a positive cash flow. Pairs can have multiple positive cash flows during the trading interval because they can be reopened after the initial convergence. A pair that is open but has no convergence will only have a cash flow on the last day of the trading interval when all positions are closed. The result for a pair trading strategy is therefore a set of positive cash flows that are randomly distributed during the trading period and a series of cash flows at the end of the trading interval that can be positive or negative. Each pair can have multiple cash flows during the trading interval. Or, if the price does not exceed two standards during the trading interval, there may not be a transaction. Because the trading profits and losses are calculated at a long-distance selling position of $ 1, the yield has an interpretation of excess return. During the trading interval, the excess return on a pair is calculated as the reinvested earnings during the trading interval. In particular, short and long portfolio positions enter the market every day. The daily returns for long and short positions are calculated as weighted revenue values in the following ways:

$$r_{P,t} = \frac{\sum_{i \in P} w_{i,t} r_{i,t}}{\sum_{i \in P} w_{i,t}} \quad \ldots\ldots (4.1)$$

$$w_{i,t} = w_{i,t-1}(1 + r_{i,t-1}) = (1 + r_{i,1}) \ldots\ldots (1 + r_{i,t-1}) \quad \ldots\ldots (4.2)$$

where r defines returns and w defines weights, and the daily returns are compounded to obtain monthly returns. This has the simple interpretation of a buy-and-hold strategy. We consider two measures of excess return on a portfolio of pairs: the return on committed capital and the fully invested return, that is, the return on actual employed capital. In this study, we set up the transaction costs with 0.01% every time we buy and sell the stock. In addition, for the Sharpe ratio, we used 3-month United States Treasury bill rate as a risk-free rate.

4.4. Empirical Results

In this section, we divided in to four different time periods. The total period for analysis is considered from 2000-01-01 to 2015-05-30. The total period is then divided into four sub-periods: Pre-crisis (from 2005-01-03 to 2007-07-31), Subprime-crisis (from 2007-08-01 to 2009-12-07), European-crisis (from 2009-12-08 to 2012-04-27), and Post-crisis (2012-04-28 to 2015-05-30). Note that this sub-periods are also considered in Paulo Horta, Sérgio Lagoa, and Luis Martins (2014) (Paulo Horta, Sergio Lagoa, & Luis Martins, 2014).

4.4.1 The Whole Periods

Table 4.1

Excess return of Pairs Trading for whole periods

	Preferred Pairs	Correlation Pairs	Co-integration Pairs
One-standard deviation diverge strategy			
Mean excess return	0.0171	0.0188	0.0210
Standard Deviation	0.0116	0.0233	0.0197
Sharpe Ratio	1.3269	0.7351	0.9800
Two-standard deviation diverge strategy			
Mean excess return	0.0091	0.0105	0.0114
Standard Deviation	0.0087	0.0139	0.0105
Sharpe Ratio	0.8541	0.6399	0.9267

Table 4.1 summarizes the excess returns for the pairs portfolios. According to the results, one-standard deviation diverge strategy has higher mean excess return and Sharpe ratio than those of two-standard deviation diverge strategy in all different Pairs Trading training methods. This represents that in Pairs Trading method, more frequent trading with limited return is more effective than less frequent trading with high return. On the other hands, if the investors are risk seeker, they would prefer using one-standard deviation diverge strategy over two-standard deviation diverge strategy because the stand deviation of two-standard deviation diverge strategy is lower than one-standard deviation diverge

strategy. The one-standard deviation diverge strategy with co-integration training method Pairs Trading has the highest six month excess return with 0.021. The one-standard deviation diverge strategy with preferred stocks training method Pairs Trading has the highest Sharpe ratio with 1.3269.

Comparing the training method, as we previous expected pairing with common stock and its preferred stock method has relatively small monthly excess return compare to correlation method or the co-integration method. This is because common-preferred pairs does not open as frequently as correlation or co-integration pairs. On the other hands, comparing the standard deviation of each Pairs Trading method, pairing with common stock and its preferred stock has the lower value than correlation and co-integration method. This indicates that using common stock and its preferred stock as a pair for the Pairs Trading has comparably lower risk than other methods. This result can interpret as Pairs Trading with common stock and its preferred stock can minimize the risk of having pairs that even diverge more and never revert until the end of operation periods. As we mentioned previously, the biggest risk of Pairs Trading is investor believes a pair will move together because they have moved together but they move different directions; this will cause the loss because when pair is diverge, investor would take short position the winner and take long position the loser, if they move different direction, investor would lose in both way. By using common stock and its preferred stock pair method, we could minimize the losing probability and the result in table 4.1 supports it. In both one-standard deviation strategy and two-standard deviation strategy, pairing with common stock and its preferred stock method has the lowest standard deviation value compare to the other pairing methods.

Even though pairing with common stock and its preferred stock method has the lowest risk, it is important to have comparably high monthly excess return. We used Sharpe ratio to measure the risk adjusted return. Comparing the Sharpe Ratio of each portfolios, portfolios with common stock and its preferred stock pair with one-standard deviation strategy has the highest value with 1.3269 and portfolios with correlation methods with two-standard deviation strategy has the lowest value with 0.6399. This results presents that Pairs Trading with common stock and its preferred stocks have significant trading performance and it is eligible to use. In addition if the investor is the risk seeker, using the co-integration pairing method with one-standard deviation strategy is recommended.

4.4.2 Pre Crisis

Table 4.2

Excess return of Pairs Trading for Pre-crisis periods

	Preferred Pairs	Correlation Pairs	Co-integration Pairs
One-standard deviation diverge strategy			
Mean excess return	0.0106	0.0098	0.0115
Standard Deviation	0.0018	0.0031	0.0029
Sharpe Ratio	5.0421	2.6369	3.3536
Two-standard deviation diverge strategy			
Mean excess return	0.0060	0.0054	0.0061
Standard Deviation	0.0013	0.0016	0.0019
Sharpe Ratio	3.2290	2.3181	2.4076

In table 4.2, it presents the monthly excess return on each Pairs Trading portfolio with different pairs formation method and strategy in pre-crisis. During the pre-crisis, the stock market bull markets and there were many opportunities for the Pairs Trading. In this section, we will find out which Pairs Trading method is optimal for the investors during the pre-crisis periods. Comparing the Pairs Trading strategy, one-standard deviation outperformed the two-standard deviation strategy in mean excess return and Sharpe ratio but the two-standard deviation strategy has lower standard deviation of excess return. This indicates that one-standard deviation diverge strategy could have higher mean excess return of each portfolio than that of using two-standard deviation diverge strategy.

For comparing pairing formation methods, among the one-standard deviation diverge strategy, Co-integration method portfolio has the highest mean excess return with 0.0115 and pairing method using correlation portfolio has the lowest return with 0.0098. For the standard deviation of mean excess return, pairing with common and its preferred stock portfolio method has the lowest standard deviation of excess return with 0.0018 and portfolio using correlation method has the highest standard deviation of excess return with 0.0031. For the Sharpe ratio, portfolio using pairing method with common stocks and its preferred stock has the highest value with 5.0421 and portfolio using pairing method with Correlation methods has the lowest value with 2.6369.

Among the two-standard deviation diverge strategy, Co-integration method portfolio has the highest mean excess return with 0.061 and pairing method using correlation portfolio has the lowest return with 0.0054. For the standard deviation of mean excess return, pairing with common and its preferred stock portfolio method has the lowest standard deviation of excess return with 0.0013 and portfolio using co-integration method has the highest standard deviation of excess return with 0.0019. For the Sharpe ratio, portfolio using pairing method with common stocks and its preferred stock has the highest value with

3.2290 and portfolio using pairing method with Correlation methods has the lowest value with 2.3181.

According to our results, if the investor is the risk-avoider, portfolio using pairing method with common stock and its preferred stock with one-standard deviation diverge strategy is the optimal Pairs Trading strategy during the Subprime-Crisis periods. If the investor is the risk-lover, portfolio using co-integration pairing method with one-standard deviation diverge strategy is the optimal Pairs Trading strategy during the Subprime-Crisis periods.

4.4.3 Subprime Crisis

Table 4.3

Excess return of Pairs Trading for Subprime-crisis periods

	Preferred Pairs	Correlation Pairs	Co-integration Pairs
One-standard deviation diverge strategy			
Mean excess return	0.0344	0.0507	0.0478
Standard Deviation	0.0152	0.0437	0.0334
Sharpe Ratio	2.1567	1.1243	1.3810
Two-standard deviation diverge strategy			
Mean excess return	0.0188	0.0293	0.0258
Standard Deviation	0.0088	0.0265	0.0174
Sharpe Ratio	1.9397	1.0427	1.3853

In table 4.3, it presents the monthly excess return on each Pairs Trading portfolios with different pairs formation method and strategy in Subprime-crisis periods. During the Subprime-crisis, the stock market bear markets and there were many opportunities for the Pairs Trading because the market was very volatile, and this leads more opportunity for pairs to be open and close. In this section, we will find out which Pairs Trading method is optimal for the investors during the Subprime-crisis periods. Comparing the Pairs Trading strategy, one-standard deviation outperformed the two-standard deviation strategy in mean excess return and Sharpe ratio, but the two-standard deviation strategy has lower standard deviation of excess return. This indicates that one-standard deviation diverge strategy could have higher mean excess return of each portfolio than that of using two-standard deviation diverge strategy.

For comparing pairing formation methods, among the one-standard deviation diverge strategy, correlation method portfolio has the highest mean excess return with 0.0507 and pairing method using common stocks and its preferred stocks portfolio has the lowest return with 0.0344. For the standard deviation of mean excess return, pairing with common and its preferred stock portfolio method has the lowest standard deviation of excess return with 0.0152 and portfolio using correlation method has the highest standard deviation of excess return with 0.0437. For the Sharpe ratio, portfolio using pairing method with common stocks and its preferred stock has the highest value with 2.1567 and portfolio using pairing method with Correlation methods has the lowest value with 1.1243.

Among the two-standard deviation diverge strategy, correlation method portfolio has the highest mean excess return with 0.0293 and pairing method using common stock and its preferred stock portfolio has the lowest return with 0.0188. For the standard deviation of mean excess return, pairing with common and its preferred stock portfolio method has the lowest standard deviation of excess return with 0.0088 and portfolio using correlation method has the highest standard deviation of excess return with 0.0265. For the Sharpe ratio, portfolio using pairing method with common stocks and its preferred stock has the highest value with 1.9397 and portfolio using pairing method with Correlation methods has the lowest value with 1.0427. According to our results, if the investor is the risk-avoider, portfolio using pairing method with common stock and its preferred stock with one-standard deviation diverge strategy is the optimal Pairs Trading strategy during the Subprime-Crisis periods. If the investor is the risk-lover, portfolio using correlation pairing method with one-standard deviation diverge strategy is the optimal Pairs Trading strategy during the Subprime-Crisis periods.

4.4.4 European Crisis

Table 4.4

Excess return of Pairs Trading for European-crisis periods

	Preferred Pairs	Correlation Pairs	Co-integration Pairs
One-standard deviation diverge strategy			
Mean excess return	0.0125	0.0152	0.0177
Standard Deviation	0.0043	0.0059	0.0059
Sharpe Ratio	2.5322	2.2921	2.7239
Two-standard deviation diverge strategy			
Mean excess return	0.0076	0.0077	0.0103
Standard Deviation	0.0023	0.0029	0.0031
Sharpe Ratio	2.5856	2.0632	2.8039

In table 4.4, it presents the monthly excess return on each Pairs Trading portfolio with different pairs formation method and strategy in European-crisis. During the European-crisis, the stock market bear markets and there were many opportunities for the Pairs Trading because the market was very volatile and this leads more opportunity for pairs to be open and close. In this section, we will find out which Pairs Trading method is optimal for the investors during the European-crisis periods. Comparing the Pairs Trading strategy, one-standard deviation outperformed the two-standard deviation strategy in mean excess return and Sharpe ratio but the two-standard deviation strategy has lower standard deviation of excess return. This indicates that one-standard deviation diverge strategy could have higher mean excess return of each portfolio than that of using two-standard deviation diverge strategy.

For comparing pairing formation methods, among the one-standard deviation diverge strategy, Co-integration method portfolio has the highest mean excess return with 0.0177 and portfolio using common stock and its preferred stock pairing method has the lowest return with 0.0125. For the standard deviation of mean excess return, portfolio using pairing method with common and its preferred stock has the lowest standard deviation of excess return with 0.0043 and portfolio using correlation method and co-integration method has the same highest standard deviation of excess return with 0.0059. For the Sharpe ratio, portfolio using co-integration method has the highest value with 2.7239 and portfolio using pairing method with Correlation methods has the lowest value with 2.2921.

Among the two-standard deviation diverge strategy, co-integration method portfolio has the highest mean excess return with 0.0103 and pairing method using common stock and its preferred stock portfolio has the lowest return with 0.0076. For the standard deviation of mean excess return, pairing with common and its preferred stock portfolio method has the lowest standard deviation of excess return with 0.0023 and portfolio using co-integration method has the highest standard deviation of excess return with 0.0031. For the Sharpe ratio, portfolio using Co-integration method has the highest value with 2.8039 and portfolio using pairing method with Correlation methods has the lowest value with 2.0632.

According to our results, if the investor is the risk-avoider, portfolio using pairing method with common stock and its preferred stock with two-standard deviation diverge strategy is the optimal Pairs Trading strategy during the European-Crisis periods. If the investor is the risk-lover, portfolio using co-integration pairing method with one-standard deviation diverge strategy is the optimal Pairs Trading strategy during the European-Crisis periods.

4.4.5 Post Crisis

Table 4.5

Excess return of Pairs Trading for Post-crisis periods

	Preferred Pairs	Correlation Pairs	Co-integration Pairs
One-standard deviation diverge strategy			
Mean excess return	0.0098	0.0099	0.0124
Standard Deviation	0.0024	0.0040	0.0057
Sharpe Ratio	3.4380	2.0625	1.9026
Two-standard deviation diverge strategy			
Mean excess return	0.0057	0.0055	0.0067
Standard Deviation	0.0018	0.0022	0.0036
Sharpe Ratio	2.2210	1.7569	1.3985

In table 4.5, it presents the monthly excess return on each Pairs Trading portfolio with different pairs formation method and strategy in Post-crisis periods. During the Post-crisis, the stock market bull markets and there were many opportunities for the Pairs Trading because the market was very volatile and this leads more opportunity for pairs to be open and close. In this section, we will find out which Pairs Trading method is optimal for the investors during the European-crisis periods. Comparing the Pairs Trading strategy, one-standard deviation outperformed the two-standard deviation strategy in mean excess return and Sharpe ratio but the two-standard deviation strategy has lower standard deviation of excess return. This indicates that one-standard deviation diverge strategy could have higher mean excess return of each portfolio than that of using two-standard deviation diverge strategy.

For comparing pairing formation methods, among the one-standard deviation diverge strategy, Co-integration method portfolio has the highest mean excess return with 0.0124 and portfolio using common stock and its preferred stock pairing method has the lowest return with 0.0098. For the standard deviation of mean excess return, portfolio using pairing method with common and its preferred stock has the lowest standard deviation of excess return with 0.0024 and portfolio using correlation method and co-integration method has the same highest standard deviation of excess return with 0.0057. For the Sharpe ratio, portfolio using common stock and its preferred stock method has the highest value with 3.4380 and portfolio using pairing method with Co-integration methods has the lowest value with 1.9026.

Among the two-standard deviation diverge strategy, co-integration method portfolio has the highest mean excess return with 0.0067 and pairing method using correlation portfolio has the lowest return with 0.0022. For the standard deviation of mean excess return, pairing with common and its preferred stock portfolio method has the lowest standard deviation of excess return with 0.0018 and portfolio using co-integration method has the highest standard deviation of excess return with 0.0036. For the Sharpe ratio, portfolio using common stock and its preferred stock method has the highest value with 2.2210 and portfolio using pairing method with Co-integration methods has the lowest value with 1.3985.

According to our results, if the investor is the risk-avoider, portfolio using pairing method with common stock and its preferred stock with one-standard deviation diverge strategy is the optimal Pairs Trading strategy during the Post-Crisis periods. If the investor is the risk-lover, portfolio using co-integration pairing method with one-standard deviation diverge strategy is the optimal Pairs Trading strategy during the Post-Crisis periods.

4.5 Conclusion

The main object of this research is that propose new Pairs Trading method which is pairing with common stocks and their preferred stocks and verify the performance and also the risk. We verify the performance by comparing with classical Pairs Trading method. Such analysis was also carried our considering different values for the threshold parameter; we used one standard diverge strategy and two standard deviation strategy to compare. In addition, we analysis the different time periods (Pre Crisis, Subprime Crisis, European Crisis, Post Crisis). In Pairs Trading, it is important to find pairs that moves together and select right strategy to operate. Pairs that moves together in the past should move together in the future in order to make profit in Pairs Trading. We suggested to use Pairs Trading method pairing with common stock and its preferred stock because pairing in that method could minimize the risk of pairs that moves different directions.

The main conclusion of this paper is that pairs-trading method that pairing with common stock and its preferred stock could minimize the risk and had higher Sharpe-Ratio compare to classical method. Considering only excess return, Co-integration paring method has the highest excess return but also with highest risk. According to result, we can proclaim that using common stock and its preferred stock as a pair to operate Pairs Trading can make a profit with minimal risk. In comparing one standard deviation diverge strategy and two standard deviation diverge strategy, one standard deviation diverge strategy has higher excess return than those of two standard deviation diverge strategy; Two standard deviation diverge strategy has lower risk than one standard deviation diverge strategy, and considering both return and risk, one-standard deviation diverge strategy has higher Sharpe ratio than two-standard deviation diverge strategy.

For the sub period results, we can conclude that during pre-crisis periods, portfolio using common stock and its preferred stock paring method with one standard deviation diverge strategy has the highest Sharpe ratio and portfolio using co-integration method with one-standard deviation diverge strategy has the highest mean excess return; during Subprime-crisis, portfolio using common stock and its preferred stock paring method with one standard deviation diverge strategy has the highest Sharpe ratio and portfolio using correlation method with one-standard deviation diverge strategy has the highest mean excess return; during European Crisis, portfolio using co-integration paring method with two standard deviation diverge strategy has the highest Sharpe ratio and portfolio using co-integration method with two-standard deviation diverge strategy has the highest mean excess return;

during Post Crisis, portfolio using common stock and its preferred stock paring method with one standard deviation diverge strategy has the highest Sharpe ratio and portfolio using co-integration method with one-standard deviation diverge strategy has the highest mean excess return. According to the results, portfolio using common stock and its preferred stock method with one-standard deviation strategy has the highest Shape ratio value except during European Crisis. Therefore, considering the risk, using common stock and its preferred stock paring method for the Pairs Trading is valid to use.

In this paper, we only tested our pairing method with one standard deviation diverge strategy and two standard deviation diverge strategy in United States Stock market. For the further research, we can test with more various diverge strategy and find the optimal threshold and we can test this Pairs Trading method in other stock market such as China or Japan.

Chapter 5

Concluding Remarks

5.1 Summary and contributions

The notion of developing a new index that can explain future stock market return is of great interest to researchers and practitioners in finance. There has been a lot of research on predicting future market returns based on the usefulness of valuation ratios such as dividend-to-price ratio, book-to-market ratio, earnings-to-price ratio and payout yield (Boudoukh et al., 2007; Campbell, 1987; Campbell & Shiller, 1988; Eugene F Fama & French, 1988; Kothari & Shanken, 1997). Even though the presence of an index having market predictive power is yet debatable, the construction of an index indicating either overvalued or undervalued market condition is still a compelling area to research. Therefore, this dissertation focuses on using preferred stocks to develop a new market index that could explain future stock market returns and introduce a new Pairs Trading method that minimizes operating risks. The results are summarized as follows.

At first, in Chapter 2, we chapter showed that the return spread between common stocks and their preferred stocks could be used to assess the stock market condition and predict the future market returns. We also introduced that how the CPS-index was built to measures the spread between cumulative returns of common stocks and their preferred stocks in the market. We investigated whether this CPS-index determined either the overvaluation or the undervaluation of common stocks against corresponding preferred stocks. In addition, we illustrated the strong negative relationship between CPS-index and future realized return of stock market, while the relationship between the future realized return and any of VIX, TED, S&P500 were not so significant. Moreover, the correlation of MSE and adjusted R squared value of the regression between CPS-index and future realized return show that CPS-index has immense capability power to explain the future realized market return up to 48 months ahead of time. It also showed that the average future realized return increases when CPS-index is below the lower threshold of -0.2 and decreases when CPS-index is above the upper threshold of 0.1, with an estimated cycle time for overvalued stock market to be undervalued is thirty months. This implies that if CPS-index is higher than 0.1, the investor should take a short position and wait for thirty months to maximize profit. Lastly, we confirmed that the past and current CPS-index affects the future realized market return using stepwise regression and Granger Causality test. In addition, we found that applying this CPS index to neural network to predict the realized market return provided better prediction results in any time horizon of more than eighteen months.

In Chapter 3, we illustrated has shown that the cumulative log return spread between common stocks and their preferred stocks could be used to assess the stock market condition and predict the future market return better by using currently existing valuation ratios such as price to earning ratio, price to book ratio, and Cyclically Adjusted price to earning ratio. We reached several conclusions using univariate regression and presented that CPS-index has better explanatory power of explaining market future return compared to currently existing valuation ratios such as price to earning ratio, price to book ratio, and price to operational earning ratio, especially predicting more than 30 months ahead. Since CPS-index is not built by any of earning or book value, this result gave us the possibility to enhance the market predictability by using it together. In the same chapter, we also used multivariate regression to present adding CPS-index to currently existing valuation ratio to enhance the market predictability. We used multivariate regression models using two independent variables to four independent variables in every multivariate regression model that have the most market predictability power having CPS-index as independent variable. According to results, using CPS-index could explain predict the long term market return better. Lastly, we outlined the recommended parameters for using the CPS-index. We demonstrated that building the CPS-index using the difference between cumulative log return of common stocks that have corresponding preferred stock with those preferred stocks from seven years explains the market future return better than that of any other parameter. Our results showed that using seven year CPS-index and eight year CPS-index have significantly high explanatory power. We showed that both seven year CPS-index and eight year CPS-index have better market predictability power than currently existed ratios. Chapter 3 concluded with the confirmation that past and current seven year CPS-index affects future realized market return, using the Granger Causality test.

Lastly, Chapter 4, proposed a new Pairs Trading method which is pairing with common stocks and their preferred stocks and verifying both the performance and the risk. This chapter verified the performance by comparing results of our method to the results of the traditional Pairs Trading method. Such analysis was also carried out considering different values for the threshold parameter; it used one standard diverge strategy and two standard deviation strategy to compare. We also analyzed different time periods (Pre Crisis, Subprime Crisis, European Crisis, Post Crisis). In Pairs Trading, it is important to find pairs that moves together and select the right strategy to operate. The main conclusion of this chapter was that a pairs-trading method that pairs a common stock with its preferred stock could minimize the risk and have higher Sharpe-Ratio compares to the classical method. Considering only excess return, Co-integration Paring method has the highest excess return but also presents the highest risk. Chapter 4 claims that according to our results, using common stock and its preferred stock as a pair to operate Pairs Trading can make a profit with minimal risks. In comparing one standard deviation diverge strategy and two standard deviation diverge strategy, one standard deviation diverge strategy has higher excess return than those of two standard deviation diverge strategy; Two standard deviation diverge strategy has lower risk than one standard deviation diverge strategy. Considering both the return and the risk, one-standard deviation diverge strategy has higher Sharpe ratio than the two-standard deviation diverge strategy. Throughout the sub periods, portfolios using common stocks and their preferred stock method Pairs Trading had the highest Sharpe Ratios.

The contribution of this dissertation is as follows. The new market index, and a CPS-index that can evaluate current market conditions and predict future market returns. Using our strategies, CPS-index investors can decide whether the market is overvalued or undervalued compared to the past and help to decide their investment position. Since CPS-index is only using the difference between cumulative log return of common stocks and those of their preferred stocks, this index can be used with the currently existing market valuation ratios. By using with the existing market valuation ratios, it significantly enhances the long term market predict power. The other benefit of CPS-index is that it is easy to use and understand. What good is it investors if a newly developed market is too sophisticated to understand and use? In addition, using parameter estimation, this dissertation developed a more robust CPS-index. This robus CPS-index is not highly affected by the starting date of CPS-index. In addition, this dissertation presented the new Pairs Trading method of using common stock and its preferred stock as a pair. By using this method, it empirically showed that investors could save time and achieve higher profit considering its risk. This research also suggested which Pairs Trading method is suggested depending on the type of financial periods, and whether investors are risk-seeking or risk-averse. In a bigger picture, this dissertation has shown that studying preferred stocks has the possibility to explain complex financial markets. Currently, there is limited research on preferred stocks, and this dissertation could link preferred stocks and its market predictability research.

5.2 Limitation and Future Work

This study can be further developed in the following ways. First, the CPS-index is the difference between cumulative returns of common stocks that have corresponding preferred stocks and cumulative returns of preferred stocks. Therefore, CPS index directly indicates how commons stocks are valued compared to their preferred stocks. There should be a better way to show this spread apart from using the cumulative log return. Second, analyzing longer periods of data, instead of the sixteen years (from 2000 to 2016) used in this paper, one might get more robust results. This paper only applied CPS-index to the US market data, so one could investigate whether the CPS-index has the equivalent prediction power in other global markets such as Europe and China. Third, the upper and lower threshold values of CPS-index could to be studied in more scientific ways in order to maximize the profit of its investment strategy. In addition, portfolios consisting of pairs of common stock and its preferred stock can be developed using CPS-index, and then verified in comparison with other benchmark portfolios. Lastly, more accurate parameters tuning and having right rule for finding optimal starting point for the CPS index could be researched using a longer period of data.

References

Ang, A., Bekaert, G., & Wei, M. (2007). Do macro variables, asset markets, or surveys forecast inflation better? *Journal of monetary Economics, 54*(4), 1163-1212.

Badshah, I., Frijns, B., Knif, J., & Tourani-Rad, A. (2016). Asymmetries of the intraday return-volatility relation. *International Review of Financial Analysis, 48*.

Blanchard, O. J., & Watson, M. W. (1982). Bubbles, Rational Expectations and Financial Markets. *National Bureau of Economic Research*.

Boudoukh, J., Michaely, R., Richardson, M., & Roberts, M. (2007). On the importance of measuring payout yield: Implications for empirical asset pricing. *The Journal of Finance, 62*(2), 877-915.

Brunnermeier, K. M. (2009). Deciphering the liquidity and credit crunch 2007-2008. *Journal of Economic perspectives, 23*, 77-100.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics, 18*(2), 373-399.

Campbell, J. Y., & Shiller, R. J. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *The Review of Financial Studies, 1*(3), 195-228.

Caraiani, P. (2017). The predictive power of local properties of financial networks. *Physica A: Statistical Mechanics and its Applications, 466*, 79-90.

Caraiani, P., & Adrian, C. C. (2018). The effects of Monetary Policy on Stock Market bubbles at Zero Lower Bound: Revisiting the Evidence. *Economic letters*.

Chacko, G., Das, S. R., & Fan, R. (2016). An index-based measure of liquidity. *Journal of Banking & Finance, 68*(1), 162-178.

Chen, S., Bao, S., & Zhou, Y. (2016). The predictive power of Japanese candlestick charting in Chinese stock market. *Physica A: Statistical Mechanics and its Applications, 457*, 148-165.

Diba, B. T., & Grossman, H. I. (1987). On the inception of rational bubbles. *The Quarterly Journal of Economics, 102*(3), 697-700.

Elliott, R., Hoek, J. V. D., & Malcolm, W. (2005). Pairs Trading. *Quantitative Finance, 5*(3), 271-276.

Emanuel, D. (1983). A Theoretical Model for valuing preferred stock. *The Journal of Finance, 38*(4), 1133-1155.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of political Economy, 96*(2).

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*(1), 23-49.

Gali, J., & Gambetti, L. (2015). The Effects of Monetray Policy on Stock Market Bubbles: Some Evidence. *American Economics Journal: Macroeconomics, 7*(1).

Gatev, E., Goetzmann, W. N., & Rouwenhorst, K. G. (2006). Pairs Trading: Performance of a relative-value arbitrage rule. *The Review of Financial Studies, 19*(3), 797-827.

Greenwood, R., Shleifer, A., & You, Y. (2017). Bubbles for Fama. *Working paper*.

Heiberger, R. H. (2018). Predicting economic growth with stock networks. *Physica A: Statistical Mechanics and its Applications, 489*, 102-111.

Hong, G., & Susmel, R. (2003). Pairs-Trading in the Asian ADR Market. *Working paper*.

Horta, P., Lagoa, S., & Martins, L. (2014). The impact of the 2008 and 2010 financial crises on the Hurst exponents of international stock markets: Implications for efficiency and contagion. *International Review of Financial Analysis, 35*, 140-153.

Horta, P., Lagoa, S., & Martins, L. (2014). The impact of the 2008 and 2010 financial crises on the Hurst exponents of international stock markets: Implications for efficiency and contagion. *International Review of Financial Analysis, 35*.

Ioannis, P., & Sermpinis, G. (2016). Modelling and trading the US implied volatility indices. Evidence from the VIX, VXN and VXD indices. *International Journal of Forecasting, 32*(4), 1268-1283.

Jegadeesh, N., & Titman, S. (1993). Returns to Buying Winners and Selling Losers: Implications for Stock Market Efficiency. *The Journal of Finance, 48*(1), 65-91.

Jorda, O., Schularick, M., & Taylar, A. (2015). Leveraged bubbles. *Journal of monetary Economics, 76*, 1-20.

Kanas, A. (2012). Modelling the risk-return relation for the S&P 100: The role of VIX. *Economic Modelling, 29*(3), 795-809.

Kindleberger, C. P., & Aliber, R. Z. (2005). *Manias, Panics and Crashes: A History of Financial Crises.* . Hoboken,NJ: John Wiley and Sons, Inc.

Kothari, S. P., & Shanken, J. (1997). Book-to-market, dividend yield, and expected market returns: A time-series analysis. *Journal of Financial Economics, 44*(2), 169-203.

Kristjanpoller, W., & Minutolo, M. C. (2015). Gold price volatility: A forecasting approach using the Artificial Neural Network-GARCH model. *Expert Systems with Applications, 42*(20), 7245-7251.

LeRoy, S. F., & Porter, R. D. (1981). The Present-Value Relation: Tests Based on Implied Variance Bounds. *The Econometric Society, 49*(3), 555-574.

Li, Y., Ng, D. T., & Swaminathan, B. (2013). Predicting market returns using aggregate implied cost of capital. *Journal of Financial Economics, 110*(2), 419-436.

Linn, S. C., & Pinegar, J. M. (1988). The effect of issuing preferred stock on common and preferred stockholder wealth. *Journal of Financial Economics, 22*(1), 155-184.

M.Garber, P. (1990). Famous Frist Bubbles. *The Journal of Ecomomic Perspectives, 4*(2), 35-54.

Nath, P. (2004). High Frequency Prairs Trading with U.S. Treasury Securities: Risks and Rewards for Hedge Funds. *London Business*

School.
Perlin, M. S. (2009). Evaluation of pairs-trading strategy at the Brazilian financial market. *Journal of Derivatives and Hedge Funds, 15*(2), 122-136.
Phillips, P. C. B., Wu, Y., & YU, J. (2011). Explosive behavior in the 1990s Nasdaq. *INTERNATIONAL ECONOMIC REVIEW, 52*(1), 201-226.
Shiller, R. (1981). Do stock prices move too much to be justified by subsequent changes in dividends? *American Economics review, 71*, 421-436.
Song, J. W., Ko, B., & Chang, W. (2018). Analyzing systemic risk using non-linear marginal expected shortfall and its minimum spanning tree. *Physica A: Statistical Mechanics and its Applications, 491*, 289-304.
Stickel, S. E. (1986). The effect of preferred stock rating changes on preferred and common stock prices. *Journal of Accounting and Economics, 8*(3), 197-215.
Tirole, J. (1982). On the possibility of speculation under rational expectations. *Econometrica, 50*(5), 1163-1181.
Yangru, W. (1997). Rational bubbles in the stock market: Accounting for the U.S. stock-price volatility *Economic Inquiry, 35*, 309-319.

초 록

최근 다양한 금융위기 이후, 시장의 위험을 측정 하고 안전한 수익을 낼 수 있는 금융 상품의 개발의 토대가 될 수 있는 금융시장 분석의 중요성은 더욱 강조 되고 있다. 특히 다양한 금융 시장의 특성을 활용한 지표들을 중심으로 시장의 상태와 미래 수익률 과의 관계를 설명 하였는데 대표적인 위험 지표로는 TED와 VIX가 있고 시장 가치 평가 지표로는 Price to Earning ratio, Price to Book ratio, CAPE (Cyclically Adjusted Price to Earning ratio), Price to Operational Earning ratio 등이 있다. 사전의 연구들은 이러한 지표들을 통하여 현재 주식시장의 상태를 측정하고, 정확한 측정을 통하여 미래 수익률을 설명 하였다. 많은 연구들이 기존에 존재하고 있는 시장 지표들을 변형하고 발전 시켜 미래 수익률을 보다 더 잘 설명 하는 지표를 개발 한 것과 달리 본 논문에서는 그동안 잘 쓰이지 않았던 주식 시장의 우선주를 활용하여 시장의 미래 수익률을 설명 할 수 있는 지표를 개발하고 이를 검증 하였다. 먼저 본 논문에서는 주식 시장에서의 우선주를 가지고 있는 본주들의 누적 수익률과 우선주들의 누적 수익률의 편차를 이용하여 CPS-index(Common Preferred Spread index)를 개발하였다. 이 CPS-index는 현재 주식시장에서 우선주를 가지고 있는 보통주 들이 우선주들에 대비하여 과거보다 얼마나 더

올랐는지 혹은 떨어졌는지를 알려준다. 본 논문에서는 CPS-index 가 높을 때 즉 본주들이 우선주보다 더 많이 올랐을 때를 시장이 고평가 되어 있다고 판단하고 CPS-index 가 낮을 때 즉 본주들이 우선주보다 더 적게 올랐을 때를 시장이 저평가 되어 있다고 판단하여 이를 미래 장기 수익률과의 상관관계 분석을 실시하였다. 그 결과 CPS-index는 미래 장기 수익률과 매우 높은 음의 상관관계를 가지고 있었고, CPS-index가 높을 때 공매도 하고 CPS-index가 낮을 때 매수하는 투자전략을 제시하여 실증적으로 이 투자 전략을 활용 시 높은 수익률을 낼 수 있다는 것을 보여주었다. 또한 granger causality test 와 신경망 분석을 활용하여 CPS-index가 미래 수익률을 보다 정확하게 예측 할 수 있다는 것을 보여주었다. 두번째로는 CPS-index를 기존의 미래 수익률을 설명할 수 있는 지표들과 비교 분석하여 기존 지표들과 함께 사용하였을 때 미래 수익률을 보다 정확하게 설명 할 수 있는지를 보여주었고, 또한 parameter tuning을 통하여 어느 정도의 과거 데이터를 활용해야 보다 정확한 CPS-index를 만들 수 있는지 제시하였다. 이 연구를 통하여 CPS-index를 기존의 지표들인 Price to Earning ratio, Price to book ratio, Price to Operational Earning ratio 등과 같이 활용하였을 때 시장의 미래 수익률 기존의 지표들만 활용하였을 때 보다 월등히 더 잘 설명 할 수 있음을 보였고 이를 통하여 CPS-index의 시장 평가 지표로의 가능성을 확인하였다. 마지막으로 우선주와 우선주를 가지고 있는 본주를

활용한 Pairs Trading 방법을 제시하였다. 기존의 방법인 같이 움직이는 두 주식을 활용한 Pairs Trading 방법과 달리 우선주와 본주를 활용하면 pair를 찾는데 필요한 시간과 계산을 줄일 수 있고, Pairs Trading 운영기간 동안 두 주식이 다른 방향으로 움직이는 위험을 줄일 수 있다고 판단하였다. 본 논문에서는 각 금융시장 상황에 따른 최적의 Pairs Trading 방법을 찾는 것이 중요하다고 생각이 들어 이를 검증하는데 초점을 맞추었고, 그 결과 우선주와 본주를 활용한 Pairs Trading 방법이 기존의 Pairs Trading 방법보다 수익률은 좋지 않았지만, 위험 대비 수익률인 Sharpe Ratio는 월등히 좋다는 것을 검증하였다. 결과적으로 본 논문에서는 기존에 연구가 진행되지 않았던 주식 시장에서의 우선주를 가지고 있는 본주와 우선주의 움직임을 분석하여 현재 주식시장의 상태를 측정하고 장기 미래수익률을 설명 할 수 있는 지표를 개발하여 투자자들이 보다 안전하게 투자할 수 있는 투자 전략을 제시였고, 안전한 투자전략인 Pairs Trading에 적용하여 pair를 찾는데 필요한 시간과 돈을 절약하고 위험 대비 수익률을 높일 수 있는 방법을 제시하였다. 이를 통하여 우선주의 관한 연구가 활발히 진행될 것이며 일반 투자자들이 쉽게 이해하고 활용할 수 있는 투자 지표와 투자방법을 제공하였다.

주요어: 우선주, Pairs Trading, Predictive index, Market valuation index, Spread return,

학 번: 2015-30238