LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

MASTER OF SCIENCE IN FINANCE

MASTER'S FINAL WORK

DISSERTATION

PORTFOLIO SELECTION: A STUDY USING PRINCIPAL COMPONENT ANALYSIS

MADALENA BAIOA PARAÍSO NUNES

OCTOBER - 2017

MASTER OF SCIENCE IN

FINANCE

MASTER'S FINAL WORK

DISSERTATION

PORTFOLIO SELECTION: A STUDY USING PRINCIPAL COMPONENT ANALYSIS

MADALENA BAIOA PARAÍSO NUNES

SUPERVISION:

SARA BÁRBARA DUTRA LOPES

OCTOBER – 2017

Abstract

In this thesis we apply principal component analysis to the Portuguese stock market using the constituents of the PSI-20 index from July 2008 to December 2016. The first seven principal components were retained, as we verified that these represented the major risk sources in this specific market. Seven principal portfolios were constructed and we compared them with other allocation strategies. The 1/N portfolio (with an equal investment in each of the 26 stocks), the PPEqual portfolio (with an equal investment in each of the 7 principal portfolios) and the MV portfolio (based on Markowitz's (1952) mean-variance strategy) were constructed. We concluded that these last two portfolios presented the best results in terms of return and risk, with PPEqual portfolio being more suitable for an investor with a greater degree of risk aversion and the MV portfolio more suitable for an investor willing to risk more in favour of higher returns. Regarding the level of risk, PPEqual is the portfolio with the best results and, so far, no other portfolio has presented similar values. Therefore, we found an equally-weighted portfolio among all the principal portfolios we built, which was the most risk efficient.

Keywords: principal component analysis; principal portfolios; principal components; efficient portfolio; risk.

Resumo

Nesta tese aplicámos a análise de componentes principais ao mercado bolsista português usando os constituintes do índice PSI-20, de Julho de 2008 a Dezembro de 2016. Os sete primeiros componentes principais foram retidos, por se ter verificado que estes representavam as maiores fontes de risco deste mercado em específico. Assim, foram construídos sete portfólios principais e comparámo-los com outras estratégias de alocação. Foram construídos o portfólio 1/N (portfólio com investimento igual para cada um dos 26 ativos), o portfólio PPEqual (portfólio com igual investimento em cada um dos 7 principal portfólios) e o portfólio MV (portfólio que tem por base a teoria moderna de gestão de carteiras de Markowitz (1952)). Concluímos que estes dois últimos portfólios apresentavam os melhores resultados em termos de risco e retorno, sendo o portfólio PPEqual o mais adequado a um investidor com maior grau de aversão ao risco e o portfólio MV mais adequado a um investidor que estaria disposto a arriscar mais em prol de maior retorno. No que diz respeito ao nível de risco, o PPEqual é o portfólio com melhores resultados e nenhum outro portfólio conseguiu apresentar valores semelhantes. Assim encontrámos um portfólio que é a ponderação de todos os portfólios principais por nós construídos e este era o portfólio mais eficiente em termos de risco.

Palavras-Chave: análise de componentes principais; portfólios principais; componentes principais; portfólio eficiente; risco.

Acknowledgements

First of all, I would like to express my deepest gratitude to my supervisor, Sara Dutra Lopes, who has been relentless from the first to the last day, giving me all the support, guidance, prompt responses and suggestions throughout the entire process. Without her this thesis would not have been possible. Sara, thank you so much!

No less important, I have to thank my parents and brother who have always given me unconditional support, unconditional love, never allowing me to give up and always doing everything to make all my dreams come true. António, mom and dad, I love you. Thank you!

Table of Contents

List of Figures

List of Tables

1 Introduction

Lowenfeld (1909) has been credited with being the first to describe the benefits of diversification, however, Markowitz (1952) carried out most of the work in this area.

According to Markowitz (1952), it is advisable to invest in securities with low correlation, being necessary to diversify between industries, preferably industries with different economic characteristics.

By announcing his theory of mean-variance (MV), Markowitz (1952) became the driving force behind modern portfolio theory. After his initial work, many others began to give importance to portfolio diversification using a heterogeneous set of assets instead of holding a single stock. This theory has greatly contributed to the advancement of the economy in general but, as it would be expected, it also has some limitations.

For example, Chopra & William (1993) have noted that mean-variance theory is sensitive to errors in the inputs estimates, which leads to large changes in the composition of portfolios that were previously considered optimal.

Security investments are characterized by the uncertainty of their returns and the correlations between security returns. Low correlations or negative correlations between securities make diversification possible, that is why the analysis of this type of investment is difficult. When dealing with only a few securities, looking at their variances and their covariances or correlations can be effortless, but if we are studying a large number of securities we will not be able to analyse their variances, nor their covariances or correlations. It was to solve this type of problems that Principal Component Analysis (PCA) was devised.

According to Jolliffe (1986), the main goal of PCA is to reduce the size of data sets, without losing much variation (information). This can be achieved by transforming the original variables into a new set: the so-called principal components (PCs), which are uncorrelated with each other. The PCs must be ordered so that the first ones can retain most of the variation contained in the original variables. This new approach will reduce the complexity of the problem previously described.

Principal Component Analysis (PCA) has already been applied in a wide variety of areas, for example, demography, biology, psychology or genetics; but studies using it in the finance sector are fairly recent and they are still uncommon, particularly in the context of portfolio management. That is precisely why we decided to focus our analysis on the application of PCA in portfolio management.

To simplify the analysis, we chose to study a portfolio consisting only of stocks.

Applying PCA to portfolio construction implies reducing the complexity of a portfolio by transforming its original variables into new ones, the principal components, which are uncorrelated and will represent uncorrelated risk sources. Once the original variables are transformed, we can consider the principal components as individual investment assets, instead of creating a portfolio based on the underlying stocks. This will facilitate the essential examination for portfolio selection.

Our research is organized as follows: Chapter 2 presents the literature review; Chapter 3 describes the data; Chapter 4 determines how many components should be retained so that we can make a correct analysis of the stock market; in Chapter 5 the construction of the portfolios based on the principal components retained in the previous chapter is presented and we compare the principal portfolios with other allocation strategies; Chapter 6 presents the conclusions and considers potential further research.

2 Literature Review

2.1. Principal Component Analysis – some definitions

PCA is a statistical method that reduces the size of a data set without losing significant information.

From a mathematical perspective, Sharma (1996) presents PCA as a technique to compose new variables as linear combinations of the original variables such that the new variables are orthogonal to each other and have maximal variance. Considering a set of n observable variables, $x_1, x_2, \dots, x_n$, we will form n linear combinations as

$$\xi_1 = w_{11}x_1 + w_{12}x_2 + \cdots + w_{1n}x_n$$

$$\xi_2 = w_{21}x_1 + w_{22}x_2 + \cdots + w_{2n}x_n$$

$$\dots$$

$$\xi_n = w_{n1}x_1 + w_{n2}x_2 + \cdots + w_{nn}x_n$$

(Equation 2.1)

where $\xi_1, \xi_2, \dots, \xi_n$ represent the n principal component and w_{ij} the weight of the j^{th} variable for the i^{th} principal component. The weights, w_{ij}, are estimated such that:

1. ξ_1, the first PC, accounts for the maximum variance in the data; ξ_2, the second PC, explains the maximum variance that has not been accounted for by the first PC, and so on.
2. $w_{i1}^2 + w_{i2}^2 + \cdots + w_{in}^2 = 1 \qquad i = 1, \dots, n$ (Equation 2.2)
3. $w_{i1}w_{j1} + w_{i2}w_{j2} + \cdots + w_{in}w_{jn} = 0$ for all $i \neq j$. (Equation 2.3)

Sill according to Sharma (1996), considering λ_1, λ_2,..., λ_n the eigenvalues (ordered from largest to smallest) of the covariance/correlation matrix and γ_1, γ_2,…, γ_n the corresponding eigenvectors, the solution to this maximization problem is given by choosing the weights of the new variables as $w_1 = \gamma_1$, $w_2 = \gamma_2$,…, $w_n = \gamma_n$ and in this case the variance of the variable ξ_i is given by the eigenvalue λ_i (also called the latent root of the principal component i).

2.2. Principal Component Analysis in Portfolio Construction

To address an old problem related to the efficient portfolio, Partovi & Caputo (2004) presented a new method based on PCA. This consists in reorganizing the original set of assets into a set of uncorrelated portfolios, called the Principal Portfolios (PPs). Through these PPs, it is possible to develop an uncorrelated asset investment environment and thus simplify the portfolio selection analysis.

Whereas it is easy to verify that actual market assets, such as stocks, can be highly correlated, by using this methodology, it is possible to transform them into a set of totally uncorrelated assets, so that the investor can choose between a set of uncorrelated portfolios.

For example, Yang (2015) applied PCA to the Australian stock market using the ASX200 index and its constituents and found there were 10 PCs that represented the major risk sources in this stock market; based on this information, 10 PPs were constructed. Thus, it was noted that when applied to PPs rather than to underlying stocks, allocation strategies reduced the risk considerably, and they could have even avoided the dramatic fall in the share prices during the 2008 financial crisis.

3 Data

Our study focuses on the Portuguese stock market, through the analysis of the components of the PSI-20 index (Portuguese Stock Index), between July 2008 and December 2016. The PSI-20, created in 1992, is composed of twenty stocks with the highest market capitalization, making it the benchmark index in Portugal.

Given that the PSI-20 index undergoes frequent changes in its composition, it was necessary to identify all the stocks that were included in the PSI-20, during the period under study. We initially compiled a list of 32 stocks, and we obtained the prices and returns for all of them using the Bloomberg terminal, but we immediately understood there were some issues associated with these data:

1. Some stocks had missing information, since some days had no pricing data, even when they were neither weekends nor holidays. Thus, we decided to assign zero returns in those occasions.
2. Some of the constituents of the PSI-20 were not quoted on the Stock Exchange during the period under study and therefore had to be removed from our analysis. They included *Brisa - Autoestradas de Portugal*, *BANIF*, *Banco Espírito Santo* and *Espírito Santo Financial Group*. Still, we tried to compensate the absence of *Banco Espírito Santo* by including *Novo Banco* in our analysis, but we could not do it, since it had no quotations because the shares are held in the Resolution Fund[1].
3. Some of the constituents of PSI-20 were only listed in the stock exchange after the beginning of the study period and therefore did not have sufficient data. Such was

[1] "The Resolution Fund was created in 2012 and its primary goal is providing financial support for the implementation of resolution measures determined by Banco de Portugal. The Resolution Fund is a public-law legal person with administrative and financial autonomy" – Novo Banco (2016).

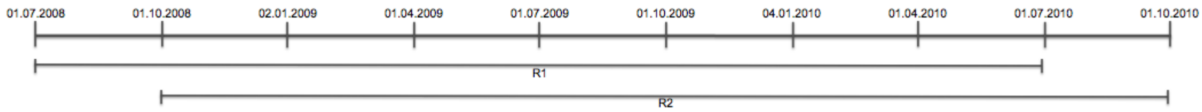
the case with *CTT - Correios de Portugal* and *Caixa Económica Montepio Geral*, which were only quoted after 2013. However, it was not possible to start the study with later data, since this was the only possible way to analyse the impact of the 2008 crisis in our study. Therefore, we were left with a total of 26 stocks (Appendix A).

Appendix B shows the characteristics of the PSI-20 return data with Figure 3.1 and Figure 3.2. In Figure 3.1 we present the plots for the daily returns of the PSI-20 index, volatility clustering, boxplot of PSI-20 index returns and the normal Q-Q plot for PSI-20 index returns. The plot for the daily returns shows that the highest absolute returns were recorded between the years 2008 and 2010, this can be attributed to the financial crisis experienced during this period. It is also relevant to mention that we can see signs of volatility clustering, as "large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes" (Mandelbrot, 1963). The Boxplot of PSI-20 Index Returns indicates the existence of some outliers and the Quantile-Quantile Plot of PSI-20 Index Returns shows that the returns distribution is not Gaussian.

Figure 3.2 shows a skewness of 0,030 indicating a positive asymmetry in the distribution, which, in turn, suggests that bad scenarios are less common than good scenarios. This circumstance is rare since the stock returns are known to be negatively asymmetric, however, it is important to remember that the PSI-20 is composed of twenty of the "best" Portuguese companies. On the other hand, kurtosis presents a value of 5,644, which indicates that we are in the presence of a leptokurtic distribution (heavy tails) and therefore there is a greater chance of extreme outcomes, when compared to the normal distribution.

4 Principal Component Analysis - Application

After selecting the set of 26 stocks with complete return information for the period under study, we applied a rolling window approach to extract the principal components.
To do so, we first had to select the span of the rolling windows. Given the size of the data set, a rolling window of two years, rebalanced quarterly, seemed the most adequate approach. The next step required testing if a rolling window with this size would be adequate for applying PCA. For this purpose, we used the Kaiser-Meyer-Olkin (KMO) test, which measures the amount of variance among stocks that can be regarded as common variance. KMO statistic varies between 0 and 1; a value close to zero indicates that the stocks do not share a lot of common variation, and in this case, the PCA will not yield helpful information. According to Kaiser (1974) values below 0,50 are considered unacceptable; values between 0,50 and 0,59 are miserable; values from 0,60 to 0,69 are mediocre; from 0,70 to 0,79 are middling; between 0,80 and 0,89 are meritorious, and values equal to or greater than 0,90 are considered marvellous.
In our case, the KMO average was 0,9383, with a maximum value of 0,9662 and a minimum value of 0,9155. Given this, we considered that the sample size selected for our study was adequate for applying PCA, so we got the following rolling windows for the first two periods under study:



The subsequent rolling windows will look similar to the first two, with starting dates separated by 3 months. Thus, we will have 27 rolling windows, each one associated with a two-year period (rebalanced quarterly), as presented in Appendix C, Table II.

As explained in 2.1, PCA is a tool that allows the construction of new variables, which are linear combinations of all the original variables. However, it is important to mention that in order to simplify the scope of the study, we chose to use a smaller number of new variables than the original ones. According to Jolliffe (1986), the key to do so is replacing the elements with a smaller number of PCs, without losing considerable information. In our case, the KMO statistic is high and we should not worry because this suggests that more variation can be accounted for by the first few principal components. Furthermore, the eigenvalues of the principal components decrease quickly, along with their importance, thus, we should select a reduced number of PCs.

There is no set number of components that should be retained; instead, the decision depends on how much information we are willing to sacrifice for the sake of simplifying the work of the analyst. Thus, we selected the minimum number of PCs suitable for all rolling windows.

There are several rules that can be taken into account when deciding on the number of principal components to be retained. We analysed our data according to these rules using the information contained in Appendix C, Table III.

4.1. Cumulative Percentage of Total Variation

According to Jolliffe (1989), the cumulative percentage of total variation is the most useful rule for choosing the number (m) of components that should be retained. This rule consists in selecting a cumulative percentage out of the total variance that we want PCs to explain. Therefore, the required number of PCs is the smallest value of m that exceeds the cumulated variance that we want to explain.

In our case, we decided that PCs should explain, at least, 50% of the total variance. In some time windows this is only possible if we select at least seven PCs. This is the case for R11, R14, R15, R16, R17 and R18.

4.2. Kaiser's Rule

Kaiser (1960) suggested the best way to solve the problem of choosing the number of components that should be retained was to select all the variables with eigenvalues greater than one; provided that, if the eigenvalues are less than that, the principal components will contain less information than the original variables.

As shown by the results presented in Table III of Appendix C, for R19, R21 and R25, only seven PCs complied with Kaiser's rule, and were thus retained.

4.3. The Scree Graph

The last criteria we considered is the so-called scree graph (Cattell, 1966). The scree graph is a plot composed of eigenvalues and a number of components, exhibiting a downward curve. To decide how many components should be retained, we should find a point in the graph beyond which the remaining eigenvalues are "small". Graphically, this corresponds to the point where the curve stops falling steeply (this point is usually called the 'elbow'), since this graph studies the marginal cumulative variance, that is, what happens if we add one more component. This rule is more subjective than the former and, because of that, we opted instead to continue with our initial strategy of selecting seven PCs.

Finally, Table IV of Appendix C, shows that the selection of the first seven PCs is appropriate.

5 Principal Portfolios

If at the beginning of our study we had 26 stocks under analysis, we now have 26 uncorrelated principal components in which the first seven represent the main risk drivers of stock returns. According to Partovi & Caputo (2004) we can transform any set of correlated assets into a set of uncorrelated assets, which can be treated as individual investments, through the construction of Principal Portfolios.

The major advantage of constructing PPs is that each portfolio will represent a different risk source, which is uncorrelated with the other risks in the market. Thus, investors can choose which principal portfolio to hold while knowing they will be exposed to a single risk source. Therefore, it is no longer a concern how the portfolio interacts with the others, and the investor can focus on the variance and the return of the selected portfolio.

Through the construction of principal portfolios, it is possible to start from an idea of independence that will be confronted with other strategies in this chapter, such as the naive allocation strategy (1/N portfolio), Markowitz's (1952) mean-variance strategy and a portfolio built based on equal investments on each of the seven principal portfolios (PPEqual), so we can verify if a PCA-based strategy is the most appropriate allocation approach for the data set we consider in our study.

5.1. Constructing Principal Portfolios

Using the first seven PCs obtained using SPSS, we verified that the principal component one had mostly positive coefficients. However, some of its values were negative for some timeframes in our analysis. This happened because some stocks had information losses (see *Chapter 3 – Data*).

Following the reasoning of Fenn et al. (2011), the principal component one is the market component, which means that its contribution to all stocks should have the same direction. Knowing this, we decided to transform the negative coefficient (which sometimes appears in *Corticeira Amorim* stocks) in a positive value like the others. Because its coefficient in PC1 is close to zero, we assumed that transforming this value into a positive would not affect the performance of the first principal portfolio.

We can now define the weights of the investments in each stock. Since Partovi & Caputo (2004) do not take into account whether the coefficients are positive or negative, and this leads to exaggerated weights for certain stocks, we followed instead Yang (2015) to calculate these weights, in order to make them more stable. To do this, and having established that a positive coefficient value indicates a long position whereas a negative value implies a short position, we divided the coefficients by the sum of all the coefficients, positive or negative, depending on their sign. In this case, the weights are obtained through Equation 5.1:

$$W_i^k = \begin{cases} \dfrac{PC_i^k}{\sum_{j=1}^{26} PC_j^k \mathbf{1}_{\left\{PC_j^k \geq 0\right\}}}, & if\ PC_i^k \geq 0 \\ \dfrac{PC_i^k}{\sum_{j=1}^{26} \left| PC_j^k \mathbf{1}_{\left\{PC_j^k < 0\right\}} \right|}, & if\ PC_i^k < 0 \end{cases}$$

(Equation 5.1)

Where PC_i^k is the i^{th} score of the Principal Component k, with $k = 1, 2, \dots, 7$ and $i = 1, 2, \dots, 26$.

5.2. Comparison with PSI-20 Index

Next, we present the results of the trajectories comparisons between the principal portfolios and the PSI-20 index value, and we consider the relative performance given by Equation 5.2 as measure.

$$Relative\ Performance\ of\ PP_n = \frac{PP_n(t) - PSI20(t)}{PSI20(t)}$$

(Equation 5.2)

Where $PP_n(t)$ is the principal portfolio n value[2] at time t, $PSI20(t)$ is the index value at time t.

Figure 5.1 of Appendix D, shows the trajectory of the seven PPs and the PSI-20 index, allowing easy comparisons between the portfolios. On the right side of this Figure it is possible to analyze the plots of the relative performances of the PPs when compared to the index.

The analysis of Figure 5.1 shows the first few years under study (up to the third quarter of 2012) were critical for the PSI-20 index, as well as for PP1 and PP2 (PP2 only began to improve in the first quarter of 2013). The poor performance of the index and the PPs for this period can be attributed to the global crisis.

Although the global crisis was felt between 2007 and 2008, its financial consequences in Portugal were felt mostly between 2010 and 2012, a period when austerity measures were imposed, and which had a great impact on the Portuguese market. Due to this, both PSI-20 and PP1 registered the lowest value of the period under review in June 2012, with a significant increase thereafter, which was recorded until its peak in April 2014. PP2 also registered its best value in the same period. Nevertheless, when looking at the

[2] For all investments, we have considered an initial value of one thousand euros.

relative performance of PP1 and PP2 with the PSI-20, it was evident that during the years with greatest financial crisis, these graphs present negative values, suggesting that the principal portfolios were more sensitive to the crisis than the index. Whereas the remaining principal portfolios were not so affected by the financial crisis.

Figure 5.1 shows that PP1 was closely related to the PSI-20 and, in fact, looks like an amplified version of the PSI-20, provided that it was more volatile than the index itself. PP2 was also related to the PSI-20 index, however, when compared to it, PP2 presented an unfavourable relative performance, since their values were lower than those verified by the PSI-20 index.

Table V of Appendix D shows the price correlations between the principal portfolios and the PSI-20 index. Looking at this table we can understand that PP1, PP2 and PP7 were more correlated with the PSI-20 index than the remaining PPs; and we can also verify that only the first two presented positive values. Whereas PP7 had a significant negative correlation (-0,743). Also, PP1 presented a value of 0,798, which is higher than the one exhibited by PP2, 0,730; this difference was expected. The remaining PPs were negatively correlated, showing opposite behaviour to the index. Yet, through the analysis of Figure 5.1 we can verify that they presented, generally, more favourable values than the index and, therefore, their relative performance charts are mostly positive.

Table VI shows us that PP1 was the only PP with high daily return correlations with the PSI-20 index (presenting a value of 0,906), and that all the remaining PPs had negative values for the daily return correlations.

Although PP2 presented high values related to its price correlation with the PSI-20 index, the value of the daily return correlation was negative (-0,165), meaning that PP2 tracked the index in the long term, but its daily movements were not correlated with the PSI-20.

Close observation of PP3 shows it is one of the most volatile portfolios under analysis. PP3 began to grow in the second quarter of 2011, a period marked by the financial crisis, for this reason, the PSI-20 index continued to decline, failing to keep pace with the evolution of this PP3, resulting in high relative performance figures for this portfolio. Still, PP3 saw its highest peak in February 2013 and, from then on, it started to decline. At this point the PSI-20 was already recovering from the crisis and, although it was growing at a slow pace, PP3 decreased so fast that the PSI-20 managed to exceed its value between March and May 2014 (although it presented fairly little relative performance values). However, the PSI-20 was already in decline and continued to fall until January 2015, when it began to recover and outperform PP3 between March and May 2015, as the year before. In addition to having exceeded PP3 for a short period of time, the relative performance of PSI-20 was insignificant and PP3 quickly recovered, having reached its second highest peak in November 2015 and decline after this peak, being, again, outperformed by PSI-20 in the last months of our study.

As Figure 5.1 shows, neither PP outperformed the PSI-20 index permanently over the full period under analysis, yet, PP4 was the closest to achieve this.

Except for the first months of our analysis, in which PP4 and the PSI-20 had very close values (with PSI-20 sometimes outperforming PP4), from June 2011 onward this PP systematically performed better than the PSI-20, and thus presented a positive relative performance chart.

For its part, PP5 was characterized by its many peaks. The market was still in decline when PP5 started to grow, reaching its first and highest peak in October 2011. From then on, PP5 had several ups and downs, but managed to always remain above the PSI-20 until February 2013, when, due to a fall in its value, PP5 dropped below the index. PP5 always managed to have values close to the index during the timespan of our study. From the second semester of 2014 onwards, PP5 always presented significantly higher values than the PSI-20, this difference, however, was not so significant when PP5 was below the PSI-20.

The last two PPs showed similar behaviours. Both fell below PSI-20 in the first months, although this value was not significative. They started to grow in the third quarter of 2011 (PP7 is the principal portfolio that had the highest growth in this period, reaching its peak in August 2012), remaining above PSI-20 by the end of 2013. Between January and July 2014, both PPs were lower than the PSI-20, but recovered immediately afterwards and remained above PSI-20 until the end of our analysis.

5.3. Allocation strategies comparison

In order to study whether the application of PCA to our data is an efficient allocation strategy, we had to compare it to other strategies by also applying them to our data set. Thus, in addition to constructing our portfolios using PCA, we also studied the behaviour of the PSI-20 index as a portfolio (as seen in 5.2). We created a portfolio with equal investment in our 26 stocks, this portfolio is commonly referred to as "naive" but during our study we will refer to it as "1/N portfolio", we also consider a portfolio with equal investment for all seven principal portfolios, which we will designate "PPEqual"; finally, we created a portfolio for Markowitz's mean-variance strategy, called "MV portfolio". Several indicators were used for the construction and selection of the MV portfolio, such as mean returns, standard deviation, and Sharpe Ratio. These three indicators, as well as Value at Risk and Expected Shortfall will be explained next since they allowed us to choose the most appropriate strategy for our data set.

For our study, we considered the historical Value at Risk (VaR) at a significance level of 95%. While standard deviation is useful for making comparisons between strategies, it does not take into account the direction of investment movements. Whenever there is a sudden gain, an increase in volatility can be observed, thus suggesting an increase in investment risk. The problem is that for investors, what counts as increased risk is the possibility of losing money, not earning it. VaR avoids the shortcomings of standard deviation by answering the question any risk adverse investor has when studying investment options: "how much can I lose with $\propto$ % probability over a given time horizon?" (RiskMetrics Group,1994).

VaR is defined as the maximum value that meets the following condition:

$$P(G < -VaR) = \propto$$

(Equation 5.3)

Where G is the gain and $\propto$ is the confidence level.

Historical VaR assumes that history repeats itself from a risk point of view, so in order to calculate the VaR, we have to find the percentile that we are seeking; in our case, 5%.

According to Nicolau (2011), one of the advantages of using VaR is that it can aggregate different types of risk in a single measure, this is something that traditional risk measures cannot do, and which simplifies its comprehension. As reported by Hull (2015), this advantage is also present when using Expected Shortfall (ES) —also known as Conditional Value at Risk (CVaR), Average Value at Risk (AVaR), or Expected Tail Loss (ETL). While VaR can answer the question “how bad can things get?”, ES can answer “If things do get bad, what is the expected loss?”. To calculate ES, we also need to define the time horizon and the confidence level and in our case, these two variables remain the same as those used for the VaR calculation. Thus, we can say that ES is the expected loss when the loss is greater than the VaR level.

We also calculated the Sharpe ratio associated with each portfolio. Sharpe ratio evaluates the performance of the portfolio through a risk-versus-return analysis, this is useful because, a greater exposure to risk is not always synonymous with higher returns. With this ratio we obtained a common (“dimensionless”) unit of measure for the various portfolios, this allowed us to compare them without taking into account the individual characteristics of each portfolio. Using this type of valuation the investor can know if the high return of a portfolio comes as a result of a greater exposure to risk or if it is simply more efficient than the rest. Unlike VaR and ES, in which smaller values of this measure indicated a lower risk, in the Sharpe ratio a higher value implies a higher return per risk unit.

The formula used to calculate the Sharpe ratio is the following:

$$Sharpe\ Ratio = \frac{\mu_p - r_f}{\sigma_p}$$

(Equation 5.4)

Where μ_p is the expected portfolio return, r_f is the risk free rate and σ_p is the portfolio standard deviation. For our study, we considered a risk free rate of zero.

As explained above, the mean returns, standard deviation and Sharpe ratio were used first to select the most efficient MV portfolio.

The MV portfolio minimizes the risk level, provided that a certain expected return value is given. Knowing that the expected return can be defined by Equation 5.5,

$$E(R_p) = \sum_i w_i E(R_i)$$

(Equation 5.5)

where R_p is the portfolio return, R_i is the return on asset i and w_i is the weight of asset i; And knowing that the portfolio return variance can be defined by Equation 5.6:

$$\sigma_p^2 = \sum_i w_i^2 \sigma_i^2 + \sum_i \sum_{j \neq i} w_i w_j \sigma_i \sigma_j \rho_{ij}$$

(Equation 5.6)

where σ_i is the sample standard deviation of asset i and ρ_{ij} is the correlation coefficient between the returns on assets i and j. To find the weights of each asset for each MV portfolio, we used the Microsoft Excel "Solver" function targeting the expected return of each MV portfolio to the expected return of a given PP, with the constraint that the sum of the weights of each MV portfolio must be one being, thus, possible to find efficient portfolios for each mean return level of each PP.

When we constructed the MV portfolios, they were initially seven. However, to simplify our analysis, and since the remaining alternative allocation strategies had only one portfolio, we selected the most efficient one as representative of the entire MV strategy.

In Table VII of Appendix E we present the performance statistics for all the MV portfolios.

A brief analysis of Table VII suggested that MV portfolio 2 should be excluded because it presented a negative mean of the returns, losing its attractiveness. Furthermore, it also presented the highest standard deviation among all portfolios, thus being the portfolio with higher associated risk.

MV portfolio 3 also showed negative mean returns and its attractiveness was also affected. Once MV portfolio 2 was excluded, MV portfolio 3 became the portfolio with the highest standard deviation. The Sharpe ratio associated with these two indicators presented a negative value.

MV portfolio 1 also presented a negative Sharpe ratio and was thus also excluded.

Of the remaining four portfolios, the one with the highest mean returns, the lowest standard deviation and, consequently, the best Sharpe ratio was MV portfolio 4, which thus became the representative ("MV Portfolio") portfolio of our entire MV strategy.

Having seven principal portfolios and a portfolio for each of the other allocation strategies, we are able to compare them by studying the performance trajectory of each PP, PSI-20 index, 1/N portfolio, PPEqual and MV portfolio, through Figure 5.2 and Figure 5.3. We also compared the strategies using the relative performances of each PP with each of the remaining strategies – this was done by using Figure 5.4 (relative performances of PPs and the 1/N Portfolio), Figure 5.5 (relative performances of PPs and the PPEqual), Figure 5.6 (relative performances of PPs and MV Portfolio). In order to simplify the comparisons, and following the same procedure used in 5.2, we constructed two Tables - Table VIII, where we present the price correlations of each PP with the 1/N portfolio, PPEqual and MV portfolio; and Table IX, which indicates the daily return correlations of each PP with the 1/N portfolio, PPEqual and MV portfolio. Figures 5.2 to 5.6 and Tables VIII and IX can be found in Appendix F.

In Figure 5.2 we can see from the "Values of Principal Portfolios and **1/N Portfolio**" that the PP1 was closely connected to 1/N portfolio and could be even confused with 1/N portfolio. This correlation between the PP1 and the 1/N portfolio can be confirmed through the Table VIII and Table IX.

Table VIII shows the PP1 assumed high values of price correlation with the 1/N portfolio, presenting a value of 0,985, a value very close to the unit. We can also see, in Table IX, that the daily return correlations of PP1 to the 1/N portfolio were similarly high and this value was roughly equal to that observed in Table VIII (0,985).

Still, as we can see from Figure 5.4, the relative performance of PP1 and 1/N portfolio was definitely negative, increasing the discrepancy between them over the years, in a generalized way. However, at its peak, the discrepancy assumes a value of -14,5%, which makes this difference only slightly relevant.

Figure 5.2 shows that PP2 follows the path of the 1/N portfolio, however, from the third quarter of 2012, the distance between PP2 and 1/N increased and, henceforth, PP2 could no longer reach the values achieved by the 1/N portfolio. This fact can be verified by looking at Figure 5.4, which shows PP2 mostly negative, and, from the mentioned period onwards, it assumes even higher negative values, verifying, at the worst time of our analysis, a value of relative performance to the 1/N portfolio of about -65%.

Table VIII and Table IX show that only PP1 and PP2 had positive values for the correlations under analysis and, although the values assumed by PP1 are much higher than those assumed by the remaining PPs, PP2 presented a price correlation to 1/N portfolio of 0,497 and daily return correlations of 0,138. This indicates that PP2 tracked the 1/N portfolio in the long term, but its daily movements were not correlated with the 1/N portfolio.

Figure 5.3 shows that this portfolio was affected by the financial crisis, exhibiting a decrease in value until the end of the third quarter of 2012. The same happened with the PP1 and PP2, which were the ones more affected by the crisis. Also, 1/N portfolio was closely related to the PSI-20, presenting very similar trajectories, although the 1/N portfolio presents more discrepancies and, therefore, a higher level of volatility.

Compared to the 1/N portfolio, PP3 was the first to present a relative performance chart with mostly positive values. We already noted that this PP presented negative values for the correlations (-0,260 for the price correlations and -0,061 for the daily return correlations), which is evident in the analysis of Figure 5.2. In PP3, the values that exceeded the ones observed in the 1/N portfolio were significantly high (reaching a relative performance value 175% higher than the 1/N portfolio); however, when PP3 was lower than 1/N portfolio, the differences in value were almost insignificant.

Figure 5.4 shows that no PP permanently outperformed 1/N over the full period under analysis, yet PP4 was close to achieving it.

If we would only have measured the mean returns of each portfolio without associating them with Standard Deviation, Sharpe ratio, VaR, or ES, PP4 would be without a doubt the most attractive portfolio for any investor.

In Figure 5.4, PP5 and PP6 present a mostly positive relative performance, compared to the 1/N portfolio, however, they appear several times with values below the 1/N portfolio.

PP7 presented a more negative value correlation with 1/N portfolio. Figure 5.4 shows that, although PP7 sometimes exhibited a trajectory contrary to the one observed in 1/N portfolio, its relative performance presented mostly positive values, which, in June and August 2012 sometimes surpassed 200%.

Turning our attention to **PPEqual portfolio**, in Figure 5.3 we can see this was the most stable portfolio, with values always in the interval between 970 and 1320 euros.

Regarding correlations, Table VIII and Table IX show mostly positive correlations, which should be expected, since this portfolio was built based on the seven PPs. Nevertheless, Table VIII shows PP2 assumed a negative price correlation and PP4 assumed the highest value observed. PP4 was also the most attractive portfolio when comparing its performance with PPEqual portfolio, as we can see in Figure 5.5.

Figure 5.5 shows that PP2 had a relatively unfavourable relative performance, followed by PP1, which despite achieving positive values in the first semester of 2014 and positive (insignificant) values in the first months of our analysis, presented (like PP2) an unfavourable relative performance.

Unlike PP3, which reached a relative performance of 80%, and PP7, which surpassed 55%, PP5 and PP6 were quite stable with respect to their relative performances; with PP5 presenting values between -37% and 45%, and PP6 placing its values between -25% and 35%. PP6 presented the highest daily return correlations (0,476) having a trajectory and value similar to PPEqual portfolio, during most of the period under analysis.

The **MV portfolio** was, without a doubt, the one that grew the most over the years and was able to reach higher values, from the second quarter of 2015. Nevertheless, and although there were some breaks over time, the MV portfolio was the one that assumed the highest values from the third quarter of 2013 and thus remained until the end of our analysis. The reason this portfolio had higher values than the others can be attributed to the fact that it was the most efficient one for this strategy. As Figure 5.6 shows, contrary to what happened in the other allocation strategies, when the relative performance plots of the PPs and the MV portfolio were studied, no PP was close to outperform the MV portfolio. On the contrary, PP1 and PP2 showed more unfavourable relative performance plots than those that have been verified so far.

Table VIII and Table IX show that PP1 had a considerably high price correlation with the MV portfolio (0,434) and an even greater daily return correlation (0,71). These correlations can be verified in Figure 5.2 where PP1 closely followed MV portfolio throughout the study period. However, the MV portfolio appears to be an extension of PP1 as they both had a very similar trajectory.

PP2 and PP3 presented negative price correlations (-0,355 and -0,461, respectively) and positive daily return correlations (0,046 and 0,018, respectively). Although they had similar trajectories, PP2 presented low values when compared to the MV portfolio resulting in an unfavourable relative performance. For its part, PP3 showed higher values than PP2, and managed to stay above the MV portfolio between the third quarter of 2011 and the first quarter of 2014, nevertheless, this portfolio now presents a poorer relative performance than those it had presented so far.

Throughout the study period, PP4 presented favourable values of relative performance when compared with the other strategies, yet, when compared to the MV portfolio PP4 was no exception and did not present such satisfactory results. Nonetheless, this portfolio presented favourable figures since the second quarter of 2011, and managed to maintain them until the first quarter of 2014, when it again lost its competitiveness — which only recovered in the first quarter of 2015. After this recovery, some ups and downs were observed, with PP4 maintaining, in the most of the time, values above those presented by MV portfolio.

PP5 presented an unfavourable relative performance in general, although it presented some stability when compared to the MV portfolio, since its relative performance ranged between -45% and 65% (approximate values).

PP6 and PP7 had similar behaviours; both showed negative price correlation values when measured against the MV portfolio (-0,068 and -0,124, respectively). Regarding daily return correlations, PP6 had a value of approximately 0 and PP7 had a positive value of 0,006. Figure 5.2, shows these two portfolios presented similar trajectories, although most of the time PP7 presented higher values than PP6. Having analysed their relative performances, we can see that both behaved favourably until the end of the third quarter of 2013; from then on, their relative performances remained unfavourable until the end of the period under analysis. However, when PP7 presented positive values, these were better than those presented by PP6, reaching 100% for PP7 and only 40% for PP6.

Appendix F, presents the behaviour of the PPs (and therefore, how single risk sources behave), as well as the behaviour of all the remaining allocation strategies studied. Appendix G presents the statistical performance of all PPs together with each of the other strategies. This involved computing the mean returns, standard deviation, Sharpe Ratio, $VaR_{95\%}$ and $ES_{95\%}$ for all seven PPs, as well as for the PSI-20, 1/N portfolio, PPEqual portfolio, and MV portfolio; these indicators can be found in Table X.

Looking at Table X it is possible to see that PP1, PP2 and PP3 presented negative mean returns values, with consequent negative Sharpe ratio, thus becoming less attractive to any investor.

Of all PPs, PP4 presented the higher mean returns and Sharpe ratio, as would be expected after comparing its behaviour against the other allocation strategies. Thus, in case a potential investor did not have access to all risk indicators, but only to the plots previously presented, he would probably choose this portfolio as the most appealing. However this becomes problematic when considering a risk-return perspective since this PP also presents the highest values for all risk indicators, from standard deviation to VaR and ES, being therefore the riskiest portfolio.

Of the remaining PPs, PP6 presents lower values for VaR and ES, being a safer bet than PP5, which presents smaller returns and lower Sharpe ratio and, additionally, still presents higher standard deviation, VaR and ES. Comparing PP6 with PP7 is, nonetheless, a rather complicated task since the comparisson depends on the degree of risk aversion of the investor. On the one hand, PP7 offers a better mean returns, a more favourable standard deviation and a better Sharpe ratio but, on the other hand, it also presents higher values for VaR and ES.

Moving on to the analysis of the alternative portfolios, the 1/N portfolio performed better than the PSI-20, since all its indicators were more favourable than the index. Although it is a “naive” allocation strategy (and hence the one probably chosen by investors without financial knowledge), through the analysis made so far, we can say that it can compete well with the other allocation strategies.

PPEqual portfolio is the best of all the alternative strategies studied so far. Its VaR and ES were, by far, the most favourable values checked for all portfolios, its standard deviation was also lower compared to the others. Nonetheless, PPEqual shows a mean return of only 0,008, which can be interpreted as insignificant for investors who are willing to risk more (less risk-averse investors).

Finally, the MV portfolio is probably the best portfolio in terms of risk-return, since it showed the second highest return, and it appeared even more competitive after measuring its Sharpe ratio, which presents the best results. Consequently, this portfolio presents the highest level of return per risk unit.

In terms of VaR and ES, the MV portfolio shows favourable figures, although not as favourable as those presented by PPEqual (portfolio with low mean returns) and PP2 (portfolio with negative returns and, therefore, unattractive).

6 Conclusions and Further Research

As discussed in Chapter 4, between July 2008 and December 2016, the Portuguese stock market presented seven main risk sources inherent within the stocks. With seven principal components, we were able to retain more than 50% of data variance in all rolling windows under study. In this work, we constructed seven Principal Portfolios, after analysing their performance we reached the following conclusions:

Although PP1 was the portfolio that best represented the market, it was PP4 that performed better when compared to alternative strategies and, although its performance was not considerably better than the other strategies, it almost achieved this goal.

PP4 was the portfolio with the highest mean returns, being the most attractive when only taking into account this characteristic, unfortunately, this was also the portfolio with the highest levels of risk according to all risk indicators (standard deviation, VaR e ES). Given this, the PP4 would only be suitable for an investor with very little aversion to risk, which is highly unusual.

Of the portfolios constructed using alternative strategies, the two most attractive were without a doubt PPEqual and MV, but the choice between the two would ultimately depend on the investor's degree of risk aversion. The reason being that while PPEqual showed more favourable values for all risk indicators; MV offered much more attractive mean returns than the previous portfolios and presented the best Sharpe ratio among all the studied portfolios.

To choose the most attractive portfolio, we should start by classifying them according to their risk levels. However, a non-risk averse investor would first look at the values of the mean returns. Given this criteria, two portfolios stand out for their values of mean returns: PP4 and MV portfolio. After selecting them, investors would notice that their differences in mean returns were not significant, and would then proceed to analyse their remaining characteristics. It would not take long for them to realise that the MV Portfolio is far less riskier than PP4. Given these two options, conscious investors would almost certainly choose the MV Portfolio.

PP1, PP2, and PP3 were excluded due to their negative mean returns values, however, despite its good performance, we also decided to exclude PP4 due to its high-risk levels.

PP5 and the "naive" 1/N portfolio, were also excluded because they were not competitive enough (they did not present mean returns that could justify their associated risks). Having excluded the former, the remaining portfolios were PP6, PP7, PPEqual portfolio and MV portfolio.

PP6 and PP7 exhibited the highest levels of risk and, when compared to the MV portfolio, presented lower mean returns, therefore, we also excluded them. The remaining two portfolios represented good options for different investment strategies: PPEqual portfolio would be suitable for a risk-averse investor; whereas MV portfolio would be more appealing to an investor willing to assume more risks in exchange of higher returns. In any case the most relevant outcome of our research was being able to obtain a portfolio (PPEqual) using the PCA approach, and which turned out to be the most risk efficient portfolio to emerge from our data set.

As further research we consider that the analysis and interpretation of the principal components and possible relation with economic variables (e.g., GDP, inflation, ...) could be interesting and give a better insight for investment strategies.

References

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate Behavioural Research* 1, 245-276.

Chopra, V. K. and William, T. Z. (1993). The effect of errors in mean, variances, and covariances on optimal portfolio choice. *Journal of Portfolio Management* 19 (2), 6-11.

Fenn, D. J., Porter, M. A., Williams, S., MacDonald, M., Johnson, N. F., and Jones, N. S. (2011). Temporal evolution of financial-market correlations. *Physical Review E* 84, 026109.

Hull, J. C. (2015). *Risk management and financial institutions*, 4th Edition, New Jersey: John Wiley & Sons.

Jolliffe, I.T. (1986). *Principal Component Analysis*, 2nd Edition, New York: Springer.

Kaiser, H. F. (1960). The application of electronic computers to factor analysis. *Educational and Psychological Measurement* 20, 141-151.

Kaiser, H. F. (1974). An index of factor simplicity. *Psychometrika* 39, 31-36.

Lowenfeld, H. (1909). *Investment, an exact science.* Financial Review of Reviews.

Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business* 36 (4), 394-419.

Markowitz, H. (1952). Portfolio Selection. *Journal of Finance* 7, 77-91.

Nicolau, J. (2011). *Modelação de Séries Temporais Financeiras*. ISEG/UTL & CEMAPRE.

Novo Banco (2016). Institutional Presentation – A Leading Franchise in Portugal. [online]https://www.novobanco.pt/site/cms.aspx?srv=207&stp=1&id=733826&fext=.pdf

Partovi, M. H. and Caputo, M. (2004). Principal Portfolios: recasting the efficient frontier. *Economics Bullentin* 7 (3), 1-10.

RiskMetrics Group (1994). Riskmetrics technical document. Technical report, JPMorgan.

Sharma, S. (1996). *Applied Multivariate Techniques*. John Wiley & Sons, Inc.

Yang, L. (2015). *An application of principal component analysis to stock portfolio management.* Department of Economics and Finance University of Canterbury.

Appendix A

Characteristics of the Returns

Table I Mean Returns and Standard Deviations for the 26 Stocks with complete data

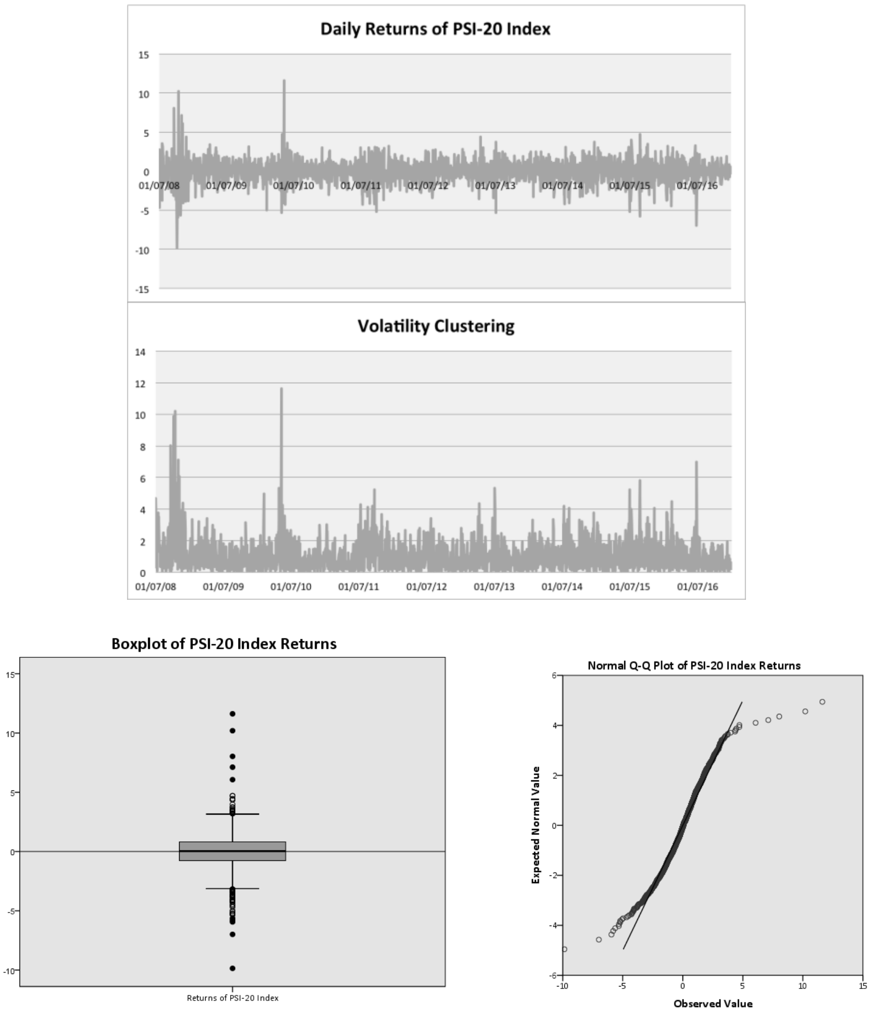
	Stocks	Mean Returns	Standard Deviation
Altri	ALTR	0,0685118	2,0188757
Semapa Sociedade de Investimento	SEM	0,0629746	1,7318959
Jerónimo Martins	JMT	0,0651285	1,8636807
Sonae	SON	0,0383836	1,9199546
Sonaecom	SNC	0,0685465	2,1182610
Sonae Indústria	SONI	-0,2477068	4,0019495
NOS	NOS	0,0612122	1,8398323
EDP - Energias de Portugal	EDP	0,0399535	1,5553932
BCP - Banco Comercial Português	BCP	-0,1016765	3,5056381
Galp Energia	GALP	0,0334921	1,8694730
Mota-Engil	EGL	0,0316884	2,6482484
The Navigator Company	NVG	0,0678303	1,5898118
Corticeira Amorim	COR	0,1840437	1,9149854
BPI - Banco Português de Investimento	BPI	0,0255176	2,7259956
Pharol	PHR	-0,1346770	3,2019309
Impresa	IPR	-0,0768497	3,1870404
Cofina	CFN	-0,0147736	3,0441490
Novabase	NBA	0,0146425	1,6241637
Cimpor - Cimentos de Portugal	CPR	-0,1360492	2,7379056
Inapa - Investimentos, Participações e Gestão	INA	-0,0259927	3,5957362
SDC Investimentos	SDCAE	-0,0728214	5,5250790
Teixeira Duarte - Engenharia e Construções	TDSA	-0,0397433	3,3757422
REN - Rede Elétrica Nacional	RENE	0,0254540	1,1225317
Sonae Capital	SONC	0,0855419	3,3361727
EDP Renováveis	EDPR	0,0298748	1,7176647
Ibersol	IBS	0,0527335	1,6903462

Source: Bloomberg

Appendix B

Characteristics of PSI-20 Return Data

Figure 3.1 Stylized Facts for PSI-20 Index



Source: Bloomberg

Figure 3.2 Kurtosis and Skewness of Returns of PSI-20 Index

Returns of PSI-20 Index

PSI 20 Index

N	Valid	2180
	Missing	0
Skewness		.030
Std. Error of Skewness		.052
Kurtosis		5.644
Std. Error of Kurtosis		.105

Source: Bloomberg

Appendix C

Reviews Related to the Rolling Windows

Table II *Periods associated with each rolling window*

R	Period	R	Period
R1	01.07.2008 to 30.06.2010	**R15**	02.01.2012 to 31.12.2013
R2	01.10.2008 to 30.09.2010	**R16**	02.04.2012 to 31.03.2014
R3	02.01.2009 to 30.12.2010	**R17**	02.07.2012 to 30.06.2014
R4	01.04.2009 to 31.03.2011	**R18**	01.10.2012 to 30.09.2014
R5	01.07.2009 to 30.06.2011	**R19**	02.01.2013 to 31.12.2014
R6	01.10.2009 to 30.09.2011	**R20**	02.04.2013 to 31.03.2015
R7	04.01.2010 to 30.12.2011	**R21**	01.07.2013 to 30.06.2015
R8	01.04.2010 to 30.03.2012	**R22**	01.10.2013 to 30.09.2015
R9	01.07.2010 to 29.06.2012	**R23**	02.01.2014 to 31.12.2015
R10	01.10.2010 to 28.09.2012	**R24**	01.04.2014 to 31.03.2016
R11	03.01.2011 to 31.12.2012	**R25**	01.07.2014 to 30.06.2016
R12	01.04.2011 to 28.03.2013	**R26**	01.10.2014 to 30.09.2016
R13	01.07.2011 to 28.06.2013	**R27**	02.01.2015 to 30.12.2016
R14	03.10.2011 to 30.09.2013		

Table III Summary of KMO values, Explained Variance, Lowest Selected Eigenvalue, Number of Factors Necessary to Explain 50% of the Variance and Number of Factors to have Eigenvalues ≥ 1, for each R.

R	KMO Value	Number of Factors	Explained Variance	Lowest Selected Eigenvalue	Nº of Factors Necessary to Explain 50% of the Variance	Nº of Factors Necessary to Have Eigenvalues ≥ 1
R1	0,966	1	43,592	11,334	3	2
		2	48,578	1,296		
		3	52,413	0,997		
R2	0,966	1	43,936	11,423	3	3
		2	48,861	1,281		
		3	52,874	1,043		
		4	56,696	0,994		
R3	0,962	1	39,941	10,385	4	4
		2	44,477	1,179		
		3	48,748	1,111		
		4	52,638	1,011		
		5	56,476	0,998		
R4	0,965	1	41,855	10,882	3	3
		2	46,420	1,187		
		3	50,670	1,105		
		4	54,446	0,982		
R5	0,965	1	42,048	10,932	3	3
		2	46,608	1,186		
		3	50,824	1,096		
		4	54,475	0,949		
R6	0,965	1	42,473	11,043	3	3
		2	47,069	1,195		
		3	51,161	1,064		
		4	54,958	0,987		
R7	0,959	1	39,591	10,294	4	5
		2	44,367	1,242		
		3	48,546	1,087		
		4	52,674	1,073		
		5	56,628	1,028		
		6	60,129	0,910		

Continued on next page

Table III *Continued from previous page*

R	KMO Value	Number of Factors	Explained Variance	Lowest Selected Eigenvalue	Nº of Factors Necessary to Explain 50% of the Variance	Nº of Factors Necessary to Have Eigenvalues ≥ 1
R8	0,956	1	36,824	9,574	4	4
		2	41,705	1,269		
		3	46,205	1,170		
		4	50,219	1,044		
		5	54,054	0,997		
R9	0,937	1	29,558	7,685	6	5
		2	34,629	1,319		
		3	39,524	1,273		
		4	43,874	1,131		
		5	47,829	1,028		
		6	51,662	0,996		
R10	0,930	1	28,652	7,449	6	5
		2	33,648	1,299		
		3	38,455	1,250		
		4	42,919	1,161		
		5	47,091	1,085		
		6	50,907	0,992		
R11	0,926	1	27,220	7,077	**7**	5
		2	32,265	1,312		
		3	37,166	1,274		
		4	41,639	1,163		
		5	45,928	1,115		
		6	49,729	0,988		
		7	53,519	0,985		
R12	0,924	1	26,934	7,003	6	6
		2	32,685	1,495		
		3	37,639	1,288		
		4	41,984	1,130		
		5	46,169	1,088		
		6	50,059	1,011		
		7	53,794	0,971		

Continued on next page

Table III *Continued from previous page*

R	KMO Value	Number of Factors	Explained Variance	Lowest Selected Eigenvalue	Nº of Factors Necessary to Explain 50% of the Variance	Nº of Factors Necessary to Have Eigenvalues ≥ 1
R13	0,933	1	27,887	7,251	6	5
		2	33,669	1,503		
		3	38,575	1,276		
		4	42,893	1,123		
		5	47,021	1,073		
		6	50,846	0,994		
R14	0,926	1	26,304	6,839	**7**	6
		2	32,127	1,514		
		3	36,970	1,259		
		4	41,328	1,133		
		5	45,529	1,092		
		6	49,594	1,057		
		7	53,332	0,972		
R15	0,920	1	25,518	6,635	**7**	6
		2	31,394	1,528		
		3	36,276	1,269		
		4	40,696	1,149		
		5	44,990	1,116		
		6	48,865	1,008		
		7	52,674	0,990		
R16	0,921	1	25,976	6,754	**7**	6
		2	31,995	1,565		
		3	36,682	1,219		
		4	41,060	1,138		
		5	45,272	1,095		
		6	49,344	1,059		
		7	53,148	0,989		
R17	0,919	1	26,100	6,786	**7**	6
		2	32,100	1,560		
		3	36,563	1,160		
		4	40,878	1,122		
		5	44,943	1,057		
		6	48,982	1,050		
		7	52,808	0,995		

Continued on next page

Table III Continued from previous page

R	KMO Value	Number of Factors	Explained Variance	Lowest Selected Eigenvalue	Nº of Factors Necessary to Explain 50% of the Variance	Nº of Factors Necessary to Have Eigenvalues ≥ 1
R18	0,925	1	27,517	7,155	**7**	6
		2	33,058	1,441		
		3	37,488	1,152		
		4	41,733	1,104		
		5	45,766	1,049		
		6	49,772	1,042		
		7	53,576	0,989		
R19	0,924	1	28,715	7,466	6	**7**
		2	34,080	1,395		
		3	38,727	1,208		
		4	43,057	1,126		
		5	47,263	1,094		
		6	51,292	1,048		
		7	55,214	1,020		
R20	0,915	1	27,773	7,221	6	6
		2	33,078	1,379		
		3	37,619	1,181		
		4	42,000	1,139		
		5	46,123	1,072		
		6	50,103	1,035		
		7	53,836	0,971		
R21	0,920	1	28,195	7,331	6	**7**
		2	33,476	1,373		
		3	38,119	1,207		
		4	42,444	1,125		
		5	46,540	1,065		
		6	50,438	1,013		
		7	54,293	1,002		

Continued on next page

Table III *Continued from previous page*

R	KMO Value	Number of Factors	Explained Variance	Lowest Selected Eigenvalue	Nº of Factors Necessary to Explain 50% of the Variance	Nº of Factors Necessary to Have Eigenvalues ≥ 1
R22	0,927	1	29,614	7,700	6	6
		2	35,055	1,415		
		3	39,662	1,198		
		4	43,907	1,104		
		5	48,068	1,082		
		6	51,937	1,006		
		7	55,674	0,971		
R23	0,937	1	30,595	7,955	6	5
		2	35,953	1,393		
		3	40,469	1,174		
		4	44,762	1,116		
		5	48,841	1,060		
		6	52,591	0,975		
R24	0,936	1	30,781	8,003	6	6
		2	36,132	1,391		
		3	40,626	1,169		
		4	44,943	1,122		
		5	49,010	1,058		
		6	52,961	1,027		
		7	56,752	0,986		
R25	0,938	1	30,815	8,012	6	**7**
		2	36,121	1,380		
		3	40,664	1,181		
		4	45,114	1,157		
		5	49,163	1,053		
		6	53,136	1,033		
		7	56,998	1,004		
R26	0,936	1	30,051	7,813	6	6
		2	35,460	1,406		
		3	40,015	1,184		
		4	44,445	1,152		
		5	48,727	1,113		
		6	52,746	1,045		
		7	56,493	0,974		

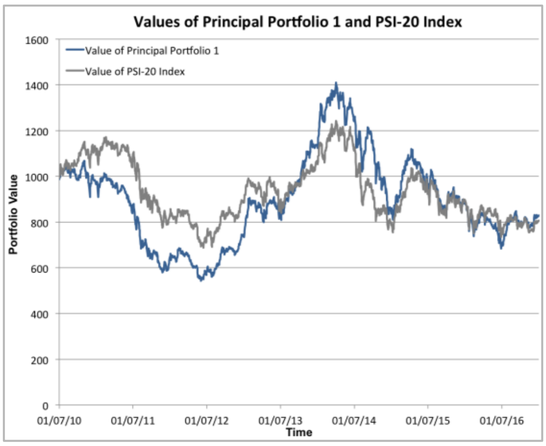
Table IV Summary of the Explained Variances and Lowest Selected Eigenvalues, considering a number of seven PCs, for each R

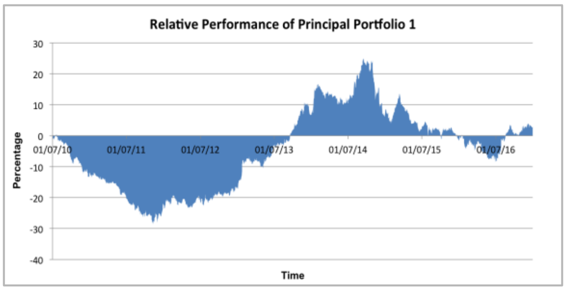
R	Number of Factors Selected	Explained Variance	Lowest Selected Eigenvalue
R1	7	65,79301323	0,750536033
R2	7	66,10696108	0,751301493
R3	7	62,92530723	0,824023745
R4	7	64,3059277	0,786683551
R5	7	64,21056407	0,806643209
R6	7	65,0691484	0,801464199
R7	7	63,39062176	0,847933286
R8	7	61,24068822	0,898916467
R9	7	55,37138576	0,964410532
R10	7	54,63938438	0,97030036
R11	7	53,51865593	0,985264071
R12	7	53,79438549	0,971294305
R13	7	54,58005096	0,970935306
R14	7	53,33214915	0,972004725
R15	7	52,67359879	0,990288909
R16	7	53,1479651	0,988985762
R17	7	52,80847057	0,994925668
R18	7	53,57604157	0,988956919
R19	7	55,2135663	1,019522124
R20	7	53,83644308	0,970660465
R21	7	54,29253446	1,002165308
R22	7	55,67359361	0,971413202
R23	7	56,29618262	0,963233992
R24	7	56,75182424	0,985596797
R25	7	56,99783644	1,004178749
R26	7	56,4926116	0,974167933

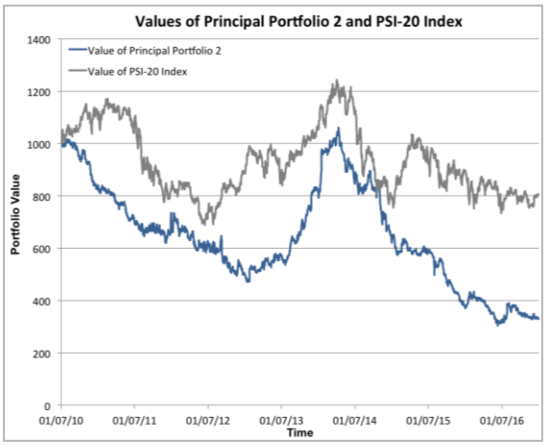
Appendix D

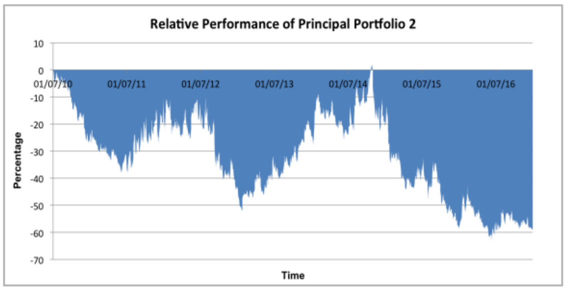
Analysis of Principal Portfolios and PSI-20 Index

Figure 5.1 *Plots of Principal Portfolios 1 to 7 with the PSI-20 index. The data set used is the 26 stocks for the whole period under analysis. Each PP were constructed on a rolling window of two years and rebalanced quarterly. The right panel shows the relative performance of PPs and the index (calculated using Equation 5.2).*



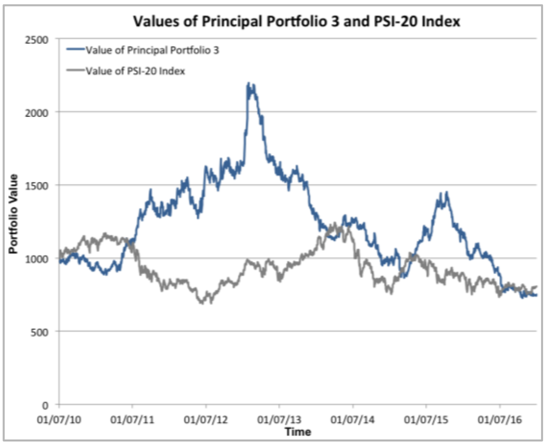


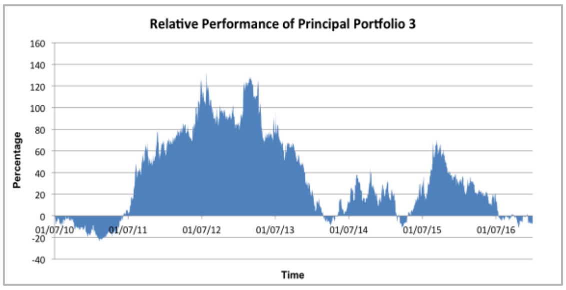


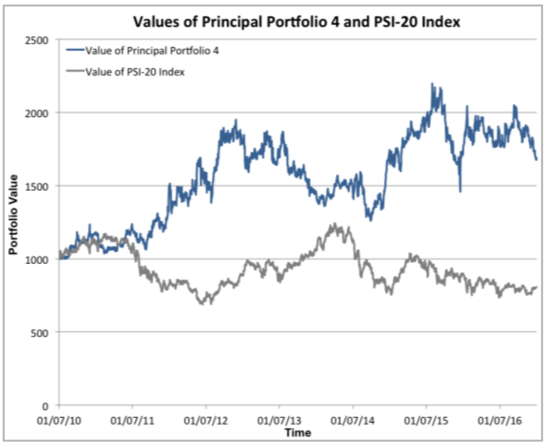


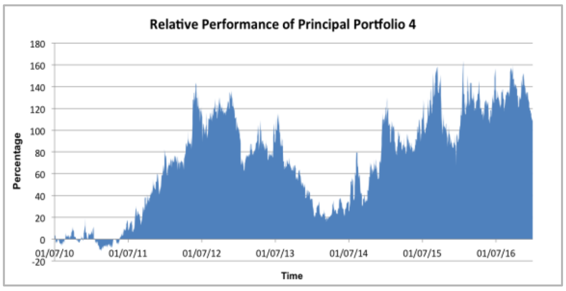
Continued on next page

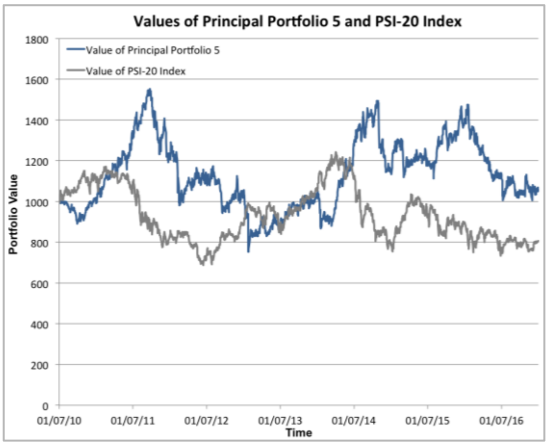
Figure 5.1 Continued from previous page

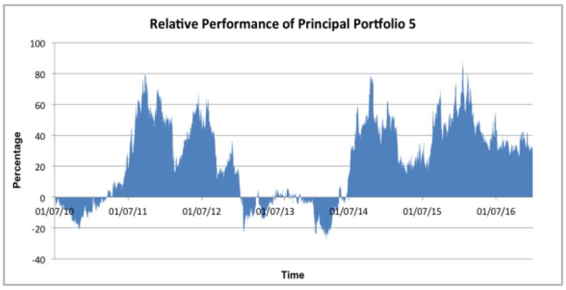






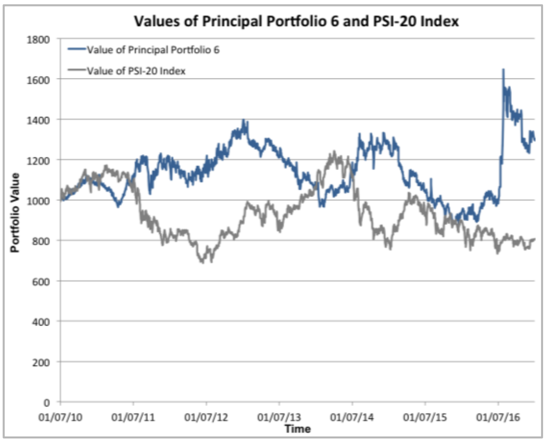


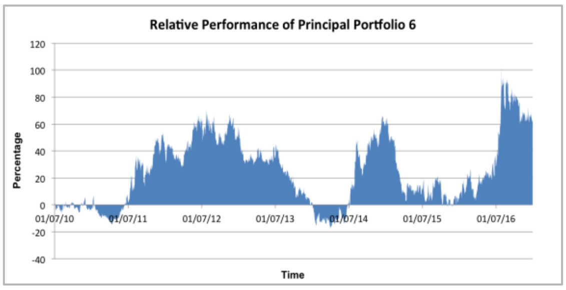


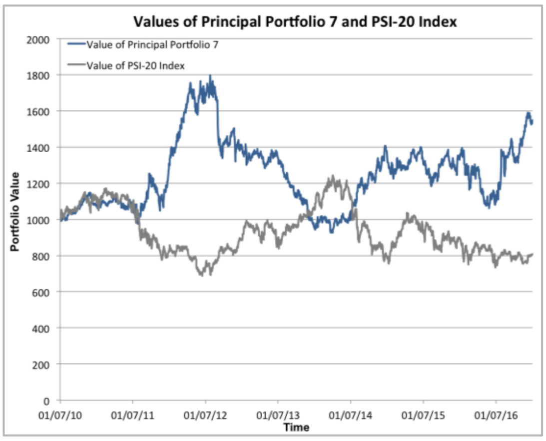


Continued on next page

Figure 5.1 *Continued from previous page*







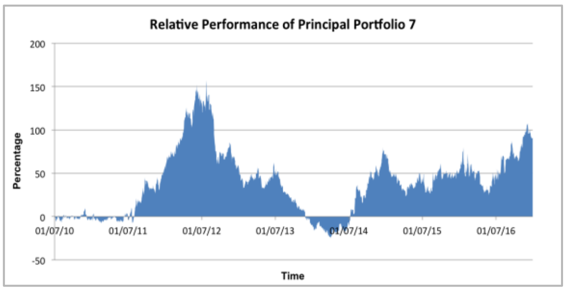


Table V Price correlations of each PP to the PSI-20 index

	PSI-20 Index
Principal Portfolio 1	0,7981248
Principal Portfolio 2	0,7300203
Principal Portfolio 3	-0,1121505
Principal Portfolio 4	-0,5194493
Principal Portfolio 5	-0,2176707
Principal Portfolio 6	-0,3436039
Principal Portfolio 7	-0,7432033

Table VI Daily return correlations of each PP to the PSI-20 index

	PSI-20 Index
Principal Portfolio 1	0,9059309
Principal Portfolio 2	-0,1646068
Principal Portfolio 3	-0,1460975
Principal Portfolio 4	-0,2815474
Principal Portfolio 5	-0,1712194
Principal Portfolio 6	-0,0277906
Principal Portfolio 7	-0,1280223

Appendix E

Reviews Related to the MV Portfolios

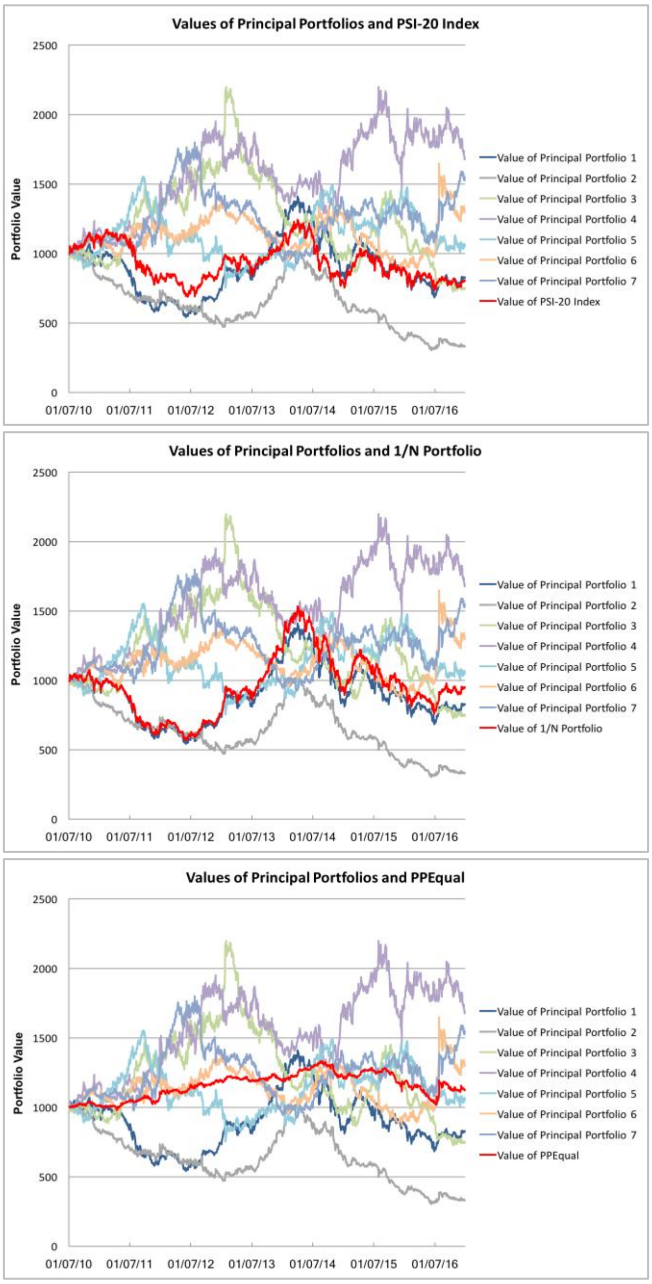
Table VII The performance statistics for the MV portfolios

Statistics	MV Portfolio 1	MV Portfolio 2	MV Portfolio 3	MV Portfolio 4	MV Portfolio 5	MV Portfolio 6	MV Portfolio 7
Mean Returns	-0,003	-0,056	-0,007	0,044	0,017	0,026	0,035
Standard Deviation	0,904	1,209	0,920	0,789	0,835	0,812	0,798
Sharpe Ratio	-0,003	-0,046	-0,008	0,056	0,020	0,032	0,043

Appendix F

Analysis of Principal Portfolios, PSI-20 Index, 1/N Portfolio, PPEqual and MV Portfolio

Figure 5.2 The performance trajectory of Principal Portfolios and PSI-20 Index, 1/N Portfolio, PPEqual and MV Portfolio



Continued on next page

Figure 5.2 Continued from previous page

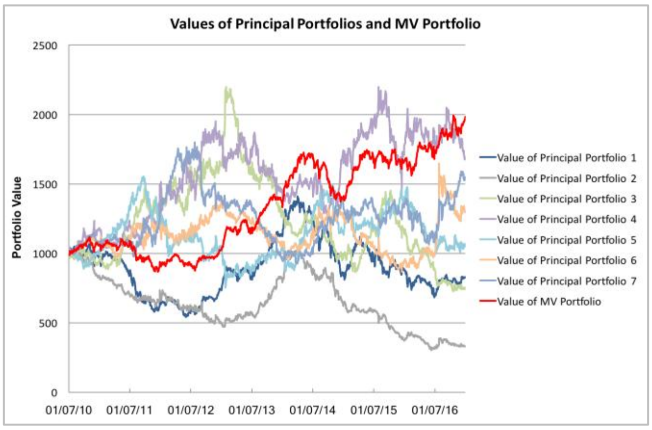


Figure 5.3 The performance trajectory of PSI-20 Index, 1/N Portfolio, PPEqual and MV Portfolio

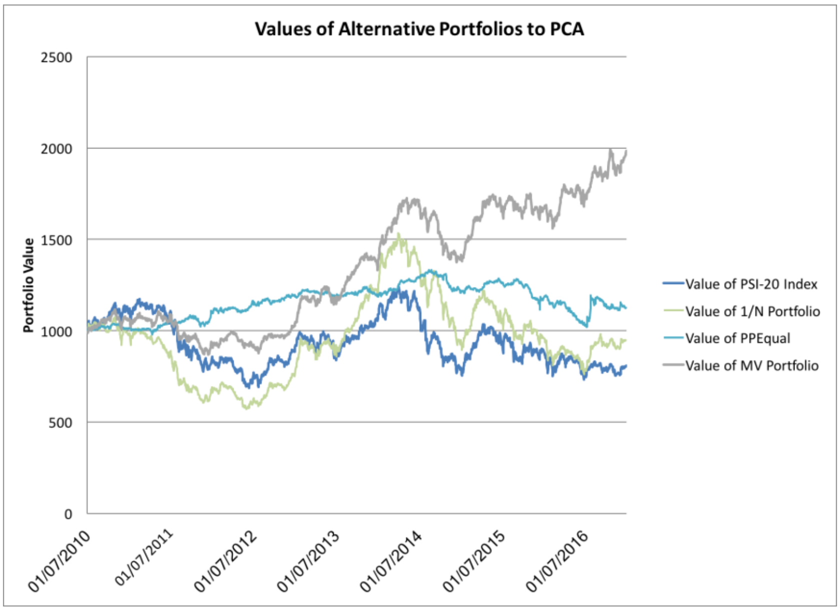


Figure 5.4 Relative performances of Principal Portfolios and the 1/N Portfolio

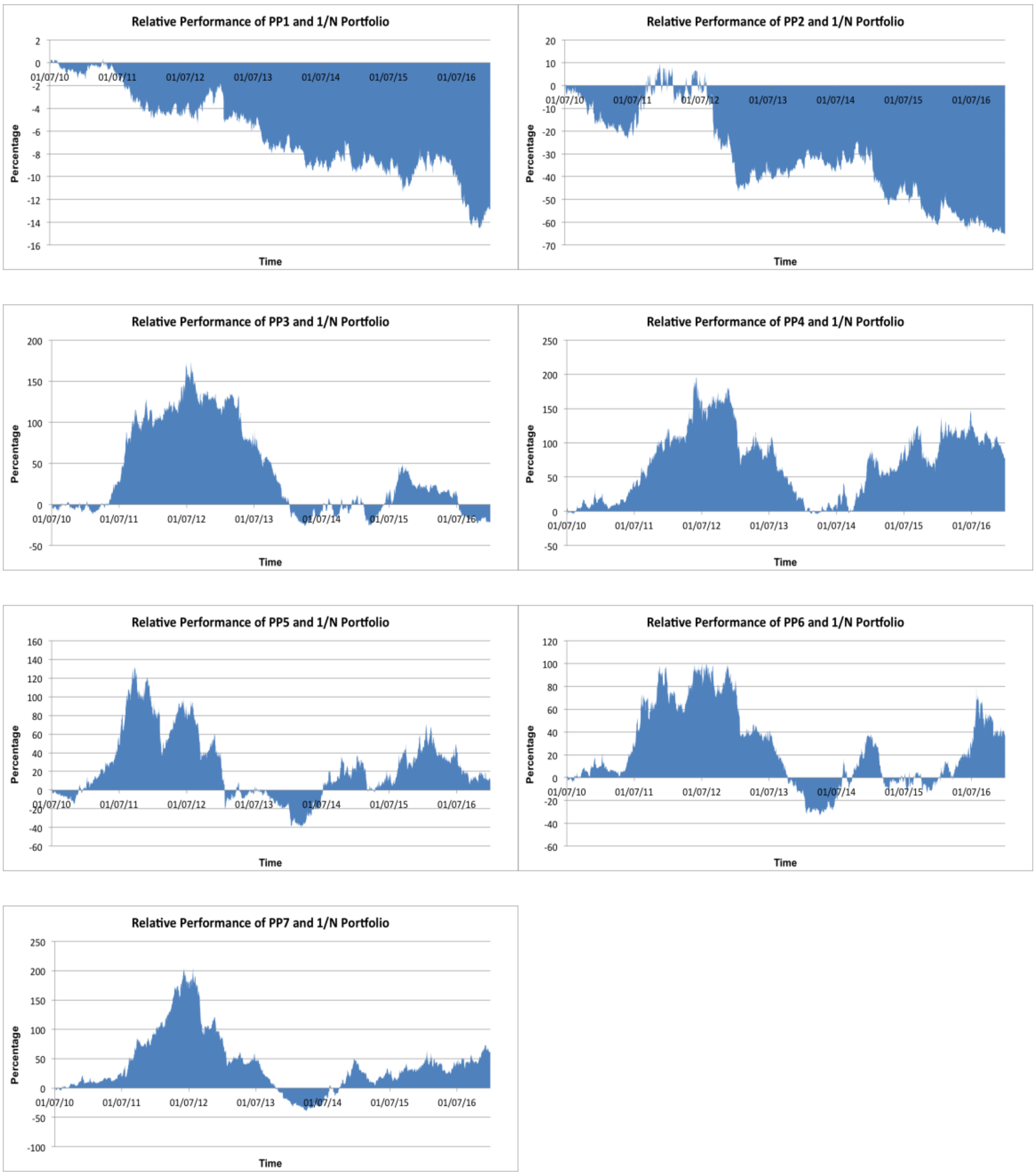


Figure 5.5 Relative performances of Principal Portfolios and the PPEqual Portfolio

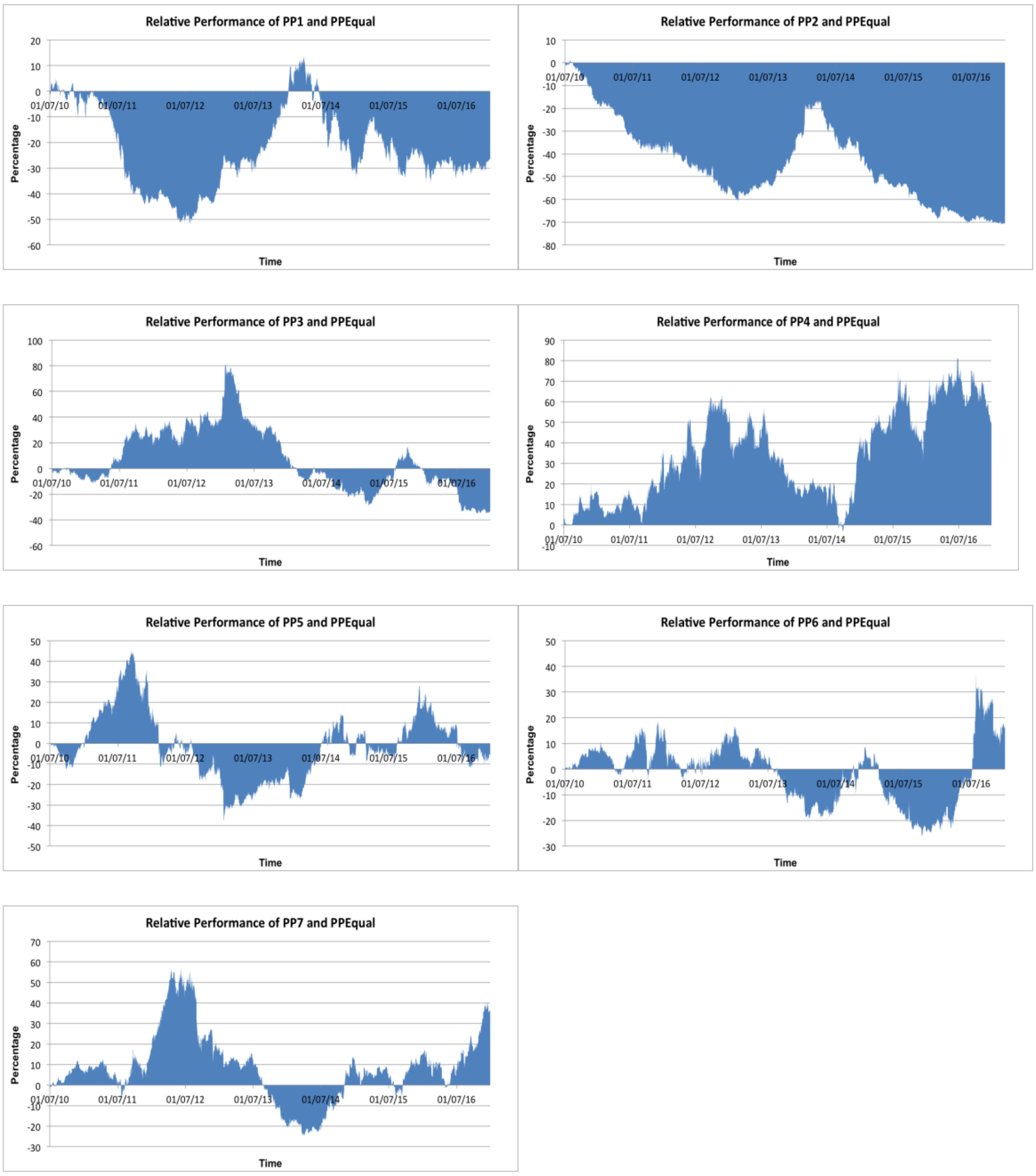


Figure 5.6 Relative performances of Principal Portfolios and the MV Portfolio

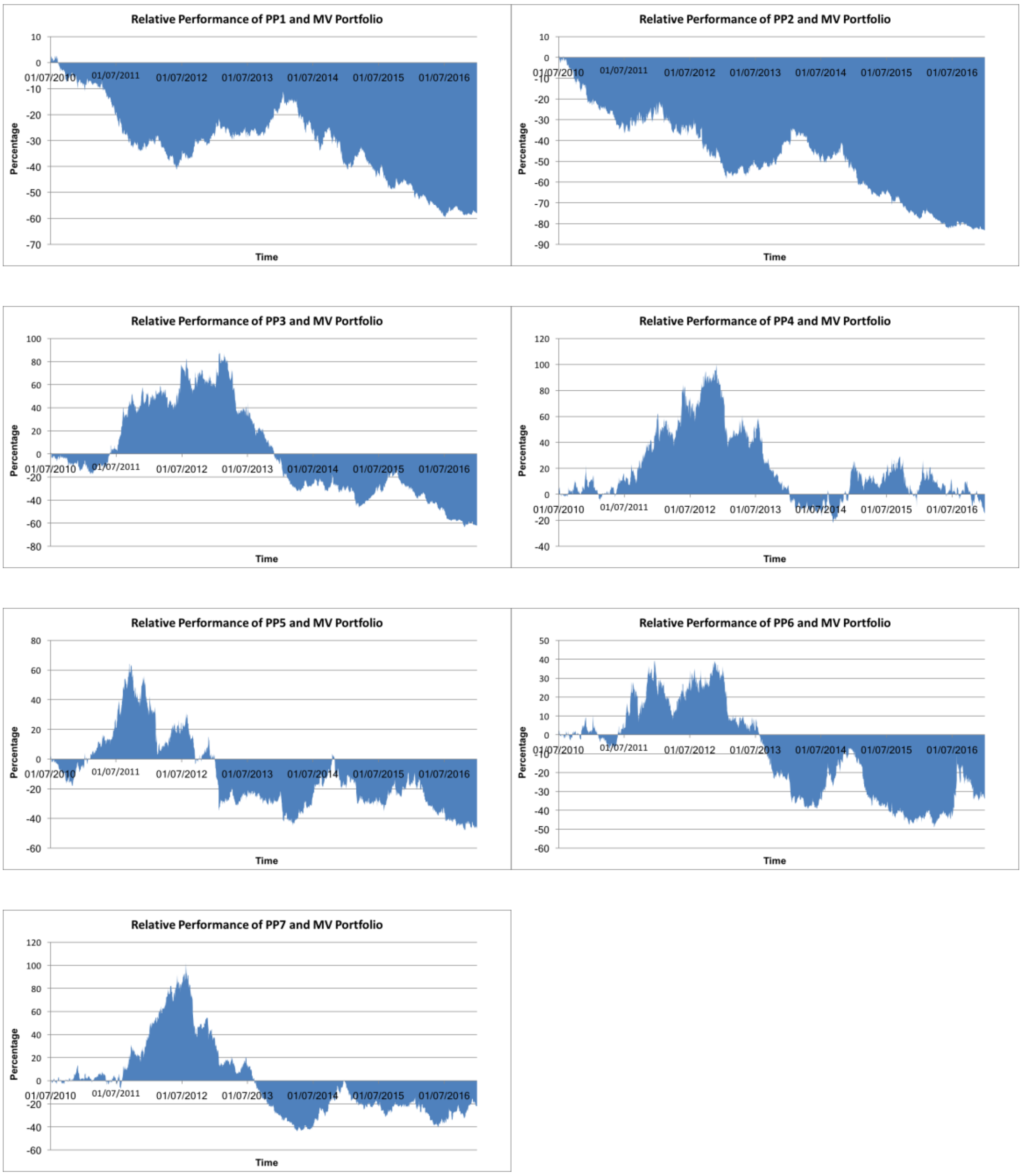


Table VIII Price correlations of each PP to the PSI-20 index, 1/N Portfolio, PPEqual and MV Portfolio

	PSI-20 Index	***Portfolio 1/N***	***PPEqual***	***Portfolio MV***
Principal Portfolio 1	0,79812	0,98498	0,40349	0,43362
Principal Portfolio 2	0,73002	0,49749	-0,07393	-0,35477
Principal Portfolio 3	-0,11215	-0,26049	0,29599	-0,46143
Principal Portfolio 4	-0,51945	-0,03830	0,58210	0,59111
Principal Portfolio 5	-0,21767	-0,08761	0,08778	0,17793
Principal Portfolio 6	-0,34360	-0,21453	0,18744	-0,06778
Principal Portfolio 7	-0,74320	-0,59623	0,16693	-0,12427

Table IX Daily return correlations of each PP to the PSI-20 index, 1/N Portfolio, PPEqual and MV Portfolio

	PSI-20 Index	***Portfolio 1/N***	***PPEqual***	***Portfolio MV***
Principal Portfolio 1	0,90593	0,98480	0,02712	0,70921
Principal Portfolio 2	-0,16461	0,13829	0,40372	0,04555
Principal Portfolio 3	-0,14610	-0,06066	0,36680	0,01805
Principal Portfolio 4	-0,28155	-0,27704	0,34268	-0,10437
Principal Portfolio 5	-0,17122	-0,20706	0,31585	-0,06180
Principal Portfolio 6	-0,02779	-0,02959	0,47608	-0,00182
Principal Portfolio 7	-0,12802	-0,10272	0,43659	0,00605

Appendix G

Statistical performance of each PP and each of the other strategies

Table X The performance statistics of PP1 to PP7, PSI-20 index, 1/N Portfolio, PPEqual and MV Portfolio

Statistics	**PP1**	**PP2**	**PP3**	**PP4**	**PP5**	**PP6**	**PP7**	**PSI-20 Index**	**1/N Portfolio**	**PPEqual**	**MV Portfolio**
Mean Returns	-0,003	-0,056	-0,007	0,045	0,017	0,027	0,035	-0,004	0,004	0,008	0,044
Standard Deviation	1,264	1,463	1,443	1,664	1,648	1,520	1,346	1,299	1,191	0,508	0,789
Sharpe Ratio	-0,003	-0,038	-0,005	0,027	0,010	0,018	0,026	-0,003	0,003	0,016	0,056
VaR (95%)	18,679	14,197	29,304	41,556	30,095	25,813	27,779	20,144	18,457	7,812	16,769
ES (95%)	27,200	19,888	41,345	65,001	44,993	38,345	41,205	28,525	27,055	11,904	25,042