Erasmus University Rotterdam
Erasmus School of Economics
18 June 2018

The Curious Case of Portfolio Selection: Analysis of Correlations and Diversification Over Time

Bora Oskay
414387bo

Supervisor
Dr. P.J.P.M. (Philippe) Versijp
Second assessor
Dr. R. (Rogier) Quaedvlieg

A thesis submitted for the degree of
Master of Science in Financial Economics

The Curious Case of Portfolio Selection:

Analysis of Correlations and Diversification Over Time

This thesis was typeset using $\LaTeX\,2_\varepsilon$

Programming languages used in this thesis are MATLAB and **R**

Abstract

This paper revisits the Modern Portfolio Theory (Markowitz, 1952) by investigating correlations between countries and for various bonds, stocks and real estate indices over time from an U.S. perspective. Analyzing the changes over time can provide information for investors on the question whether they should construct portfolios dynamically by taking possible changes into account. The goal of this thesis is to come to a generalized statement about (1) the degree of stability of correlations over time, (2) the pattern of correlations over time, (3) the movement of the efficient frontier over time and (4) the diversification benefits over time. First, stability of correlations is tested with the Jennrich-Test (Jennrich, 1970). Subsequently, the pattern of the correlations is examined with three methods: rolling-window, EWMA and the DCC(1,1)-GARCH(1,1) model. In addition, the impact over time of these correlations on the efficient frontier is analyzed in two ways. In one way, sub-period efficient frontiers are created in order to investigate the effect of possible unstable correlations in a dynamic setting. In the other way, several variants of the mean-variance spanning tests, as laid out by Kan & Zhou (2008), are computed in order to gain information about the significance of the movements of efficient frontiers and possible diversification gains over time.

Results show that correlations of stocks and real estate seem to be stable over time, but have the tendency to increase during a crisis. However, the strong increase seems to stabilize over time. These findings confirm the fact that correlations tend to be higher during bear markets than in bull markets, implying stabilization on the long-term. In contrast, correlations of bonds are not stable over time and could be the result of country-differences in the dynamics of risk characteristics of bonds and monetary policy. Correlations of bonds did not substantially increase during the crisis. Correlations of bonds have in fact fallen for a short period of time during the recession, although they do show small upward shocks. This implies possible short-term demand shifting from high (sovereign debt) risk to low risk government bonds.

With regard to the effects of correlations on the risk-return trade-off, it has been found that this trade-off resembles the pattern of the correlation coefficients. Only stocks and real estate have an increase in risk over time and a substantial upward shock in risk during a recession. In contrast, bonds appear to be relatively resistant to a possible increase in risk, especially during a recession. Adding a short-selling restriction does not change the main findings, but it did show that a short-selling restriction limits (potential) diversification gains.

Finally, on the basis of various spanning tests, it can be indicated that spanning tests need to be corrected for normality, as the spanning tests under normality can lead to distorted results and conclusions due to the nature of financial data in this study and in general. Results of the asymptotic spanning tests show that (1) an investor benefits from diversification (over time), (2) these benefits have the tendency to decline during recessions and that (3) a relatively large portfolio could make diversification gains quite resistant during recessions.

Keywords: Correlations; Recession; Diversification; Portfolio; Efficient-frontier; Mean-Variance Spanning Tests; Jennrich-Test; Rolling; EWMA; DCC-GARCH; GMM-Wald; SDF;

Contents

List of Figures

List of Tables

Part I

Introduction

Chapter 1

Background

Globalization and technological progress have ensured that investing is no longer as difficult as it was by (1) making markets more accessible, (2) integrating markets and (3) lowering transaction costs. The increasing accessibility to investing has led roughly to more retail investors and thus to larger heterogeneity among investors. Even though the marginal investor differs across time and markets, the marginal investor constantly tries to maximize the portfolio return while minimizing its risk. Lowering portfolio risk is a challenge that could be achieved by numerous ways. One way is by investing in less risky assets. However, the aforementioned method will lead to lower returns, since it is scientifically accepted that risk and return are positively correlated with each other (Markowitz, 1952). It appears that an investor chooses to only hold his risky asset if the return is high enough to offset it (Hillier, Ross, Westerfield, Jaffe, & Jordan, 2013). Therefore, investors will always consider and weigh the trade-off between risk and returns of each security and can ask themselves one (or both) of the following two questions (Markowitz, 1952): "*How much extra x% risk do I accept for an extra y% return?*" or "*how much y% return am I prepared to submit for x% lower risk?*".

Any investor can make the trade-off easier by considering *diversification* as a serious option. Bergstrom (1975), for example, reveals that diversification can reduce portfolio risk by, up to, 40 percent without giving in any returns. Empirical evidence also suggests that there is continuous existence of gains from especially international diversification and that the country's perspective does not seem to matter (Madura & Soenen, 1992). Thus, one can benefit by diversifying the portfolio with international assets (B. H. Solnik, 1995). The benefit comes from the fact that international markets are not fully integrated with each other. This has a positive effect on the diversification possibilities because assets from different countries are not perfectly correlated with each other.

Even though markets are not yet fully integrated, the trend of globalization, which is still continuing, has most likely ensured that correlations between international markets have increased. The increased correlations over time may limit the future diversification possibilities (Elton & Gruber, 1977; Ford, 2001; Hight, 2010). This negative impact on diversification can be severe in times of high volatility (e.g. recession), where markets tend to behave as one (Junior & Franca, 2012). It seems that no investor can overcome possible severe diversification limits in times of high volatility. Therefore, there is a stain on high correlations for risk reduction, especially when short-selling is not costless.

Research that has been done so far seems to confirm the hypothesis that correlation coefficients increase/fluctuate over time and that markets are more integrated than ever before (Koch & Koch, 1991; Erb, Harvey, & Viskanta, 1994; B. Solnik, Boucrelle, & Le Fur, 1996). One author that not only analyzes the stability and pattern of correlation coefficients over time but also its impact on efficient portfolios by means of the efficient frontier, is Eichholtz (1996a). His analysis considers multiple asset classes: *bonds*, *stocks* and *real estate* indices. Eichholtz (1996a) does not only use the Jennrich-test (Jennrich, 1970) to test the stability of unconditional correlation matrices over time, but he also uses the Modern Portfolio Theory (Markowitz, 1952) to compare efficient frontiers across asset classes. Results of Eichholtz (1996a) show two important results. First, international real estate returns have significant lower correlation between each other in comparison with stock and bond returns. Secondly, international diversification should relatively lead to lower risk for a real estate portfolio than for a bond or stock portfolio.

The study by Eichholtz (1996a) is a fairly static analysis in which no account is taken of possible changes of the efficient frontier and diversification benefits over time. It is all the more interesting to extend the line from Eichholtz (1996a) to a more dynamic setting, as rational investors should take these possible changes into account in the construction of an efficient portfolio. Therefore, this paper aims to investigate the current state of affairs regarding correlation coefficients, efficient frontiers and diversification benefits of bonds, stocks and real estate indices with more recent data, between a wide range of countries and in a more detailed manner. This paper adds value by also taking into account the potential changes of the efficient frontier and diversification benefits over time. The ultimate goal is to come to a generalized statement about (1) the degree of stability of correlation coefficients over time, (2) the pattern of correlation coefficients over time, (3) the movement of the efficient frontier over time and (4) the diversification benefits over time. More specifically, the research question of this paper can be stated as follows:

"*Analyzing bonds, stocks and real estate indices, what effect does the pattern and degree of stability of correlation coefficients between countries over time have on the efficient frontier and diversification benefits over time?*"

The focal point of this study is the research within the asset classes, between different countries and over time. In order to show the changes over time more specifically, the focus will not only be placed on the full-sample but also on three sub-periods (pre-crisis, crisis and post-crisis).

Mainly, this paper's methodology consists of four parts. Firstly, the stability of the correlation coefficients for different sub-periods will be investigated by means of the Jennrich-test. Next, the pattern of correlation coefficients will be analyzed on the basis of three methods: Rolling Correlations, an EWMA-model and a DCC-GARCH model. After the correlation analysis has been completed, this paper will examine the possible effects of changing correlation coefficients on the risk-and-return trade-off and diversification benefits over time. On the one hand, efficient frontiers will be made over different sub-periods in order to map the return-risk trade-off over time. This will be done in both the case of unlimited short-selling and restricted short-selling. On the other hand, various mean-variance spanning tests, as described in, among others, Kan & Zhou (2008), will be carried out for different sub-periods in order to show the diversification benefits over time. As each spanning test has its own assumptions, advantages and disadvantages, the application of different mean-variance spanning tests should lead to amplification in the robustness of the conclusion regarding diversification benefits over time.

Results in this paper show that correlation coefficients of the risky assets (stocks and real estate), are increasing, but stable, over time. These correlations have the tendency to specifically increase (more strongly) during a crisis and stabilize afterwards. Analogous to these results are the main findings of Von Furstenberg, Jeon, Mankiw, & Shiller (1989); King & Wadhwani (1990); Longin & Solnik (1995, 2001); Ang & Bekaert (2002); Goetzmann, Li, & Rouwenhorst (2005); Hyde, Bredin, & Nguyen (2007). Bonds, on the other hand, show a deviating pattern. Although the correlations of bonds are unstable, they seem to be resistant to recessions. This is likely to be due to the fact that the collective uncertainty surrounding bonds during recessions is lower in comparison to stocks and real estate, because bonds are considered safe assets. Furthermore, constructing efficient frontiers for different (sub-)periods generates results in line with the pattern and stability of correlation coefficients. The results of the efficient frontier analysis imply that the risk-and-return trade-off worsen during the crisis for both stocks and real estate, but stabilize after the crisis. Bonds, on the other hand, appear to be relatively resistant to changes in the risk-and-return trade-off during the crisis, verifying that bonds are considered as safe assets and are therefore not heavily impacted by negative investor sentiment or collective uncertainty during recessions. Making comparisons between frontiers with and without a short-selling restriction does not change these findings. However, the findings of the efficient frontier analysis support the general consensus that the risk-and-return trade-off worsen when short-selling constraints are added. Not entirely illogical if one accounts for the fact that short-selling has the advantage to make use of high positive correlations between assets. Finally, this paper generates various spanning tests because the significance of the movement of the efficient frontiers need to be tested in order to make conclusions regarding the diversification benefits over time. Surprisingly, based on unlimited short-selling, diversification benefits do not tend to decline over time, regardless of the portfolio size, but do get tight during recessions when the size of the portfolio is small. Most papers develop or use models based on a static analysis of efficient portfolios per certain point in time. However, insight in the background of possible changes over time or during time is mostly not given. In comparison with most of the literature, this thesis wants to provide insight into the changes of efficient portfolios and its background over time. This allows investors, academics and other stakeholders to take into account the fact that efficient portfolios, from an optimal perspective, should be made dynamically. In addition, changes in the type of model or assumptions should also be considered in order to increase robustness of investment and portfolio choices.

Following this introductory section, the literature review will discuss existing literature in chapter 2. Subsequently in chapters 3-5, the methodologies, which function as the backbone of this paper, will be elaborated upon in detail. Chapter 6 will describe the data that will be used, whereas chapters 7-9 will focus on the results based on the acquired data. Finally, this paper will end with chapter 10, which, respectively, includes the conclusion and final remarks of this research paper.

Chapter 2

Literature Review

To date, previous studies have been published that have investigated correlation coefficients, in particular its stability and pattern. That is why this chapter will focus on a number of findings from previous studies. This chapter will consist of three sections. The first section focuses not only on the definition of correlation coefficients but also on the findings of literature over time. Subsequently, a section is devoted to the concept of the efficient frontier in which it is shown to what extent correlation coefficients influence the risk-and-return trade-off. Finally, this chapter will conclude with a brief section on diversification benefits over time.

2.1 Correlation Coefficients

The correlation coefficient is a simple way to gain insight into the strength of the linear relationship between two variables (Sedgwick, 2012). Hall (2015) mentions that one of the most popular methods for measuring this is through the Pearson Correlation Coefficient (short: PCC)[1].

Because the PCC helps quantifying the degree of the association between two variables (in this paper: assets), the PCC is bounded to range between -1 and +1. A negative value implies a negative linear relationship between the two variables, whereas a positive value implies the opposite. If the PCC is equal to $|1|$, then these two variables are perfectly (positive or negative) correlated with each other. The formula for the PCC can be visualized as follows, where, for example, A stands for asset A and B for asset B and where $\bar{x}$ represents the mean of either asset (variable) A or B:

$$\rho_{A,B} = \frac{\sum_{i=1}^{n}(A_i - \bar{x}_A)(B_i - \bar{x}_B)}{\sqrt{\sum_{i=1}^{n}(A_i - \bar{x}_A)^2}\sqrt{\sum_{i=1}^{n}(B_i - \bar{x}_B)^2}} \tag{2.1}$$

The correlation matrix between asset A and asset B could then be defined as follows:

$$\begin{bmatrix} \rho_{A,A} & \rho_{A,B} \\ \rho_{B,A} & \rho_{B,B} \end{bmatrix} \tag{2.2}$$

[1]The method of calculation is subordinate to the interpretation. Whatever calculation is used, the interpretation of the correlation coefficients remains the same.

Logically, the correlations $\rho_{A,A}$ and $\rho_{B,B}$ are equal to 1 and $\rho_{A,B}$ yields the same correlation as $\rho_{B,A}$. Knowing this, the correlation matrix can be generalized for multiple variables (assets):

$$\begin{bmatrix} 1 & \rho_{A,B} & \cdots & \rho_{A,n} \\ \rho_{B,A} & 1 & \cdots & \rho_{B,n} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{n,A} & \cdots & \cdots & 1 \end{bmatrix} \tag{2.3}$$

Early studies on correlation coefficients of (risky) assets reveal that correlation coefficients seem to (1) be unstable and (2) increase over time, especially during high volatility periods. Von Furstenberg, Jeon, Mankiw, & Shiller (1989), for example, analyzed daily stock data and find that the pattern of correlations is more variable over time and that the co-movement between these stocks are more pronounced since the October 1987-crash. Longin & Solnik (1995) extend this analysis with the use of monthly excess returns data by not only modeling the multivariate asset return process but by also considering a longer time period.

More specifically, King & Wadhwani (1990) find that contagion effects are more strongly in times of high volatility and conclude that the increasing correlation coefficients on equity markets after the October 1987-crash are evidence of this. There are studies that imply that there is a relationship between correlation coefficients and the state of the economic market, as correlation coefficients seem to increase (decrease) in bear (bull) markets (Longin & Solnik, 2001; Ang & Bekaert, 2002; Hyde, Bredin, & Nguyen, 2007). Increases in correlation coefficients are however also often devoted to the increase in market integration (Goetzmann, Li, & Rouwenhorst, 2005; Hyde, Bredin, & Nguyen, 2007).

If correlation coefficients increase in bear markets and periods of high volatility, whereas correlation coefficients decrease in periods of bull markets, then this would imply possible instability of correlation coefficients over time. With these findings, it is useful to start with a number of hypotheses that can contribute to the answer of the research question. Since many papers assume that correlation matrices are unstable, this would be a reasonable hypothesis to investigate first[2]:

H_{01}: Cross-country correlation-matrices are significantly unstable over time

[2] Of course, in terms of statistics, the null hypothesis would be that one would test that there is no effect (cross-country correlation-matrices are stable over time).

After assessing the stability, it is important to look at the pattern. As indicated earlier, and according to a large number of papers, correlation coefficients often increase over time due to globalization and market integration, especially in times of high volatility (e.g., recessions)[3]:

H_{02}: *Cross-country correlation coefficients increase substantially over time*

2.2 The Efficient Frontier

The previous section revealed that correlations may be unstable and increasing over time. The impact that correlations have on the risk-and-return trade-off (or diversification) may well be underestimated. Not only does empirical evidence suggest that correlations are unstable over time (Kaplanis, 1988), it also stresses the importance of stable correlations in order to achieve diversification gains (Jorion, 1985; Eun & Resnick, 1988). In addition, Markowitz (1952) highlights the fact that substantial risk reduction can be achieved by adding a pool of assets that are low correlated with each other. Markowitz (1952) is certainly not the only one who documents a negative relationship between correlation coefficients and risk reduction[4] (Roy, 1952). Besides increasing the pool of assets, an investor can specifically exploit this negative relationship by diversifying internationally. According to Grauer & Hakansson (1987), international diversification should result in a higher return-risk ratio than domestic diversification. A possible reason for this is that correlations between foreign assets and domestic assets are lower than between domestic assets (Pfau et al., 2007). The statement of Pfau et al. (2007) could well be the consequence of variation of monetary, fiscal and industrial policies between countries which in turn implies differences between country returns dynamics (Lagoarde-Segot & Lucey, 2007). Following up, Grubel (1968) presents results that are based on the time-period 1959-1966 and show that international diversification outperform domestic portfolios in terms of obtaining welfare gains and lower variance[5]. Just like the paper of Grubel (1968), the paper of Levy & Sarnat (1970) is also based on the work of Markowitz (1952) and includes data with a fairly similar time-period. These authors also underline that internationally diversified portfolios are an improvement over domestic portfolios.

Derived from Markowitz (1952), one can construct efficient frontiers in order to compare (efficient) portfolios. The efficient frontier concerns the display of a curve representing various portfolios that maximize the expected return given a corresponding risk or minimize risk given a corresponding return, where risk is measured by the standard deviation or covariance. As the model is heavily influenced by the historical estimates of the (1) mean return of, (2) standard deviation (or: variance) of and (3) correlation between assets' returns (Levy & Post, 2005), one can easily visualize the effects of changing correlation coefficients on the risk-and-return relationship. Like any other model, the efficient frontier model is subject to a

[3]Of course, in terms of statistics, the null hypothesis would be that one would test that there is no effect (cross-country correlation coefficients remain unchanged over time).

[4]The implicit assumption is positive portfolio weights.

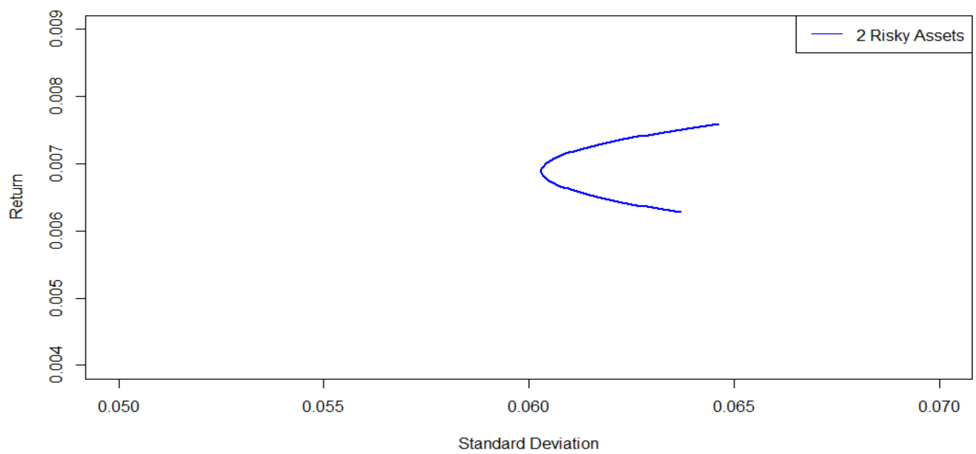
[5]The variance is one way to measure risk. Another common measure is the standard deviation.

number of assumptions (Markowitz, 1952; Jensen, 1968; Bodie, Kane, & Marcus, 2013), where assumption 8 does not always have to apply in practice and as a result can usually be left in theory without it seriously harming the efficient frontier[6]:

1. The investor's planning horizon is one period (for example two months, one year or fifty years).
2. Investors maximize the expected utility across all possible portfolios.
3. The expected usefulness of an investment is only a function of the expected return $\mathbf{E}(R)$, and the risk $\sigma(R)$ of that investment.
4. Investors prefer a higher return compared to a lower return and are also risk averse.
5. There is no tax, no inflation and there are no transaction and other costs.
6. Investors can have all information at no cost.
7. All investment objects are infinitely divisible.
8. Investors can do free of charge and unlimited short-selling.

Figure 2.1 below shows an example of an efficient frontier based on two risky assets. The (expected) portfolio return is marked on the y-axis and the measurement of risk, standard deviation, is represented at the x-axis. A frontier normally has two curves. The upper-curve (portion of the hyperbola with a positive slope) is essentially the 'efficient' frontier, as this frontier consists of portfolios that yield a higher expected return for the same amount of risk in comparison with the 'sub-optimal' portfolios on the lower-curve (portion of the hyperbola with a negative slope).

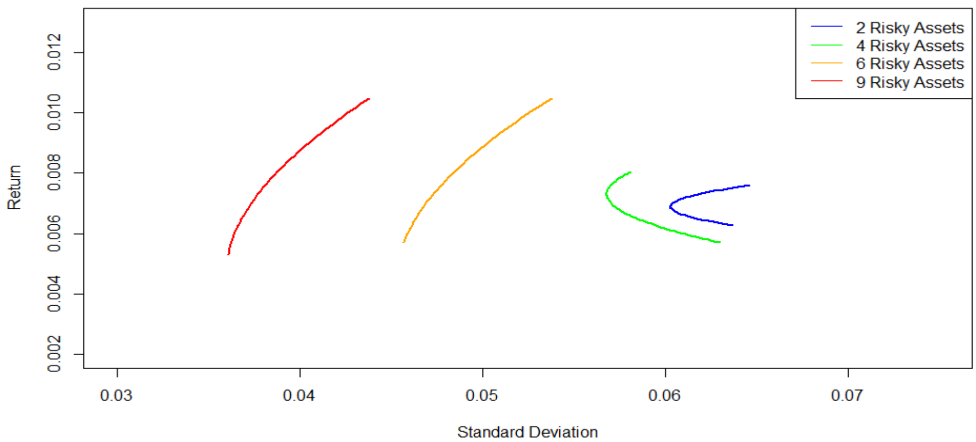
Figure 2.1: Example of an Efficient-Frontier based on Two Risky Assets



[6]In practice, assumption 5 is also fairly weak. Important to note is that violations of assumptions 5 (e.g. transaction costs) and 8 (short-selling limitations) can both lead to movement of the efficient frontier.

The efficient frontier shifts when the risk-and-return trade-off changes. For example, diversification, especially international diversification, can lead changes in the trade-off (Markowitz, 1952). A general consensus is that the risk can be severely reduced by increasing the amount of assets (Levy & Post, 2005). One interpretation is that more assets should lead to a more and well diversified portfolio due to the fact that the total portfolio risk is then spread over more assets (Markowitz, 1952; B. H. Solnik, 1995). Another interpretation is that more assets bring more and different correlation coefficients in the calculation and therefore lead to more diversification possibilities. In turn, the efficient frontier should bulge or/and shift to left as there are possibly more lower correlation coefficients[7] (Levy & Post, 2005). Figure 2.2 reflects this notion:

Figure 2.2: Different Efficient Frontiers (n assets)

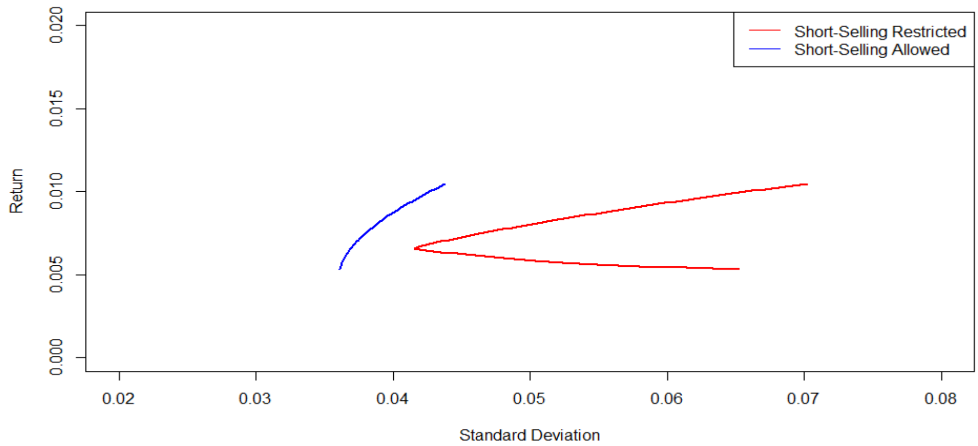


Even though the general consensus is that portfolio risk usually declines when, for example, more and more stocks (or other assets) are added to the portfolio, it is worth mentioning that it happens on a decreasing rate (Brigham & Houston, 2012). Figure 2.2 also serves as an example for the fact that increasing portfolio size has diminishing returns for risk reduction. B. H. Solnik (1995) provides evidence that risk reduction of having a portfolio with a size that is beyond some optimal portfolio size will only lead to marginal reduction in risk. Therefore, when the portfolio size passes the optimal size, the benefits of diversification will not significantly outweigh its costs. As one can observe from Figure 2.2, increasing the portfolio size from 2 to 6 (risky) assets results in a higher risk reduction than increasing the portfolio size from 6 to 9 (risky) assets. This example also shows that at some point the shape of the frontier does not change significantly. As can be seen, the frontier dramatically changes form when the amount of assets increase from 2 to 4 or from 2 to 6 in comparison to a increase in assets from 6 to 9.

[7]More assets does not necessarily imply that all assets have a low correlation with each other. An example of this is a portfolio that consists of only chip manufacturers (for example: *AMD*, *Intel*, *Qualcomm*, *Nvidia*, *Texas Instruments* and *Samsung*).

Another possible way in which the risk-and-return trade-off can change, is the restriction of short-selling[8]. In particular, a short-selling puts restriction on the weight allocation to each asset. The restriction ensures that each asset can only get a weight greater than or equal to zero. Levy & Post (2005) reason that the short-selling restriction leads to deterioration of the efficient frontier, because the highest possible (expected) return (given the risk) decreases in comparison to the fact that the lowest possible variance, or standard deviation, increases (given the expected return). Figure 2.3 shows the efficient frontier with and without the short-selling constraint. When there are no restrictions on the weights (short-selling is allowed), the marginal investor has more investment opportunities (Levy & Post, 2005) and, therefore, the efficient frontier without restrictions should clearly dominate.

Figure 2.3: Efficient-Frontier - Short-Sales Allowed/Restricted



In line with the previous mentioned empirical findings, Levy & Post (2005) also demonstrate that higher correlations between assets will lead to smaller gains from diversifications, ceteris paribus. This means that as the gains from diversifications diminish, the risk reduction will also be smaller (Levy & Post, 2005), implying a higher (portfolio) standard deviation and covariance for the same expected return. In addition, an unstable correlation matrix over time suggests that the corresponding covariance matrix is equally unstable over time. As the Modern Portfolio Theory (Markowitz, 1952) teaches us that both matrices are important inputs that are related to each other and influence the efficient frontier through the portfolio weights, it makes it all the more interesting to investigate the following hypothesis with the help of the efficient frontier[9]:

H_{03}: Over time, the risk (standard deviation) increases for the same level of return, meaning that the efficient frontier shifts to the right

[8]In practice, transaction costs occur (frequently). Since transaction costs lead to a reduction in net returns, this will also have to lead to deterioration of the efficient frontier. However, transaction costs are not taken into account in this paper and, therefore, this is not discussed in more detail.

[9]Of course, in terms of statistics, the null hypothesis would be that one would test that there is no effect (Over time, the risk, standard deviation, remains unchanged for the same level of return, meaning that the efficient frontier does not shift).

2.3 Diversification benefits

The majority of literature agrees that correlations rise over time as a result of globalization. At first glance, low correlations between markets seem to be the only necessary assumption to benefit from diversification. However, stable correlations between markets are as equally important (Jorion, 1985; Eun & Resnick, 1988). The latter assumption is also implied by Hyde, Bredin, & Nguyen (2007). The authors examine Asia-Pacific, EU and US equity markets with the help of the asymmetric generalized dynamic conditional correlation GARCH model (AG-DCC-GARCH) and provide evidence that the fluctuation of correlation and covariance between these markets suggest a significant time variation in the opportunities of international portfolio diversification. To test the diversification benefits over time, mean-variance spanning tests have been used by Driessen & Laeven (2007) and Kan & Zhou (2008). These methods essentially test if any movement of the efficient frontier (over time) is significant. The goal with a mean-variance spanning test is to test the null hypothesis that the improvement of the global minimum variance or/and tangency portfolio is not statistically significant. Technically stated, it tests the null hypothesis that a frontier with K (benchmark) assets 'spans' (is identical to) a frontier with K+N assets, where N are additional (test) assets. Throughout time and literature, mean-variance spanning tests are mainly used with the purpose for either testing diversification between assets (adding different N assets to a set of K benchmark assets) or testing the diversification over time (adding N assets to a set of K benchmark assets for multiple subperiods). The findings in both papers (Driessen & Laeven, 2007; Kan & Zhou, 2008) indicate a decrease in the diversification benefits over time.

Testing hypothesis H_{03} can help to identify the developments of the risk-and-return trade-off over time and thus to say something about the diversification benefits over time. Nevertheless, increasing risk over time does not always mean that the diversification benefits or diversification options will decrease *significantly* over time. The reason is that visualizing any movement of the efficient-frontier does not necessarily imply that the movement is significant. Therefore, it is necessary to test the actual significance of the movement in order to obtain insight regarding the significance of gains in diversification over time. That is why it is important to also consider the spanning tests, as described in Kan & Zhou (2008), and draw up a fourth hypothesis[10]:

H_{04}: *Over time, international diversification gains have significantly decreased*

[10] Of course, in terms of statistics, the null hypothesis would be that one would test that there is no effect (over time, international diversification gains have not significantly changed).

Part II

Methodology and Data

Chapter 3

Correlation Analysis

This chapter emphasizes the stability and pattern of correlation coefficients over time by applying different methodologies. In order to first explore the stability of the correlation coefficients and to answer hypothesis H_{01}, the Jennrich-test (Jennrich, 1970) will be used for different sub-periods (pre-crisis, crisis and post-crisis). Subsequently, the correlation analysis will be completed with three methods (Rolling-method, the EWMA-model and the DCC-GARCH model) for checking the pattern of the correlation coefficients throughout time. The use of three different methods gives the advantage that one can place sight on the pattern of correlation coefficients over time on the basis of three different perspectives. Analyzing the pattern of correlation coefficients will help to formulate an argument against or in favor of hypothesis H_{02}.

3.1 Testing for Stability

3.1.1 Jennrich Correlation Matrix Equality Test

An answer to the question of whether correlation coefficients have remained stable over time can be given when it is clear whether unconditional correlation matrices, based on the Pearson Correlation Coefficient, differ significantly between different sub-periods. In the distant past, various tests have been designed that put sight on this. However, a test can not simply be chosen on arbitrariness. It is important that the right test is chosen that takes into account the type of data that is underlying the research.

Since this paper will be based on financial data, it is not illogical to consider that financial data are skewed and heavy-tailed in most of the cases (Upadhyay, Singh, Dey, & Loganathan, 2015). One test that is suitable to tackle this, is the Jennrich-test (Jennrich, 1970). This asymptotic test relies on the χ^2-distribution and thus does not rely on the assumption of normality of observations (Omelka & Pauly, 2012). A feature of this test that suits well with this research is that the length of the time series is not ignored and that therefore the reliability of the test increases when the correlations are calculated with a large(r) length of data series (Schindler, 2009)[11]. This test involves another advantage because it gives the opportunity to compare the results with other papers, since the majority of the literature uses this test (Kaplanis, 1988; Longin & Solnik, 1995; Eichholtz, 1996a)[12].

[11] Stephen (2005) demonstrates this by also considering different lengths when testing the stability of correlation coefficients, based on UK Real Estate data.

[12] Some papers, like the one of Kaplanis (1988), also use the Box M (Box, 1949) test in order to also assess the stability of covariance matrices for the purpose of making comparisons.

Mathematically, the Jennrich-test can be defined as follows, in line with Jennrich (1970); Eichholtz (1996a); Schindler (2009):

$$\chi^2 = \frac{1}{2} tr(Z^2) - diag'(Z) S^{-1} diag(Z) \tag{3.1}$$

Z = $c^{1/2} \cdot R^{-1} \cdot (R_1 - R_2)$
c = $n_1 \cdot n_2 / (n_1 + n_2)$
n_1 = obs. of 1st correlation matrix
n_2 = obs. of 2nd correlation matrix
R = $(n_1 \cdot R_1 + n_2 \cdot R_2)/(n_1 + n_2)$
R_1 = 1st correlation matrix with n_1 obs.
R_2 = 2nd correlation matrix with n_2 obs.

S = $(\delta_{ij} + r_{ij} r^{ij})$
δ_{ij} = Kronecker delta ($\delta_{ij} = 1$ for $i = j$)
r_{ij} = elements of R
r^{ij} = elements of R^{-1}
tr = trace of matrix
$diag$ = diagonal of matrix

If D is equal to the dimension of the correlation matrices, then the degrees of freedom of the Jennrich-test is equal to $\frac{D(D+1)}{2}$. Finally, the statistical hypotheses which is being test is[13]:

$$H_0 : R_1 = R_2$$

$$H_a : R_1 \neq R_2$$

Preference is given to investigate unconditional correlation matrices of the following three specific sub-periods from the full sample: pre-crisis, crisis and post-crisis[14].

3.2 Development of Correlations Over Time

3.2.1 Rolling Analysis of Correlation Coefficients

One way to assess how correlation coefficients between indices have developed, is the calculation of correlation coefficients on a rolling basis. Despite its simplicity, a rolling analysis is very powerful because it allows correlation coefficients to be time-varying.

As mentioned in Buyuksahin, Haigh, & Robe (2008), if x_1 and x_2 are assumed to be deviations from the means of the particular two random variables, with a mean equal to zero, then the rolling correlation can be calculated by means of this formula:

$$\hat{\rho}_{12,t+1} = \frac{\sum_{s=t-k}^{t} x_{1,s} x_{2,s}}{\sqrt{(\sum_{s=t-k}^{t} x_{1,s}^2)(\sum_{s=t-k}^{t} x_{2,s}^2)}} \tag{3.2}$$

[13] Note that the statistical hypothesis is the opposite of hypothesis H_{01}.
[14] See chapter 6 (Data) for the specific time-frame of each sub-period.

Where $\hat{\rho}_{12,t+1}$ stands for the correlation coefficient at time t+1 between variable 1 and 2 and where k stands for the length of estimation. Ideally, the length of the window-size (estimation-window) should be as large as possible. The larger the estimation-window, the greater the number of observations that are included in the estimation. An advantage of this is that this improves the accuracy of the estimation and leads to less noisy estimates. However, it is not realistic to choose indiscriminately for a random large estimation-window. A choice must be made for a length in which this reflects reality as well as possible. In this paper, k = 12 as the rolling analysis will be done on the basis of 12 months. The reason that 12 months have been chosen, is that most of the investors re-balance annually (per 12 months).

3.2.2 Exponentially Weighted Moving Average (EWMA)

A major disadvantage of the rolling analysis, discussed in subsection 3.2.1, is that the weights are evenly distributed over the observations that are included in the estimation-window. As (un)expected changes in international markets, for example monetary or regulatory policies, are not rare, there could be a delayed impact on correlation coefficients[15]. The EWMA-method tackles this disadvantage because its calculation is a moving average where the weights attached to past observations *decay* exponentially (Hendricks, 1997). Therefore, this method is also used in this paper. As the data in this paper is monthly, the decay factor, λ, is set fixed on 0.97[16], in line with RiskMetrics (Brooks, 2014).

Important to understand is that the EWMA-method does essentially the same as the standard method of rolling correlations. However, the difference between the two methods lies in the fact that the EWMA-method makes use of a certain decay-factor instead of a specified window-size of k length. The mathematical definition of the EWMA-method below shows both the resemblance and difference between the standard rolling approach and EWMA in a clear way (Buyuksahin, Haigh, & Robe, 2008):

$$\hat{\rho}_{12,t+1} = \frac{\sum_{s=1}^{t} \lambda^{t-s} x_{1,s} x_{2,s}}{\sqrt{(\sum_{s=1}^{t} \lambda^{t-s} x_{1,s}^2)(\sum_{s=1}^{t} \lambda^{t-s} x_{2,s}^2)}} \tag{3.3}$$

Where $\hat{\rho}_{12,t+1}$ stands for the correlation coefficient at time t+1 between variable 1 and 2 and where x_1 and x_2 are assumed to be deviations from the means of the particular two random variables, with a mean equal to zero and λ equals the smoothing parameter that ensures that past observations decay exponentially (Hendricks, 1997).

[15]Normally one can also choose to make the window size as small as possible to ensure that only the most 'recent' observations are included in the estimation window. However, a too small estimation-window can lead to inaccurate and noisy correlation coefficients.

[16]$\lambda = 0.94$ is also common and has been used by, for example, Engle (2002). However, his analysis is based on daily returns.

The difference between the rolling method and EWMA procedure can also be demonstrated by looking at the weights of each method on each observation. As indicated earlier, the rolling method is based on an estimation-window of k = 12. This means that, each time, the last 12 observations are used in the estimation. At first, the 1st to the 12th observations are used, followed by the 2nd to the 13th and then the 3rd to the 14th and so on. Each time the weights are evenly distributed (1/k), in this case 1/12. Thus, each of the 12 observations has a weight of 0.0833 (8.33%).

In contrast, as mentioned, the EWMA-procedure does not rely on any estimation window-size and therefore all 'past' observations are taken into account. However, the EWMA procedure is biased towards more recent observations. The weights for the EWMA procedure can be determined on the basis of $(1-\lambda)*(\lambda)^{b_i}$, where b_i is equal to (n_i - 1) and n_i is equal to the i-th observation. For example, the weight for the 1st and most recent observation can be calculated as $(1-0.97)*(0.97)^0 = 0.0300$. The same calculation can be used to calculate the weight for the 2nd, 24th and 60th observation: $(1-0.97)*(0.97)^1 = 0.0291$, $(1-0.97)*(0.97)^{23} = 0.0149$ and $(1-0.97)*(0.97)^{59} = 0.005$. A tabular representation of the weight allocation difference between the rolling and EWMA method for the 12 most recent observations can be observed in Table 3.1:

Table 3.1: Weights of Rolling- and EWMA Method for The Most Recent 12 Observations

Observation	**Rolling Correlations**	**EWMA**
1	0,0833	0,0300
2	0,0833	0,0291
3	0,0833	0,0282
4	0,0833	0,0274
5	0,0833	0,0266
6	0,0833	0,0258
7	0,0833	0,0250
8	0,0833	0,0242
9	0,0833	0,0235
10	0,0833	0,0228
11	0,0833	0,0221
12	0,0833	0,0215

The table presents the weights of the 12 most recent observations for both the rolling correlation method and the EWMA-procedure in order to make the differences transparent. The 1st observation is the most recent observation relative to the 12th.

3.2.3 DCC-GARCH

The final method to investigate the pattern of correlation coefficients over time, is the Dynamic Conditional Correlation GARCH (DCC-GARCH) model of Engle (2002), which is (1) based on the GARCH-model (Bollerslev, 1986) and (2) a generalization of the Constant Conditional Correlation GARCH model (CCC-GARCH) of Bollerslev (1990). Preference is given to the DCC-GARCH instead of the CCC-GARCH because the former allows the correlation matrix to be time-dependent (Orskaug, 2009).

One may wonder to what extent the DCC-GARCH model differs from the previous two methods (Rolling and EWMA). According to Buyuksahin, Haigh, & Robe (2008) the DCC-GARCH is better able to account for changes in volatility. The author affirms that both the Rolling and the EWMA method are sensitive to volatility shocks and that these two methods, as discussed in subsection 3.2.1 and 3.2.2, make the interpretation of the estimated correlation coefficients difficult during periods of high volatility.

To which Equation 3.2 and 3.3 have great similarities with each other, the formula of the DCC-GARCH model has that to a very limited extent. This is mainly due to the dynamic structure and calculation of the model concerned, as the estimation is based on two stages. In the first stage, the conditional volatility is calculated with the help of a GARCH model. The corresponding standardized residuals from this stage are then used, for the second stage, to estimate the dynamic correlation matrix (Buyuksahin, Haigh, & Robe, 2008)

Mathematically, the correlation estimator of the DCC-GARCH can be denoted in a straightforward way. However, the DCC-GARCH deserves some explanation about its specification, where the mathematical structure and definition is based on Engle (2002) and Buyuksahin, Haigh, & Robe (2008). If there is a normally-distributed $nx1$ vector with mean zero, return series r_t of n assets and a corresponding covariance matrix H_t, then the following baseline can be made:

$$r_t \sim N(0, H_t) \tag{3.4}$$

$$H_t = D_t R_t D_t \tag{3.5}$$

H_t = Conditional covariance matrix
D_t = diagonal matrix of time-varying standard deviations
R_t = time-varying correlation matrix

D_t can also be interpreted as $D_t = diag\sqrt{E_{t-1}(r_{i,t}^2)} = diag\sqrt{h_{i,t}}$, where $h_{i,t}$ can be seen as a generalized GARCH(p,q) model (Peters, 2008):

$$h_{i,t} = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{p} \gamma_i h_{t-i} \tag{3.6}$$

ω = weighted long run variance
$\sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2$ = moving-average term
α_i = weight i
ε_{t-i}^2 = squared-innovations
$\sum_{i=1}^{p} \gamma_i h_{t-i}$ = Sum of p lagged variances multiplied by weight γ_i, $i = 1, ..., p$

Therefore, assuming that $\varepsilon_{i,t} \sim N(0, R_t)$, the standardized disturbance can be denoted as $\varepsilon_{i,t} = r_{i,t}/\sqrt{h_{i,t}} = D_t^{-1} r_{i,t}$. Based on this, conditional correlations can be expressed as:

$$\rho_{ij,t} = \frac{E_{t-1}[r_{i,t} r_{j,t}]}{\sqrt{E_{t-1}[r_{i,t}^2 r_{j,t}^2]}} \tag{3.7}$$

Equation 3.7 can also be re-expressed in terms of the standardized residuals from the GARCH-model:

$$\rho_{ij,t} = E_{t-1} \varepsilon_{i,t} \varepsilon_{j,t} \tag{3.8}$$

Further specification is possible with the following dynamics of the DCC-GARCH model (Engle, 2002):

$$R_t = Q_t^{*-1} Q_t Q_t^{*-1} \tag{3.9}$$

$$Q_t = (1 - \alpha - \beta)\bar{Q} + \alpha(\varepsilon_{i,t-1} \varepsilon_{j,t-1}) + \beta Q_{t-1} \tag{3.10}$$

$Q_t = q_{ij,t}$ = covariance between $r_{i,t}$ and $r_{j,t}$
$\bar{Q}_t = E[\varepsilon_t \varepsilon_t^{'}] =$ unconditional correlation matrix of standardized residuals
Q_t^* = diagonal matrix composed of the square roots of the diagonal elements of Q_t
R_t = time-varying conditional correlation matrix of returns

Interestingly to note is the fact that if $\alpha + \beta < 1$ holds in Equation 3.10, then there will be mean-reversion.

With the knowledge about the structure of the DCC-GARCH model, the dynamic conditional correlation between the two assets (i and j) at a certain time point (t) can subsequently be calculated with the use of the following formula:

$$\rho_{ij,t} = \frac{q_{ij,t}}{\sqrt{q_{ii,t}q_{jj,t}}} \tag{3.11}$$

Generalization of Equation 3.10 can be expressed in terms of DCC(m, n) as follows (Peters, 2008):

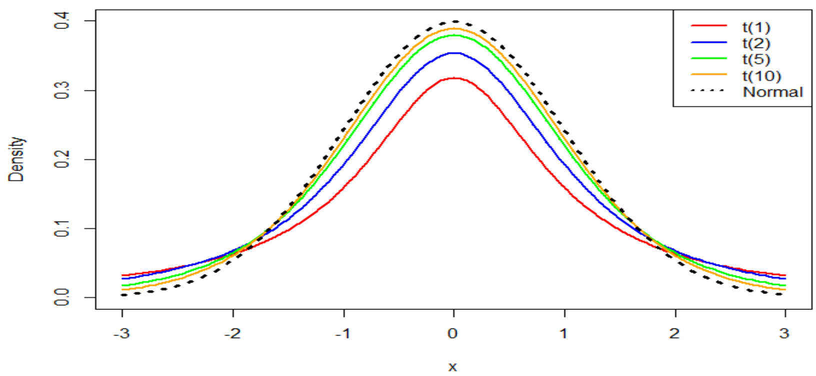
$$Q_t = (1 - \sum_{i=1}^{m} \alpha_i - \sum_{j=1}^{n} \beta_j)\bar{Q} + \sum_{i=1}^{m} \alpha_i \varepsilon_{i,t-i} \varepsilon'_{i,t-i}) + \sum_{j=1}^{n} \beta_j Q_{t-j} \tag{3.12}$$

The following loglikelihood function is with the purpose of estimation and maximization:

$$L = -\frac{1}{2}\sum_{t=1}^{T}(nlog(2\pi) + 2log(|D_t|) + log(|R_t|) + \varepsilon' R_T^{-1} \varepsilon) \tag{3.13}$$

In this paper, the DCC(1,1)[17] is considered because it is not only simple and highly effective but also because it is frequently used in literature. The DCC-GARCH model will be based on the student's t-distribution, instead of the standard normal distribution. This choice has been made due to the fact that the majority of the literature agrees that the t-distribution performs better than the normal distribution. Dube (2016), for example, has come to the conclusion that (1) the normal-DCC model is not capable of taking into account fat-tails in financial time series data, where the t-DCC model has no issues with these fat-tails and (2) that t-DCC model outperforms the normal-DCC model[18]. As example, Figure 3.1 shows a visual comparison between the normal distribution and the student's t-distributions based on 1, 2, 5 and 10 degrees of freedom.

Figure 3.1: Normal Distribution vs. Student's t-distribution



[17] Specifically: the DCC-GARCH (1,1,1,1), also known as DCC(1,1)-GARCH(1,1).

[18] Important mentioning is that the same author remarks that the t-DCC model has difficulties passing the Kolmogorov-Smirnov goodness-of-fit test.

Chapter 4

Dynamic Efficient Frontiers

After analyzing correlation coefficients, light needs to be shed on the effects of correlation coefficients on the trade-off between risk and (expected) return in order to form an answer regarding whether hypothesis H_{03} will be rejected or not. A common way to show these effects over time is to use graphs showing efficient frontiers. As hypothesis H_{03} is based on what happens over time, efficient frontiers will be created for the full-sample period and three sub-periods (pre-crisis, crisis and post-crisis)[19]. As transaction costs influence the risk-and-return tradeoff and the efficient frontier, transaction costs are assumed to be constant and equal to zero in this paper.

Although the theoretical approach and the underlying general assumptions of the model have been discussed in section 2.2, the efficient frontier still requires explanation regarding the mathematical approach and the way the efficient frontiers will be constructed and utilized in this paper. Firstly, it is important to know that to find the efficient frontier, one needs to find portfolios that minimizes the portfolio's variance:

$$\sigma_p^2 = \omega' V \omega \tag{4.1}$$

where, $\omega = (\omega_1, ..., \omega_n)$ represents the weights allocated to each asset and V represents the covariance matrix of n assets;

$$V = \begin{pmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_{nn} \end{pmatrix} \tag{4.2}$$

$\sigma_{i,j}$ stands for the covariance between asset i and j, whereas σ_i^2 stands for the covariance of only asset i (or: variance of asset i). Even though it is logical, it is still worth noting that $\sigma_{i,j}$ is identical to $\sigma_{j,i}$ as both represent the covariance of assets i and j.

[19]chapter 6 defines these sub-periods.

As mentioned, Equation 4.1 can be optimized by finding portfolios that minimize risk given a particular expected (desired) portfolio return. Therefore, for a range of possible expected portfolio returns, the Langrangian will subject to the following constraint:

$$\bar{r}_p = \omega' \bar{r} \tag{4.3}$$

Where, the expected return of the portfolio ($\bar{r}_p$) equals the expected return of the assets ($\bar{r} = r_1, ..., r_n$) multiplied by the portfolio weights (ω).

As this paper will only focus on the assets (data) that are going to be investigated, no specific risk-free asset will be included. Therefore, an additional constraint must be added that ensures that the total allocated weight is equal to 1. This means that no lending or borrowing is allowed due to the fact that the risk-free asset is non-existent. The constraint can be stated as follows, where 'e' is a vector consisting solely of ones:

$$\omega' e = 1 \tag{4.4}$$

Finally, with the help of the following Langrangian, the variance is minimized by adjusting the weights, subject to the constraints that the expected portfolio return equals $\bar{R}_p$ and weights (ω) sum to 1:

$$\underset{\omega}{min} \frac{1}{2} \omega' V \omega + \lambda \left[\bar{R}_p - \omega' \bar{r} \right] + \gamma \left[1 - \omega' e \right] \tag{4.5}$$

Optimization should yield to the following formula for constructing the optimal weights (ω^*):

$$\omega^* = \frac{e' V^{-1} e \bar{R_p} - \bar{r}' V^{-1} e}{\bar{r}' V^{-1} \bar{r} e' V^{-1} e - \left(\bar{r'} V^{-1} e \right)^2} V^{-1} \bar{r} + \frac{\bar{r}' V^{-1} \bar{r} - \bar{r}' V^{-1} e \bar{R_p}}{\bar{r}' V^{-1} \bar{r} e' V^{-1} e \left(\bar{r'} V^{-1} e \right)^2} V^{-1} e \tag{4.6}$$

The formula for the optimal weights (ω^*) show that the optimal portfolio gets influenced by three factors: the average returns of the assets ($\bar{r}$), the co-movement between the assets (V) and the answer on how much risk the particular investor is willing to bear ($\bar{R_p}$). Equation 4.6 and point (3) illustrates exactly the hypothesis that the marginal investor chooses to only hold his (risky) asset if the return is high enough to offset it (Hillier et al., 2013). According to Levy & Post (2005), additional constraints can be considered in real life applications such as short-sales restrictions. Therefore, it is useful to also investigate hypothesis H_{03} for the case in which short-selling is not allowed. This helps to make comparisons with the non-restricted frontiers. Equation 4.7, below, shows the constraint that needs to be applied to the weights in order to restrict short-selling:

$$\omega_i \geq 0 \tag{4.7}$$

Chapter 5

Tests of Mean-Variance Spanning

In the previous chapter, the focus lies in the risk-and-return trade-off over time. However, as mentioned earlier in this paper, diversification can influence this trade-off. Therefore it is interesting to find out whether diversification benefits have decreased over time (hypothesis H_{04}). And thus, this chapter will shed light on this particular question with the help of the Mean-Variance Spanning Tests. The main idea behind the Mean-Variance Spanning Test is quite conventional. It basically tests the null-hypothesis that a set of K assets has the same Minimum-Variance Frontier (short: MVF) as the set of $K+N$ assets (Kan & Zhou, 2008). If and only if both MVFs are equal to each other, then one can say that a set of K assets spans a larger set of $K+N$ assets (Kan & Zhou, 2008). However, if the null-hypothesis is rejected, then $K+N$ assets will statistically improve the MVF and thus add value in terms of diversification. Because hypothesis H_{04} is based on what happens *over time*, the tests will be computed for multiple sub-periods of the sample. The same sub-periods will be used as for the methodologies in subsection 3.1.1 and chapter 4.

The original, regression-based, Mean-Variance Spanning Test (Huberman & Kandel, 1987) dates from many years ago . Nowadays, there are several variants of the Mean-Variance Spanning Test. It may well be that the use of a certain variant can lead to a different conclusion than when another variant is used. Therefore it is important to apply at least a couple variants of the spanning test in order to check for robustness under, for example, non-normality. All notations, variants and brief explanations in this paper will be based on the paper of Kan & Zhou (2008). Therefore, in the case of theoretical or practical ambiguities that require more explanation, it is advisable to take a look at Kan & Zhou (2008).

5.1 Regression Approach

Huberman & Kandel (1987) are the first ones to introduce the Mean-Variance Spanning Test, based on the regression framework. Because this variant is based on the regression framework, it is necessary that the assumptions of linear regression hold. At first, let R_{1t} be a K-vector of returns on the K benchmark assets and R_{2t} an N-vector of the returns on the N test assets. Based on this, the raw returns on $K+N$ assets, R_t, can be mathematically written as $[R_{1t}^{'}, R_{2t}^{'}]'$ The corresponding expected returns on $K+N$ assets can be written as follows, where the subscript 1 refers to the K benchmark assets and subscript 2 to the N test assets:

$$\mu = \mathbf{E}[R_t] = \begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix} \tag{5.1}$$

If the covariance matrix is nonsingular, then it can be defined as:

$$V = Var[R_t] = \begin{bmatrix} V_{11} & V_{12} \\ V_{21} & V_{22} \end{bmatrix} \tag{5.2}$$

As the name implies, the regression approach is based on the fact that a regression is the underlying of the statistical test in this approach:

$$R_{2t} = \alpha + \beta R_{1t} + \varepsilon_t \tag{5.3}$$

$E[\varepsilon_t] = 0_N$, where 0_N is an N-vector of zeros
$E[\varepsilon_t R_{1t}^{'}] = 0_{N\times K}$, where $O_{N\times K}$ is an N by K matrix of zeros
$\alpha = \mu_2 - \beta\mu_1$
$\beta = V_{21}V_{11}^{-1}$
$\delta = 1_N - \beta 1_K$, where 1_N and 1_K are respectively a N- and K-vector that consist of ones

The necessary hypotheses that are going to be tested, is stated as follows:

$$H_0\colon\ (1)\ \alpha = 0_N,\ (2)\ \delta = 0_N$$

Essentially, this tests whether the intercept(s) and slope(s) are equal to zero. Because when this null hypothesis holds, the portfolio of K benchmark assets has the same mean, but a lower variance than the portfolio of test assets, which means the K benchmark assets dominate the N test assets[20]. Another way stated, this approach has the null hypothesis that (1) the portfolio weights of N additional test assets in the tangency portfolio (Q_{w1}) and (2) the global minimum-variance portfolio (Q_{w2}) are equal to zero. The mathematical definitions of (Qw_1) and (Qw_2) can lead to a more elegant way of understanding why α and δ need to be zero[21]:

$$Qw_1 = \frac{\Sigma^{-1}\alpha}{1_{N+K}^{'}V^{-1}\mu'} \tag{5.4}$$

$$Qw_2 = \frac{\Sigma^{-1}\delta}{1_{N+K}^{'}V^{-1}1_{N+K}} \tag{5.5}$$

With, $\Sigma = V_{22} - V_{21}V_{11}^{-1}V_{12}$ and 1_{N+K} equal a N+K matrix that includes only ones and $Q = [O_{N\times K}, I_N]$, where I_N is an N by N identity matrix. Also, $w_1 = \frac{V^{-1}\mu}{1_{N+K}^{'}V^{-1}\mu}$ and $w_2 = \frac{V^{-1}1_{N+K}}{1_{N+K}^{'}V^{-1}1_{N+K}}$. Important to mention is that this study will focus on using the regression-approach in a multivariate setting, as discussed in Kan & Zhou (2008).

[20]As Kan & Zhou (2008) state, this is true because it is assumed that R_{1t} and ε_t are are uncorrelated and $Var[\varepsilon]$ has positive definiteness.

[21]The fractions can only be zero if the numerators are zero. Since Σ^{-1} consists of covariances, it is nevertheless logical that only α and δ can ensure that the numerators (and therefore the fractions) are equal to zero.

Kan & Zhou (2008) mention three widely used asymptotic tests that can be combined with the regression-approach of the mean-variance spanning test: Likelihood Ratio test (LR-test), Wald-test (W) and the Lagrange Multiplier test (LM)[22]. However, there is one major disadvantage associated with the use of an asymptotic test. The disadvantage is that the use of an asymptotic test could lead to less reliable results in finite samples (Gibbons, Ross, & Shanken, 1989; Kan & Zhou, 2008). Therefore, based on Huberman & Kandel (1987); Jobson & Korkie (1989); Kan & Zhou (2008), a corrected F-test will be applied in order to improve the reliability of the results. The corrected F-test, that will be used in this paper, consists of two possible test-statistics. The first test statistic is valid for $N = 1$, whereas the second test statistic is only valid for $N \geq 2$. In Equation 5.6 and 5.7, $U = \frac{|\hat{\Sigma}|}{|\tilde{\Sigma}|}$ and $\frac{1}{U} = \frac{|\tilde{\Sigma}|}{|\hat{\Sigma}|}$ where $\tilde{\Sigma}$ is the constrained and $\hat{\Sigma}$ unconstrained maximum likelihood estimator of $\Sigma = V_{22} - V_{21}V_{11}^{-1}V_{12}$.

$$\left(\frac{1}{U} - 1\right)\left(\frac{T-K-1}{2}\right) \sim F_{2N,T-K-1} \tag{5.6}$$

$$\left(\frac{1}{U^{\frac{1}{2}}} - 1\right)\left(\frac{T-K-N}{N}\right) \sim F_{2N,2(T-K-N)} \tag{5.7}$$

5.2 Step-Down Procedure

In contrast to the regression-approach, a sequential test with two hypotheses, defined as the step-down procedure, provides results that yield more useful information. Essentially, the step-down procedure tests exactly the same as the joint F-test in the regression-approach. There is, however, one major difference. That is that the step-down procedure considers two separate (F-)tests instead of 1 joint test. The first test (F_1) tests that $\alpha = 0_N$ and the second test (F_2) tests whether $\delta = 0_N$, conditional on the fact that $\alpha = 0_N$. As Kan & Zhou (2008) mention, this test has two major advantages due to the fact that it is sequential. The first one is that this test should give (more) information about the reasoning behind the rejection of the null hypothesis. If the sequential test does not pass due to the F_1, then the reason is that the two tangency portfolios are statistically different from each other. However, if the rejection is the result of F_2, then the two global minimum-variance portfolios do significantly differ from each other. There is also a possibility that the rejection is due both the F_1 and F_2. Logically, this implies that both the two tangency and the two global minimum-variance portfolios differ significantly from each other. The second advantage is that, based on the relative economic significance, one can give both *seperate* tests different significance levels[23].

[22] In a multivariate test setting, the popular assumption is to assume that α and β are constant over time while regressing $R_{2t} = \alpha + \beta R_{1t} + \varepsilon_t$, where $t = 1, 2, ..., T$, and where T is the length of time series.

[23] In this paper however, preference is given to simplicity by using the same significance levels (1%, 5% and 10%) for both seperate tests.

As the step-down procedure will be considered in this paper, it won't do any harm to mathematically define the two F-test statistics[24]:

$$F_1 = \left(\frac{T-K-N}{N}\right)\left(\frac{\bar{\Sigma}}{\hat{\Sigma}} - 1\right) \tag{5.8}$$

$$F_2 = \left(\frac{T-K-N+1}{N}\right)\left(\frac{\tilde{\Sigma}}{\bar{\Sigma}} - 1\right) \tag{5.9}$$

Where, first, $\hat{\Sigma}$ is the unconstrained maximum likelihood estimator of $\Sigma = V_{22} - V_{21}V_{11}^{-1}V_{12}$. In addition, $\bar{\Sigma}$ is the constrained maximum likelihood estimator of Σ by imposing the condition that $\alpha = 0_N$. Finally, $\tilde{\Sigma}$ is the constrained maximum likelihood estimator of Σ by imposing the conditions that $\alpha = 0_N$ and $\delta = 0_N$ (Kan & Zhou, 2008).

5.3 Spanning Tests Under Non-Normality

Due to the nature of financial data, it will not come as a surprise if the error-term (ε_t) of Equation 5.3 exhibits non-normality. There are two types of non-normalities that can occur frequently: (1) conditional homoskedasticity and (2) conditional heteroskedasticity. If ε_t is conditionally homoskedastic, then ε_t is still independently and identically distributed, conditional on R_{1t}. When ε_t exhibits the behavior of being conditional heteroskedasticity however, then the variance of ε_t can be time-varying as a function of R_{1t} (Kan & Zhou, 2008).

In the former case, the asymptotic tests can still be valid due to its very good approximation. Even though, it was mentioned that the F-test has the favor over asymptotic tests due to more reliable results, this paper will also consider an asymptotic test. The two main reasons are that (1) asymptotic tests do not rely on the normality assumption, where the F-test does and (2) asymptotic tests can handle conditional homoskedasticity. Thus, this paper will use the Wald-test as asymptotic test as stated in Equation 5.10[25]. In Equation 5.10, it is assumed that λ_1 estimates the maximum distance between the two ex-post frontiers, whereas λ_2 measures the minimum distance. Important to note is that these estimations are done in terms of the squared sample Sharpe Ratios (Kan & Zhou, 2008).

$$W = T(\lambda_1 + \lambda_2) \overset{A}{\sim} \chi^2_{2N} \tag{5.10}$$

[24]Kan & Zhou (2008) present and introduce the constants in order to clarify the geometrical interpretation.

[25]Kan & Zhou (2008)) normally recommend, in the sake of completeness, to use all three tests. Their reason is that not all of three asymptotic tests (LR-, LM- and W-test) are equal to each other. Conflicting results can be a logical consequence. Nevertheless, just like Kan & Zhou (2008), only the Wald test is chosen in this paper. There is one reason for this, which is the same as the reasoning of Kan & Zhou (2008). That is that the GMM Wald test does not require a specification of the initial weighting matrix and the number of iterations.

However, in the case of conditional heteroskedasticity it is a reasonable idea to possibly use an alternative. This alternative is the GMM Wald Test (Hansen, 1982), which corrects for conditional heteroskedasticity and is valid for all distributions, as it does not depend on the normality assumption. The exact test statistic (W_a) can be derived with following the approach of Kan & Zhou (2008) and Ferson, Foerster, & Keim (1993). First, a couple of descriptions and definitions must be stated:

If, $x_t = [1, R_{1t}^{'}]'$ and $\varepsilon_t = R_{2t} - B'x_t$, then the moment conditions for the GMM estimation of B are, assuming $0_{(K+1)N}$ is a (K+1)N vector of zeros:

$$E[g_t] = E[x_t \otimes \varepsilon_t = 0_{(K+1)N}] \tag{5.11}$$

The sample moments can be stated as follows, assuming that R_t is stationary with finite fourth moments:

$$\bar{g}_T(B) = \frac{1}{T}\sum_{t=1}^{T} x_t \otimes (R_{2t} - B'x_t) \tag{5.12}$$

Minimization of $\bar{g}_T(B)'S_T^{-1}\bar{g}_T(B)$ yields to the GMM estimate of B, where S_T is the consistent estimate of $S_0 = E[g_t g_t^{'}]$, where g_t is assumed to be serial uncorrelated. Kan & Zhou (2008) state the test statistic as follows:

$$W_a = T vec(\hat{\Theta}')' \left[(A_T \otimes I_N) S_T (A_T^{'} \otimes I_N)\right]^{-1} vec(\hat{\Theta}') \overset{A}{\sim} \chi^2_{2N} \tag{5.13}$$

vec = vectorization

I_N = N by N identity matrix

$A_T = \begin{bmatrix} 1+\hat{a}_1 & -\hat{u}_1 \hat{V}_{11}^{-1} \\ \hat{b}_1 & -1_K' \hat{V}_{11}^{-1} \end{bmatrix}$

$\hat{\mu}_1 = \frac{1}{T}\sum_{t=1}^{T} R_{1t}$

$\hat{a}_1 = \hat{\mu}_1^{'} \hat{V}_{11}^{-1} \hat{\mu}_1$

$\hat{b}_1 = \hat{\mu}_1^{'} \hat{V}_{11}^{-1} 1_K$

$\hat{V}_{11} = \frac{1}{T}\sum_{t=1}^{T}(R_{1t} - \hat{\mu}_1)(R_{1t} - \hat{\mu}_1)'$

$\hat{\Theta} = [\hat{\alpha}, \hat{\delta}]'$

At first glance this would be a good alternative to support a robustness check. Nevertheless, there is another alternative in the case of a multivariate elliptical distribution. The meaning behind this distribution is that the conditional variance of ε_t is not constant but a function of R_{1t}, unless the returns are multivariate distributed (Kan & Zhou, 2008). This form of the GMM-Wald test statistic (W_a^e) can be build up as follows:

$$\hat{H} = \hat{\Theta}\hat{\Sigma}^{-1}\hat{\Theta}' = \begin{bmatrix} \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\alpha} & \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\delta} \\ \hat{\alpha}'\hat{\Sigma}^{-1}\hat{\delta} & \hat{\delta}'\hat{\Sigma}^{-1}\hat{\delta} \end{bmatrix} \tag{5.14}$$

$$\hat{G}_a = \begin{bmatrix} 1+(1+\hat{k})\hat{a}_1) & (1+\hat{k})\hat{b}_1) \\ (1+\hat{k})\hat{b}_1) & (1+\hat{k})\hat{c}_1) \end{bmatrix} \tag{5.15}$$

$\hat{b}_1$ and $\hat{a}_1$ in Equation 5.15 are the same as in Equation 5.13. $\hat{c}_1 = 1_K'\hat{V}_{11}^{-1}1_K$ and $\hat{k}$ is the consistent estimate of the kurtosis parameter k:

$$k = \frac{E[((R_t - \mu)'V^{-1}(R_t - \mu))^2]}{(N+K)(N+K+2)} - 1 \tag{5.16}$$

The test statistic is as follows, where tr = trace (Kan & Zhou, 2008):

$$W_a^e = T\,tr(\hat{H}\hat{G}_a^{-1}) \overset{A}{\sim} \chi^2_{2N} \tag{5.17}$$

Precisely because this test statistic takes a specific situation into consideration, this test statistic is only valid when there is a multivariate elliptical distribution. And as mentioned before, W_a version is fairly valid for every distribution. In order to improve the reliability and check for robustness, both the W_a and W_a^e will be used in this paper.

5.4 Tests Based on The Stochastic Discount Factor (SDF)

An innovative view of the mean-variance spanning test is presented by De Santis (1993) and Bekaert & Urias (1996). The innovation that they apply, is that they focus on the stochastic discount factor (SDF) and combine this with the GMM[26]. Both SDF-approaches test if N additional test assets could describe the variance of the SDF (Kan & Zhou, 2008). Even though both approaches have similarities, they are significantly different from each other. Both SDF-approaches will be used in this paper[27]. Note that the SDF gives the opportunity to go outside the mean-variance framework, as the SDF is equivalent to the inter-temporal marginal rate of substitution and can therefore be linked to, for example, (the marginal utility of) consumption. The reason for this is that the SDF can be derived from the utility function.

5.4.1 DeSantis (1993)

De Santis (1993) has devised a way to test the spanning hypothesis with the use of the stochastic discount factor (m_t). As formulated in Kan & Zhou (2008), the goal of De Santis (1993) is to analyze whether N additional (test) assets contribute to explaining the variance of m_t. He defines the stochastic discount factor as follows, where c equals the mean on the returns of $K+N$ assets and where 1_{N+K}, under the law of one price, is equal to $E[(1_{N+K} + R_t)m_t]$:

$$m_t = c + (1_{N+K} + R_t)'\gamma(c) + \varepsilon_t \tag{5.18}$$

[26]This approach assumes that the weights of the frontier portfolios are constant over time (Kan & Zhou, 2008).

[27]For specific details regarding both the SDF-approaches, including the differences and similarities between each other, it is advisable to check (Kan & Zhou, 2008).

$$\gamma(c) = U^{-1}[(1-c)1_{N+K} - c\mu] \tag{5.19}$$

$$U = E[(1_{N+K} + R_t)(1_{N+K} + R_t)'] \tag{5.20}$$

The hypotheses that are being tested, can be denoted as follows, where 0_N is an N-vector of zeros:

$$H_0\colon \ (1)\ Q\gamma(c_1) = 0_N,$$
$$(2)\ Q\gamma(c_2) = 0_N$$

Where, $Q = [O_{N+K}, I_N]$ and $\gamma(c1)$ and $\gamma(c2)$ are two different linear combinations of the weights of two frontier portfolios of $K+N$ assets (Kan & Zhou, 2008). Note that $Q\gamma(c_1)$ and $Q\gamma(c_2)$ have the same meaning as $Q\omega_1 = 0_N$ and $Q\omega_2 = 0_N$.

According to Kan & Zhou (2008), the GMM Wald test statistic can be build up as follows. First assume that the sample moment conditions are equal to:

$$\bar{m}_T(\gamma(c_1), \gamma(c_2)) = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T}(1+R_t)(c_1 + (1+R_t)'\gamma(c_1)) - 1_{N+K} \\ \\ \frac{1}{T}\sum_{t=1}^{T}(1+R_t)(c_2 + (1+R_t)'\gamma(c_2)) - 1_{N+K} \end{bmatrix} \tag{5.21}$$

Based on this and assuming that $\hat{U} = \frac{1}{T}\sum_{t=1}^{T}(1+R_t)(1+R_t)'$, the unconstrained estimates of $\gamma(c_1)$ and $\gamma(c_2)$ are:

$$\hat{\gamma}(c_1) = \hat{U}^{-1}[1_{N+K} - c_1(1_{N+K} + \hat{\mu})] \tag{5.22}$$

$$\hat{\gamma}(c_2) = \hat{U}^{-1}[1_{N+K} - c_2(1_{N+K} + \hat{\mu})] \tag{5.23}$$

If $\hat{g} = (\hat{\gamma}(c_1)', \hat{\gamma}(c_2)')'$ and if $\hat{S}_m$ is equal to the consistent estimate of the asymptotic variance of $\bar{m}_T$, then the GMM Wald test statistic for this approach is as follows:

$$J_3 = T\left(\hat{g}Q_2^{'}\left[\left(I_2 \otimes Q\hat{U}^{-1}\right)S_m\left(I_2 \otimes \hat{U}^{-1}Q'\right)\right]^{-1} Q_{2\hat{g}}\right) \overset{A}{\sim} \chi^2_{2N} \tag{5.24}$$

Where $c_1 = 0$ and $c_2 = 1$, in line with Kan & Zhou (2008). The authors state that this choice will not harm the generality, as J_3 is not numerically dependent of the choice of c_1 and c_2

5.4.2 Bekaert and Urias (1996)

The approach of Bekaert & Urias (1996) does not differ much from the methodology of De Santis (1993). Theoretically, Bekaert & Urias (1996) also check whether N additional (test) assets contribute to explaining the variance of m_t. However, there is a significant difference between both methods. Bekaert & Urias (1996) mathematically and theoretically define the stochastic discount factor in a different way:

$$m_t = c + (R_t - \mu)'\beta(c) + \varepsilon_t \tag{5.25}$$

Where $\beta(c)$ is:

$$\beta(c) = V^{-1}[(1-c)1_{N+K} - c\mu] \tag{5.26}$$

The corresponding hypotheses are as follows, where 0_N is an N-vector of zeros:

$$H_0\text{: (1) } Q\beta(c_1) = 0_N,$$
$$\text{(2) } Q\beta(c_2) = 0_N$$

Where, $Q = [O_{N+K}, I_N]$ and $\beta(c1)$ and $\beta(c2)$ are two different linear combinations of the weights of two frontier portfolios of $K+N$ assets (Kan & Zhou, 2008). In Bekaert & Urias (1996), the hypotheses are based on $\beta(c1)$ and $\beta(c2)$ instead of $\gamma(c1)$ and $\gamma(c2)$. For the specific differences between γ and β, see Kan & Zhou (2008). In contrast to De Santis (1993), the expected return (μ) is part of the formula of the stochastic discount factor (m_t).

The GMM Wald test statistic(s) for the approach of Bekaert & Urias (1996) are based on the following propositions and assumptions. Let's first assume that the sample moment conditions are:

$$\bar{h}_T(\beta(c_1), \beta(c_2)) = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} R_t(c_1 + (R_t - \hat{\mu})'\beta(c_1)) - (1-c_1)1_{N+K} \\ \\ \frac{1}{T}\sum_{t=1}^{T} R_t(c_2 + (R_t - \hat{\mu})'\beta(c_2)) - (1-c_2)1_{N+K} \end{bmatrix} \tag{5.27}$$

Kan & Zhou (2008) mention, based on Bekaert & Urias (1996), that the standard way is to parameterize μ and use the following sample moment conditions instead:

$$\bar{h}^*_T(\beta(c_1), \beta(c_2)) = \begin{bmatrix} \frac{1}{T}\sum_{t=1}^{T} R_t(c_1 + (R_t - \mu)'\beta(c_1)) - (1-c_1)1_{N+K} \\ \\ \frac{1}{T}\sum_{t=1}^{T} R_t(c_2 + (R_t - \mu)'\beta(c_2)) - (1-c_2)1_{N+K} \\ \\ \frac{1}{T}\sum_{t=1}^{T}(R_t - \mu) \end{bmatrix} \tag{5.28}$$

Next, the unconstrained estimates of $\beta(c_1)$ and $\beta(c_2)$ are defined as:

$$\hat{\beta}(c_1) = \hat{V}^{-1}[(1-c_1)1_{N+K} - c_1\hat{\mu}] \tag{5.29}$$

$$\hat{\beta}(c_2) = \hat{V}^{-1}[(1-c_2)1_{N+K} - c_2\hat{\mu}] \tag{5.30}$$

Subsequently, assume that $S = Avar(\bar{h}_T^*)$ with Avar = *Asymptotic Variance*. In addition, also assume that S_{11} relates to the first two blocks of sample moments of the pricing equation and S_{11} relates to the sample moments for estimating the expected return (Kan & Zhou, 2008). With these assumptions, the following definition can be written:

$$S = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \tag{5.31}$$

Assuming that $\hat{b}$ reflects the GMM estimator of $(\beta(c_1)', \beta(c_2)')'$, then the corresponding asymptotic variance (Avar) of $\hat{b}$ can be summarized as:

$$Avar(\hat{b}) = C^{-1}(S_{11} - F'S_{21} - S_{12}F + F'S_{22}F)C^{-1} \tag{5.32}$$

Note that Equation 5.32 assumes that $C = I_2 \otimes V$ and $F = [\beta(c_1)\mu', \beta(c_2)\mu']$. Finally, the GMM Wald test statistics J_1 and J_2 can be stated as:

$$J_1 = T\left(\hat{b}'Q_2'\left[Q_2\hat{C}^{-1}\left(\hat{S}_{11} - \hat{F}'\hat{S}_{21} - \hat{S}_{12}\hat{F} + \hat{F}'\hat{S}_{22}\hat{F}\right)\hat{C}^{-1}Q_2'\right]^{-1}Q_2\hat{b}\right) \overset{A}{\sim} \chi^2_{2N} \tag{5.33}$$

$$J_2 = T\left(\hat{b}'Q_2'\left(Q_2\hat{C}^{-1}\hat{S}_{11}\hat{C}^{-1}Q_2'\right)^{-1}Q_2\hat{b}\right) \overset{A}{\sim} \chi^2_{2N} \tag{5.34}$$

Where $Q_2 = I_2 \otimes Q$ with $Q = [O_{N+K}, I_N]$ and where $\hat{C}$, $\hat{F}$, $\hat{S}$ are the consistent estimators of C, F and S. According to Kan & Zhou (2008), J_1 includes the Errors-In-Variables adjustment and J_2 does not. Both versions of the SDF approach of Bekaert & Urias (1996) are considered in this paper for the purpose of checking robustness. Important to mention is that J_1 and J_2 are also numerically independent of the choice of c_1 and c_2. Therefore, also for this approach, the choice is to use $c_1 = 0$ and $c_2 = 1$ without the risk of having loss of generality (Kan & Zhou, 2008).

Chapter 6

Data

With Thomson Reuters DataStream, monthly (unhedged) price data of three types of asset classes for various countries have been acquired: stock market indices, government bond indices and real estate indices[28]. The choice for monthly data is not random. On the one hand, monthly data is preferred to annual data because of the simple fact that monthly data gives more data points and thus leads to a more accurate and less noisy estimate of, among other things, correlation coefficients[29]. On the other hand, monthly data is preferred above, for example, daily data. The reason for the latter is that previous studies have shown that monthly data, relatively speaking, is approximately more normal distributed and has a lower leptokurtosis[30] than daily data (Fama, 1976; Blattberg & Gonedes, 2010). The preference for monthly frequency data implicitly sets the investment horizon, and thus monthly returns imply that the marginal investor optimizes for an investment for 1 month. The market indices from the following countries are chosen to investigate: *Australia, Belgium, Canada, France, Germany, Japan, Netherlands, Sweden* and *United States*. The time-period of the data spans from 09/2002 to 12/2017. For all these indices, the monthly prices are obtained and denominated in U.S. dollars[31], consistent with De Roon, Nijman, & Werker (2001), as this eliminates the risk of portfolio returns being influenced by changing exchange rates (Eun & Resnick, 1988; Sukumaran, Gupta, & Jithendranathan, 2015). Although the methodologies in this paper are based on return data, the acquired data consists of price data. Therefore, prices have to be converted to returns. Transformation from prices to returns will logically result in excluding the first month of observation. Therefore, the final sample spans from 10/2002 to 12/2017. Transformation will be done on the basis of the *simple returns method*:

$$R_{i,t} = \frac{P_{i,t} - P_{i,t-1}}{P_{i,t-1}} \tag{6.1}$$

where,
$R_{i,t}$ = Return of index i at time t
$P_{i,t}$ = Price of index i at time t
$P_{i,t-1}$ = Price of index i at time $t-1$

To a large extent the results depend on the underlying data. After all, a methodology must be applied to the data. It is therefore not unimportant to explore the data on the basis of a

[28] Appendix A.1 shows the exact corresponding Datastream symbols for each of the retrieved variables.

[29] In relation to the spanning tests, Kan & Zhou (2008) state that asymptotic tests could be *grossly misleading*.

[30] Distributions are defined as leptokurtotic when the value of the kurtosis is greater than 3 (or: excess kurtosis is greater than 0), implying that the particular distribution has fatter tails than a normal distribution.

[31] With this, this research implicitly focuses on investors who focus on pure U.S. Dollars returns. However, there are also investors who value local currency returns. One paper that highlights both sides is that of Driessen & Laeven (2007)

number of descriptive statistics[32] before obtaining the results. First of all, the sample standard deviation (s_x), variance (s_x^2) and mean ($\bar{x}$) of the returns are obtained. The standard deviation and variance are both derivatives of the average and provide an insight into the background of the dataset. The standard deviation relates to the spread of the data around the mean, where the value can be either negative or positive. The variance, on the other hand, can only take a positive value because it is calculated as the square of the standard deviation. The variance is also known as the average of the squared deviations from the mean[33]

Table 6.1 shows the results for each of the indices of each asset class. One can conclude that the returns of bonds appear to be less risky than both stock and real estate returns, since the standard deviation and variance are lower for bond indices. By observing the standard deviations and variances of all three asset classes, one can also sense that the degree of risk of both stocks and real estate are more in line with each other than with bonds. The notion that risk and return seem to have a positive relationship with each other is confirmed by the fact that the two riskier assets (stocks and real estate) have higher average returns than bonds (Ball & Bowers, 1986).

Table 6.1: Mean, Standard Deviation and Variance of Dataset Variables

	Bonds			**Stocks**			**Real Estate**		
Country	$\bar{x}$	s_x	s_x^2	$\bar{x}$	s_x	s_x^2	$\bar{x}$	s_x	s_x^2
Australia	0,004	0,037	0,001	0,008	0,065	0,004	0,004	0,066	0,004
Belgium	0,004	0,033	0,001	0,006	0,064	0,004	0,004	0,048	0,002
Canada	0,003	0,027	0,001	0,008	0,059	0,004	0,006	0,059	0,003
France	0,004	0,032	0,001	0,006	0,063	0,004	0,011	0,069	0,005
Germany	0,004	0,031	0,001	0,010	0,070	0,005	0,008	0,094	0,009
Japan	0,002	0,031	0,001	0,006	0,049	0,002	0,009	0,068	0,005
Netherlands	0,004	0,031	0,001	0,005	0,065	0,004	0,004	0,068	0,005
Sweden	0,003	0,033	0,001	0,009	0,066	0,004	0,015	0,083	0,007
United States	0,001	0,022	0,000	0,007	0,045	0,002	0,007	0,075	0,006

This table reports the mean, standard deviation and variance of each index, based on the full-sample period 10/2002 to 12/2017, where the values are expressed in decimals and not percentages.

After assessing the background of the data, it is equally important to judge if the return distributions of the variables are normal. Describing distributions can be done on the basis of two parameters: *Skewness* and *Kurtosis*. The former measures the degree of symmetry. If the skewness is equal to 0, then the distribution is assumed to be symmetrical (characteristic of normality), whereas a skewness higher (lower) than 0 implies a positively (negatively) skewed distribution. Kurtosis on the other hand, measures the degree of fat tails. The distribution is assumed to exhibit no fat tails (characteristic of normality) when the kurtosis equals 3, where a value higher (lower) than 3 refers to thicker (thinner) tails. To test the normality of a distribution, for each variable, based on these two parameters, the Jarque-Bera chi-square

[32]Appendix A.2 provides the formulas of the (descriptive) statistics in Table 6.1 and Table 6.2.

[33]Because there is no unambiguous criterion for measuring risk, the two most common measures (variance and standard deviation) are used.

goodness-of-fit test will be used. The results of the three distribution describing statistics are shown in Table 6.2. Most of the statistics are consistent with the notion that, in general, financial instruments are (negatively) skewed[34], fat-tailed and fail to pass the Jarque-Bera test on all common levels of significance (1%, 5% and 10%). In contrast to the majority of the statistics, the bond indices Sweden and United States and the real estate indices Germany and Sweden have a positive skewness. In addition, it is striking that the the bond indices Japan and Sweden take a form of normal distribution, since the Jarque-Bera's null hypothesis of normality can not be rejected for these two bond indices.

Table 6.2: Skewness, Kurtosis and Jarque-Bera Test p-value of Dataset Variables

	Bonds			**Stocks**			**Real Estate**		
Country	$s_{kewness}$	$k_{urtosis}$	$JB_{p-value}$	$s_{kewness}$	$k_{urtosis}$	$JB_{p-value}$	$s_{kewness}$	$k_{urtosis}$	$JB_{p-value}$
Australia	-0,386	4,671	0,000***	-0,563	5,171	0,000***	-1,244	8,725	0,000***
Belgium	-0,239	4,588	0,000***	-0,657	5,361	0,000***	-0,673	5,903	0,000***
Canada	-0,378	4,463	0,000***	-0,490	6,354	0,000***	-1,186	9,616	0,000***
France	-0,077	4,437	0,000***	-0,383	4,042	0,002***	-0,474	4,361	0,000***
Germany	-0,010	4,589	0,000***	-0,294	4,333	0,000***	1,355	16,741	0,000***
Japan	-0,040	3,241	0,782	-0,562	4,551	0,000***	-0,095	4,365	0,001***
Netherlands	-0,192	4,520	0,000***	-0,631	4,997	0,000***	-0,456	4,090	0,000***
Sweden	0,137	3,463	0,332	-0,051	4,616	0,000***	0,603	7,770	0,000***
United States	0,542	6,063	0,000***	-0,758	5,786	0,000***	-0,127	13,511	0,000***

This table reports the skewness, kurtosis and p-value of the JB (Jarque-Bera) chi-square test of each index, based on the full-sample period 10/2002 to 12/2017, where the values are expressed in decimals and not percentages. Ideally, a variable exhibits the form of a normal distribution if the values of skewness and kurtosis are respectively, approximately, close to 0 and 3. The Jarque-Bera test of goodness-of-fit tests the null hypothesis that the skewness and kurtosis matches a normal distribution. Significance levels: *** $p<0.01$, ** $p<0.05$, * $p<0.10$.

Finally, it is useful to make some notes regarding the relationship between the data and methodology section. Firstly, in this paper, the diversification benefits will be analyzed from an U.S. perspective. Therefore, the U.S. bonds, stocks and real estate indices will serve as benchmark portfolios for bonds, stocks and real estate. Another note is based on the specification of the sub-periods (sub-samples), as the methodologies are based on three sub-periods. The choice was made to look at the sub-periods: pre-crisis, crisis and post-crisis. This will provide more insight into the comparison of whether there are differences or similarities in the periods before, during and after the crisis. Thus, this research offers the prospect of saying more than just what happens over time. Specifically, the next three sub-periods will be investigated, each containing exactly 5 years[35]:

- October 2002 – October 2007 (Pre-crisis)
- November 2007 – November 2012 (Crisis)
- December 2012– December 2017 (Post-crisis)

[34]The general consensus is that the marginal investor prefers positive skewness.

[35]Although it was part of reality, it was decided to avoid the internet bubble (1997-2000) and subsequent after-effects (2001). It could have severely biased the results, while the likelihood of a similar event is extremely small.

Part III

Results

Chapter 7

Correlation Analysis

7.1 Testing for Stability

Testing the stability of correlations over time seems rather descriptive in nature. Nevertheless, this chapter starts with this because corresponding results can create expectations about the upcoming sections. If correlations over time are, for example, unstable, one would expect that over time: (1) fluctuating patterns will occur, (2) risk may fluctuate and (3) diversification benefits will fluctuate. Table 7.1 on below shows the results of the Jennrich-test.

Table 7.1: Jennrich-Test of Equality of Correlation-Matrices

Subsamples	Jennrich χ^2	p-Value
Bonds		
10/2002 - 10/2007 vs. 11/2007 - 11/2012	99,92	0,00***
11/2007 - 11/2012 vs. 12/2012 - 12/2017	76,03	0,00***
10/2002 - 10/2007 vs. 12/2012 - 12/2017	167,30	0,00***
Stocks		
10/2002 - 10/2007 vs. 11/2007 - 11/2012	29,79	0,76
11/2007 - 11/2012 vs. 12/2012 - 12/2017	14,21	1,00
10/2002 - 10/2007 vs. 12/2012 - 12/2017	23,71	0,94
Real Estate		
10/2002 - 10/2007 vs. 11/2007 - 11/2012	18,13	0,99
11/2007 - 11/2012 vs. 12/2012 - 12/2017	31,65	0,68
10/2002 - 10/2007 vs. 12/2012 - 12/2017	17,40	1,00

This table reports the Jennrich χ^2 statistics with its corresponding p-values. Comparisons of matrices is done within each asset class and between time periods. The considered time periods are: pre-crisis (10/2002 - 10/2007), crisis (11/2007 - 11/2012) and post crisis (12/2012 - 12/2017).
Significance levels: *** p<0.01, ** p<0.05, * p<0.10.

Table 7.1 shows that correlation-matrix of bonds are unstable over time, because all three the p-values are significant on all levels and thus can not reject H_{01}, for each sub-period comparison, that the correlation-matrices are unstable over time. Instability may be the result of country-differences in the dynamics of risk characteristics of bonds, that in turn may also be driven by country-differences in monetary policy. According to Fabozzi (2007), these risk characteristics include, among others, *interest rate risk*, *credit/sovereign risk*, *liquidity risk*, *exchange rate risk* and *inflation risk* and influence the price of bonds. As the monetary policy and degree of exposure to sovereign debt risks are subject to change over time and differ between countries, it is not entirely illogical that the correlation-matrix of bonds is unstable over time.

Both stocks and real estate on the other hand, can reject H_{01} and therefore show opposite results, as all p-values of both asset classes are insignificant for all levels. This is in conflict with the consensus of literature that correlations of these two asset classes over time should be unstable. Nevertheless, the results are not uncommon, as Kaplanis (1988) also wants to add the note that correlation-matrices are quite stable compared to covariance matrices. In addition, Tang (1998) finds that the correlation-matrix of stock returns is more stable intertemporally than the corresponding covariance-matrix, considering four different investment horizons. Based on real estate returns, Eichholtz (1996b) also concludes that correlations appear to be more stable than the variances. He states that correlation is, in fact, a reflection of market integration. According to Eichholtz (1996b), markets are gradually integrating over time due to the increasing capital flows. This should explain why correlations do not immediately change and do not show a high degree of instability. Adams, Füss, & Glück (2017) provide a very noteworthy and recent article regarding the dynamics of correlations. Their study is based on the DCC model (Engle, 2002) and focuses on the structure of correlation dynamics among financial assets. The authors acknowledge the main findings of previous literature, that correlations can vary over time. However, they provide evidence that the results from the dynamic conditional correlation estimates show no significant variation when controlling for correlation breaks. According to them, the correlations are time-varying in response to (correlation) breaks and that these breaks are a consequence of financial and economic shocks. In the same paper, it is stated that these breaks can lead to distorted correlation estimates. This should in turn result in a higher portfolio variance. However, they advise that the marginal investor should optimize its portfolio (weights) frequently (e.g. with rolling window) in order to overcome this problem.

To conclude and summarize this section, the p-values stated in Table 7.1 implicate that hypothesis H_{01} can only be rejected for the stock and real estate indices.

7.2 Development of Correlations Over Time

The previous section can only provide an answer if correlations are significantly (un)stable over time. In contrast, this section aims to expose the pattern of correlation coefficients (1) over time and (2) per sub-period (pre-crisis, crisis and post-crisis). Appendix B, C and D show, respectively, the charts for the Rolling-method, EWMA and DCC-GARCH method of measuring correlations per asset class and index. Figure 7.1 presents the main results. Figure 7.1a, 7.1b and 7.1c represent, respectively, the average correlations of bond, stock and real estate indices per method (Rolling, EWMA and DCC-GARCH(1,1,1,1).

Firstly, note that the rolling-method shows strongly fluctuating correlations for all three indices. At first glance these seem to be natural shocks, but this is largely because the rolling window-size is shifting each point in time. As the window-size 'rolls', each point in time includes a new observation. These new observations immediately get the same weight as all other observations. Therefore, the rolling-method proves to be inaccurate due to the possibility of having volatile estimates of correlations. The intuition is that past shocks still have a significant influence on the correlation coefficients until the window-size has shifted to such an extent that the particular observations fall outside the estimation-window (Isogai, 2016). The EWMA seems to minimize this problem because the EWMA is biased to more recent observations. The improvement entailed by the EWMA is not illogical if account is taken of the fact that the EWMA theoretically takes the form of a restricted I-GARCH model. However, even the EWMA seems to have a problem to fully take into account the short-term volatility shocks. The DCC(1,1)-GARCH(1,1) allows for mean-reversion and seems to be the most accurate in demonstrating significant (short-term) shocks. This is confirmed by Akgiray (1989). Based on daily data, the relevant author compares the degree of forecast accuracy between the benchmark (simple historical average), EWMA, ARCH and GARCH with the actual data. Akgiray (1989) concludes that the GARCH(1,1) seems to provide the best fit and forecast accuracy.

Next, excluding the rolling-method due to its chance of providing inaccurate estimates, one can observe interesting results. Number one is that stock and real estate are more exposed to the volatility-shock during the recession. The strong exposure leads to a strong short-term increase in correlations. These findings correspond with, among others, Erb et al. (1994); Longin & Solnik (1995) and could be the consequence of *an endogenous increase in the uncertainty about the global economy* (Model, Ribeiro, & Veronesi, 2002). However, the short-term substantial rise in correlations stabilize at the end of the crisis. Furthermore, Figure 7.1 indicates that stock and real estate indices exhibit a long-term trend of stability, consistent with section 7.1.

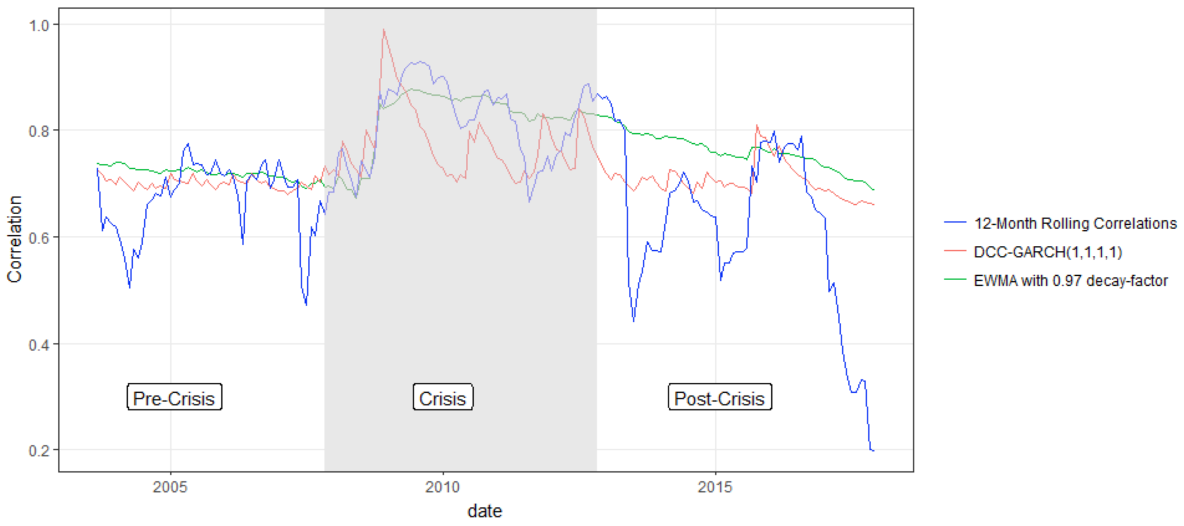
In contrast, bond correlations reflect the notion that bonds are considered as safe assets and are therefore not heavily impacted by negative investor sentiment or collective uncertainty during recessions, as Hartmann, Straetmans, & de Vries (2001) document that the probability that equity markets crash simultaneously is two times larger when compared to probability for bond markets. It also seems that bond indices are less stable over time due to small upward shocks over time. The pattern of correlation instability between bonds is in line with section 7.1. In contrast to risky assets, correlations of bond indices do not exhibit a (substantial) increase during the recession but rather a short-term decreasing trend with small upward shocks. The downward trend in correlations seems contradictory, but that is not the case. During recessions, some countries are more exposed to sovereign debt risks than other countries. This in turn leads to short-term shifting in demand to government bonds with relatively less exposure to sovereign debt risk. Specifically, the process is as follows: investors reallocate their investments to bonds with relatively lower sovereign debt risk with the result that the aggregate demand for these bonds increases, which in turn increase (decrease) prices (yields). The opposite is the case for countries with a relatively higher sovereign debt risk. And therefore, prices and returns across government bonds are less correlated during recessions.

Even though volatility shocks may have its impact on stocks and real estate, and in a lesser degree on bonds, Figure 7.1 shows that the correlations among all three assets do not increase substantially over time. Thus, hypothesis H_{02} can be rejected for the bond, stock and real estate indices.

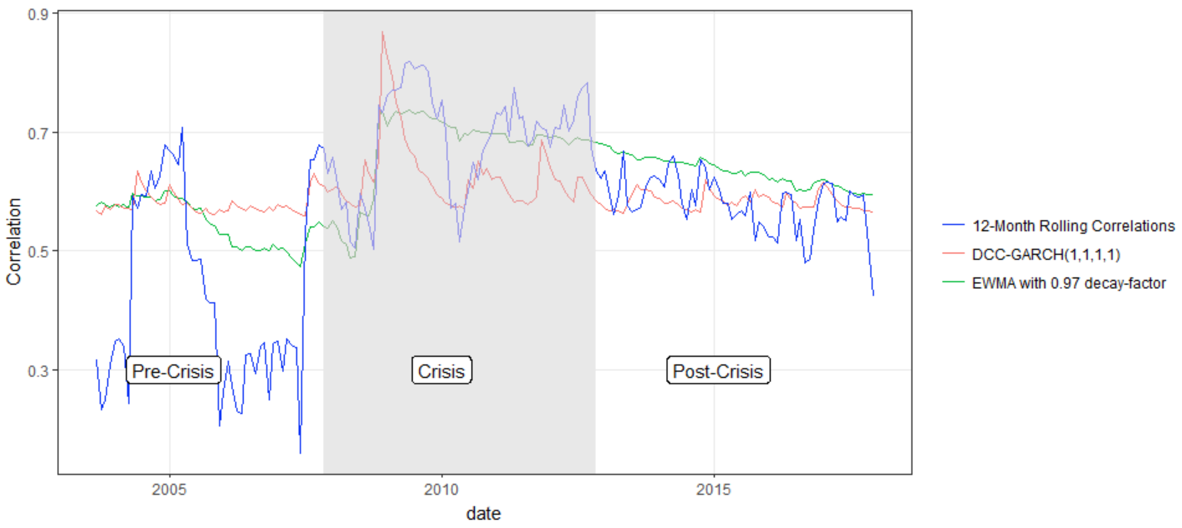
Figure 7.1: Average Correlation Coefficients of Rolling-method, EWMA and DCC-GARCH Over Time

(a) Bonds

(b) Stocks



(c) Real Estate



Chapter 8

Dynamic Efficient-Frontiers

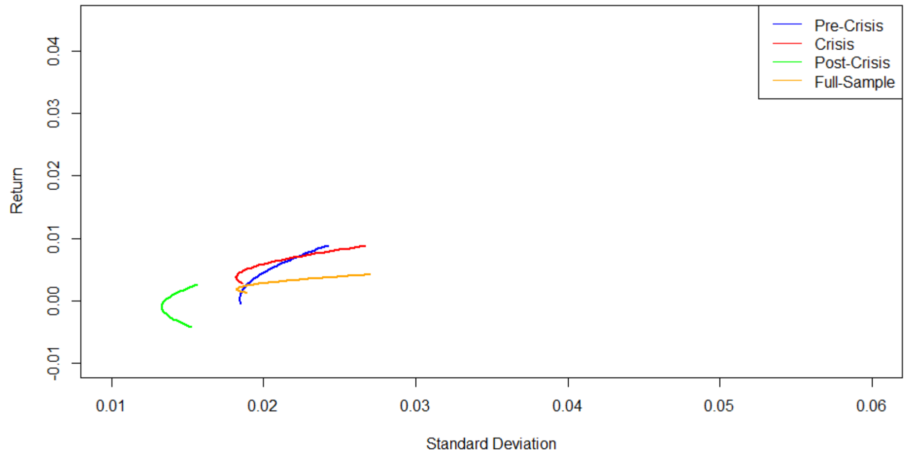
Efficient frontiers are based on without and with the short-selling restriction and can be found in, respectively, Figure 8.1 and 8.2. Both figures include efficient frontiers for the full-sample and sub-periods. The figures identify that, over time, the efficient frontiers of stocks and real estate have shifted to right. This can be observed by comparing the pre-crisis and post-crisis frontiers. The movement of the frontiers of these two asset classes are consistent with the findings in section 7.2 and the notion that increased correlation coefficients impact the risk-and-return trade-off negatively over time, especially when there are short-selling restrictions. Where these effects are clearly visible for stocks and real estate, this is to a lesser extent for bonds. The small movements of the bonds frontier are probably more reflected by changes in macroeconomic policy and confidence in the economy of the bond's country, in line with the findings regarding bond's instability (section 7.1) and small shocks over time (section 7.2).

Both figures also show a interesting pattern of the efficient frontier at the time of a recession, for at least stocks and real estate. The crisis-period brings two effects for these two assets: the first effect is that the efficient frontiers of these two asset classes shift right, implying a worsening in the risk-and-return trade-off due to a strong increase in the correlation coefficients when there is a period of high volatility (Longin & Solnik, 1995). The second effect is that the movement of these efficient frontiers are rather short-term than long-term, as the post-crisis frontier shifts strongly to the left. The post-crisis frontier shifts even far more left than the pre-crisis frontier when there are short-selling restrictions, referring to a strong stabilization of correlation coefficients and probably a slightly excessive investor's confidence for the foreseeable future. The stabilization of the efficient frontiers of stocks and real estate is in line with that correlation coefficients are higher in bear markets than in bull markets, implying stabilization (Longin & Solnik, 2001; Ang & Bekaert, 2002; Hyde et al., 2007). Bonds, on the other hand, do not show the identical pattern. As can be seen, the crisis affects its efficient frontier to a lesser degree. This pattern is in line with section 7.2, which presents evidence of a decreasing trend during crisis and small upward shocks over time of correlations between bonds. The consequences for diversification when short selling is unlimited or restricted, is observable by comparing figure Figure 8.1 and 8.2. The comparison shows that the risk, measured in standard deviation, is higher when a short-selling restriction is applied. Therefore, this finding is no different from what has been accepted within the science of asset-pricing. As it is clear in Levy & Post (2005) and Lee & Lee (2010), short-selling restrictions restrict diversification possibilities and therefore making it more difficult to achieve lower risk.

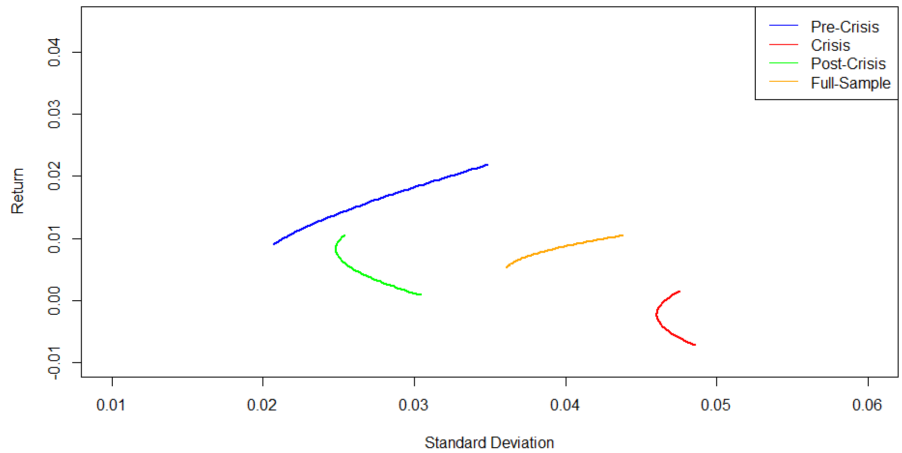
Summarizing, Hypothesis H_{03} can only be rejected for bonds. It is therefore important to be cautious about the fact that, over time, higher correlations over time and/or short-selling restrictions lead to higher risk for the same level of return. At least for stocks and real estate indices.

Figure 8.1: Efficient-Frontiers Without Short-Selling Restrictions

(a) Bonds



(b) Stocks



(c) Real Estate

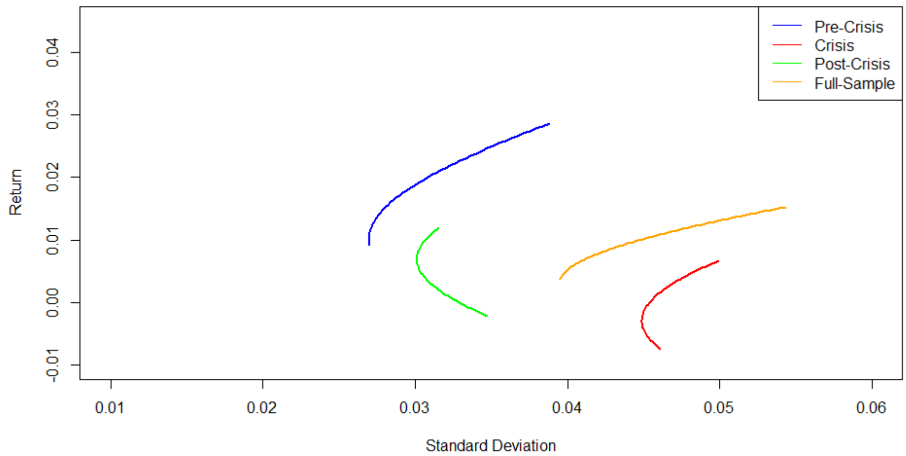
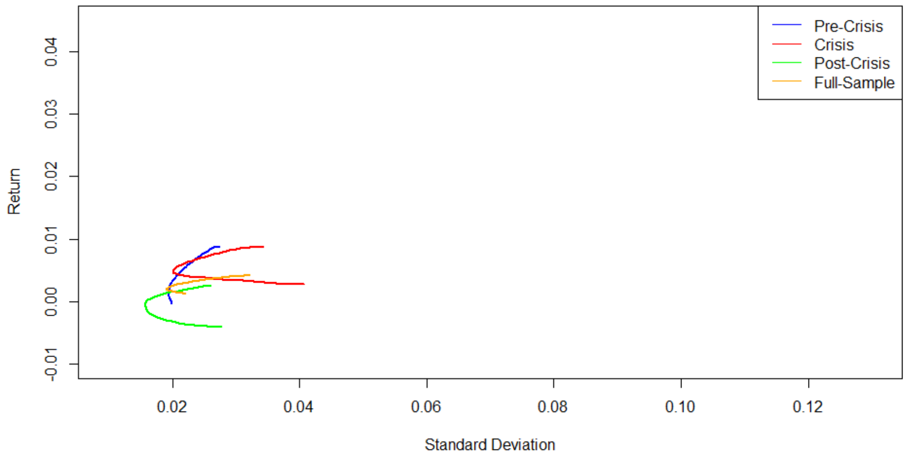
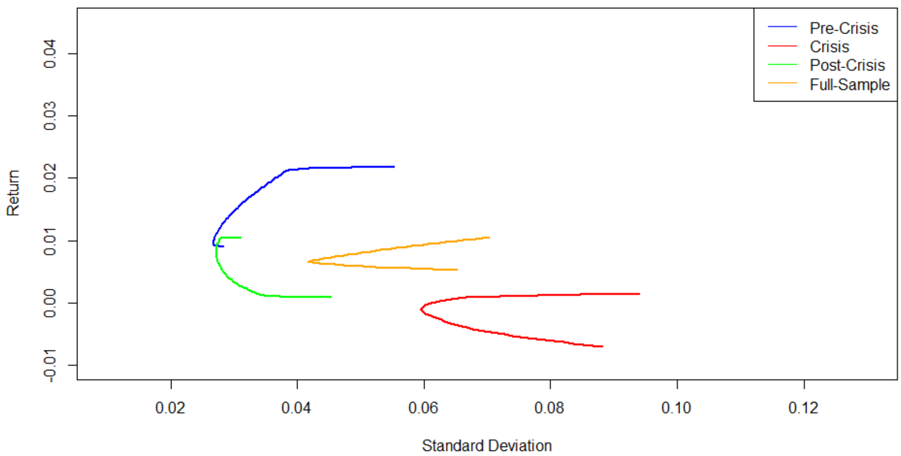


Figure 8.2: Efficient-Frontiers With Short-Selling Restrictions

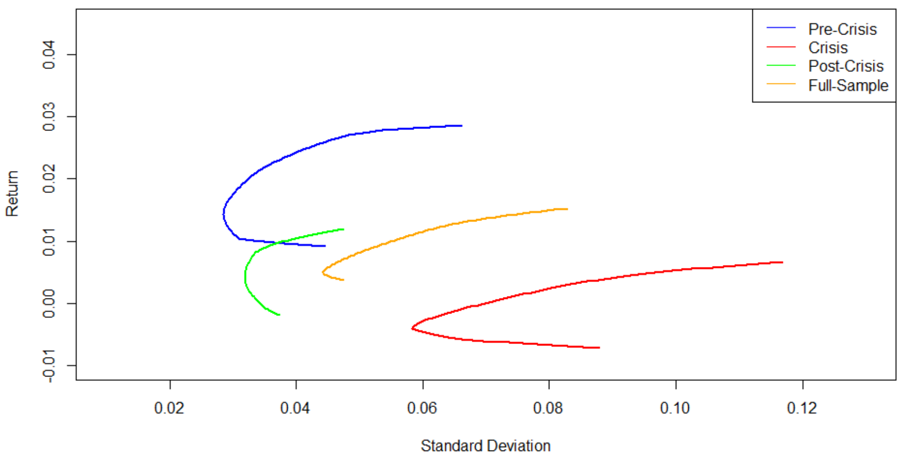
(a) Bonds



(b) Stocks



(c) Real Estate



Chapter 9

Tests of Mean-Variance Spanning

The previous section elaborated on the risk-and-return trade-off over time. However, visual observations do not imply any significance of efficient frontier movements. Therefore, it is necessary to test the significance of these movements. Testing the significance with the help of various mean-variance spanning tests can yield opportunities to make conclusions regarding the diversification benefits over time. In order to stay in line with the methodology of the previous sections, all considered spanning tests were performed on both the full sample and the three sub-periods (pre-crisis, crisis and post-crisis).

Before the results will be brought to light, it is important to repeat the difference between rejection and non-rejection. As stated earlier, the spanning tests have the null-hypothesis that a set of K benchmark assets 'spans' a larger set of $K+N$ assets. If it does, then it is assumed that the Minimum-Variance Frontier of K benchmark assets is statistically equal to that of $K+N$ assets. If the null-hypothesis is rejected however, then adding N (test) assets to one or more K (benchmark) assets will statistically improve the Minimum-Variance Frontier and thus add value in terms of diversification.

This chapter is divided into three sections. In section 9.1, the results of the mean-variance spanning tests under the normality assumption will be reported. Subsequently, in section 9.2, the asymptotic versions of the mean-variance spanning tests will be treated. Finally, section 9.3 will discuss the results stated in section 9.1 and 9.2 and answer hypothesis H_{04}.

9.1 Spanning Tests Under Normality

Table 9.1, 9.2 and 9.3 are related to the spanning tests under normality and represent, respectively, bonds, stocks and real estate. Important to note is that the first four columns of the tables present the corrected traditional F-test, whereas the last five columns show the step-down procedure. If there are any test assets that can not reject the traditional F-test, then the step-down procedure can help to obtain insight in the reason. If the F-test can not reject due to the F_1, then the reason is that the tangency portfolio of the test assets is not significantly different from the benchmark asset's tangency portfolio. No rejection due to F_2 implies that the global minimum-variance portfolio of the test assets is not significantly different from the benchmark asset's global minimum-variance portfolio. It is possible that non-rejection comes from the fact that both F_1 and F_2 cannot be rejected. This implies then that both the tangency and global minimum-variance portfolio of the test assets can be improved. Important to mention is the possibility that the joint- F-test can be rejected while the traditional F-test can not. This can happen because, in general, joint-tests exhibit higher significance than individual tests.

According to Table 9.1 (bonds), there are a couple of test assets that can not be rejected on any common significance level for the traditional F-test. For the pre-crisis period, the non-rejected test assets are Japan and Sweden. Japan can not be rejected due to the F_1, whereas Sweden can not be rejected due to F_2. The post-crisis period also has some non-rejections. Australia and, again, Japan. Both test assets can not be rejected because of both F_1 and F_2.

Referring to Table 9.2 (stocks), there are also some test assets that can not be rejected on the basis of the traditional F-test. For the full-sample period, these are Australia (due to F_1) and Canada (due both F_1 and F_2). The crisis period shows one non-rejection and that is Canada, which is the result of not rejecting both F_1 and F_2. Finally, Belgium, France, Germany and Netherlands are not rejected for the post-crisis period. Where Belgium can not be rejected because of F_1, the fault of not rejecting for France, Germany and Netherlands lies at both F_1 and F_2. For Table 9.3 (real estate), only Germany is rejected for the full-sample and crisis period due to failing to reject both F_1 and F_2.

Finally, on the basis of the F-tests, it can be noted that, over time, a large part of the test assets remains significant. This applies to all three asset classes (bonds, stocks and real estate). This means that diversification benefits have not decreased over time. It is striking that this also holds for the crisis period. This points to the fact that diversification benefits are not strongly negatively affected by upward-shocks in correlations during the crisis. This is contradictory with the findings of chapter 8, which showed that the risk-and-return trade-off worsens during a period of high volatility for at least stocks and real estate.

However, there are two possible causes for this. A reason for this could be that the spanning tests are based on unlimited short selling and that, therefore, high correlations do not have to pose a problem. Short-selling actually offers an advantage when assets are highly correlated with each other. Another reason may be that the spanning tests have to be corrected for non-normality. Not entirely illogical if account is taken of the fact that in chapter 6 it has emerged that this dataset is subject to non-normality. In the next section, emphasis will be placed on spanning tests that take non-normality into account in order to check for robustness of these results.

Table 9.1: Mean-Variance Spanning Tests Under Normality - Bond Indices

					Step-Down Test				
Country (test assets)	α	δ	F-test	p-value	F_1	p-value	F_2	p-value	Joint p-value
Full-Sample (10/2002 - 12/2017)									
Australia	0,003	0,414	6,724	0,002***	1,282	0,259	12,147	0,001***	0,000***
Belgium	0,003	0,452	10,334	0,000***	2,275	0,133	18,266	0,000***	0,000***
Canada	0,003	0,689	31,224	0,000***	2,120	0,147	59,959	0,000***	0,000***
France	0,003	0,414	9,609	0,000***	2,522	0,114	16,556	0,000***	0,000***
Germany	0,003	0,393	9,685	0,000***	2,487	0,117	16,745	0,000***	0,000***
Japan	0,001	0,212	2,924	0,056*	0,243	0,623	5,629	0,019**	0,012**
Netherlands	0,003	0,394	9,404	0,000***	2,491	0,116	16,183	0,000***	0,000***
Sweden	0,003	0,516	12,372	0,000***	1,208	0,273	23,509	0,000***	0,000***
All countries			4,689	0,000***	0,460	0,883	9,779	0,000***	0,000***
Pre-crisis (10/2002 - 10/2007)									
Australia	0,008	0,106	2,697	0,076*	4,967	0,030**	0,400	0,530	0,016**
Belgium	0,008	0,075	2,969	0,059*	5,651	0,021**	0,266	0,608	0,013**
Canada	0,009	0,343	7,283	0,001***	9,002	0,004***	4,909	0,031**	0,000***
France	0,008	0,070	3,005	0,057*	5,762	0,200	0,230	0,633	0,012**
Germany	0,007	0,068	2,823	0,067*	5,408	0,024**	0,223	0,639	0,015**
Japan	0,001	0,306	2,200	0,120	0,092	0,763	4,374	0,041**	0,031**
Netherlands	0,008	0,069	3,101	0,052*	5,953	0,018**	0,230	0,633	0,011**
Sweden	0,008	0,071	2,295	0,110	4,410	0,040**	0,171	0,681	0,027**
All countries			1,773	0,043**	1,935	0,074*	1,645	0,134	0,010**
Crisis (11/2007 - 11/2012)									
Australia	0,006	0,723	5,800	0,005***	1,144	0,289	10,431	0,002***	0,001***
Belgium	0,001	0,641	5,718	0,005***	0,037	0,848	11,585	0,001***	0,001***
Canada	0,004	0,996	22,969	0,000***	1,128	0,293	44,715	0,000***	0,000***
France	0,001	0,561	5,189	0,008***	0,025	0,875	10,524	0,002***	0,002***
Germany	0,001	0,499	4,534	0,015**	0,077	0,783	9,131	0,004***	0,003***
Japan	0,005	0,274	2,639	0,080*	2,054	0,157	3,168	0,080*	0,013**
Netherlands	0,001	0,509	4,559	0,014**	0,043	0,837	9,223	0,004***	0,003***
Sweden	0,003	0,784	9,714	0,000***	0,288	0,593	19,370	0,000***	0,000***
All Countries			3,440	0,000***	0,494	0,855	7,775	0,000***	0,000***
Post-crisis (12/2012 - 12/2017)									
Australia	-0,004	0,139	0,831	0,441	1,205	0,277	0,456	0,502	0,139
Belgium	0,003	0,469	3,922	0,025**	0,837	0,364	7,025	0,001***	0,004***
Canada	-0,003	0,437	5,537	0,006***	1,738	0,193	9,223	0,004***	0,001***
France	0,003	0,483	3,976	0,024**	0,809	0,372	7,166	0,010**	0,004***
Germany	0,002	0,552	6,102	0,004***	0,367	0,547	11,963	0,001***	0,001***
Japan	-0,003	0,000	0,379	0,686	0,758	0,388	0,001	0,980	0,380
Netherlands	0,002	0,523	5,092	0,009***	0,425	0,517	9,854	0,003***	0,001***
Sweden	-0,001	0,474	3,904	0,026**	0,250	0,619	7,654	0,008***	0,005***
All Countries			2,745	0,001***	1,246	0,292	4,621	0,000***	0,000***

The table presents the mean-variance spanning tests, that hold under normality, on eight bond indices, using the U.S. bond index as benchmark asset. The first test is an F-test of $H_0 : \alpha = 0_N$ and $\delta = 0_N$. The second test is a step down test where F_1 is an F-test of $\alpha = 0_N$, and F_2 is an F-test of $\delta = 0_N$ conditional on $\alpha = 0_N$. The two tests are performed on each bond index as well as jointly on all bond indices. The reported p-values are exact under the normality assumption on the residuals. The results are presented for the three sub-periods (pre-crisis, crisis and post-crisis). Significance levels: *** p<0.01, ** p<0.05, * p<0.10.

Table 9.2: Mean-Variance Spanning Tests Under Normality - Stock Indices

Country (test assets)	α	δ	F-test	p-value	Step-Down Test F_1	p-value	F_2	p-value	Joint p-value
Full-Sample (10/2002 - 12/2017)									
Australia	0,000	-0,137	2,155	0,119	0,003	0,958	4,330	0,039**	0,037**
Belgium	-0,002	-0,158	3,270	0,040**	0,335	0,564	6,227	0,013**	0,008***
Canada	0,001	-0,061	0,631	0,533	0,090	0,764	1,178	0,279	0,213
France	-0,003	-0,205	6,953	0,001**	1,018	0,314	12,886	0,000***	0,000***
Germany	0,001	-0,344	16,776	0,000***	0,232	0,630	33,461	0,000***	0,000***
Japan	0,002	0,343	13,765	0,000***	0,365	0,547	27,260	0,000***	0,000***
Netherlands	-0,003	-0,258	11,059	0,000***	1,708	0,193	20,331	0,000***	0,000***
Sweden	0,000	-0,243	8,379	0,000***	0,028	0,868	16,821	0,000***	0,000***
All countries			5,883	0,000***	1,007	0,432	11,871	0,000***	0,000***
Pre-crisis (10/2002 - 10/2007)									
Australia	0,013	0,067	4,886	0,011**	9,528	0,003***	0,214	0,646	0,002***
Belgium	0,008	-0,188	2,704	0,075*	2,488	0,120	2,850	0,097*	0,012**
Canada	0,012	-0,031	5,159	0,009***	8,889	0,004***	1,263	0,266	0,001***
France	0,005	-0,326	6,147	0,004***	1,644	0,205	10,537	0,002***	0,000***
Germany	0,004	-0,859	21,660	0,000***	0,959	0,331	42,390	0,000***	0,000***
Japan	0,006	0,525	3,781	0,029**	1,271	0,264	6,263	0,015**	0,004***
Netherlands	0,000	-0,513	7,632	0,001***	0,014	0,906	15,505	0,000***	0,000***
Sweden	0,008	-0,539	10,329	0,000***	3,015	0,088*	17,069	0,000***	0,000***
All countries			4,636	0,000***	1,847	0,089*	8,517	0,000***	0,000***
Crisis (11/2007 - 11/2012)									
Australia	0,000	-0,265	4,427	0,016**	0,004	0,950	8,999	0,004***	0,004***
Belgium	-0,008	-0,232	5,080	0,009***	2,367	0,129	7,620	0,008***	0,001***
Canada	0,000	-0,133	1,181	0,314	0,005	0,946	2,398	0,127	0,120
France	-0,007	-0,247	6,111	0,004***	1,919	0,171	10,148	0,002***	0,000***
Germany	0,001	-0,321	8,508	0,001***	0,011	0,915	17,289	0,000***	0,000***
Japan	-0,003	0,277	5,698	0,005***	0,299	0,587	11,228	0,001***	0,001***
Netherlands	-0,006	-0,276	7,243	0,002***	1,627	0,207	12,725	0,001***	0,000***
Sweden	0,000	-0,288	6,369	0,003***	0,004	0,953	12,950	0,001***	0,001***
All Countries			3,119	0,000***	0,691	0,698	6,490	0,000***	0,000***
Post-crisis (12/2012 - 12/2017)									
Australia	-0,007	0,229	2,937	0,061*	1,943	0,169	3,871	0,054*	0,009***
Belgium	0,000	0,237	2,030	0,140	0,012	0,912	4,115	0,047**	0,043**
Canada	-0,007	0,215	5,409	0,007***	3,638	0,061*	6,877	0,011**	0,001***
France	-0,003	0,111	0,894	0,414	0,493	0,485	1,307	0,257	0,125
Germany	-0,002	-0,017	0,09	0,914	0,179	0,674	0,000	0,988	0,665
Japan	0,005	0,522	7,101	0,002***	1,408	0,240	12,707	0,001***	0,000***
Netherlands	-0,003	0,053	0,597	0,553	0,625	0,432	0,573	0,452	0,195
Sweden	-0,005	0,183	3,197	0,048**	1,830	0,181	4,501	0,038**	0,007***
All Countries			2,358	0,004***	1,028	0,428	3,998	0,001***	0,000***

The table presents the mean-variance spanning tests, that hold under normality, on eight stock indices, using the U.S. stock index as benchmark asset. The first test is an F-test of $H_0 : \alpha = 0_N$ and $\delta = 0_N$. The second test is a step down test where F_1 is an F-test of $\alpha = 0_N$, and F_2 is an F-test of $\delta = 0_N$ conditional on $\alpha = 0_N$. The two tests are performed on each stock index as well as jointly on all stock indices. The reported p-values are exact under the normality assumption on the residuals. The results are presented for the three sub-periods (pre-crisis, crisis and post-crisis). Significance levels: *** p<0.01, ** p<0.05, * p<0.10.

Table 9.3: Mean-Variance Spanning Tests Under Normality - Real Estate indices

					Step-Down Test				
Country (test assets)	α	δ	F-test	p-value	F_1	p-value	F_2	p-value	Joint p-value
Full-Sample (10/2002 - 12/2017)									
Australia	0,000	0,429	37,475	0,000***	0,012	0,913	75,347	0,000***	0,000***
Belgium	0,001	0,631	137,219	0,000***	0,230	0,632	275,373	0,000***	0,000***
Canada	0,002	0,392	56,334	0,000***	0,547	0,461	112,402	0,000***	0,000***
France	0,007	0,371	28,866	0,000***	3,194	0,076*	53,889	0,000***	0,000***
Germany	0,002	0,072	0,740	0,479	0,196	0,659	1,290	0,258	0,170
Japan	0,006	0,590	48,885	0,000***	1,699	0,194	95,703	0,000***	0,000***
Netherlands	0,000	0,425	34,405	0,000***	0,002	0,966	69,187	0,000***	0,000***
Sweden	0,010	0,219	9,228	0,000***	5,476	0,020**	12,667	0,000***	0,000***
All countries			21,773	0,000***	2,266	0,025**	57,454	0,000***	0,000***
Pre-crisis (10/2002 - 10/2007)									
Australia	0,014	0,637	26,097	0,000***	11,652	0,001***	34,430	0,000***	0,000***
Belgium	0,009	0,745	32,485	0,000***	4,355	0,041**	57,404	0,000***	0,000***
Canada	0,012	0,402	14,992	0,000***	11,000	0,002***	16,272	0,000***	0,000***
France	0,021	0,308	8,509	0,001***	13,908	0,000***	2,559	0,115	0,000***
Germany	0,012	0,440	4,428	0,016**	2,582	0,113	6,113	0,016**	0,002***
Japan	0,019	0,611	5,947	0,004***	4,597	0,036**	6,884	0,011**	0,000***
Netherlands	0,014	0,494	10,653	0,000***	6,759	0,012**	13,272	0,001***	0,000***
Sweden	0,020	0,117	4,259	0,019**	8,492	0,005***	0,024	0,878	0,004***
All countries			6,409	0,000***	3,388	0,003***	10,553	0,000***	0,000***
Crisis (11/2007 - 11/2012)									
Australia	-0,010	0,423	14,158	0,000***	1,141	0,290	27,110	0,000***	0,000***
Belgium	-0,007	0,636	71,684	0,000***	1,164	0,285	141,818	0,000***	0,000***
Canada	0,000	0,377	25,701	0,000***	0,002	0,969	52,271	0,000***	0,000***
France	-0,002	0,386	16,422	0,000***	0,057	0,812	33,310	0,000***	0,000***
Germany	-0,009	-0,030	0,406	0,668	0,724	0,398	0,088	0,768	0,306
Japan	-0,004	0,576	35,267	0,000***	0,205	0,653	71,273	0,000***	0,000***
Netherlands	-0,009	0,432	17,700	0,000***	1,192	0,279	34,099	0,000***	0,000***
Sweden	0,003	0,207	3,224	0,047**	0,109	0,743	6,435	0,014**	0,010**
All Countries			11,771	0,000***	1,282	0,273	37,099	0,000***	0,000***
Post-crisis (12/2012 - 12/2017)									
Australia	-0,002	0,270	4,274	0,018**	0,228	0,635	8,429	0,005***	0,003***
Belgium	0,002	0,484	11,520	0,000***	0,270	0,605	23,049	0,000***	0,000***
Canada	-0,005	0,513	14,766	0,000***	1,388	0,244	27,962	0,000***	0,000***
France	0,000	0,357	3,796	0,028**	0,003	0,955	7,716	0,007***	0,007***
Germany	0,009	0,438	6,657	0,002***	2,551	0,116	10,492	0,002***	0,000***
Japan	0,003	0,703	10,869	0,000***	0,184	0,670	21,852	0,000***	0,000***
Netherlands	-0,004	0,326	3,365	0,041**	0,417	0,521	6,375	0,014**	0,007***
Sweden	0,007	0,449	5,839	0,005***	1,647	0,204	9,924	0,003***	0,001***
All Countries			3,842	0,000***	2,012	0,063*	6,183	0,000***	0,000***

The table presents the mean-variance spanning tests, that hold under normality, on eight real estate indices, using the U.S. real estate index as benchmark asset. The first test is an F-test of $H_0 : \alpha = 0_N$ and $\delta = 0_N$. The second test is a step down test where F_1 is an F-test of $\alpha = 0_N$, and F_2 is an F-test of $\delta = 0_N$ conditional on $\alpha = 0_N$. The two tests are performed on each real estate index as well as jointly on all real estate indices. The reported p-values are exact under the normality assumption on the residuals. The results are presented for the three sub-periods (pre-crisis, crisis and post-crisis). Significance levels: *** p<0.01, ** p<0.05, * p<0.10.

9.2 Asymptotic Spanning Tests

Since the normality assumption is not a realistic assumption with financial data, it makes sense to look at asymptotic versions of the mean-variance spanning test in order to test the results and interpretations for robustness. Table 9.4, 9.5 and 9.6 reflect the spanning tests based on non-normality for, respectively, the bonds, stocks and real estate indices. The first six columns relate to the Wald tests. W is the standard asymptotic Wald test, W_a^e is the Wald test based on the specific situation where there would be a multivariate elliptical distribution and the W_a test is the improved GMM version of the W-test that is valid for each distribution[36]. The last six columns present the two SDF based GMM Wald Tests. J_1 and J_2 reflect the version of Bekaert & Urias (1996), where J_1 incorporates the Errors-In-Variables adjustment and J_2 does not[37]. J_3 is the version of De Santis (1993)[38].

Table 9.4 (bonds) presents a hand full of non-rejections. For the full-sample, only Japan is not rejected. Besides the standard W-test (only rejection on 1% level), all other five asymptotic spanning tests can not reject Japan on all common significance levels. For the pre-crisis period, Japan can yet again not be rejected. This time, besides W_a^e, all other asymptotic tests can not reject the spanning hypothesis for Japan on all common significance levels. Looking at the crisis period, 6 of the 9 test assets can not be rejected by all asymptotic tests: Australia (W_a^e), Belgium (W_a^e), France (W_a^e), Germany (W_a^e, J_1, J_2, J_3), Japan (W_a^e, W_a J_1, J_2, J_3) and Netherlands (W_a^e). According to the results of the post-crisis period, Australia (W, W_a^e, W_a, J_1, J_2, J_3), Germany (W_a^e), Japan (W, W_a^e, W_a, J_1, J_2, J_3) and Netherlands (W_a^e) can not be rejected.

Regarding Table 9.5 (stocks), all (sub)-periods have non-rejections. First, the full-sample can not be rejected for Australia (W, W_a^e, W_a, J_1, J_2, J_3), Belgium (W_a^e, J_1, J_2, J_3) and Canada (W, W_a^e, W_a, J_1, J_2, J_3). For the pre-crisis period, only Belgium (J_1) can not be rejected. The crisis period shows that 4 of the 9 test assets can not be rejected: Australia (J_1, J_2), Canada (W, W_a^e, W_a, J_1, J_2, J_3), Japan (W_a^e) and Sweden (W_a^e, J_1, J_2). Finally, there is the post-crisis period which has non-rejections for the following countries: Belgium (W, W_a^e), France (W, W_a^e, W_a, J_1, J_2, J_3), Germany (W, W_a^e, W_a, J_1, J_2, J_3), Netherlands (W, W_a^e, W_a, J_1, J_2, J_3) and Sweden (J_1).

Lastly, Table 9.6 (real estate), reports that there are a couple non-rejections across all (sub)-periods. For the full-sample, Australia (W_a^e) and Germany (W, W_a^e, W_a, J_1, J_2, J_3) are not rejected completely on the basis of all asymptotic tests. For the pre-crisis period, Germany (W_a^e) can not be rejected. For the crisis period, the test assets that are not rejected by all

[36]Kan & Zhou (2008) report that the W_a is inflated in small samples and that therefore W_a should be higher than W_a^e.

[37]Kan & Zhou (2008) do not report the J_1 because it is numerically very close to J_2. Even though the results show that both tests are numerically very close to each other, both are still reported.

[38]According to Kan & Zhou (2008) and their simulation results, J_3 should reject the spanning hypothesis more often than J_2 because the test-statistic of J_3 should be much larger.

asymptotic tests are Australia (W_a^e), Canada (W_a^e), Germany (W, W_a^e, W_a, J_1, J_2, J_3) and Sweden (W_a^e, J_2). Last but not least, the post-crisis period has just one non-rejection and that is Japan (J_1, J_2).

Table 9.4: Asymptotic Mean-Variance Spanning Tests - Bond Indices

	Regression Based						**SDF Based**					
Country (test assets)	W	p-value	W_a^e	p-value	W_a	p-value	J_1	p-value	J_2	p-value	J_3	p-value
Full-Sample (10/2002 - 12/2017)												
Australia	13,596	0,001***	6,986	0,030**	8,104	0,017**	6,485	0,039**	6,725	0,035**	7,145	0,028**
Belgium	20,897	0,000***	9,201	0,010**	14,580	0,001***	9,776	0,008***	10,162	0,006***	11,995	0,002***
Canada	63,138	0,000***	30,854	0,000***	41,100	0,000***	23,150	0,000***	24,121	0,000***	33,546	0,000***
France	19,430	0,000***	8,585	0,014**	14,757	0,001***	10,180	0,006***	10,576	0,005***	12,291	0,002***
Germany	19,583	0,000***	8,988	0,011**	14,068	0,001***	10,037	0,007***	10,411	0,005***	12,144	0,002***
Japan	5,913	0,052*	3,466	0,177	4,027	0,134	3,985	0,136	4,014	0,134	3,468	0,177
Netherlands	19,015	0,000***	8,381	0,015**	14,070	0,001***	9,686	0,008***	10,042	0,007***	11,427	0,003***
Sweden	25,017	0,000***	19,474	0,000***	13,752	0,001***	12,469	0,002***	12,945	0,002***	14,237	0,001***
All countries	86,398	0,000***	64,551	0,000***	91,046	0,000***	53,331	0,000***	54,547	0,000***	78,156	0,000***
Pre-crisis (10/2002 - 10/2007)												
Australia	5,576	0,062*	5,396	0,067*	5,294	0,071*	5,066	0,079*	5,963	0,051*	5,284	0,071*
Belgium	6,139	0,046**	6,057	0,048**	6,449	0,040**	7,697	0,021**	9,117	0,010**	7,766	0,021**
Canada	15,059	0,001***	13,108	0,001***	13,960	0,001***	11,267	0,004***	14,624	0,001***	19,244	0,000***
France	6,214	0,045**	6,144	0,046**	6,568	0,037**	7,692	0,021**	9,112	0,011**	7,732	0,021**
Germany	5,838	0,054*	5,774	0,056*	6,136	0,047**	7,193	0,027**	8,450	0,015**	7,216	0,027**
Japan	4,549	0,103	17,456	0,000***	2,762	0,251	3,171	0,205	3,154	0,207	3,383	0,184
Netherlands	6,413	0,041**	6,348	0,042**	6,749	0,034**	7,918	0,019**	9,419	0,009***	7,962	0,019**
Sweden	4,746	0,093*	4,706	0,095*	4,731	0,094*	5,739	0,057*	6,699	0,035**	5,776	0,056*
All countries	33,434	0,006***	31,884	0,010**	44,230	0,000***	35,288	0,004***	47,012	0,000***	61,496	0,000***
Crisis (11/2007 - 11/2012)												
Australia	11,994	0,002***	4,238	0,120	11,004	0,004***	5,208	0,074*	5,360	0,069*	5,830	0,054*
Belgium	11,824	0,003***	3,889	0,143	11,255	0,004***	5,268	0,072*	5,283	0,071*	5,969	0,051*
Canada	47,494	0,000***	13,758	0,001***	66,290	0,000***	12,526	0,002***	12,653	0,002***	19,923	0,000***
France	10,730	0,005***	3,305	0,192	10,361	0,006***	4,806	0,090*	4,814	0,090*	5,413	0,067*
Germany	9,375	0,009***	2,946	0,229	7,803	0,020**	4,135	0,127	4,158	0,125	4,453	0,108
Japan	5,456	0,065*	3,706	0,157	3,773	0,152	3,235	0,198	3,563	0,168	3,370	0,185
Netherlands	9,428	0,009***	2,898	0,235	9,283	0,010**	4,767	0,092*	4,782	0,092*	4,923	0,085*
Sweden	20,087	0,000***	9,010	0,011**	16,329	0,000***	7,873	0,020**	8,051	0,018**	9,323	0,009***
All Countries	76,986	0,000***	46,970	0,000***	199,368	0,000***	26,833	0,043**	28,470	0,028**	86,394	0,000***
Post-crisis (12/2012 - 12/2017)												
Australia	1,719	0,423	1,524	0,467	1,563	0,458	1,538	0,464	1,567	0,457	1,490	0,475
Belgium	8,110	0,017**	109,912	0,000***	6,407	0,041**	6,257	0,044**	6,772	0,034**	10,490	0,005***
Canada	11,450	0,003***	32,091	0,000***	6,433	0,040**	6,959	0,031**	6,821	0,033**	8,647	0,013**
France	8,221	0,016**	60,724	0,000***	6,296	0,043**	5,997	0,050*	6,458	0,040**	10,108	0,006***
Germany	12,617	0,002***	-96,336	1,000	8,191	0,017**	8,193	0,017**	8,614	0,013**	16,151	0,000***
Japan	0,784	0,676	0,784	0,676	0,786	0,675	0,711	0,701	0,728	0,695	0,711	0,701
Netherlands	10,530	0,005***	-177,256	1,000	7,766	0,021**	7,899	0,019**	8,345	0,015**	13,680	0,001***
Sweden	8,073	0,018**	8,761	0,013**	6,422	0,040**	5,975	0,050*	5,770	0,056*	7,606	0,022**
All Countries	54,420	0,000***	52,335	0,000***	63,054	0,000***	46,549	0,000***	52,683	0,000***	100,192	0,000***

The table presents the asymptotic versions of the mean-variance spanning tests on eight bond indices, using the U.S. bond index as benchmark asset. The first two tests, W_a^e and W_a are regression based GMM Wald tests. W_a is valid under general distribution whereas W_a^e is only valid when returns follow a multivariate elliptical distribution. The other three tests, J_1, J_2 and J_3, are SDF based GMM Wald tests. J_1 and J_2 are the versions used by Bekaert and Urias (1996), where J_1 includes the Errors-In-Variables adjustment and J_2 does not. J_3 is the version used by DeSantis (1993). The tests are performed on each bond index as well as jointly on all bond indices. All tests have an asymptotic χ^2_{2N} distribution, where N is the number of test assets, and the reported p-values are asymptotic ones. The results are presented for the entire sample period as well as for the sub-periods (pre-crisis, crisis and post-crisis). Significance levels: *** p<0.01, ** p<0.05, * p<0.10.

Table 9.5: Asymptotic Mean-Variance Spanning Tests - Stock Indices

	Regression Based						**SDF Based**					
Country (test assets)	W	p-value	W_a^e	p-value	W_a	p-value	J_1	p-value	J_2	p-value	J_3	p-value
Full-Sample (10/2002 - 12/2017)												
Australia	4,357	0,113	1,344	0,511	3,032	0,220	2,667	0,264	2,667	0,264	2,570	0,277
Belgium	6,612	0,037**	1,828	0,401	4,616	0,099*	4,200	0,122	4,148	0,126	3,870	0,144
Canada	1,276	0,528	0,468	0,792	0,992	0,609	1,098	0,577	1,103	0,576	1,061	0,588
France	14,059	0,001***	5,692	0,058*	11,968	0,003***	10,770	0,005***	10,887	0,004***	13,462	0,001***
Germany	33,923	0,000***	16,686	0,000***	23,460	0,000***	22,713	0,000***	22,606	0,000***	39,685	0,000***
Japan	27,834	0,000***	16,566	0,000***	15,108	0,001***	15,616	0,000***	15,833	0,000***	24,737	0,000***
Netherlands	22,363	0,000***	8,109	0,017**	13,691	0,001***	12,693	0,002***	12,556	0,002***	16,340	0,000***
Sweden	16,944	0,000***	7,085	0,029**	9,280	0,010**	7,472	0,024**	7,461	0,024**	9,845	0,007***
All countries	107,785	0,000***	61,740	0,000***	72,571	0,000***	43,253	0,000***	44,008	0,000***	108,219	0,000***
Pre-crisis (10/2002 - 10/2007)												
Australia	10,104	0,006***	9,997	0,007***	10,294	0,006***	7,058	0,029**	9,175	0,010**	7,149	0,028**
Belgium	5,591	0,061*	5,011	0,082*	5,367	0,068*	4,596	0,100	5,044	0,080*	4,989	0,083*
Canada	10,668	0,005***	10,651	0,005***	11,500	0,003***	7,523	0,023**	9,261	0,010**	7,510	0,023**
France	12,710	0,002***	10,707	0,005***	8,566	0,014**	5,230	0,073*	5,596	0,061*	7,326	0,026**
Germany	44,789	0,000***	25,203	0,000***	21,985	0,000***	15,652	0,000***	16,625	0,000***	43,937	0,000***
Japan	7,818	0,020**	7,481	0,024**	7,002	0,030**	6,034	0,049**	6,437	0,040**	8,963	0,011**
Netherlands	15,781	0,000***	9,167	0,010**	13,661	0,001***	7,819	0,020**	7,852	0,020**	10,824	0,004***
Sweden	21,358	0,000***	16,588	0,000***	14,956	0,001***	9,678	0,008***	11,183	0,004***	16,441	0,000***
All countries	95,933	0,000***	68,465	0,000***	81,466	0,000***	38,814	0,001***	46,999	0,000***	137,034	0,000***
Crisis (11/2007 - 11/2012)												
Australia	9,154	0,010**	8,967	0,011**	7,665	0,022**	4,494	0,106	4,505	0,105	6,316	0,043**
Belgium	10,505	0,005***	11,164	0,004***	11,037	0,004***	7,924	0,019**	8,056	0,018**	9,376	0,009***
Canada	2,443	0,295	1,575	0,455	2,318	0,314	2,881	0,237	2,883	0,237	2,713	0,258
France	12,637	0,002***	113,302	0,000***	11,626	0,003***	8,611	0,013**	9,225	0,010**	17,167	0,000***
Germany	17,593	0,000***	20,820	0,000***	14,868	0,001***	11,981	0,003***	11,896	0,003***	32,869	0,000***
Japan	11,782	0,003***	-8,354	1,000	8,677	0,013**	11,946	0,003***	11,649	0,003***	20,892	0,000***
Netherlands	14,976	0,001***	31,322	0,000***	11,071	0,004***	8,728	0,013**	8,848	0,012**	16,717	0,000***
Sweden	13,170	0,001***	-640,553	1,000	8,003	0,018**	4,385	0,112	4,373	0,112	8,588	0,014**
All Countries	66,612	0,000***	61,946	0,000***	78,585	0,000***	32,230	0,009***	34,062	0,005***	147,401	0,000***
Post-crisis (12/2012 - 12/2017)												
Australia	6,074	0,048**	5,688	0,058*	6,333	0,042**	8,676	0,013**	9,106	0,011**	8,937	0,011**
Belgium	4,197	0,123	2,881	0,237	4,992	0,082*	5,503	0,064*	5,525	0,063*	6,801	0,033**
Canada	11,185	0,004***	10,137	0,006***	12,967	0,002***	7,274	0,026**	7,845	0,020**	7,321	0,026**
France	1,850	0,397	1,501	0,472	2,238	0,327	2,391	0,303	2,462	0,292	2,593	0,273
Germany	0,185	0,912	0,176	0,916	0,200	0,905	0,203	0,904	0,204	0,903	0,202	0,904
Japan	14,682	0,001***	8,785	0,012**	12,726	0,002***	9,278	0,010**	10,357	0,006***	24,947	0,000***
Netherlands	1,235	0,539	1,155	0,561	1,274	0,529	1,465	0,481	1,503	0,472	1,490	0,475
Sweden	6,610	0,037**	5,229	0,073*	5,348	0,069*	4,548	0,103	4,968	0,083*	5,071	0,079*
All Countries	46,489	0,000***	43,689	0,000***	43,381	0,000***	43,631	0,000***	49,449	0,000***	69,585	0,000***

The table presents the asymptotic versions of the mean-variance spanning tests on eight stock indices, using the U.S. stock index as benchmark asset. The first two tests, W_a^e and W_a are regression based GMM Wald tests. W_a is valid under general distribution whereas W_a^e is only valid when returns follow a multivariate elliptical distribution. The other three tests, J_1, J_2 and J_3, are SDF based GMM Wald tests. J_1 and J_2 are the versions used by Bekaert and Urias (1996), where J_1 includes the Errors-In-Variables adjustment and J_2 does not. J_3 is the version used by DeSantis (1993). The tests are performed on each stock index as well as jointly on all stock indices. All tests have an asymptotic χ^2_{2N} distribution, where N is the number of test assets, and the reported p-values are asymptotic ones. The results are presented for the entire sample period as well as for the sub-periods (pre-crisis, crisis and post-crisis). Significance levels: *** p<0.01, ** p<0.05, * p<0.10.

Table 9.6: Asymptotic Mean-Variance Spanning Tests - Real Estate Indices

	Regression Based						SDF Based					
Country (test assets)	W	p-value	W_a^e	p-value	W_a	p-value	J_1	p-value	J_2	p-value	J_3	p-value
Full-Sample (10/2002 - 12/2017)												
Australia	75,779	0,000***	-16,392	1,000	7,769	0,021**	13,505	0,001***	13,530	0,001***	31,400	0,000***
Belgium	277,470	0,000***	47,380	0,000***	114,897	0,000**	42,775	0,000***	42,935	0,000***	105,157	0,000***
Canada	113,914	0,000***	53,037	0,000***	19,249	0,000***	16,492	0,000***	16,707	0,000***	44,068	0,000***
France	58,371	0,000***	11,940	0,003***	32,578	0,000***	25,039	0,000***	26,423	0,000***	18,484	0,000***
Germany	1,496	0,473	0,211	0,900	0,614	0,736	0,563	0,755	0,557	0,757	0,613	0,736
Japan	98,849	0,000***	17,776	0,000***	87,903	0,000***	30,047	0,000***	30,853	0,000***	25,982	0,000***
Netherlands	69,570	0,000***	14,634	0,001***	41,158	0,000***	31,048	0,000***	31,071	0,000***	19,030	0,000***
Sweden	18,659	0,000***	5,853	0,054*	16,118	0,000***	12,290	0,002***	11,839	0,003***	11,288	0,004***
All countries	500,224	0,000***	162,242	0,000***	692,480	0,000***	88,982	0,000***	91,317	0,000***	684,503	0,000***
Pre-crisis (10/2002 - 10/2007)												
Australia	53,963	0,000***	34,212	0,000***	49,067	0,000***	28,588	0,000***	35,048	0,000***	45,560	0,000***
Belgium	67,172	0,000***	52,213	0,000***	64,320	0,000***	30,421	0,000***	33,633	0,000***	53,004	0,000***
Canada	31,001	0,000***	13,775	0,001***	17,843	0,000***	13,476	0,001***	18,427	0,000***	47,137	0,000***
France	17,594	0,000***	14,943	0,001***	18,137	0,000***	11,702	0,003***	15,727	0,000***	15,072	0,001***
Germany	9,156	0,010**	3,655	0,161	7,311	0,026**	4,656	0,097*	4,770	0,092*	5,313	0,070*
Japan	12,297	0,002***	7,858	0,020**	12,670	0,002***	10,544	0,005***	12,274	0,002***	17,653	0,000***
Netherlands	22,028	0,000***	16,067	0,000***	22,192	0,000***	18,031	0,000***	21,791	0,000***	26,078	0,000***
Sweden	8,807	0,012**	8,485	0,014**	8,220	0,016**	6,828	0,033**	8,461	0,015**	7,194	0,027**
All countries	132,807	0,000***	101,707	0,000***	156,715	0,000***	55,217	0,000***	79,533	0,000***	161,448	0,000***
Crisis (11/2007 - 11/2012)												
Australia	29,275	0,000***	-6,416	1,000	5,267	0,072*	10,746	0,005***	10,797	0,005***	26,194	0,000***
Belgium	148,228	0,000***	43,990	0,000***	86,187	0,000***	21,669	0,000***	21,567	0,000***	114,996	0,000***
Canada	53,145	0,000***	-18,410	1,000	11,570	0,003***	7,736	0,021**	7,731	0,021**	34,197	0,000***
France	33,957	0,000***	12,037	0,002***	24,249	0,000***	14,558	0,001***	14,489	0,001***	15,587	0,000***
Germany	0,840	0,657	0,742	0,690	0,943	0,624	1,017	0,602	1,063	0,588	1,078	0,583
Japan	72,924	0,000***	25,662	0,000***	66,521	0,000***	20,163	0,000***	19,992	0,000***	39,097	0,000***
Netherlands	36,600	0,000***	16,014	0,000***	32,011	0,000***	17,854	0,000***	17,687	0,000***	17,240	0,000***
Sweden	6,666	0,036**	1,602	0,449	6,496	0,039**	4,769	0,092*	4,601	0,100	4,921	0,085*
All Countries	355,954	0,000***	176,966	0,000***	1005,203	0,000***	48,693	0,000***	50,152	0,000***	2070,664	0,000***
Post-crisis (12/2012 - 12/2017)												
Australia	8,839	0,012**	12,869	0,002***	11,215	0,004***	10,143	0,006***	9,917	0,007***	10,705	0,005***
Belgium	23,820	0,000***	25,537	0,000***	28,898	0,000***	14,387	0,001***	14,215	0,001***	21,188	0,000***
Canada	30,532	0,000***	37,367	0,000***	38,527	0,000***	17,557	0,000***	17,680	0,000***	25,060	0,000***
France	7,848	0,020**	8,165	0,017**	7,091	0,029**	6,159	0,046**	6,153	0,046**	6,885	0,032**
Germany	13,765	0,001***	14,535	0,001***	14,190	0,001***	8,970	0,011**	8,869	0,012**	10,240	0,006***
Japan	22,475	0,000***	5,338	0,069*	22,199	0,000***	4,076	0,130	4,096	0,129	8,387	0,015**
Netherlands	6,958	0,031**	29,918	0,000***	5,335	0,069*	5,953	0,051*	6,100	0,047**	6,736	0,034**
Sweden	12,074	0,002***	10,960	0,004***	13,876	0,001***	8,917	0,012**	9,078	0,011**	10,459	0,005***
All Countries	76,756	0,000***	65,384	0,000***	115,116	0,000***	48,742	0,000***	70,066	0,000***	79,291	0,000***

The table presents the asymptotic versions of the mean-variance spanning tests on eight real estate indices, using the U.S. real estate index as benchmark asset. The first two tests, W_a^e and W_a are regression based GMM Wald tests. W_a is valid under general distribution whereas W_a^e is only valid when returns follow a multivariate elliptical distribution. The other three tests, J_1, J_2 and J_3, are SDF based GMM Wald tests. J_1 and J_2 are the versions used by Bekaert and Urias (1996), where J_1 includes the Errors-In-Variables adjustment and J_2 does not. J_3 is the version used by DeSantis (1993). The tests are performed on each real estate index as well as jointly on all real estate indices. All tests have an asymptotic χ^2_{2N} distribution, where N is the number of test assets, and the reported p-values are asymptotic ones. The results are presented for the entire sample period as well as for the sub-periods (pre-crisis, crisis and post-crisis). Significance levels: *** p<0.01, ** p<0.05, * p<0.10.

9.3 Meaning Of The Results From The Spanning Tests

Section 9.2 shows that the results of section 9.1 are not resistant to non-normality and that the traditional and joint F-test have the tendency to over-reject. This especially applies for the crisis period. Interestingly, data of Kan & Zhou (2008) also shows that their results weakened when the asymptotic tests were used to correct for non-normality (e.g. conditional heteroskedasticity and fat-tails). This study confirms that correction for non-normality is extremely important as the tests under normality can yield distorted conclusions, especially when most authors agree that financial data is skewed and fat-tailed.

From the results of the asymptotic spanning tests, it can be observed that during the crisis and over time there are fluctuations in the significance of test assets. In some sub-periods, test assets are not significant to which they were in different sub-periods. These fluctuations imply instability in the diversification benefits. Fluctuations of test assets in whether or not they are significant are clearly visible in bonds. These fluctuations are stronger within bonds than within stocks and real estate, reflecting the findings of section 7.1 and 7.2. These sections showed that correlations between bonds are more unstable than the correlations between risky assets. Results, therefore, confirm the findings of earlier papers, which stress the importance of stable correlations in order to achieve diversification gains (Jorion, 1985; Eun & Resnick, 1988). Results also show that there are, overall, less insignificant test assets for real estate. This verifies that, relatively, real estate is a better diversifying asset (Eichholtz, 1996a).

Nevertheless, diversification benefits, for bonds, stocks and real estate, possibly fluctuate over time and tend to only decline in periods of high volatility (recession), in line with findings in chapter 8. But this only applies for individual test assets. When all eight countries are included in the portfolio of test assets (*all countries*), diversification gains do not fluctuate and does not have the tendency to decrease in recessions (and also not over time). This is in line with general consensus that diversification is mostly substantial when a bunch of assets is added to the portfolio, as more assets should lead to a more and well diversified portfolio due to the fact that the total portfolio risk is then spread over more assets (Markowitz, 1952; B. H. Solnik, 1995; Levy & Post, 2005). This also gives the advantage that a large diversified portfolio is more resistant to periods of high volatility. Important to note is that these spanning tests did not include any short-selling restriction. Therefore, it is quite reasonable to predict that the results will further worsen when there are short-selling restrictions, as short-selling restrictions help the marginal investor to make use of high positive correlations.

This chapter can be summarized in three points: (1) an investor benefits from diversification over time, (2) these benefits have the tendency to decline during recessions and (3) diversification benefits do not tend to fall consistently over time, especially not when the portfolio is sufficiently large. These three points jointly result that hypothesis H_{04} can be rejected for bonds, stocks and real estate.

Part IV

Conclusion and Limitations

Chapter 10

Conclusion and Final Remarks

10.1 Conclusion

As stated in chapter 1, the main goal of this paper was to investigate the (1) degree of stability of correlation coefficients over time, (2) pattern of correlation coefficients over time, (3) movement of the efficient frontier over time and (4) diversification benefits over time. Analyzing the changes over time can provide information for investors on the question whether they should construct efficient portfolios dynamically by taking possible changes into account (e.g. changing correlation coefficients over time or during a crisis). The purpose of investigating the aforementioned four points was to ultimately form an answer to the following research question:

"*Analyzing bonds, stocks and real estate indices, what effect does the pattern and degree of stability of correlation coefficients between countries over time have on the efficient frontier and diversification benefits over time?*"

This study therefore has some interesting findings to conclude this paper with. First of all, correlation coefficients of stocks and real estate seem to be quite stable over time. The pattern of the correlations of these two asset classes have the tendency to increase over time, especially during a crisis. However, it must be indicated that that the very strong increase seems to stabilize after the end of the period of high volatility, implying stability on the long-term. These findings implicate the confirmation of the fact that correlation coefficients tend to be higher during bear markets than in bull markets, implying stabilization on the long-term (Longin & Solnik, 2001; Ang & Bekaert, 2002; Hyde et al., 2007).

In contrast, correlations of bonds do not appear to be (1) stable over time and (2) strongly influenced by the crisis. Unstable correlations between government bonds could be due to differences between countries in changes of macroeconomic factors, such as monetary policy. Regarding point (2), correlations of bonds have in fact fallen slightly during the recession, although they do show small upward shocks during the recession and over time. The downward trend in correlations of bonds could well be the result of short-term shifting in demand from government bonds with relatively high exposure to sovereign debt risk to government bonds with relatively less exposure to sovereign debt risk. During recessions, investors reallocate their investments to bonds with relatively lower sovereign debt risk with the result that the aggregate demand for these bonds increases, which in turn increase (decrease) prices (yields). The opposite should then be the case for countries with a relatively higher sovereign debt risk. And therefore, prices and returns across government bonds are less correlated during recessions.

Analysis of efficient frontiers showed the heavy influence of correlation coefficients on the risk-and-return trade-off, as the movements of the efficient frontier resemble the pattern of correlation coefficients. The results of the efficient frontier analysis implied that the risk-and-return trade-off worsen during the crisis for both stocks and real estate, but stabilize after the crisis. Bonds, on the other hand, appear to be relatively resistant to changes in the risk-and-return trade-off during the crisis, verifying that bonds are considered as safe assets and are therefore not heavily impacted by negative investor sentiment or collective uncertainty during recessions. Adding a short-selling restriction does not change the main findings of the efficient frontier analysis, but it did show that a short-selling restriction indeed limits diversification gains (Levy & Post, 2005; Lee & Lee, 2010).

Finally, this research included various spanning test in order to gain information about the significance of the movements of efficient frontiers and therefore the significance of possible diversification gains. Results of these spanning tests first put in sight that one must use spanning tests that correct for normality in order to make robust conclusions, as the spanning tests under normality can lead to distorted results and conclusions. Second, the findings of the spanning test indicate that the marginal investor benefits from diversification and can therefore expand its investment opportunities over time by diversifying internationally. Also, the diversification benefits do not decrease consistently over time. However, these diversification gains seem to worsen in times of crisis if the portfolio of (test) assets is not sufficiently large. Finally, one must note that short-selling restrictions were not included in these spanning tests and that therefore the results, regarding diversification benefits, can worsen dramatically when short-sales constraints are added. Anyone who thinks of investing or examining investment choices must take into account that changes over time can affect models. Only examining changes of, for example, efficient frontiers is not enough. It is advisable to also examine the underlying variables such as correlation coefficients to put sight on the effect in a broader perspective. Finally, it should certainly not be forgotten to take into account that each model has its own assumptions. That is why it is always wise to use different models and assumptions to keep the robustness of the results and conclusions intact.

10.2 Limitations and Final Remarks

Every research has its limitations. Unfortunately, this paper can not escape having limitations. One of the biggest shortcomings of this research is that the spanning tests did not take into account possible short sales constraints, whereas Chapter 8 (Efficient-Frontiers) highlights the differences between an accessible and unaccessible short-selling market. De Roon, Nijman, & Werker (2001) and Glabadanidis (2017) offer, respectively, alternative spanning and portfolio efficiency tests that also take into account short sales constraints. There is a deeper underlying reason for not adding this. First of all, based on De Roon, Nijman, & Werker (2001), Driessen & Laeven (2007), show that the addition of short-sales constraints can lead to adverse effects.

The same authors mention that the power of spanning tests could be lower when short-sales constraints are added, as the number of short-selling restrictions influence the small sample properties of the spanning tests. The authors therefore choose to minimize these adverse effects by applying the short-sales constraint to just one asset. Glabadanidis (2017) also acknowledges possible adverse effects of adding the short-sales restriction on the N test assets and therefore proposes to only apply short-sales constraints to only the K benchmark (base) assets. The addition of a short-selling restriction on one or a small number of assets makes the differences between constrained and unconstrained frontier negligible. This is because there is a risk that the test statistics of the spanning tests and corresponding Sharpe Ratio's will not differ significantly[39].

Therefore, adding short-selling constraints would not be suitable for this research as this paper only considers one benchmark asset per asset class. This provides an opportunity for follow-up research, because this research is quite isolated by only analyzing within asset classes (and between countries). Future research can therefore expand this paper's analysis by also considering an investigation between assets (and again, between countries). Thus, with different asset classes, multiple benchmark assets can be considered which in turn leads to possibilities of introducing short-sales restricted spanning/efficiency tests with less risk of adverse effects. Another point that can be taken for follow-up research is that the Black-Litterman Model (Black & Litterman, 1992) can be used to construct efficient frontiers, as this is a refinement of the Modern Portfolio Theory (Markowitz, 1952) and offers new perspective by also giving weight to views and expectations on assets' returns instead of only to historical data. Moreover, one can add value to show the diversification benefits in terms of local returns as a large part of the investors value 'total' returns and currency risk. Thus, new insights in the results of the efficient frontier analysis and spanning tests can be obtained by incorporating currency risk and local currency returns. With regard to the efficient frontier, efficient frontiers can also be computed with correlation and covariance matrices estimated from the EWMA or DCC-GARCH model. In addition, the DCC-GARCH model can be refined by changing the underlying distribution because this paper only considered the standard student's t-distribution. However, the Johnson's SU distribution, Skewed generalized t-distribution, Laplace distribution or the Inverse-Gaussian distribution may also be possible.

All of the above listed points of improvement should certainly be considered for those who want to continue, expand or refine this research, as this topic offers many opportunities for follow-up research.

[39] This research did in fact replicate the GMV and GRS portfolio efficiency tests, thanks to the provided source code by Glabadanidis (2017) on the dataset used in this paper. However, adding short-selling restrictions on just one benchmark did not lead to other values and in fact lead to adverse effects. The adverse effect is that extending the short-selling restrictions to all or test assets made the unconstrained portfolio perform equal or even worse than the constrained portfolio, in terms of the GRS/GMV-test and Sharpe Ratio's.

Part V

Appendices

Appendix A

Description of Data

A.1 Datastream Symbols

Table A.1: Retrieved Symbols From DataStream

Country	Government Bond	Index Code	Stock	Index Code	Real Estate	Index Code
Australia	BMAU10Y	CI	ASXAORD	PI	SBBPAUL	PI
Belgium	BMBG10Y	CI	BGBEL20	PI	SBBPBEL	PI
Canada	BMCN10Y	CI	TTOSP60	PI	SBBPCAL	PI
France	BMFR10Y	CI	FRCAC40	PI	SBBPFRL	PI
Germany	BMBD10Y	CI	DAXINDX	PI	SBBPDEL	PI
Japan	BMJP10Y	CI	JAPDOWA	PI	SBBPJPL	PI
Netherlands	BMNL10Y	CI	AMSTEOE	PI	SBBPNLL	PI
Sweden	BMSD10Y	CI	SWEDOMX	PI	SBBPSEL	PI
United States	BMUS10Y	CI	S&PCOMP	PI	SP4GRES	PI

The table presents the symbols and corresponding index code for each country index per asset class, where CI and PI, respectively, reflect the *Clean Price Index* and *Price Index*.

A.2 Formulas of Each Descriptive Statistic

In this appendix, the exact formulas of each descriptive statistic are stated, where x_i equals observation i and n equals the total amount of observations:

Mean:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \tag{A.1}$$

Standard Deviation:

$$s_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1}} \tag{A.2}$$

Variance:

$$s_x^2 = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^2}{n-1} \tag{A.3}$$

Skewness:

$$s_{kewness} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^3}{(1/n)\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^{3/2}} \tag{A.4}$$

Kurtosis:

$$k_{urtosis} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})^4}{(1/n)\left[\sum_{i=1}^{n}(x_i - \bar{x})^2\right]^2} \tag{A.5}$$

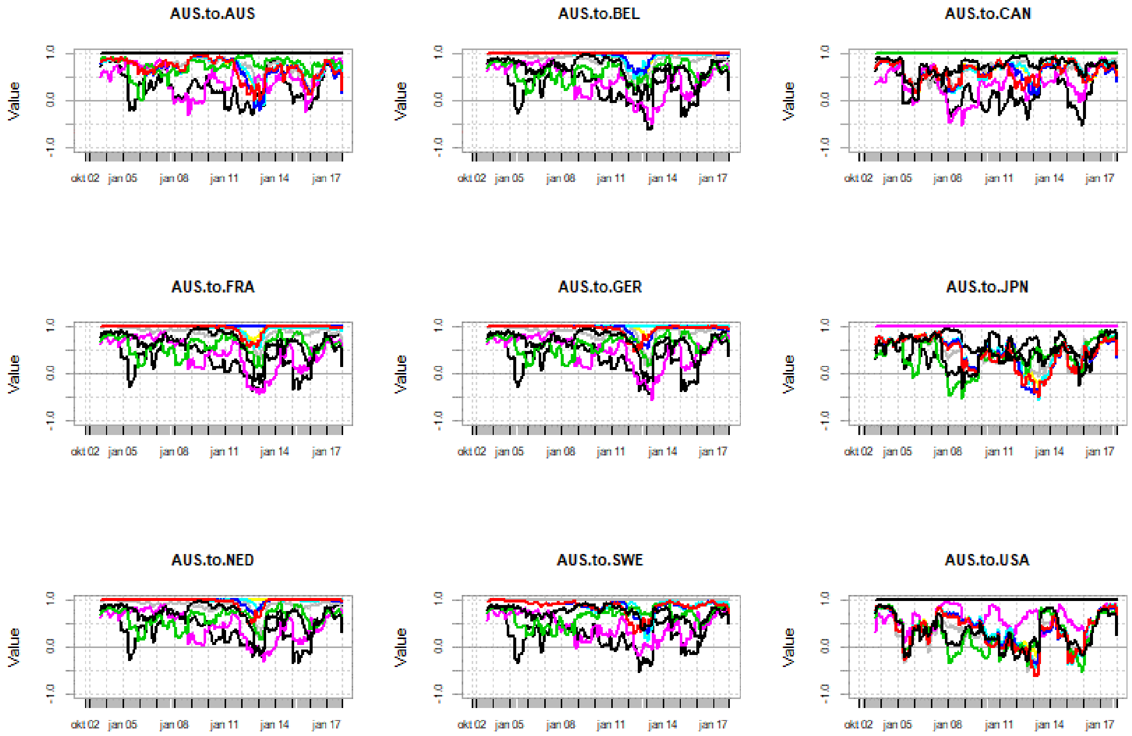
Jarque-Bera Chi-Square Goodness-of-Fit Test with degrees of freedom equal to 2:

$$JB = \frac{n}{6}\left(s_{kewness}^2 + \frac{1}{4}(k_{urtosis} - 3)^2\right) \tag{A.6}$$

Appendix B

Rolling Correlations - Charts

Figure B.1: Rolling Correlation Charts of Bond Indices



Note: The charts of figure B.1. starts with the rolling correlations of the country AUS (top-left) and end with USA (bottom-right)

Figure B.2: Rolling Correlation Charts of Stock Indices

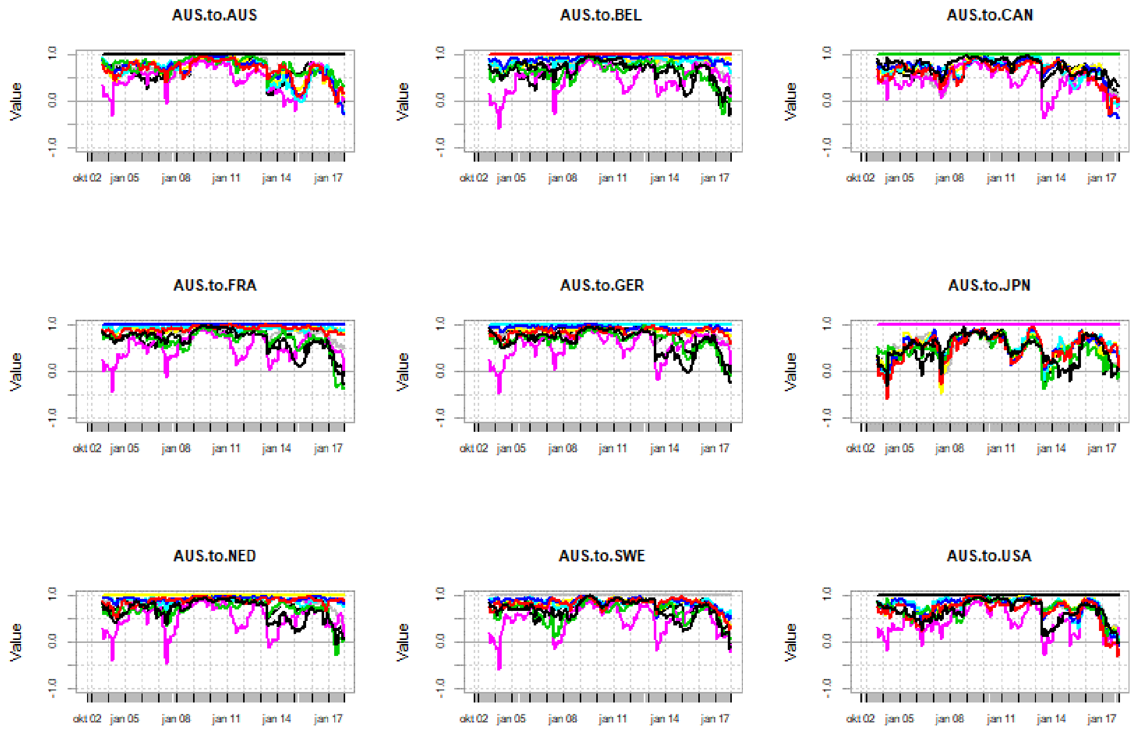
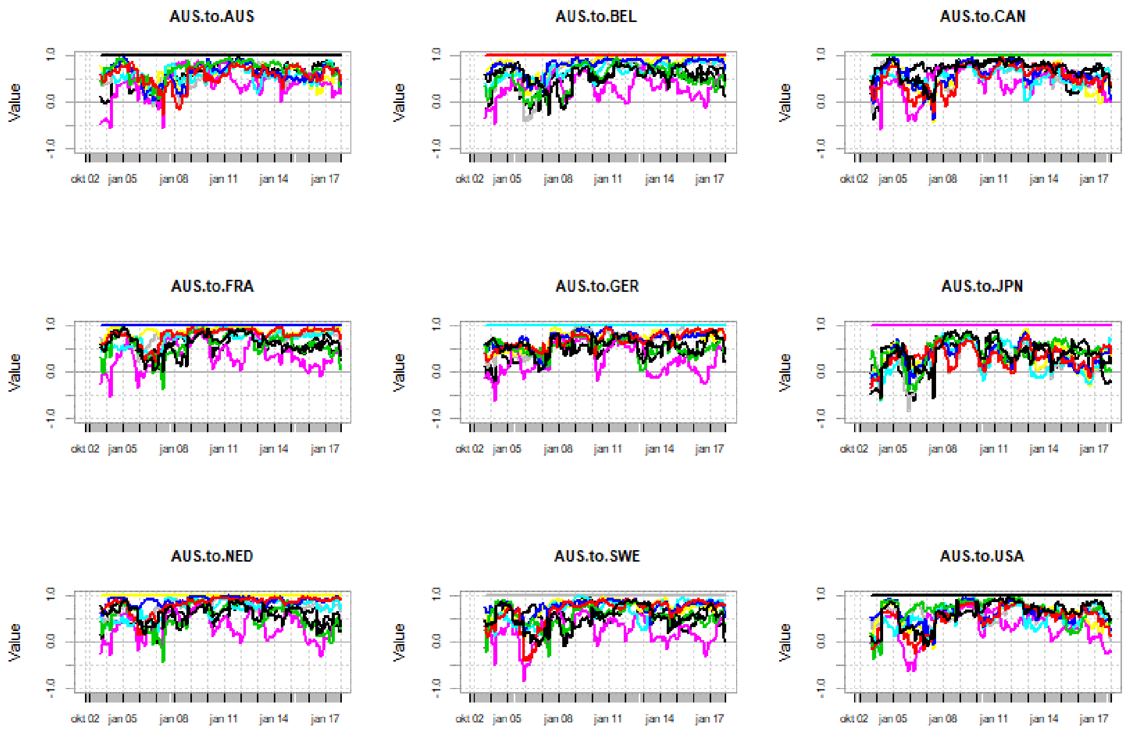


Figure B.3: Rolling Correlation Charts of Real Estate Indices

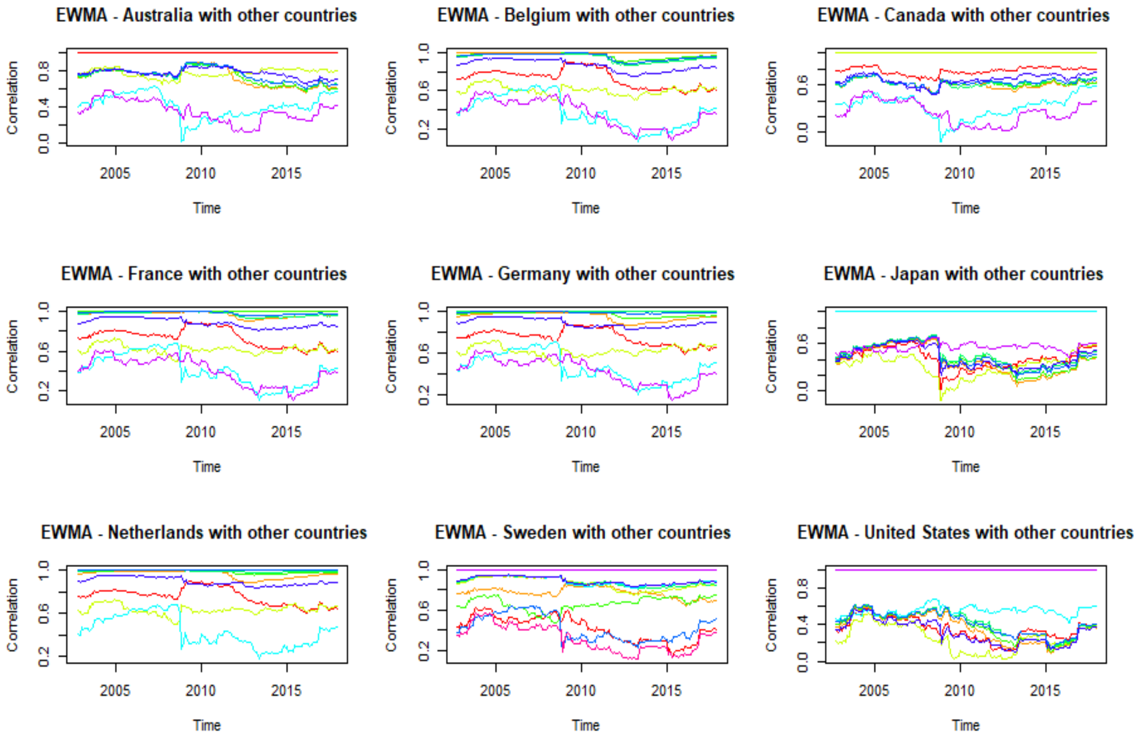


Note: The charts of both figures B.2. and B.3. start with the rolling correlations of the country AUS (top-left) and end with USA (bottom-right).

Appendix C

EWMA Correlations - Charts

Figure C.1: EWMA Correlation Charts of Bond Indices



Note: The charts on figure C.1. starts with the EWMA correlations of the country AUS (top-left) and end with USA (bottom-right)

Figure C.2: EWMA Correlation Charts of Stock Indices

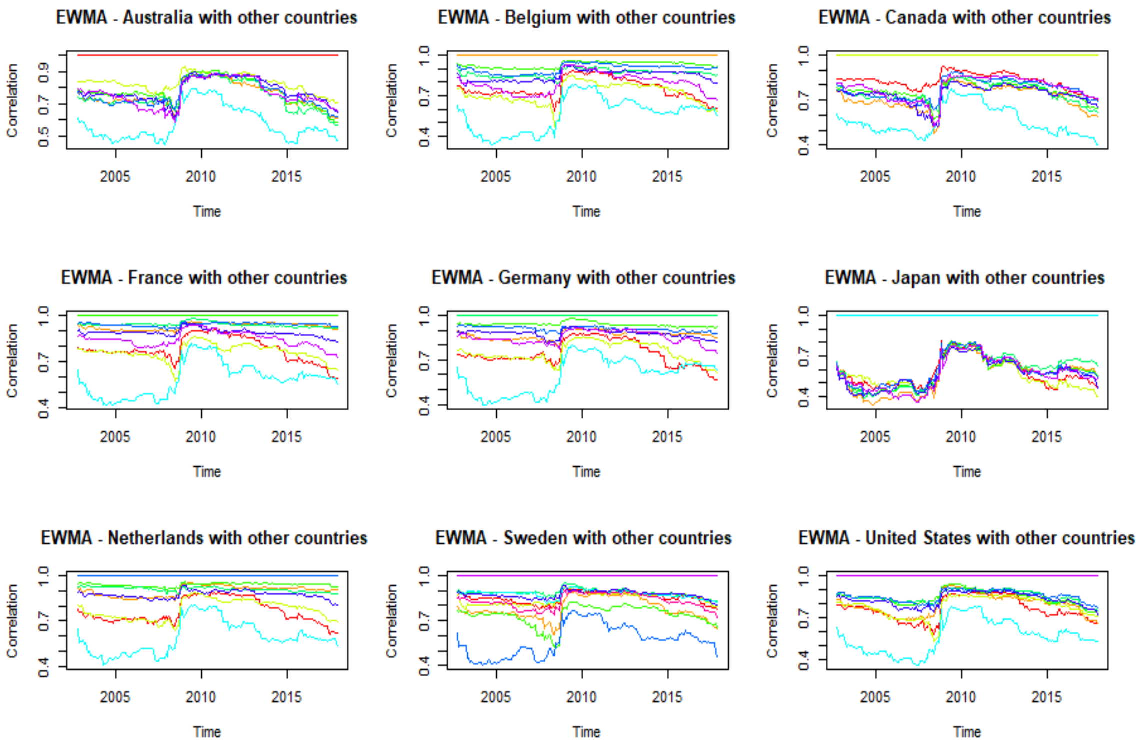
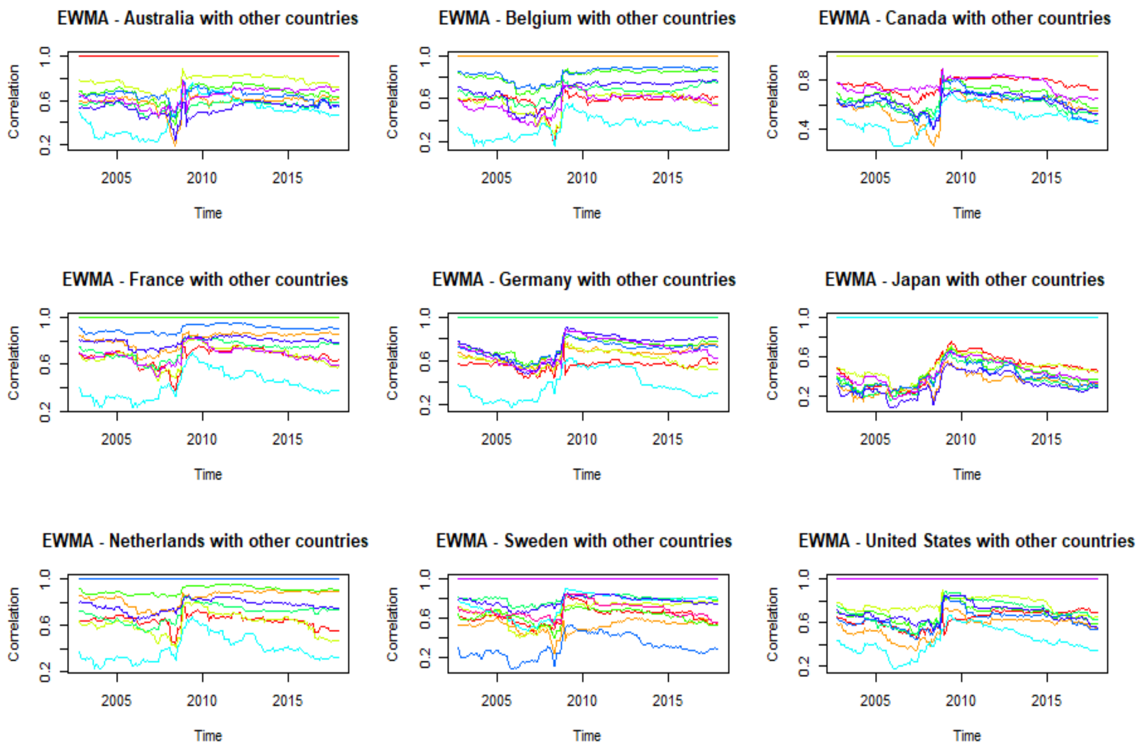


Figure C.3: EWMA Correlation Charts of Real Estate Indices

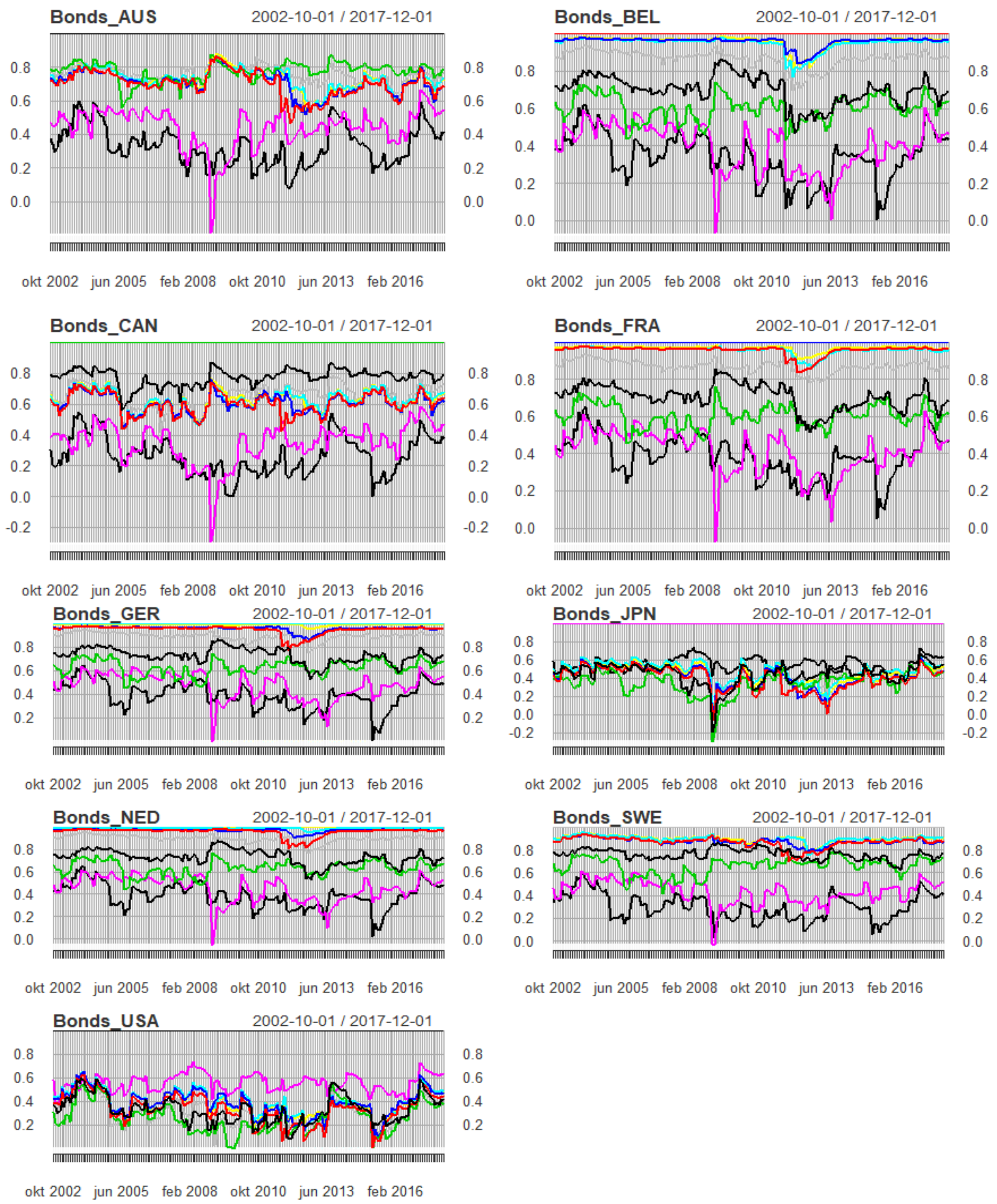


Note: The charts on both figures C.2. and C.3. start with the EWMA correlations of the country AUS (top-left) and end with USA (bottom-right).

Appendix D

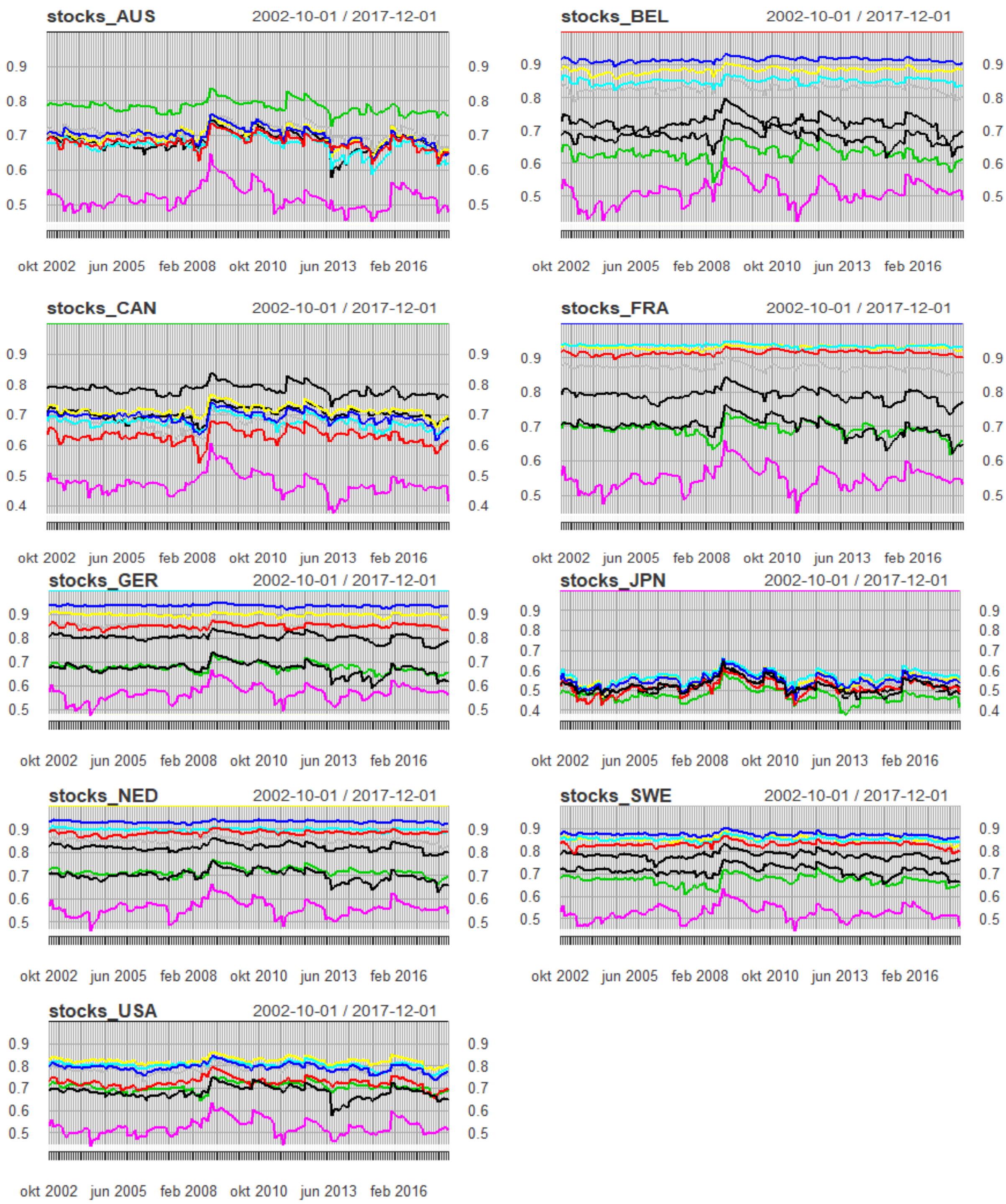
DCC-GARCH Correlations - Charts

Figure D.1: DCC-GARCH Charts of Bond Indices



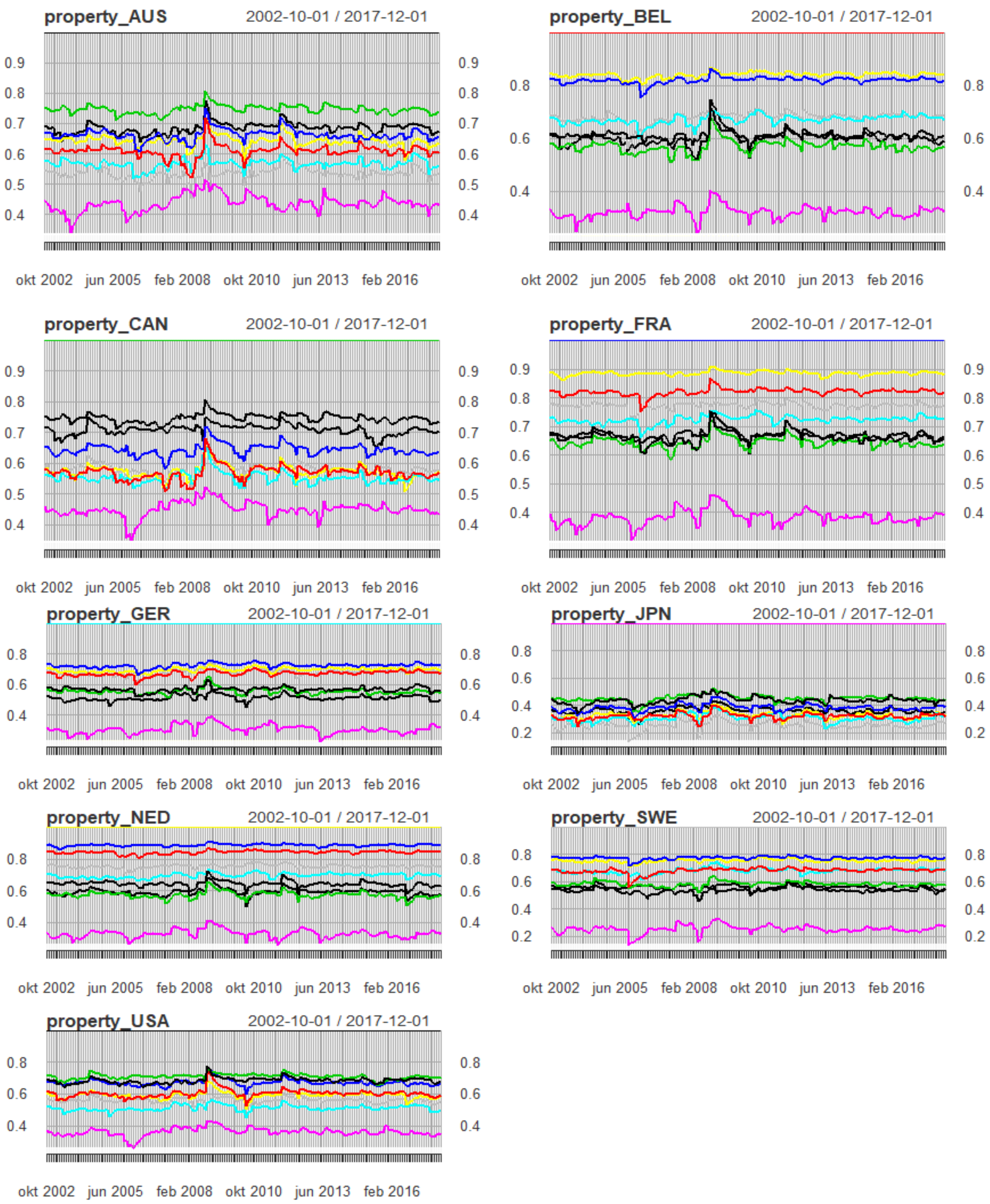
Note: The charts on figure D.1. represent the correlations of a specific bond index with other bond indices according to the DCC(1,1)-GARCH(1,1).

Figure D.2: DCC-GARCH Charts of Stock Indices



Note: The charts on figure D.2. represent the correlations of a specific stock index with other stock indices according to the DCC(1,1)-GARCH(1,1).

Figure D.3: DCC-GARCH Charts of Real Estate Indices



Note: The charts on figure D.3. represent the correlations of a specific real estate index with other real estate indices according to the DCC(1,1)-GARCH(1,1).

References

Adams, Z., Füss, R., & Glück, T. (2017). Are correlations constant? empirical and theoretical results on popular correlation models in finance. *Journal of Banking & Finance*, *84*, 9–24.

Akgiray, V. (1989). Conditional heteroscedasticity in time series of stock returns: Evidence and forecasts. *Journal of business*, 55–80.

Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *The Review of Financial Studies*, *15*(4), 1137–1187.

Ball, R., & Bowers, J. (1986). Shares, bonds, treasury notes, property trusts and inflation: Historical returns and risks, 1974-1985. *Australian Journal of Management*, *11*(2), 117–137.

Bekaert, G., & Urias, M. S. (1996). Diversification, integration and emerging market closed-end funds. *the Journal of Finance*, *51*(3), 835–869.

Bergstrom, G. L. (1975). A new route to higher returns and lower risks. *The Journal of Portfolio Management*, *2*(1), 30–38.

Black, F., & Litterman, R. (1992). Global portfolio optimization. *Financial analysts journal*, *48*(5), 28–43.

Blattberg, R. C., & Gonedes, N. J. (2010). A comparison of the stable and student distributions as statistical models for stock prices. In *Perspectives on promotion and database marketing: The collected works of robert c blattberg* (pp. 25–61). World Scientific.

Bodie, Z., Kane, A., & Marcus, A. J. (2013). *Essentials of investments* (Vol. 6). McGraw-Hill/Irwin.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of econometrics*, *31*(3), 307–327.

Bollerslev, T. (1990). Modelling the coherence in short-run nominal exchange rates: a multivariate generalized arch model. *The review of economics and statistics*, 498–505.

Box, G. E. (1949). A general distribution theory for a class of likelihood criteria. *Biometrika*, *36*(3/4), 317–346.

Brigham, E. F., & Houston, J. F. (2012). *Fundamentals of financial management*. Cengage Learning.

Brooks, C. (2014). *Introductory econometrics for finance*. Cambridge university press.

Buyuksahin, B., Haigh, M. S., & Robe, M. A. (2008). Commodities and equities:'a market of one'?

De Roon, F. A., Nijman, T. E., & Werker, B. J. (2001). Testing for mean-variance spanning with short sales constraints and transaction costs: The case of emerging markets. *The Journal of Finance*, *56*(2), 721–742.

De Santis, G. (1993). *Volatility bounds for stochastic discount factors: tests and implications from international financial markets* (Unpublished doctoral dissertation). University of Chicago, Department of Economics.

Driessen, J., & Laeven, L. (2007). International portfolio diversification benefits: Cross-country evidence from a local perspective. *Journal of Banking & Finance*, *31*(6), 1693–1712.

Dube, S. (2016). Multivariate t-distribution and garch modelling of volatility and conditional correlations on brics stock markets 1. *Journal of Applied Finance and Banking*, *6*(2), 53.

Eichholtz, P. M. (1996a). Does international diversification work better for real estate than for stocks and bonds? *Financial analysts journal*, *52*(1), 56–62.

Eichholtz, P. M. (1996b). The stability of the covariances of international property share returns. *Journal of Real Estate Research*, *11*(2), 149–158.

Elton, E. J., & Gruber, M. J. (1977). Risk reduction and portfolio size: An analytical solution. *The Journal of Business*, *50*(4), 415–437.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics*, *20*(3), 339–350.

Erb, C. B., Harvey, C. R., & Viskanta, T. E. (1994). Forecasting international equity correlations. *Financial analysts journal*, *50*(6), 32–45.

Eun, C. S., & Resnick, B. G. (1988). Exchange rate uncertainty, forward contracts, and international portfolio selection. *The Journal of Finance*, *43*(1), 197–215.

Fabozzi, F. J. (2007). *Fixed income analysis* (Vol. 6). John Wiley & Sons.

Fama, E. F. (1976). *Foundations of finance: portfolio decisions and securities prices*. Basic Books (AZ).

Ferson, W. E., Foerster, S. R., & Keim, D. B. (1993). General tests of latent variable models and mean-variance spanning. *the Journal of Finance*, *48*(1), 131–156.

Ford, J. (2001). *Asset allocation: Balancing financial risk.* Financial Planning Association.

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica: Journal of the Econometric Society*, 1121–1152.

Glabadanidis, P. (2017). An exact test of the improvement of the minimum variance portfolio. *International Review of Finance*. Retrieved from https://onlinelibrary.wiley.com/doi/abs/10.1111/irfi.12173 doi: 10.1111/irfi.12173

Goetzmann, W. N., Li, L., & Rouwenhorst, K. G. (2005). Long-term global market correlations. *Journal of Business*, *78*(1).

Grauer, R. R., & Hakansson, N. H. (1987). Gains from international diversification: 1968–85 returns on portfolios of stocks and bonds. *The Journal of Finance*, *42*(3), 721–739.

Grubel, H. G. (1968). Internationally diversified portfolios: welfare gains and capital flows. *The American Economic Review*, *58*(5), 1299–1314.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica: Journal of the Econometric Society*, 1029–1054.

Hartmann, P., Straetmans, S., & de Vries, C. (2001). *Extreme market linkages in crisis periods.*

Hendricks, D. (1997). Evaluation of value-at-risk models using historical data. *Economic Policy Review*, *2*(1).

Hight, G. N. (2010). Diversification effect: Isolating the effect of correlation on portfolio risk. *Journal of Financial Planning*, *23*(5), 54.

Hillier, D., Ross, S., Westerfield, R., Jaffe, J., & Jordan, B. (2013). *Corporate finance*. McGraw Hill.

Huberman, G., & Kandel, S. (1987). Mean-variance spanning. *The Journal of Finance*, *42*(4), 873–888.

Hyde, S., Bredin, D., & Nguyen, N. (2007). Chapter 3 correlation dynamics between asia-pacific, eu and us stock returns. In *Asia-pacific financial markets: Integration, innovation and challenges* (pp. 39–61). Emerald Group Publishing Limited.

Isogai, T. (2016). Building a dynamic correlation network for fat-tailed financial asset returns. *Applied Network Science*, *1*(1), 7.

Jennrich, R. I. (1970). An asymptotic $\chi 2$ test for the equality of two correlation matrices. *Journal of the American Statistical Association*, *65*(330), 904–912.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of finance*, *23*(2), 389–416.

Jobson, J. D., & Korkie, B. (1989). A performance interpretation of multivariate tests of asset set intersection, spanning, and mean-variance efficiency. *Journal of financial and quantitative analysis*, *24*(2), 185–204.

Jorion, P. (1985). International portfolio diversification with estimation risk. *Journal of Business*, 259–278.

Junior, L. S., & Franca, I. D. P. (2012). Correlation of financial markets in times of crisis. *Physica A: Statistical Mechanics and its Applications*, *391*(1-2), 187–208.

Kan, R., & Zhou, G. (2008). Tests of mean-variance spanning. (Available at SSRN: https://www.ssrn.com/abstract=231522)

Kaplanis, E. C. (1988). Stability and forecasting of the comovement measures of international stock market returns. *Journal of international Money and Finance*, *7*(1), 63–75.

King, M. A., & Wadhwani, S. (1990). Transmission of volatility between stock markets. *The Review of Financial Studies*, *3*(1), 5–33.

Koch, P. D., & Koch, T. W. (1991). Evolution in dynamic linkages across daily national stock indexes. *Journal of International Money and Finance*, *10*(2), 231–251.

Lagoarde-Segot, T., & Lucey, B. M. (2007). International portfolio diversification: Is there a role for the middle east and north africa? *Journal of Multinational Financial Management*, *17*(5), 401–416.

Lee, C.-F., & Lee, J. (2010). *Handbook of quantitative finance and risk management*. Springer Science & Business Media.

Levy, H., & Post, T. (2005). *Investments*. Pearson Education.

Levy, H., & Sarnat, M. (1970). International diversification of investment portfolios. *The American Economic Review*, *60*(4), 668–675.

Longin, F., & Solnik, B. (1995). Is the correlation in international equity returns constant: 1960–1990? *Journal of international money and finance*, *14*(1), 3–26.

Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *The journal of finance*, *56*(2), 649–676.

Madura, J., & Soenen, L. (1992). Benefits from international diversification: Across time and country perspectives. *Managerial Finance*, *18*(2), 1–14.

Markowitz, H. (1952). Portfolio selection. *The journal of finance*, *7*(1), 77–91.

Model, A. R. E. E., Ribeiro, R., & Veronesi, P. (2002). The excess comovement of international stock markets in bad times.

Omelka, M., & Pauly, M. (2012). Testing equality of correlation coefficients in two populations via permutation methods. *Journal of Statistical Planning and Inference*, *142*(6), 1396–1406.

Orskaug, E. (2009). *Multivariate dcc-garch model:-with various error distributions* (Unpublished master's thesis). Institutt for matematiske fag.

Peters, T. (2008). *Forecasting the covariance matrix with the dcc garch model*. Citeseer.

Pfau, W. D., et al. (2007). Asset allocation for the pakistan pension system: A role for international diversification? *GRIPS Policy Information Center Discussion Paper*, 07–06.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica: Journal of the econometric society*, 431–449.

Schindler, F. (2009). Correlation structure of real estate markets over time. *Journal of Property Investment & Finance*, *27*(6), 579–592.

Solnik, B., Boucrelle, C., & Le Fur, Y. (1996). International market correlation and volatility. *Financial analysts journal*, *52*(5), 17–34.

Solnik, B. H. (1995). Why not diversify internationally rather than domestically? *Financial analysts journal*, *51*(1), 89–94.

Stephen, L. (2005). The return due to diversification of real estate to the us mixed-asset portfolio. *Journal of Real Estate Portfolio Management*, *11*(1), 19–28.

Sukumaran, A., Gupta, R., & Jithendranathan, T. (2015). Looking at new markets for international diversification: frontier markets. *International Journal of Managerial Finance*, *11*(1), 97–116.

Tang, G. Y. (1998). The intertemporal stability of the covariance and correlation matrices of hong kong stock returns. *Applied financial economics*, *8*(4), 359–365.

Upadhyay, S. K., Singh, U., Dey, D. K., & Loganathan, A. (2015). *Current trends in bayesian methodology with applications*. CRC Press.

Von Furstenberg, G. M., Jeon, B. N., Mankiw, N. G., & Shiller, R. J. (1989). International stock price movements: links and messages. *Brookings Papers on Economic Activity*, *1989*(1), 125–179.