Aalto University
School of Economics

Diversification benefits for bond portfolios - Evidence from the euro-denominated investment grade bond market

Economics

Master's thesis

Katariina Tanhua-Tyrkkö

2011

Department of Economics
Aalto University
School of Economics

AALTO UNIVERSITY SCHOOL OF ECONOMICS ABSTRACT
Master's thesis in Economics November, 2011
Katariina Tanhua-Tyrkkö

DIVERSIFICATION BENEFITS FOR BOND PORFOLIOS - Evidence from the euro-denominated investment grade bond market

OBJECTIVES
The study examines diversification benefits for euro-denominated investment grade bond portfolios. The objective of the study is to investigate the minimum portfolio size at which the marginal benefits measured in terms of risk reduction or reward-to-risk improvement from additional diversification become very small. It is examined how such diversification benefits and minimum portfolio sizes vary when differentiating the investment opportunity sets by the sector of the issuer. Furthermore, the study investigates whether diversification benefits seem to have varied in different time periods in 2002 – 2010. The study also gives an overview of European bond markets and the characteristics of bonds as financial instruments and presents measures that can be used to evaluate portfolio performance and diversification benefits.

DATA AND METHODOLOGY
The bond sample is extracted from the Bank of America Merrill Lynch Global Index System database. The data consists of monthly returns of bonds belonging to the Merrill Lynch Euro Broad Market Index. The data is collected from 01/2002 – 12/2010. Also bond-specific descriptive data on the bond sector and domicile of the issuer of the bond was collected. Three metrics are used in the study to examine diversification benefits. The measures used are excess standard deviation or mean derived deviation, Sharpe ratio and cross-sectional kurtosis.

The study was conducted by first completing the return data of the bonds, majority of which exiting the index during the period under review due to their time to maturity rolling below 1 year, into a table format with no empty observations. The same was repeated for the set of all bonds as well as for each sector and each time sub-period separately so as to allow sector- and time-specific examination. The study is implemented by selecting bonds randomly in order to create 1000 portfolios for each investment opportunity set and each portfolio size. A portfolio size ranging from 2 to 100 is examined for each set and the optimal portfolio size is thought to have been achieved when the marginal improvement in the measure decreases below 1%.

RESULTS
The minimum portfolio size is found to vary by issuer type, time period under review and the metric used to measure the marginal benefits of further diversification. While the marginal benefits of further diversification are generally achieved with PSs of 25-55 bonds when using the mean derived deviation or Sharpe ratio as the criteria, the untapped benefits of full diversification are still sizeable. When examining the Sharpe ratio, with the optimal portfolio size the unrealized diversification benefits are still over 50%. When using the MDD metric the overall diversification benefits range between 63 and 95 percent. When examining kurtosis the required portfolio size seems to be higher, mostly in the range of 55 to 80 bonds, at which overall diversification benefits are roughly 68 to 92 percent.

KEYWORDS
Bond, Diversification, Investment Grade credit rating, performance measure

AALTO-YLIOPISTON KAUPPAKORKEAKOULU TIIVISTELMÄ
Taloustieteen pro gradu -tutkielma Marraskuu, 2011
Katariina Tanhua-Tyrkkö

DIVERSIFICATION BENEFITS FOR BOND PORFOLIOS - Evidence from the euro-denominated investment grade bond market

TAVOITTEET

Tutkielma käsittelee investment grade -luottoluokitelluista euromääräisistä joukkovelkakirjalainoista muodostetun salkun hajautushyötyjä. Tarkoituksena on tutkia, millä määrällä velkakirjoja lisähajautuksen tuoma rajahyöty laskee hyvin pieneksi. Hajautuksen marginaalihyötyjä mitataan kolmella eri mittarilla. Lisäksi tutkitaan, poikkeavatko eri sektorien liikkeeseenlaskijoiden joukkovelkakirjojen hajautushyödyt toisistaan. Tutkimuksessa tarkastellaan myös, ovatko hajautushyödyt poikenneet toisistaan eri ajanjaksoina vuosina 2002 – 2010. Lisäksi tutkielma esittelee lyhyesti eurooppalaista joukkovelkakirjamarkkinaa ja joukkovelkakirjalainojen erityispiirteitä rahoitusinstrumentteina. Tutkielma esittelee myös portfolion suoritumismittareita, joilla voidaan arvioida salkun suoriutumista sekä hajautushyötyjä.

AINEISTO JA SEN KÄSITTELY

Tutkielman aineisto koostuu Bank of America Merrill Lynchin Euro Broad Market -indeksiin kuuluvien joukkolainojen kuukausittaisista tuottohavainnoista. Myös lainakohtaiset ominaisuudet (toimialasektori ja liikkeeseenlaskijan kotimaa) kerättiin Merrill Lynchiltä. Aineisto on kerätty 01/ 2002 – 12/2010. Tutkimus toteutettiin täydentämällä aineisto ensin taulukkomuotoon ilman tyhjiä havaintoja siten, että jokaisen tarkastelujaksolla erääntyneen lainan tuottohavaintosarjan jatkeeksi arvottiin toinen laina, joka ei ollut vielä erääntynyt. Sama toteutettiin sekä koko indeksin lainoille että sektorikohtaisille alaryhmille. Lisäksi periodi jaettiin kolmeen 3 vuoden mittaiseen tarkastelujaksoon. myös näille ajanjaksoille koottiin omat datataulukot, joille varsinaiset laskelmat ajettiin. Tutkielmassa käytettiin kolmea mittaria, jolla arvioitiin lainasalkun hajautuksesta saatavia hyötyjä ja pyrittiin identifioimaan minimisalkkukoko, jonka jälkeen lisähajautuksesta saatava marginaalihyöty laskee hyvin pieneksi. Kriteerinä käytettiin marginaalihyödyn laskemista alle yhden prosentin valituilla mittareilla. Tutkitut mittarit olivat ylimääräinen keskihajonta yli koko instrumenttiavaruudesta kootun salkun, Sharpen suhdeluku ja poikkileikkauskurtoosi. Hajautushyötyjen tarkastelu tapahtui arpomalla satunnaisesti 1000 salkkua eri salkkuko'oilla kahdesta (2) sataan (100). Optimaalinen salkkukoko katsottiin saavutetuksi, kun käytetyn mittarin arvon paraneminen laski alle yhden prosentin.

TULOKSET

Tämän työn merkittävin tutkimustulos on, että salkkukoko, jolla lisähajautuksen marginaalihyöty laskee hyvin pieneksi, vaihtelee riippuen velkakirjan sektorista, käytetystä hajautushyötymittarista sekä tarkasteluperiodista. Tulosten mukaan rajahyötykriteerillä riittävä hajautus saavutetaan noin 20 – 55 joukkolainalla, kun kriteerinä on ylimääräinen keskihajonta tai Sharpen suhdeluku. Sharpen suhdelukua käytettäessä toteutumatta jäävä hajautushyöty jää kuitenkin merkittäväksi (yli 50%). Poikkileikkauskurtoosia käytettäessä minimisalkkukoko saavutettiin yleisesti huomattavasti suuremmalla salkkukoolla, n. 55 – 80 velkakirjalla. Löydetyn minimisalkkukoon kokonaishajautushyödyt mitattuna sekä poikkileikkauskurtoosilla että ylimääräisellä keskihajonnalla olivat noin 63 –95 prosenttia.

AVAINSANAT

Joukkovelkakirjalaina, Hajautus, Investment grade -luottoluokitus, suoriutumismittari

Table of Contents

1. Introduction

1.1 Background and motivation for the study

Bond markets play a crucial role in the financial system. They bring lenders and borrowers together and are important in determining prices of other assets. Bank interest rates usually depend on bond rates determined in the market. The bond market also affects the saving decisions by households and investment decisions by corporations. In Europe, at year-end 2010 over three quarters (78%) of the outstanding amount of securities issued by euro area residents, excluding financial derivatives, were debt securities (ECB 2011).

The basis of modern portfolio theory is Markowitz' study (1952), which is a mathematical formulation of the concept of diversification in investing, i.e. forming portfolios with assets whose prices are less than perfectly positively correlated with each other and thus have collectively lower risk than any of the assets individually. A lot of studies have been done concerning asset allocation between asset classes, which is natural considering the consensus among most financial professionals is that asset allocation is one of the most important decisions that investors make. However, some investors are focused on one asset class only, such as an equity mutual fund on stocks or central banks on fixed income assets.

A lot of studies since Markowitz in the 1950s have concentrated on diversification in an equity context. Different studies reach different conclusions as to what is the required minimum number of assets in a portfolio in order to achieve a 'well-diversified portfolio'. Studies have found such portfolio sizes ranging anywhere from 8 to 10 (Evans and Racher, 1968; Latane and Young, 1969) to as many as hundreds of stocks (e.g. Statman, 2004, Statman and Scheid 2005).

Despite the importance of bond markets and the widely acknowledged importance of diversification there seems to have been rather few empirical studies investigating how many bonds are needed to achieve sufficient diversification. One such study has been conducted by McEnally and Boardman (1979) who investigate how many bonds are needed to achieve a certain level of diversification benefits in terms of overall volatility risk reduction. They find that eight to sixteen bonds notably reduce volatility when differentiating the investment opportunity sets by credit ratings.

Another, more recent study on diversification benefits for bond portfolios is the paper by Dbouk and Kryzanowski (2009). The authors examine the minimum portfolio size required to capture significant diversification benefits differentiating the investment opportunity set by issuer type, credit rating, and time-to-maturity. Their results indicate that these characteristics of bonds as well as the metric used to measure the marginal benefits of additional diversification determine the minimum portfolio size. According to Dbouk and Kryzanowski the marginal benefits of further diversification are in most cases optimized with portfolio sizes of 25 to 40 investment grade bonds.

In much the same manner that most diversification studies focus on either pure stock or combined stock – bond portfolios, the focus of research has to a large extent been on the US rather the euro-denominated market. This is understandable as the US bond market was long the largest and the most liquid bond market in the world. The structure of the international bond market saw a considerable change with the establishment of the European Monetary Union (EMU). In fact, the Euro government bond market has surpassed the US market in terms of size and number of issues (Fabozzi 2007). Since the Euro introduction the euro-denominated market experienced such a notable growth in size, scope and efficiency, that it appears interesting and meaningful to examine it more closely.

The focus of research has to a large extent been on the US rather the euro-denominated market and on equities rather than bonds. This study seeks to combine these areas that have been researched less and thus investigates diversification benefits in the European, euro-denominated bond market.

This study investigates diversification benefits for investment grade bonds and attempts to find a portfolio size at which marginal benefit from extra diversification becomes very small using various risk measures. As many a fixed income investor is focused on either investment grade or high yield credit quality debt instruments, it seems meaningful to focus only on one of them. This study deals with investment grade credit quality bonds. Furthermore, the data on investment grade bonds is likely to be more reliable than that for high-yield bonds (Biais et al. 2006). This study examines portfolios consisting of government and agency as well as private sector bonds, which seems plausible as many institutional investors invest in both types of bonds.

1.2 Objectives of the study

The purpose of this study is to examine the diversification benefits for euro denominated investment grade bond portfolios in the period between January 2002 and December 2010 using three risk and performance measures to assess the benefits.

The theoretical part of the study presents bonds as financial instruments in terms of the sector of the bond issuer and risks related to investing in bonds. It also presents various portfolio performance measures some of which are also suited for measuring diversification benefits. Furthermore, a brief literature review related to diversification in a bond portfolio context is provided.

The study follows the methodology used by by Dbouk and Kryzanowski (2009) and aims at identifying a portfolio size at which marginal benefit from extra diversification becomes very small (below 1%) when investing in euro-denominated investment-grade bonds. It is also investigated whether the choice of bond sample period has an impact on the optimal portfolio size, i.e. whether the diversification benefits have differed in different three-year periods in the 2000s.

1.3 Structure of the study

This thesis is organized in eight main sections. Sections from one to five form the theoretical part of the thesis. First, the introduction section provides the background of and the motivation for examining the topic of diversification benefits for bond portfolios and states the objectives of the study. The second section gives an overview of the European bond market focusing on the euro-denominated bond market created with the introduction of the single currency in 1999.

The third section provides an overview of bonds as financial instruments by giving an overview of different types of bonds differentiated by the sector of the issuer and and discusses the risks related to investing in bonds. The fourth part of the study presents various performance measures for portfolio management. The fifth and the last theoretical section provides a literature review on diversification with a focus on the bond portfolio context.

The sixth and seventh sections form the empirical part of the study. In section six the data, sample and methodology are presented, while section seven presents the results of the analysis. The eighth and last section summarizes the findings and presents conclusions.

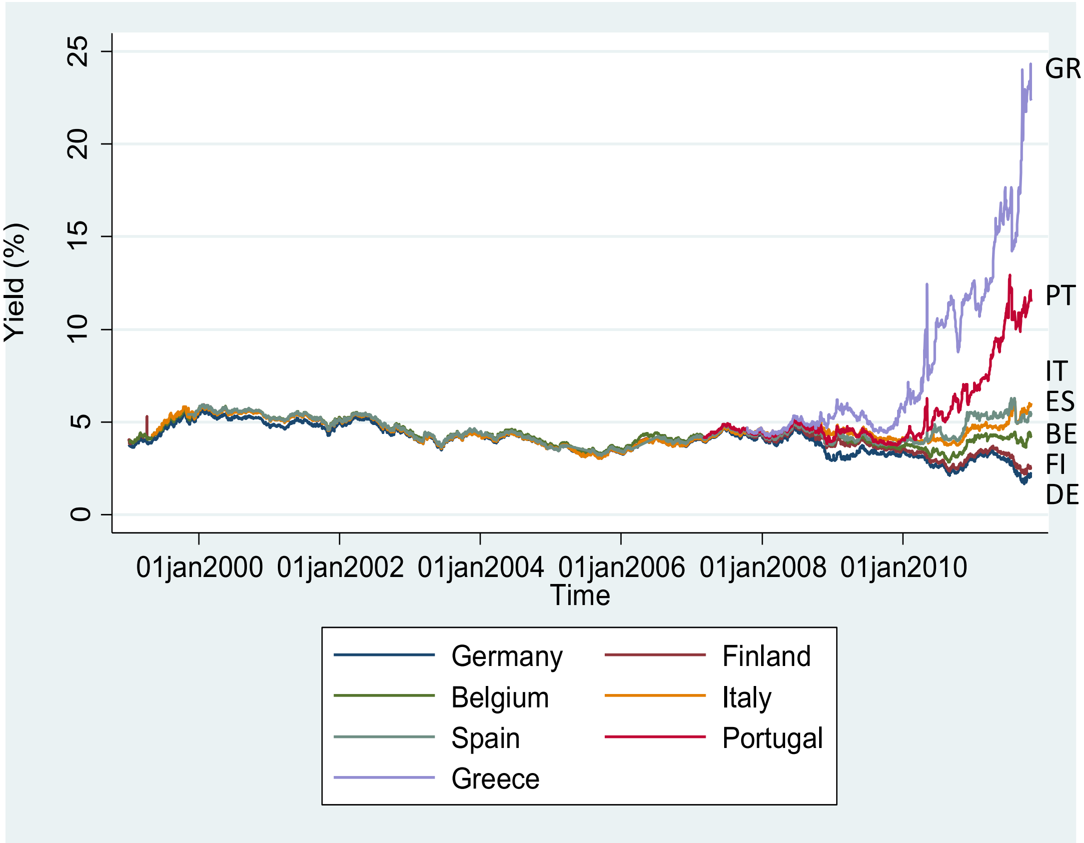
2. European bond markets

The US market was long the largest and the most liquid bond market in the world. Before the introduction of the euro in 1999 the bond markets of the current European Monetary Union (EMU) countries were disconnected and limited in size and scope. The pools of potential investors were also rather small. The euro introduction led to a dramatic growth and development in European bond markets. Before the euro corporations in need of financing often either borrowed directly from banks or issued US dollar denominated debt to attract a larger investor base (Biais et al. 2006). The creation of the euro led to a new institutional setting which combines a centralized monetary policy with decentralized fiscal and structural policies (Abad et al. 2009).

According to the Bank for International Settlements (2011), at the end of 2010 the outstanding amount of domestic debt securities of the eurozone countries amounted to 28 367 billion euro whereas the corresponding figure for the US was 8 346 billion euro. It should be noted that in the context of the euro bond market all euro-denominated issues by entities domiciled in the euro area are considered domestic, although the issuers may reside in several different eurozone countries. Furthermore, non-eurozone issuers also choose to issue securities denominated in euro to attract investor demand from the eurozone.

The Association for Financial Markets in Europe (2011) has stated that although individual countries in Europe still have domestic bond markets where governments, sub-sovereign entities and corporations in residence issue bonds, and individual investors participate, the European bond markets are increasingly acting like a single market. This view seemed indeed plausible for almost the first 10 years of the existence of the monetary union; corporate bond yields and spreads relative to government bond yields showed a declining trend and low interest rates and abundant liquidity had contributed to a high risk appetite among investors.

Figure 1. Generic 10 year government bond yields of some eurozone countries (1999 – 2010)



Source: Bloomberg generic government rates, accessed 25 Oct 2011

The period before the global financial crisis can now be characterized as underpricing of risk (Johnson 2007) and even more evidently the so-called European sovereign debt crisis which came to a head in 2010 and continues critical as of this writing have shown that the eurozone is still in many ways far from a single financial market despite its common currency.

Within the euro-denominated financial markets bond markets play an important role. According to the European Central Bank (2011) at year-end 2010 78% of the outstanding securities (excluding financial derivatives) issued by euro area residents were debt securities and an overwhelming majority of those were denominated in euro. According to the Bank for International Settlements (2011), at the end of 2010 the outstanding amount of domestic debt securities of the US was 8 346 billion euro whereas the corresponding figure for the Eurozone was €15 873 billion (ECB 2011).

Table 1. Securities issued by euro area residents (2008 – 2010), amounts outstanding

Year	Total	Debt securities					Equity securities
			of which:			of which:	
			Short-term	Long-term fixed rate	Long-term variable rate	issued in euro (%)	Quoted shares
2008	16953	13444	1627	7710	3594	90	3510
2009	19685	15276	1638	8829	4371	89	4409
2010	20464	15873	1530	9513	4369	89	4592

Source: European Central Bank Pocket Book March 2011: Amounts are outstanding amounts in EUR billions at year-end.

In Europe, bond markets are dominated by government bond markets and bonds issued by financial intermediaries whereas in the US the proportion of bonds issued by the non-financial corporate sector is much larger (Biais et al. 2006). In addition, the tax exempt municipal bonds as well as agency bonds play a major role in the US role.

A major reason behind the much less academic attention devoted to bond markets is probably data availability. Not only is the bond market less researched than the stock market, it is also very different. While the bulk of stock trading is conducted on electronic limit-order books (such as e.g. Xetra or Nasdaq), bonds are to a large extent traded over-the-counter (OTC) and often rely on bilateral non-anonymous communication, especially in the case of corporate bonds. The European corporate bond market, like its US counterpart, revolves around dealers and brokers. Since it is a dealer market, the European bond market is decentralized. This market is self-regulated by the International Capital Market Association or ICMA (Biais et al. 2006). While some dealers are based in continental Europe (e.g. in Zurich, Paris or Madrid), most operate from London. All the London-based ICMA members and all members of the Council of Reporting Dealers, have to report their trades to this self-regulatory organization, through a system called TRAX. The system captures most of the professional business in continental Europe and the UK (Biais et al. 2006). The information goes to national regulators such as the Financial Services Authority FSA in the United Kingdom or the National Bank of Belgium and so forth. The role of regulation develops all the time and its future evolvement will depend also on the regulatory ramifications of the ongoing financial crisis, as there have been calls to increase and enhance supervision and regulation in the financial market all over the developed world.

depends is and will probably be affected by the recent financial crisis. Short-selling of bonds, for example, is difficult and costly especially in Europe.

A study by Center for Economic Policy Research CEPR (Biais et al. 2006) finds that euro-denominated corporate bonds have tighter spreads than English pound sterling denominated bonds and also compare favourably to that of post-trade data of US counterparts. The study maintains that this is consistent with the presence of a large pool of potential buyers and sellers in the unified euro-area market. It is worth noting, however, that the findings might be different if a similar study would be conducted now when the global financial market but especially that of the eurozone is in a state of distress. Yet, interviews with market participants conducted by the CEPR for their study (Biais et al. 2006) suggests that, in the US for dollar-denominated bonds or in the UK for sterling ones, the number of active dealers is typically around five or six, while for euro-denominated bonds it is much larger. Some liquidity problems are specific to the bond market, as for example institutional investors such as insurance companies sometimes buy a large fraction of an issue and hold it until maturity, thereby reducing the liquidity of that issue.

Direct holdings of fixed income instruments by households vary considerably across countries in Europe. In Italy, Belgium and Germany household direct holdings of fixed income securities are significant, for example as high as 20% of total financial holdings in Italy and between 10 and 15 percent in Germany (Biais et al. 2006), but in other European countries such investments take place primarily through funds.

3. Bonds as financial instruments and bond portfolio management

There are several important differences between the microstructure of the bond market and that of the stock market, many of which are related to the characteristics of bonds as financial instruments. Bonds tend to attract a specific type of investors, many of which are institutional investors following a buy-and-hold strategy. The differences among bonds, however, are perhaps less marked than those among different stocks. A typical way to categorize bonds is to do it according to the type of the entity which issues it. Often an investor may desire to trade a certain category of bond, say an investment grade rated corporate bond. The following subsection provides such a categorization of bonds differentiated by the type of their issuer. The subsection after that presents an overview of risks related to investing in bonds.

3.1 Classification by the sector of the bond

A key feature of a bond is the nature of its issuer, also called the debtor or borrower. Bonds are sold in the primary market by government units, municipalities, or companies. Government bond issues are sold in the primary market in government bond auctions. New municipal bonds in the US are sold by one of three methods: competitive bid, negotiation, or private placement (Reilly & Brown, 2003). The issuance of municipal bonds varies country by country. Corporate bond issues, in turn, are almost always sold through a negotiated arrangement with a so-called underwriter who originates the bond, acquires the total bond issue and distributes, i.e. sells it to the investors.

Investing in bonds Europe (2011), a partnership between the Association for Financial Markets in Europe (AFME), European AFME members and European market maker banks, categorizes European bond markets in bond market sectors: government, sub-sovereign, corporate and mortgage-backed or asset-backed collateralised bonds. Strictly speaking there is overlap in these categorization, as collateralized bonds tend to be issued by financial corporations, but the categorization is also based on the risk profile of the bond. A stylized categorization based on the InvestinginbondsEurope.org classification and market practice can be defined as follows:

Table 2. A stylized classification of bonds based on the sector of the issuer

1	Government bonds	
2	SSA (sub-sovereign, supranational and agency) bonds	
		Sub-sovereign
		Supranationals
		Agency bonds
3	Corporate bonds	
4	Collateralized bonds	
		Structured debt
		Covered bonds

Based on Investing in bonds Europe (2011), Fabozzi (2007), Reilly and Brown (2003)

3.1.1 Government bonds

A government bond is a bond issued by a national government in the country's own currency. Bonds issued by national governments in foreign currencies are typically called sovereign bonds (Fabozzi 2007). Government bonds have traditionally been referred to in financial theory as being risk-free, as a government can raise taxes or the national central bank can issue more money so that the bond can be redeemed at maturity. The concept of being risk-free thus refers to free of credit risk. However, some countering examples exist. For example Russia has defaulted on its domestic currency debt during the so-called ruble crisis in 1998. Even more recently the global financial crisis since 2007 and especially its latest phase, often referred to as the European sovereign debt crisis or euro crisis as of this writing, has brought the default risk of developed countries back into speculation. The crisis that culminated in 2010 (Blackstone et al. 2010, Wall Street Journal) has further challenged the concept of risk free. The confidence of investors in the ability of some governments to meet their debt obligations was shaken; this led to a widening of the bond yield spreads between these countries and so-called core EU member countries[1], most importantly Germany. Yet, a common perception of government bonds is still to regard them as the 'most risk-free' asset class within the fixed income universe.

3.1.2 Sub-sovereign, supranational and agency (SSA) bonds

An investor can buy various securities issued by government-owned or government-sponsored organizations that are strictly speaking not a part of the government. Such agencies or government related entities are affiliated with, but separate from the government (Fabozzi, 2007). A common practice among financial professionals is to classify such bonds as sub-sovereign, supranational and agency (SSA) bonds. Sovereign bonds, i.e. bonds issued by national governments but in foreign currencies are sometimes classified as belonging to SSA bonds. Government-issued foreign-denominated bonds entail an additional risk factor compared to domestic currency denominated bonds: if a government experiences difficulties in its ability to pay its debt, it has (at least in theory) the ability to print more money and thus ease its domestic currency denominated debt burden. For foreign-currency denominated debt,

[1] Eurozone countries with high credit quality. The exact definition of core varies. Core can be seen to include only Germany or to include other countries with high credit quality, such as other AAA-rated countries (Germany, France, the Netherlands, Austria, and Finland as of 30 September 2011 by Standard & Poor's)

however, such an option is not available except through a currency swap transaction, which may not be a plausible option if the market does not trust the country in question.

The exact definition and understanding of differences in different SSA bonds varies and especially the terms sub-sovereign and agency are often used in a rather loose manner. According to the definition by the Association for Financial Markets in Europe (AFME 2011), the sub-sovereign bond market can be defined first as any level of government below the national or central government, (including regions, provinces, states, municipalities, etc.) that issues bonds. In addition to sub-sovereign and agency entities and foreign currency denominated sovereign issues also intergovernmental organizations issue debt. Such debt issues are usually referred to as supranational bonds.

Sub-sovereign bonds

The Association for Financial Markets in Europe (AFME 2011) notes that although central governments in Europe have been more likely to issue bonds than sub-sovereign entities in Europe, sub-sovereign entity participation in the bond markets is expanding dramatically in part because effective government needs to finance day to day operations of public services and capital infrastructure investments at the same time as some sovereign governments have already high debt burden and sometimes explicit debt ceiling limits. Government and sub-sovereign bond markets thus exist side by side as different and serve as independent sources of financing for the national or central government as well as for the local governments. Examples of sub-sovereign issuers are local governments such as the states (e.g. the German state of Baden-Württemberg), provinces (e.g. British Columbia in Canada), municipalities and cities. In the US municipal bonds issued by states, cities, or other political subdivisions to raise funds (Reilly and Brown 2003) play an important role. In the United States, interest income received by holders of municipal bonds is often exempt from the federal income tax and from the income tax of the state in which they are issued.

Provincial or state governments issue debt, depending on their constitutional ability to do this. Most investors consider provincial or state issuers to be very strong credits because they have the power to levy income, property and sales taxes to support their debt payments (Reilly and Brown, 2003). Yet, since they cannot control monetary policy like national governments, they are considered poorer credit quality than national governments.

Supranational bonds

Supranational bonds are bonds issued by supranational organizations where governments or member states have delegated power to the organization. The only union generally truly recognized as having achieved the status of a supranational union is the European Union (Bauböck 2007), but in financial marketcontext the term supranational is generally used to refer to international organizations which levy assessments or fees against its member governments. Ultimately, like for sub-sovereign bonds, it is this support and the taxation power of the underlying national governments that allow these organizations to make payments on their debts. Examples of such organizations include the World Bank Group with its subsidiaries and the European Investment Bank (EIB). Single issuers tend to constitute a large part of the overall market. For example, the EIB forms a considerable part of the European supranational and the whole SSA market: to illustrate, the weight of the EIB in the Merrill Lynch euro-denominated government-related (SSA) index was as high as 11.6% on 31 September 2011 (Bank of America Merrill Lynch Global Index System 2011)

Agency bonds

Agency bonds are bonds issued by government agencies that are traditionally seen as very secure, but their debt is also not considered direct obligation of the government. Since they are indirectly or directly guaranteed by the government, these bonds are called quasi-government or sub-sovereign or agency bonds. Such agencies can be publicly or privately owned. Examples of such issuers in Europe includes the German special purpose bank KFW (formerly Kreditanstalt für Wiederaufbau), a domestic development bank type of an agency making makes housing and small business development loans. The US government has had a mixed strategy, guaranteeing some mortgage bonds (GNMA, Ginnie Mae) and not others (FNMA, Fannie Mae and FHLMC Freddie Mac) (AFME 2011). In the case of Fannie Mae and Freddie Mac, however, the concerns about the ability of the agencies to raise capital and debt threatened to disrupt the U.S. housing financial market, so the US Treasury explicitly placed Fannie Mae and Freddie Mac into conservatorship in September 2008. Hence, the view of agency bonds as almost risk free has been under reconsideration during the financial crisis.

3.1.3 Corporate bonds

The major nongovernmental issuer of debt is the corporate sector. The importance of the sector differs vastly among countries. By far the largest corporate issuer type are financial corporations, but also the amount of non-financial bonds outstanding was almost 2,300 billion euros at the end of September 2010. Corporate bonds represent an especially large portion of the US market but a rather small one in Japan where corporate financing has traditionally been more loan based and the corporate funding more reliant on bank debentures (Reilly and Brown 2003) and the financial system far more bank-based. In Europe the role of corporate bond issuance in company funding has been increasing as in the past. Before the introduction of the euro corporations often either borrowed directly from banks or issued US dollar or British pound denominated debt in order to attract a larger investor base (Biais et al. 2006).

Corporate bonds are debt obligations issued by industrial, financial and service companies to finance capital investment and operating cash flow. Corporate bonds are typically further subdivided into industrial, utility and financial corporate bonds. Sometimes also transportation sector corporate bonds are considered a category of its own (Reilly & Brown, 2003). The corporate sector provides very diverse issues that can range from the highest investment-grade credit-rated firms to very high-risk bonds that are rated non-investment grade. Although, as discussed previously, government bonds are typically viewed as being the least risky fixed income instruments there are, according to some recent views (such as Carver, 2010), high grade, non-financial corporate bonds would have even replaced sovereign bonds as a safe haven for insurers and pension funds previously invested in the troubled sovereign bonds of the European periphery. It has been argued that many companies possess strong, verifiable balance sheets and ample cash positions whereas many government are heavily indebted and even the reliability of their national accounts may also prove questionable. However, the credit spread widening faced by majority of corporations during the financial crisis of the late 2000s shows that this view has not translated in extremely low spreads except for some selected investment grade companies and mostly for short maturities only.

3.1.4 Collateralized bonds

Structured and securitized debt

Many different types of debt obligations are to at least some extent structured. A structured debt instrument usually refers to any type of an obligation that is not a straightforward secured or unsecured government or corporate obligation (Investing in bonds Europe 2011). These types of transactions are usually but not always issued through special purpose vehicles or SPVs. Securitizations are one type of structured product. Securitization is the process of creating securities by pooling together various cash-flow producing financial assets. These securities are then sold to investors. Securitization, in its most basic form, is a method of financing assets (Reilly and Brown 2003). In principle any asset may be securitized as long as it is cash-flow producing. Some bond issue indentures (contracts between the issuer and the bondholder) specify that the issuer has to retire a predetermined amount of the issue periodically. The two major types of amortizing securities are mortgage-backed securities (MBS) and asset-backed securities (ABS) (Fabozzi, 2007). Another type of structured product refers to a packaging or repackaging of bonds together with various types of interest rate swaps and/or credit derivatives to change the interest and principal payment stream, in order to provide an investor with a particular risk profile that they want (Investing in bonds Europe 2011).

Covered bonds

Covered bonds are bonds backed by cash flows from mortgages or public sector loans and are thus in many ways similar to ABS and MBS created in securitization, but as opposed to those of ABS and MBS, covered bond assets remain on the issuer's consolidated balance sheet which is a major advantage from the investor's point of view; the interest is paid from an identifiable source of projected cash flow versus out of other financing operations (Reilly & Brown, 2003). A covered bond is in effect a financial corporate bond with an enhancement in the form of entitlement to an asset pool securing or "covering" the bond, should the originator (usually a financial institution) become insolvent. Hence, covered bonds are typically assigned a very high credit rating. Issuers of covered bonds must ensure that the pool consistently backs the covered bond. In the event of default, the investor has recourse to both the pool and the issuer.

3.2 Risks associated with investing in bonds

A bondholder is exposed to several kinds of risks. Market risk, often referred to as systematic risk refers to the risk resulting from movements in market prices. In addition to being exposed to market risk especially in the form of interest rate changes, an investor's risk position is propagated by other forms of financial risks. These risks include credit risk, liquidity risk, foreign exchange risk, event risk, industry risk, inflation risk, reinvestment risk, call risk, volatility risk, tax risk and risk risk (Fabozzi 2007, Reilly and Brown 2003), among others.

In financial theory systematic risk or market risk cannot be diversified away as by definition it affects all investments. Hence, it is sometimes called undiversifiable risk or aggregate risk. A specific risk, in contrast, is a risk affecting a small number of assets. In fact, most unsystematic risk is removed by holding a diversified portfolio and it is sometimes called diversifiable or residual risk. An interesting phenomenon is that during times of financial distress or crises the correlation of many normally-uncorrelated asset classes suddenly rises. Thus, even asset classes the returns of which normally bear little resemblance to each other are inclined to all decrease simultaneously. Silvapulle and Granger (2001) find that stock diversification benefits diminish with large negative movements in overall stock returns. Demirer and Lien (2004) attribute this to a heightened firm-specific return volatility when market returns are largely negative. Sancetta and Satchell (2007) state that correlations are of a bigger importance during market downturns, which has to do with the market's tail distribution.

This section provides an overview of risk factors for a fixed income investor, starting from market or interest rate risk and proceeds to give a general picture of other risks.

3.2.1 Interest rate risk

The price of a plain vanilla bond will change in the opposite direction from an interest rate change. As the interest rate rises, the price of a bond falls, and vice versa. Duration is a composite measure of the interest rate sensitivity of a bond. The concept and its development as a tool in bond analysis and portfolio management have existed for over 70 years (Macauley, 1938). Several specifications of duration have been derived over the course of decades. First, Macauley duration (Macaulay, 1938) is a measure of time flow of cash from a bond. Second, a modified version of Macaulay duration intended to indicate the price volatility of a bond in response to interest rate changes. Third, effective duration is a direct measure of the interest rate sensitivity of a bond in cases where price changes are possible to estimate using a valuation model. Finally, empirical duration measures directly the percentage price change of a bond (or other asset) for an actual change in interest rates. The popularity of effective duration and empirical duration have grown with the development of many new financial instruments which have very unique cash flows that change with interest rates (Le Sourd 2007).

The relationship between the price and interest rate is convex, which means that a graph depicting the relationship is a convex curve as opposed to a straight line. The convexity arises from the fact that as rates change, the weights given to the timing of cash flows change disproportionately to the change in rates (Grantier 1988). As rates rise, nearer coupons have a greater influence than distant coupons. The opposite holds when rates fall. The more a bond's duration changes for a given change in interest rates, the greater its convexity. Convexity protects a bond investor as a high-convexity bond will outperform an otherwise similar but low-convexity bond.

The interest-rate sensitivity of the bond price depends on various features of the bond, such as coupon, maturity or any options embedded in the issue (Fabozzi 2007). Interest rate risk can be further divided into level risk, which refers to the effect on the bond's value due to shifts in the interest rate level, and yield curve risk, which refers to the effect on the bond value due to changes in the shape of the yield curve. Interest rate risk of a bond is also referred to as market risk which has to do with the fact that it is a risk only if the investor wants to sell the bond in the secondary market or has to mark their positions to market.

3.2.2 Credit risk

Credit risk may be defined as the potential that an obligor is unwilling or unable to meet his financial obligations in a timely manner. Credit risk consists of three types of risk: default risk, credit spread risk, and downgrade risk (Reilly & Brown, 2003). Default risk is the risk that the issuer of a bond will fail to satisfy the terms of the obligation with respect to the timely payment of interest and repayment of the amount borrowed (Fabozzi 2007). Important in this respect are the so-called credit ratings assigned to a bond issue (see Table 3.).

Unicredit Bank (2011) takes a slightly different, a more of an institutional investor approach to classifying credit risk; the bank classifies credit risk components into credit default, concentration and country risk. Credit default risk refers to the risk of loss when the bank considers that the obligor is unlikely to pay its credit obligations in full or the obligor is more than 90 days past due on any material credit obligation; default risk may impact all credit-sensitive transactions, including loans, securities and derivatives. Concentration risk, in turn, is the risk associated with any single exposure or group of exposures with the potential to produce large enough losses to threaten the bank's core operations. It may arise in the form of single name concentration or industry concentration. Finally, country risk refers to the risk of loss arising when a sovereign state freezes foreign currency payments (transfer/conversion risk) or when it defaults on its obligations (sovereign risk).

Even in the absence of default an investor is concerned that the market value of a bond issue will decline in value and/or the relative performance to some benchmark will worsen. The yield of a bond consists of (1) the yield on a similar maturity government issue and (2) a premium to compensate the investor for bearing the risks that are absent in government issues. This premium is called the yield spread or credit spread and the risk that a bond price declines because of an increase or widening of the spread is called credit spread risk. Studies have shown that the main source of credit risk for investment grade bonds are related to specific credit events, while in the high yield category credit risk is mainly driven by recession. Interestingly, although diversification diversifies the credit risk of the portfolio in terms of a potential loss, a large number of bonds in the portfolio may increase the probability of at least one default or credit event taking place.

Table 3. Credit rating classes according to Standard & Poor's, Moody's Investors Service and Fitch

S&P		Moody's		Fitch		
Long-term	Short-term	Long-term	Short-term	Long-term	Short-term	Credit risk (Moody's characterization)
AAA		Aaa		AAA		Minimal
AA+		Aa1		AA+		
AA		Aa2	P-1	AA		Very low
AA-	A-1+	Aa3		AA-	F1+	
A+		A1		A+		
A	A-1	A2		A	F1	Low
A-		A3		A-		
BBB+	A-2	Baa1	P-2	BBB+	F2	
BBB		Baa2		BBB		Moderate
BBB-	A-3	Baa3	P-3	BBB-	F3	
BB+		Ba1		BB+		
BB		Ba2		BB		Substantial
BB-	B	Ba3		BB-		
B+		B1		B+		
B		B2		B		High
B-		B3	Not prime	B-	B	
CCC+		Caa1				
CCC		Caa2				Very high
CCC-	C	Caa3				
CC						In or very near default
C		Ca		CCC	C	In default with little prospect for recovery
		C		DDD		
	/	/		DD	/	In default
D		/		D		

Source: Moody's Investor Services (2011) Bankers Almanac Credit Rating Definitions (2011)

Credit ratings are evaluations of the creditworthiness of an issuer or a bond. If a bond issue is assigned a credit rating, the rating agency will monitor its credit quality. An improved credit quality is rewarded with an upgrade whereas a deterioration of credit quality is penalized with a downgrade. An unexpected downgrade of a bond or an issuer increases the aforementioned credit spread resulting in a price decline. This aspect of credit risk is called downgrade risk. Credit rating institutions have faced considerable critique during the financial crisis. They have been blamed for not downgrading companies promptly enough. Large corporate rating agencies have also been criticized for having too familiar a relationship with company management. The best-known rating agencies are Standard and Poor's, Moody's and Fitch.

Credit risk has traditionally been the single most important risk that differentiates corporate and government debt and the major reason why corporate bonds usually trade at a premium or spread to a similar maturity government bond. If market is assumed to be efficient enough for credit spreads to reflect all publicly available information, credit spreads themselves should be the best estimates for future returns. The financial crisis that began in 2007 has to some extent challenged this view. In 2010 and early 2011 the spreads of some Eurozone government bonds against German government bonds (called bunds) have skyrocketed to levels never seen before during the euro-era as previously depicted in Figure 1. Some very high credit quality corporations, in turn, have been able to get financing almost as cheaply as the governments in which they are domiciled. For example the US consumer products company Colgate-Palmolive issued bonds 3 November 2011 with record low spreads compared to similar-maturity Treasuries; at the time of issuance their 3-year notes yielded only 35 basis points more than their benchmark Treasuries. As noted also e.g. a recent Bloomberg article (Catts 2011), non-cyclical 'safe names' like Colgate benefit from the fact that investors are showing preference for the highest-rated corporate bonds. Furthermore, as the all the more increasingly uncertain economic and financial situation in Europe as of 2011 puts financial companies under pressure, diversified industrials can enter the market and refinance their debt coming due.

After issuance bond spreads to government bonds have been seen to occasionally become negative in the past few years. Bhanot and Guo (2011) use a sample of investment-grade corporate bond yields for the period September 2009 to September 2010 and find a total of 67 instances distributed among 10 companies where the credit spread is negative based on reported trade prices. As often happens with sought after bonds, their spreads tighten right after the issuance which was also the case with the Colgate bonds issue.

3.2.3 Liquidity risk

Liquidity risk, also called marketability risk refers to the ease with which an issue can be sold at or near its value. The primary measure of liquidity is the bid-ask spread i.e. the spread between the bid and ask quotes by a dealer. The wider the spread, the larger the risk. For an individual investor who plans and has the ability to hold a bond until maturity, liquidity risk is unimportant. In contrast, institutional investors have the obligation to mark their position to market, or simply mark to market, means that the portfolio manager must periodically determine the market value of each bond. To get prices that reflect market value, the bonds must trade with enough frequency (Fabozzi 2007).

Corporate bond market is relatively less liquid than government bond market. As the size of a market increases the liquidity becomes better, too. This is what happened in the euro-denominated bond market as well, although the current financial crisis has led to a worsened liquidity and at times, also to almost complete freeze of liquidity not only in the corporate but also in the government bond market (see Table 4.). At the end of October 2011, a 10-year German government bond bid-ask spread was a very tight 0.003% whereas the 10-year Greek bond yield spread was a staggering 1.947%. To contrast this with non-crisis times, CEPR (2006) found that AAA-rated corporate bonds tended to have a bid-ask yield spread of around 0.10 to 0.20 % in 2003 – 2005 and other investment grade rated bonds a bid-ask spread of around 0.25 to 0.40%. Furthermore for some of the worst-distressed eurozone sovereigns even the very wide bid-ask spreads quoted by market makers most probably do not even represent true levels at which bonds could be bought or sold.

Table 4. Comparison of bid-ask spreads of German and Greek 10-year government bonds at 10 November 2011

Country	Bond	Price		Price		Bid-ask yield spread
		Bid	Ask	Bid	Ask	
Germany	DBR 2.25 09/2021	101.985	102.01	2.025 %	2.022 %	0.003 %
Greece	GGB 6.25 06/2020	36.81	40.31	24.376 %	22.429 %	1.947 %

Source: Bloomberg generic pricing quotes

Credit risk and liquidity risk are considered to be the two main justifications for the existence of yield spreads above benchmark government notes and bonds. Liquidity risk is often name and instrument-specific of nature, but especially in market downturns larger portions or even whole markets can suffer from a liquidity freeze, in which cases the risk cannot be fully diversified away. Two issues from the same issuer may have very different liquidity. Bonds are at their most liquid immediately after issuance and rapidly lose their marketability as a larger proportion of the issues becomes locked into portfolios (Ericsson and Renault 2001).

3.2.4 Other risks

From the perspective of a euro investor, the cash flows of a non-euro-denominated bond are not known. The euro cash flows depend on the exchange rate at the time when the payments are received. If the currency in which the bond is denominated (say, US dollar) depreciates relative to the euro, fewer euros will be received. This risk is called *exchange-rate or currency risk* which is another source of market risk for international portfolios if not hedged. Exchange rate risk used to be a major concern in the European bond market, but with the introduction of the single currency the risk has disappeared within the eurozone. Currency risk exists for investors outside the euro area investing in euro-denominated bonds, but the risk is possible to hedge with currency derivatives.

Event risk refers to the the risk that an unforeseen circumstance will result in a loss. Another example might be a tobacco company carrying the event risk that the government will ban or start taxing tobacco heavier. Company restructurings also pose an event risk (Fabozzi 2007). For example, a leveraged buyout that entails large amounts of new debt will add to the debt burden of the company and potentially lead to a loss to the bondholder.

Inflation risk, also referred to as purchasing-power risk arises from the variation in the real value of the cash flows due to inflation. If the inflation rate exceeds the rate of return, the purchasing power of the money invested actually declines. If the interest rate of the bond is fixed, the bondholder is exposed to inflation rate. Inflation protected bonds, the coupons of which are usually tied to a consumer price index or another inflation measure are designed to protect investors from inflation risk. A floating-rate bond exposes the investor to less inflation risk to the extent that interest changes reflect the expected inflation rate (Reilly and Brown 2003). It is worth noting that even within the eurozone the inflation risk has not fully disappeared as inflation rates and thus purchasing power of investors may still vary.

The yield calculation of a bond assumes reinvestment of the cash flows. This additional, interest-on-interest income depends on the prevailing interest rate levels at the time of receiving the cash flow as well as on the reinvestment strategy. The risk is that the interest rate at which interim cash flows can be reinvested will fall. This variability in the reinvestment rate of a given strategy due to market interest changes is called the *reinvestment risk* (Reilly & Brown 2003). The risk is greater, the longer the holding period is and the larger early cash flows the issue pays (e.g. high-coupon bonds). It should be taken into account that interest rate risk and reinvestment risk have offsetting effects, as interest rate risk is the risk that interest rates increase whereas reinvestment rate is the risk of an interest rate decrease (Fabozzi, 2007).

Call risk is the risk that the issuer redeems all or part of the issue before the maturity date. A separate call provision has to be included in the bond indenture for this to be possible. The issuer usually wants to have this right in order to have refinancing flexibility. From the bondholder's perspective call risk comes from three sources (Fabozzi 2007). First, the cash flow pattern of a callable bond is not known with a certainty. Second, the issuer is likely to call the bond when interest rates have decreased in order to refinance its own activities cheaper. This exposes the investor to reinvestment risk. Third, the capital appreciation potential is lower for a callable bond than a plain vanilla one because the price of the bond is unlikely rise much above the price at which the issuer will call the bond. Naturally the call risk to the bondholder is usually compensated by a lower price when buying the bond, but it is hard to determine a fair compensation for bearing the call risk. Call risk is by many market participants regarded as a very important risk when investing in bonds (Fabozzi 2007).

The price of a bond with embedded options depends on the level of interest rates and factors that influence the value of the embedded option. One such factor is the expected volatility of interest rates; if the expected interest rate volatility increases the value of an option rises. In the case of a callable bond the price of the security falls because the bondholder has granted the issuer an option that is now more valuable than before the rise in interest rates. The risk that a change in volatility will affect the price of a bond is called *volatility risk* (Fabozzi 2007).

Tax risk is a risk that a especially a buyer of a US tax-exempt municipal bonds is exposed to (Fabozzi, 2007). There are two types of tax risk. The first is the risk that the income tax rate will be reduced. The higher the marginal tax rate, the greater the value of the tax exemption feature. The second type of a tax risk is that a municipal bond issued as a tax-exempt issue may eventually be declared taxable by the government.

Risk risk is defined as not knowing what the risk of a financial instrument is. There have been new and innovative structures introduced into the bond market, the risks of which have not been clearly understood by portfolio and money managers (Fabozzi 2007).

4. Measures of performance and risk

The increase in professionally managed funds and the competition in the financial market has added to the need for clear and accurate portfolio performance measurement and analysis. This has resulted in seeking methods to provide investors with the information they need. Investors want to know e.g. whether portfolio management results in sufficient return relative to the risks taken, how the portfolio does compared with the overall market or peers or whether portfolio management results arise thanks to luck or because of identifiable, repeatable skills of the portfolio manager. Hence, there is a link between performance evaluation and investigation of diversification benefits as the goal of diversification is usually to decrease risk and/or improve the reward-to-risk profile of the portfolio. Not all performance measures are suitable for assessing diversification benefits.

Performance evaluation and risk are closely related concepts. Performance measurement combines components and techniques originating in modern portfolio theory introduced by Markowitz (1952). Sharpe produced the first performance indicators, whether risk-adjusted ratios (Sharpe ratio and information ratio) or relative returns compared to a benchmark (alpha). The Capital Asset Pricing Model (CAPM) developed by Sharpe (1964) himself highlighted the notion of rewarding risk. Sharpe ratio explains portfolio returns with the market index as the only risk factor. Factor models were developed as an alternative to the CAPM so as to provide a more accurate description of risks and performance evaluation. Especially in the past few decades the research has extended into consideration of actual market conditions (Ferson & Schadt 1996) or are otherwise not reliant on the restrictions of the CAPM. Indeed, in the recent years, the research in the area of performance measurement has extended to consideration of market conditions and developing techniques that would fit situations in which the portfolio theory hypotheses are not observed.

The choice of a performance evaluation measure should balance the ease of implementation and the accuracy and comprehensibility of the resulting information (Le Sourd, 2007). The same applies for the examination of diversification benefits, as measures to evaluate diversification can all be seen to be performance measures of some sort. However, not all performance measures are equally appropriate for assessing diversification benefits. Some measures, such as Jensen's alpha, are designed to measure the portfolio manager's ability to hand-pick the best-performing assets.

This section of the study provides an overview on performance and risk measures and indicators. The categorization is a combination of those described by Le Sourd as well as Dbouk and Kryzanowski in their studies(Le Sourd 2007 and Dbouk and Kryzanowski 2009, respectively) and roughly proceeds from a simple portfolio return measurement towards more sophisticated measures and techniques. A brief description and discussion is provided on each measure type and individual measure. All metrics presented herein are not appropriate for measurement of diversification. The suitability of the indicators for diversification benefit analysis is also mentioned. The formulae for all the measures are not included as a detailed evaluation of all measures is beyond the scope of this study. For the measures used in the empirical part of the study the formulae are provided in the respective section in the empirical part.

4.1 Portfolio return and volatility

4.1.1 Return taking capital changes into account

The standard return calculation formula is defined by:

$$R_{Pt} = \frac{V_t - V_{t-1} + D_t}{V_{t-1}} \quad (1)$$

where:

V_{t-1} denotes the value of the portfolio at the beginning of the period

V_t denotes the value of the portfolio at the end of the period

D_t denotes the cash flows generated by the portfolio during the evaluation period

The straightforward return calculation for an asset or an individual portfolio becomes more complex when it involves e.g. mutual funds that are subject to contributions and withdrawals of capital by the investors. The basic return formula is only valid for a portfolio with a fixed composition throughout the evaluation period. Le Sourd (2007) lists and discusses three calculation methods that have been developed to take into account the volume of capital and the time that capital is present in a portfolio, each of which assesses a certain aspect of the return The methods are the capital weighted rate of return, the internal rate of return, and the time-weighted rate of return.

The capital-weighted rate of return refers to the relationship between the portfolio value variation and the average of the capital invested during the period. The metric is very intuitive in nature. Because of capital averaging it is an approximation of the actual internal rate of portfolio return (Le Sourd 2007). The internal rate of return, in turn, is based on an actuarial calculation; it is the discount rate at which the discounted final value of the portfolio equals the initial portfolio value and the capital flows taking place during the period under review. If there are a lot of capital in- and outflows of varying sizes, the internal rate of return is more accurate a measure that the capital-weighted rate of return, but it is more burdensome to implement (Le Sourd, 2007).

Time-weighted rate of return is based on dividing the return evaluation period into shorter sub-periods during which the portfolio composition is maintained fixed, assuming reinvestment. The sub-period returns are translated into the period return by taking the geometric mean of those returns weighted by the length of the sub-periods. Gray and Dewar (1971) show that the measure is the only well-behaved rate of return that is not influenced by contributions or withdrawals.

To summarize, a return value should always be evaluated in the light of the return calculation method, the measurement frequency and the length of time for the historical data used in the calculation. They may also serve different information and evaluation purposes. Campisi (2004) notices that time-weighting is appropriate for calculating the active return, whereas money-weighting is well suited for analysing a portfolio manager's contribution to return and attribution return.

4.1.2 Arithmetic versus geometric mean return

As highlighted by Le Sourd (2007) the simple arithmetic average of the returns for the sub-periods under consideration results in overestimation of the result. The result may be even quite distorted in situations where the sub-period returns are very volatile. The choice of sub-periods also affects on the result. The arithmetic mean does, however, have the interesting interpretation of providing an unbiased return estimate for the forthcoming period (Le Sourd, 2007).

The geometric mean, also referred to as the compound geometric rate of return allows a linkage between the arithmetic rates of return for the different periods so as to determine the real rate of return observed over the period. The calculation assumes reinvestment of the intermediate income. If the interest of the investor is over the long-term return (as opposed to just the following period), the geometric average should be considered (Le Sourd 2007). Furthermore, geometric returns accurately capture historic volatility and are therefore the appropriate measure of historical performance (Le Sourd 2007).

Jacquier, Kane and Marcus (2003) find that many finance practitioners seem to prefer geometric averages over arithmetic ones. The notice that the arithmetic average is always an upwardly biased forecast and the geometric average will generally be biased downward.[2] The biases they find are empirically significant. The conclude that for typical investment horizons the appropriate compounding rate is between the arithmetic and geometric values, more precisely a weighted average of the two. They state that in order to create as non-biased a

[2] The geometric average is unbiased only in the special case in which the sample period equals the investment horizon.

measure as possible, the shorter the investment horizon, the higher the weight of the arithmetic average; as the horizon approaches the estimation period, the weight of the geometric average should increase.

4.1.3 Choice of performance evaluation frequency

The technological development has made the portfolio performance monitoring easier on a high frequency basis. A clear advantage is that it may some forms of distorted portfolio management behavior such as so-called window-dressing of performance at the end-of-year or end-of-quarter. DiBartolomeo (2003) notices that daily observation frequency has become more common, which he suspects is believed to provide a more accurate reflection on actual investment returns. DiBartolomeo asserts, however, that such beliefs are based on a series of operational, mathematical and statistical assumptions that are demonstrably false. He details the argument by highlighting that portfolio managers are usually evaluated using time-weighted returns (TWRs) that remove the cash in- and outflow effects; thus daily attribution analysis is not of real use to investors wanting to understand their actual results. Darling and MacDougall (2002) also argue, perhaps somewhat surprisingly, that the TWR leads to some information loss, and the higher the TWR calculation frequency, the more information is potentially lost. They state that most such problems disappear when using monthly observations.

Le Sourd (2007) points out that imperfect assumptions on the asset returns (normally distributed returns, identically distributed returns and no serial correlation between subsequent returns) as another argument against overly frequent performance measurement. Le Sourd refers to results from studies by Dimson and Jackson (2001), among others, who find misinterpretations when using high-frequency performance data: Dimson and Jackson find a clearly increased probability of observing seemingly extreme observations with more frequent measurement. At worst such observations have the potential to lead to unnecessary or detrimental strategy revisions or portfolio manager terminations and thereby to increased transaction costs.

4.1.4 Volatility

One of the best-known measures of risk is the variance and its square root, the standard deviation of expected returns (Reilly and Brown 2003). The idea is that the more disperse the expected returns, the greater the uncertainty of future returns. Usually from performance evaluation point of view especially in a high credit quality context, the lower the dispersion, the better the performance, ceteris paribus. Reilly and Brown (2003) point out that advantages of using variance or standard deviation of returns are that i) they are somewhat intuitive, ii) they are correct and widely recognized measures, and iii) they have been used in most of the theoretical asset pricing models.

Various volatility and variability-based measures are probably the most common types of metrics used to evaluate diversification benefits. Volatility is often considered as longitudinal (time series) volatility of returns or wealth, but sometimes also cross-sectional volatility is investigated. For example Dbouk and Kryzanowski (2009) consider cross-sectional standard deviation, which sometimes is referred to as the mean realized dispersion. When cross-sectional variations are high, a fund manager is operating in a high-risk environment where likelihood of market over- and underperformance are high. Consequently, risk-averse managers seek to reduce their exposure to higher mean realized dispersions.

4.2 Absolute risk-adjusted performance measures

4.2.1 Sharpe ratio

The Sharpe ratio, also referred to as the Sharpe index or measure and originally called reward-to-variability ratio measures the excess return or risk premium per total risk of the investment measured by its standard deviation (Sharpe 1966). The ratio is perhaps one of the best-known performance measures. The ratio is defined by:

$$S_P = \frac{E(R_P) - R_F}{\sigma(R_P)} \tag{2}$$

$E(R_P)$ denotes the expected return of the portfolio
R_F denotes the return on the risk-free asset
$\sigma(R_P)$ denotes the standard deviation of the portfolio returns

As the measure is based on the total risk of the investment or portfolio, it allows the performance of not very diversified portfolios to be evaluated. The ratio is also well suited for assessing performance of of a portfolio representing the investor's total investment. It carries no reference to a market index. The Sharpe ratio has been subject to extensions and enhancements after its introduction. It allows significant possibilities for evaluating portfolio performance while staying simple to calculate. Due to accounting for the total risk and (excess) return of the portfolio, the Sharpe ratio is a suitable metric for evaluating diversification benefits.

4.2.2 Treynor ratio

The Treynor ratio (1965), named after Jack L. Treynor, is a measure of the relationship between the return on the portfolio, above the risk-free rate, and its systematic risk measured by the portfolio beta. The ratio is defined by:

$$T_P = \frac{E(R_P) - R_F}{\beta_P} \tag{3}$$

$E(R_P)$ denotes the expected return of the portfolio

R_F denotes the return on the risk-free asset

β_P denotes the beta of the portfolio

The Treynor ratio is drawn directly from the Capital Asset Pricing Model and its calculation presumes a reference market index to be chosen to estimate the beta of the portfolio. The ratio is thus subject to Roll's (1977) famous critique that the market portfolio is not directly observable.

The Treynor ratio is suitable especially for performance measurement of a well-diversified portfolio or a portfolio which represents only a part of the investor's assets. The reasoning behind this is that an investor who has diversified his investments well only cares about the systematic risk of his holdings. Hence, the metric is not very well suited for evaluating diversification benefits.

4.2.3 A Value-at-Risk based measure

Le Sourd (2007) presents an absolute risk-adjusted performance measure where the risk of the portfolio is denoted by Value-at-Risk (VaR). The Value-at-Risk is a widely used risk measure of the risk of loss. It is defined as a threshold value such that the probability that the mark-to-market loss on the portfolio over the given time horizon exceeds this value (assuming normal markets and no trading in the portfolio) is the given probability level.

Retrospective analysis has found some VaR-like concepts in the history, but the indicator did not emerge as a concept in its own right until the late 1980s after the stock market crash of 1987 (Jorion, 2006). The VaR indicator enables summarizing the set of risks inherent to a portfolio in a single value. It measures the risk of a portfolio as the maximum amount of loss the portfolio can sustain at a certain level of confidence (e.g. 95, 99 or 99.9%). As the VaR calculation requires choosing a confidence level, with this indicator portfolios can only be compared for a same level of confidence.

In Le Sourd (2007) the VaR is used to calculate a risk-adjusted return indicator for portfolio performance evaluation. So as to determine a logical risk measure, the author divides the VaR by the starting value of the portfolio and then proceeds to calculate a Sharpe-type indicator where the standard deviation of the original Sharpe ratio is replaced by the aforementioned relative VaR. The indicator is defined by:

$$VaR - based\ measure = \frac{R_P - R_F}{\frac{VaR_P}{V_P^0}} \qquad (4)$$

where:

R_P denotes the return of the portfolio

R_F denotes the return on the risk-free asset

VaR_P denotes the VaR of the portfolio

V_P^0 denotes the initial value of the portfolio

Value-at-Risk based measures seem suitable for assessing diversification benefits as the concentration risk and the potential of a large loss should diminish with a larger portfolio size. If returns are not perfectly positively correlated, the probability of simultaneous negative returns from all or many investments in the portfolio should decrease.

4.2.4 Critique and other extensions of the Sharpe ratio

The CAPM assumes normally distributed and thus symmetrical asset return distribution and also that investors have mean-variance preferences and thus ignore skewness (Le Sourd, 2007). Leland (1999) shows that the CAPM and risk measures drawn from the model are invalid, if only two assumptions are made: that the return on the market portfolio is independently and identically distributed and that markets are frictionless. He finds that the CAPM alpha mismeasures the value added by investment managers, which implies that the CAPM beta is also flawed.

Le Sourd notes that Cvitanic, Lazrak and Wang (2004) show in their working paper that even if the typical mean-variance efficiency justification for using the Sharpe applies in a static setting, it typically fails when reviewing a multi-period setting. Strategy resulting in the most desirable portfolio is different from the strategy providing the highest Sharpe on a yearly level. Le Sourd also provides an overview of the work of Vinod and Morey (2001) in which they highlight that one problem with using the Sharpe ratio is the randomness of its denominator. According to the authors the randomness is due to the denominator being calculated from a return sample on a certain history and not the total population of returns. Hence, the risk estimation is challenging to evaluate. Vinod and Morey (2001) have proposed a modified Sharpe, also referred to as the Double Sharpe ratio. The indicator is designed to take into account estimation risk and it is calculated by dividing the traditional Sharpe by the standard deviation of the Sharpe ratio estimate, or the estimation risk. To determine the latter they use bootstrapping to create a large number of resamples and derive a series of associated Sharpe ratios (Le Sourd 2007).

Dowd (2000) considers an investor who is thinking of modifying their portfolio and will change it by e.g. introducing a new asset if the incremental increase in risk is sufficiently low compared with the expected return increase. The Sharpe ratios of the initial portfolio and the Sharpe of the intended modified portfolio are compared to determine whether the planned adjustment is attractive. Dowd also examines the case where the reference point is a benchmark portfolio instead of a risk-free rate and defines a decision rule for the attractiveness of a considered modification to the portfolio.

Israelsen (2005) notes that both the Sharpe and the information ratio, the latter of which is presented in the following section regarding relative risk-adjusted performance measures, may result in deceptive rankings of portfolios when risk premiums or excess returns are negative. In such cases a higher Sharpe (or information) ratio does not always point to the best portfolio. Le Sourd (2007) demonstrates this by showing an example of this with realized excess returns of two funds. The example illustrates that a fund with a larger negative excess return and a higher tracking error relative to a benchmark index may actually show a less negative information (and Sharpe) ratio than a fund with a smaller loss and smaller tracking error relative to a benchmark. Israelsen (2005) suggests correcting such a situation by modifying the standard deviation in the denominator in the Sharpe (and/or information) ratio formula. The adjusted ratios suggested by Israelsen enable the consistent ranking of funds regardless of the sign of the excess return. Israelsen points out that as the modification leads to very large range for the size of the indicator, the levels of the values should not be taken as giving any further information than the ranking order.

4.3 Relative risk-adjusted performance measures

Relative risk-adjusted performance measures assess portfolios' risk-adjusted returns in reference to a benchmark.

4.3.1 Jensen's alpha

Jensen's alpha (1968), also referred to as Jensen's performance index or ex-post alpha, is used to determine the differential between the return on the portfolio in excess of the risk-free rate and the return explained by the CAPM. It is computed by carrying out the following regression:

$$R_{Pt} - R_{Ft} = \alpha_P + \beta_P(R_{Mt} - R_{Ft}) + \varepsilon_{Pt} \tag{5}$$

where:

$\beta_P(R_{Mt} - R_{Ft})$ measures the return on the portfolio forecast by the model

α_P is the share of additional return attributable to the manager's choices

The statistical significance of alpha can be assessed by calculating the t-statistic of the regression, which is equal to the estimated value of the alpha divided by its standard deviation (Le Sourd, 2007).

The absolute risk-adjusted performance measures presented in the previous section (with the exception of the Treynor ratio) do not contain the investment benchmark whereas the Jensen alpha does. The Jensen measure considers only the systematic risk of the portfolio as does the Treynor measure. Hence, it is not a good indicator for measuring diversification. The alpha method does not enable the comparison of portfolios with differing levels of risk but it can be used to compare portfolios within peer groups, i.e. portfolios managed in a similar manner and thus containing comparable levels of risk. In fact, the Jensen alpha is proportional to the beta risk of the portfolio (Le Sourd, 2007). In order to compare portfolios with different risk levels, Treynor and Black (1973) have defined the so-called Black-Treynor ratio which is the alpha divided by the beta of the investment.

4.3.2 Critique and extensions to Jensen's alpha

Black's zero-beta CAPM

Black (1972) developes a zero-beta version of the CAPM because two assumptions of the traditional model were called into question: firstly, the assumption of a risk-free asset and secondly, the assumption of a single rate for borrowing and lending. He shows the validity of the CAPM even without the existence of a risk-free rate and replaces it with an investment with a zero beta. Instead of the investor borrowing at the risk-free rate, it is possible to possible to take short positions on the risky assets.

Brennan's tax-adjusted CAPM

The basic version of the CAPM assumes that there are no taxes. However, taxation of dividends and capital gains generally differs from one another, which in turn affect the composition of the investors' portfolio of risky assets and equilibrium prices of assets. Brennan (1970) developed a CAPM version allowing the tax impact to be accounted for by adding a term describing the effect of taxation to the CAPM formula.

Elton and Gruber total risk alpha

Le Sourd (2007) presents in her article an overview on the work by Elton and Gruber (1995) who describe a performance indicator based on the same principle as the Jensen alpha, but they replace the reference portfolio located on the Security Market Line with a portfolio on the Capital Market Line containing the same level of risk than the portfolio being replaced. According to Le Sourd, this measure has later been referred to as the total risk alpha or TRA by Scholz and Wilkens (2005) Security Market Line (SML) displays the expected rate of return of an individual security as a function of its systematic or beta risk whereas the Capital Market Line (CML) is the tangent line drawn from the point of the risk-free asset to the feasible region for risky assets and the tangent point represents the market portfolio (Reilly and Brown 2003). The Elton and Gruber measure evaluates a manager with a total risk budget of a certain standard deviation. That risk level can be achieved by dividing the investment between the risk-free asset and the market portfolio. If the portfolio manager has specific asset-picking skills, they are able to construct a portfolio with higher return at a given level of risk; the return difference when comparing combination of risk-free asset and market portfolio to the portfolio picked by the portfolio manager measures the manager's skills in picking assets.

Alpha measures suitable for market timing strategies

The Jensen alpha has also been subject to the same criticism as the Treynor measure: namely on the grounds of the result depending on the choice of reference index. Furthermore, Jensen alpha often becomes negative when evaluating market timing portfolio management strategies, i.e. strategies involving a time-varying beta according to market movements expected by the manager. Treynor and Mazuy (1966) as well as Henriksson and Merton (1984) have developed Jensens alpha based performance analysis models that are able to account for variations in beta. Furthermore, Grinblatt and Titman (1989b) defines a decomposition of the Jensen alpha which enables evaluation of market timing. As can be expected, alpha measures focused on market-timing are not appropriate for diversification benefit evaluation.

The Treynor and Mazuy model was formulated empirically by the authors in 1966 and, as referred to by Le Sourd (2007, later theoretically validated by others such as Jensen (1972) amd Bhattacharya (1983). The model allows evaluating a manager's market timing capacity using a modified, quadratic version of the CAPM. If a manager anticipates market evolutions correctly, they will lower their portfolio's beta when the market falls and increase it when anticipating a rise in the market. Such a portfolio should rise more than the market does when market rises and lose less than the market portfolio when the market goes down. Hence, the relationship between the portfolio return and market return over the risk-free rate should be better described by a curve than a straight line. If the coefficient estimating the market timing skill is positive and significantly different from zero, it can be concluded that the manager possesses timing skills.

The Henriksson and Merton model refers in fact to two models (Le Sourd, 2007): a non-parametric model developed by Merton (1981) and a more widely used parametric model. They are based on the same principle that an investor can split their investments between a risky and a risk-free asset and modifies this division over time as they anticipate market movements. In the non-parametric version the portfolio is based on options theory by modeling an investment in cash and a call on the better-performing asset (Le Sourd, 2007). The parametric model (Henriksson and Merton, 1984) is based on the same idea but the measure is formulated as a modified version of CAPM which contains a coefficient describing the market timing skills of the manager as well as a parameter that adjusts the formula for the cases where the market portfolio either over- or underperforms the risk-free rate

Grinblatt and Titman (1989) present a decomposition of the Jensen measure in three terms: a term measuring the bias in the beta evaluation, a timing term and a selectivity term. They do so to counter the problem of the Jensen measure sometimes attributing a negative performance to a manager practicing market timing. The model is based on the basic CAPM assumptions of a frictionless market with no transaction costs, taxes or short selling restrictions. If the exact portfolio weightings are known, the abovementioned terms can be investigated separately. The above two models are valid for any two categories of assets, one riskier than the other (Le Sourd, 2007).

Alpha-type measures for international portfolios

Researchers have seeked to extend the Jensen's alpha to accommodate international portfolios. The McDonald's model (1973) allows attribution the contribution of different geographic markets to the total portfolio performance. It thus enables examining the portfolio manager's capacity to invest in the best-performing international markets. Like other alpha-type measures, it isn't suited for diversification benefit evaluation. The model by McDonald's can be generalized for a situation of several international markets and several asset classes. Pogue, Solnik and Rousselin (1974). proposed a year after the McDonald's model another extension of the Jensen measure. Their model measures the manager's capacity to choose the best-performing markets and their skills in choosing the best securities in each market.

Information ratio

The information ratio, sometimes referred to as the appraisal return measures the residual return of the portfolio compared to its residual risk. The residual return and risk are here understood as those not explained by the benchmark. The residual or diversifiable risk is the tracking error of the portfolio, in other words the standard deviation of the return difference between the portfolio and its benchmark. The information ratio is basically a generalization of Sharpe ratio; in the information ratio the risk-free rate in the Sharpe is replaced by the return of the benchmark and the standard deviation is replaced by the abovementioned return difference standard deviation. The information ratio is not suited for comparing the performance of a highly diversified portfolio with that of a portfolio with a low level diversification (Le Sourd), which makes it particularly inappropriate for evaluating diversification benefits of diversification for bond or other portfolios.

Modigliani and Modigliani risk-adjustment performance measure

Modigliani and Modigliani (1997) argue that the portfolio and its benchmark must carry the same risk to be compared in terms of basis points of risk-adjusted performance. So they propose leveraging or deleveraging the portfolio under review using the risk-free asset. The measure evaluates the annualised risk-adjusted performance, or RAP of a portfolio in comparison with the market benchmark, expressed in percentage terms. They maintain that the measure is more intuitive than the Sharpe ratio. Sharpe ratio and the RAP measure are proportional and they lead to the same ranking of portfolios as RAP can actually be expressed as the Sharpe times the standard deviation of the benchmark portfolio.

Scholtz and Wilkens style/risk adjusted performance measure

Scholtz and Wilkens (2005) note that, as the RAP is not appropriate for investors to whom the portfolio represents only part of their investments, because it measures risk with the standard deviation. They suggest replacing the measurement of returns relative to total risk with the return difference to market risk. The proposed measure, the market risk-adjusted performance (MRAP) suits investors who place money in several different assets and thus doesn't really suit investors evaluating diversification benefits. The work of Modigliani and Modigliani (1997) has also inspired Lobosco (1999) to propose a risk-adjusted performance measure including the portfolio management style. The SRAP (style/risk-adjusted performance) is calculated by identifying the weighting of indices that represent the portfolio management style. This style benchmark is then used to calculated the SRAP in the same manner as RAP. Furthermore, Le Sourd (2007) points out that employing style analysis enables creating more stable estimates of risk, a result also found by Sortino et al. (1997).

Muralidhar M3 measure for performance comparison within a peer group

Muralidhar (2001) has developed a risk-adjusted performance measure that allows the comparison of the performance of portfolios managed similarly or a peer group of managers. The contribution to the Modigliani and Modigliani measure is that not only the standard deviations of portfolios but also the correlations of the portfolios with respect to the benchmark as well as the correlations between the portfolios. The Muralidhar or M3 measure enables the construction of portfolios split in an optimal way between a risk-free asset, the benchmark and managers who are compared. When doing this, it takes into account the goals of the investor in terms of absolute and relative risk towards the benchmark. Le Sourd (2007) argues that the Muralidhar measure is certainly useful compared to the risk-adjusted performance measures presented previously. Le Sourd maintains that it is necessary to include the correlations between the portfolios and between the portfolios and the benchmark to provide an appropriate risk-adjusted performance measure. The M3 measure results in a ranking different from that obtained with measures such as the Sharpe ratio or the information ratio, as it employs additional information. Le Sourd also highlights that the Muralidhar measure indicates how to construct portfolios that satisfy the investors' objectives. It also enables an institutional investor to find the optimal allocation between active and passive management.

Muralidhar SHARAD measure

Muralidhar (2001, 2002) also proposes a new measure with all the properties of the M3 measure, but which also allows differences in data history to be taken into account. He names this measure SHARAD for Skill, History and Risk-Adjusted. It has all the properties of M3 but also allows taking confidence level as regards the manager skills to be examined.

AP Index

Le Sourd (2007) presents the somewhat less well-known measure by Aftalion and Poncet (1991), also called the AP index is defined as the difference between the average expected return of the portfolio and that of its benchmark, from which is deduced the product of the difference between the portfolio risk and the benchmark risk multiplied by the price of risk. The AP index enables ranking portfolios with the same benchmark index. It is also an alternative to the Sharpe ratio especially in cases of negative risk premia.

Two performance measures by Graham and Harvey

Graham and Harvey (1997) have developed two measures to counteract two problems related to the Sharpe ratio. First, they find that estimates are not precise enough when fund volatilities are too different. Second, the calculation of the Sharpe ratio is made assuming that the risk-free rate is constant and not correlated to risky asset returns. The first measure (GH1) is obtained by searching for the point with the same volatility as the investment under analysis and computing the difference between the return of this portfolio and that of the portfolio being analysed provides us with the GH1 measure. The second measure (GH2) is similar to the M^2 measure proposed by Modigliani and Modigliani (1997) but the GH2 allows for curvature in the efficient frontier.

Cantaluppi and Hug efficiency ratio

Cantaluppi and Hug (2000) measure seeks to find how efficient a portfolio is by comparing the portfolio to the ex-post efficient frontier portfolio with the same volatility. The Cantaluppi and Hug efficiency ratio is measured by the distance from the efficient frontier.

Scholtz and Wilkens investor-specific performance measure

Scholtz and Wilkens (2004) consider the case of an investor holding a certain portfolio and wanting to invest additional money to another, predefined portfolio without changing the composition of the initial portfolio. The Investor-Specific Performance Measure (ISM) is used to come up with the optimal weights for the initial portfolio and the additional portfolio, the latter of which is made up of a fund and the risk-free rate. According to derivation conducted originally by Scholtz and Wilkens (2004) the ISM value is determined by the Sharpe ratio and the Treynor ratios of the evaluated funds. The higher the Sharpe and Treynor ratios, the higher the ISM.

4.4 Measures based on downside risk and higher moments

Measures based on downside risk are appropriate for investors who are solely or more interested in the return of their investment generating loss, either in absolute terms or relative to a benchmark. Hence, such measures are generally particularly suitable for evaluating diversification benefits for bond portfolios. Firstly, investors are usually interested in diversification because it reduces the risk of the portfolio. Secondly, especially investment grade bond investors, possessing a large downside risk but only a limited or fixed upside potential in their investments, should be very interested in reducing this downside risk through diversification.

Higher moments, in turn, are loosely speaking quantitative measures of the shape of a distribution. Any distribution can be characterized by its moments, the first moment being the mean, the second moment the variance, the third moment skewness and the fourth moment kurtosis. When portfolio returns are not normally and independent and identically distributed (iid), the concepts of skewness and kurtosis come in to play to allow for adjusting for the non-normality (Le Sourd 2007).

4.4.1 Actuarial approach

Melnikoff (1998) measures risk-adjusted return by calculating a gain short-fall equilibrium, i.e. the relation between the expected return seeked by the investor to make up for a fixed shortfall risk. Melnikoff argues that for an average individual the weight of the gain-shortfall constant is such that the individual will agree to invest if the expected gain is double the shortfall risk. Sortino and Price have also defined a utility theory based performance measure reflective of downside risk, referred to as the Fouse index (Le Sourd 2007).

4.4.2 Semi-variance

Semi-variance is the same as the variance, apart from the fact that only the returns that are lower than the mean are taken into account (Reilly and Brown 2003). It thus accounts for the asymmetry of risk. It is suitable for investors who are only interested in their portfolio not losing value. Semi-variance, semi-standard deviation and measures based on them can be seen as especially appropriate for measuring diversification benefits for bond portfolios.

4.4.3 Lower partial moments

The lower partial moment (the average of the squared deviations below a target return) is a more general measure of downside risk than the semi-variance which is computed as the average of the squared deviations below the mean return.

4.4.4 Sortino ratio

A Sharpe ratio type of standard deviation based measure does not tell whether the differentials compared to the mean were produced above or below the mean. To measure the downside risk the mean return can also be replaced by the value of the target return below which the investor does not wish to drop (Le Sourd, 2007). The Sortino ratio (Sortino and Van der Meer, 1991) is a ratio defined on the same principle as the Sharpe ratio, but the risk-free rate is there replaced by the minimum rate accepted by the investor (MAR, minimum acceptable return) and the standard deviation of the Sharpe ratio is replaced wit the standard deviation of the returns below the MAR. Much like for the Sharpe ratio there are also different versions or adjustments of the Sortino ratio, for example Dbouk and Kryzanowski (2009) use a definition where risk-free rate is used instead of the MAR in the numerator, but risk in the denominator is still defined by the semi standard deviation. Both versions differentiate between ‘good’ and ‘bad’ volatility and only penalizes portfolios the returns of which fall short from their mean return.

The Sortino ratio is defined by:

$$Sortino\ ratio = \frac{E(R_p)-MAR}{\sqrt{\frac{1}{T}\sum_{\substack{t=0 \\ R_{Pt}<MAR}}^{T}(R_{Pt}-MAR)^2}} \tag{6}$$

where

$E(R_p)$	denotes the expected return of the portfolio
MAR	denotes the minimum acceptable return
$\frac{1}{T}\sum_{\substack{t=0 \\ R_{Pt}<MAR}}^{T}(R_{Pt}-MAR)^2$	denotes the semi-variance

4.4.5 The upside potential ratio

Sortino, Van der Meer and Plantinga (1999) have also developed a measure that combines the upside potential and the downside risk of the portfolio. The Upside potential ratio (UPR) is the probability-weighted average of returns over the reference rate. The numerator of the ratio is the expected return above the MAR and may be regarded as the potential for success. The denominator is downside risk as calculated in Sortino and van der Meer (1991) and can be thought of as the risk of failure. Le Sourd (2007) highlights that a notable benefit of using the UPR is the consistent wat of using the reference rate for evaluating both gains and losses. The relevance of the UPR for a fixed income investor examining diversification is perhaps questionable due to the limitedness of upside.

4.4.6 Ziemba downside-adjusted extension on Sharpe ratio

Ziemba (2005) also argues that the Sharpe ratio has the disadvantage that it is based on mean-variance theory and thus is valid basically only for quadratic preferences or normal distributions. To overcome the reliance of the Sharpe ratio on mean-variance theory and so as to penalize a fund manager for losing but not for winning, Ziemba (2005) adjusts the Sharpe ratio using downside variance from the returns that are negative (i.e. using a zero return as reference to measure downside risk). The total variance is calculated as double the downside variance to assume that the upside deviation is identical to the downside risk. Ziemba argues that with this modification of the Sharpe ratio gives more realistic results. It seems that such measure would also be suitable for diversification benefits evaluation.

4.4.7 Skewness

Lo (2001), among others, argues the Sortino ratio can be manipulated by transferring part of the risk from the first and second-order moments (mean and variance) to the third and fourth-order moments (skewness and kurtosis). Hence, when portfolio returns are not normally and independent and identically distributed (iid), the concepts of skewness and kurtosis come in to play to allow for adjusting for the non-normality (Le Sourd 2007). Dbouk and Kryzanowski (2009), for example, use a cross-sectional skewness measure.

In statistics, skewness is a measure of the asymmetry of a probability distribution. Conceptually, a negative skew indicates that the left tail of the probability density function is longer than the right side and the bulk of the values lie to the right of the mean. A positive or rightward skew indicates that the long tail is on the right. Right skewness is common when a variable is bounded on the left but unbounded on the right. In finance this is the case with for example with bond prices. Mathematically, skew is usually measured by the third standardized moment. Bond returns, however, can take values both above and below zero and thus be skewed to the left. The third standardized moment can take any positive or negative value, although in practical settings it rarely exceeds 2 or 3 in absolute value. Because it involves cubed values, the third standardized moment is sensitive to outliers (Le Sourd 2007).

A zero skew indicates that the values are relatively evenly distributed on both sides of the mean, however, this does not necessarily imply a symmetric distribution. Dbouk and Kryzanowski highlight the observation documented in literature by Harvey and Siddique (2000) and Premaratne and Tay (2002) that investors prefer portfolios with positive skewness. Many studies referred to by Dbouk and Kryzanowski, such as French et al. (1987); Brown et al. (1993) and Campbell and Hentschel (1992) find negative skewness for an index.

4.4.8 Kurtosis

As skewness, kurtosis is also a descriptor of the shape of a probability distribution. It measures the peakedness of the probability distribution. Just as for skewness, there are different ways of quantifying it for a theoretical distribution and corresponding ways of estimating it from a sample from a population.

As skewness, kurtosis is also a descriptor of the shape of a probability distribution. It measures the peakedness of the probability distribution. Just as for skewness, there are

different ways of quantifying it for a theoretical distribution and corresponding ways of estimating it from a sample from a population. If risk-averse investors weigh potential downside returns more than potential upside returns, then these investors will prefer a distribution with lowkurtosis, since the tails are more likely to fall closer to the mean (Dbouk and Kryzanowski 2009).

4.4.9 L-moments

The kurtosis index originally proposed by Karl Pearson in 1905 and commonly introduced in statistical textbooks is based on a scaled version of the fourth moment of data. The index proposed by Pearson in 1905 has been a subject of considerable debate (e.g. Fiori and Zenga 2009) and it has been argued that this measure really measures heavy tails, and not peakedness.. For this measure, higher kurtosis means that a larger part of the variance stems from infrequent but extreme deviations instead of more modest but frequent ones. An alternative measure, the so-called L-kurtosis has been suggested (e.g. Hosking 1990). L-moments are expectations of certain linear combinations of order statistics (Hosking, 1990). Hosking argues that L-measures are more robust than conventional ones to data outliers. He also maintains that L-moment suffer less from effects of sampling variability and thus allow for more secure inferences to be made based on a small sample.

4.4.10 Hwang and Satchell higher-moment CAPM

In the case of non-normal returns traditional CAPM models fail and a higher-moment CAPM should prove more appropriate (Le Sourd, 2007). Hence, a performance measure should also ne based on higher moments. Le Sourd presents a performance measure originally defined by Hwang and Satchell[3] using the three-moment CAPM. The model takes advantage of skewness and kurtosis of market returns as well as the beta and coskewness of return distributions. In a world of normally distributed returns the measures equals Jensen’s alpha. Hence, the Hwang and Satchell measure is able to account for non-gaussian distributions but suffers from other limitations of Jensen’s alpha (Le Sourd, 2007).

[3] Hwang S. and Satchell S., “Evaluation of Mutual Fund Performance in Emerging Markets”, Emerging Markets Quarterly, vol. 2, nÅã3, 1998, pp. 39-50.

4.4.11 Keating and Shadwick Omega measure

Keating and Shadwick (002) note that taking into account only the two first moments (mean and variance) of returns causes inaccuracies when measuring performance and that higher moments are needed for the purpose. The authors propose and examine a measure called Omega, which is related with downside risk, lower partial moments and gain-loss literatures. They also maintain that a return level reference is useful. The Omega measure by Keating and Shadwick captures all of the higher moments of returns distribution and incorporates a reference rate of return to evaluate gains and losses against. The calculation process of the measure consists of grouping profit and loss returns compared to the minimum acceptable return (MAR) of the investor. The authors the consider the probability weighted ratio of returns above and below the MAR threshold. In facet, Omega is a function of the MAR threshold. Ranking portfolio managers is possible using Omega. Because of the of the additional information incorporated in it, the measure provides different rankings that for example Sharpe or alpha measures. Le Sourd (2007) states that Omega can be specifically recommended for portfolios not showing normally distributed returns. Hence, it often appears in a hedge fund setting.

5. Diversification benefits in a bond portfolio context

The basis of modern portfolio theory is Markowitz' modern portfolio theory study (1952), which illustrates the benefits of diversification, i.e. forming portfolios with assets whose prices are less than perfectly positively correlated with each other. The subsequent literature since Markowitz has mainly focused on the benefits and deficiencies of diversification in the context of equity portfolios. Different studies reach different conclusions of the required minimum portfolio size in order to achieve a 'well-diversified portfolio', henceforth referred to as PS in this study. Studies have found such portfolio sizes ranging anywhere from 8 to 10 (Evans and Archer 1968, Latane and Young 1969) to more than 300 (Statman 2004).

5.1 Studies on achieving a well-diversified portfolio

The relationship between portfolio performance and the number of stocks held in the portfolio has been of interest to many investors and financial economists since the famous studies of Markowitz (1952, 1959) and Sharpe (1964). Although portfolio risk can be effectively decreased by increasing the number of less than perfectly correlated assets in the portfolio, over-diversification is likely to increase costs unnecessarily due to higher staff resource needs and transaction costs incurred from portfolio rebalancing. Hence, rational investors want to avoid unnecessary over-diversification and try to find the optimal level of diversification. For retail investors in particular, diversifying a portfolio of individual assets to a high number of instruments can be a costly and inefficient process. The costs include brokerage fees associated with buying relatively small amounts of many different stocks as well as the record-keeping associated with owning enough names to effectively eliminate unsystematic risk.

A classic paper by Evans and Archer (1968) examines how the portfolio risk for randomly selected portfolios can be reduced as a function of the number of securities included in the portfolio. To test this they develop a model using the 470 stocks that had complete data for the period 1958 to 1967. They find that a portfolio of about 8 to10 randomly chosen and equally weighted (i.e. naively diversified) stocks would have similar risk, as measured by standard deviation, to the market as a whole. The authors raise doubts concerning the economic justification of increasing portfolio sizes beyond 10 or so securities. Their study has been widely cited in subsequent diversification literature.

Fisher and Lorie (1970) provide support for Evans and Archer (1968) with simulation results of returns for different portfolio sizes for New York Stock Exchange listed stocks. Latané and Young use a method similar to that by Evans and Archer (1969) and find that an appropriate number of stocks in a portfolio from the diversification point of view is 8 to 16 stocks; they argue that such portfolio sizes are sufficient to achieve 85 to 95 per cent of the possible diversification gains. Latané and Young use ex post distributions of annual holding period returns to test portfolio building rules, but Cheng and Deets (1971) criticize that the usefulness of the findings of Latané and Young are compromised because of the bias created by their overlap of subperiods in their data and the resulting, potentially biased volatility results.

Elton and Gruber (1977), in turn, extend the work by Markowitz and find that the conventional rule of thumb that most diversification benefits could be reaped with 10 – 20 stocks might be misleading. They show that the portfolio variance can be significantly reduced by increasing the number of stocks from 15 to 100. Statman (1987) exhibits that a portfolio of 30 – 40 stocks can lead to effective diversification. The author shows that a well-diversified stock portfolio should include at least 30 stocks for a borrowing and 40 stocks for lending investor, contradicting the earlier findings of a requirement of 10 to 15 stocks. Statman also argues that the level of diversification benefits in terms of reduction of unsystematic risk is meaningless but that the marginal net benefit is what matters; diversification should thus be increased as long as the marginal benefits exceed the marginal costs. The marginal benefit approach has later been used by Dbouk and Kryzanowski (2009) in their bond diversification study, among others.

Studies during the last couple of decades have concluded that because of increasing volatility, a well-diversified portfolio should have a considerably higher number of stocks than what indicated by the earliest studies. Campbell (2000) states that thirty or forty years earlier (approximately in the 1960s) a single randomly selected stock had a standard deviation 35 percentage points higher than a portfolio invested in an equally weighted index of all available stocks; a portfolio of 20 individual stocks could reduce this excess risk to a modest level of about 5 percentage points. During the 1990s, however, a single randomly selected stock has had a standard deviation 50 percent higher than an equally weighted index. Campbell finds that it takes a portfolio of as many as 50 stocks to reduce the excess risk to the same 5 percent. He argues that the old rule of thumb is no longer adequate as a 20-stock portfolio has excess risk of 10 percent. Campbell (2000) also maintains that a typical individual stock is more volatile than before, but that it also has a lower correlation with other stocks; more of the individual-stock volatility is idiosyncratic, less is shared with the market as a whole. According to him the trend reflects changes in the market, including the trend away from conglomerates towards companies focused on one or two core competencies and the tendency to list companies earlier in their life-cycle, when their futures are still very uncertain. Bennett and Sias (2006) also attribute the need for an increased number of stocks to achieve diversification to a change in the size and the structure of industries.

Statman (2004) argues that the optimal level of diversification, measured by the rules of mean-variance portfolio theory exceeds as many as 300 stocks. Shawky and Smith (2005), in turn, empirically investigate the U.S. equity mutual funds from Morningstar data during 1992-2000 and find that the optimal portfolio size can actually be as high as 481 stocks in a portfolio and that any significant deviations from that number are not optimal. De Wit (1998) suggests that one can still significantly benefit from diversification even when adding stocks to an already-large portfolio by showing that diversification benefits can still exist by adding stocks to an equally weighted portfolio that already contains as many as 700 stocks.

Various authors document improved return-to-risk portfolio profiles from international diversification in stocks or in stocks and bonds (e.g. Solnik et al. 1996; Chollerton et al. 1986; Jorion 1987, 1989). Hunter and Simon (2004, 2005) find that the average correlation in the international bond market has increased over time but not to the same extent as observed in the equity market by Goetzman et al. (2005), among others. Hunter and Simon (2004) find that US investors who hold a well-diversified portfolio of domestic fixed-income and equity investments can obtain incremental diversification benefits from investing in international government bonds if currency risk is hedged.

Odier and Solnik (1993) find that correlations among US and international markets were higher when markets dropped than when they rose. Hence, diversification offered smaller benefits in down markets, precisely at the time when the diversification would be most welcome. Statman and Scheid (2005) observes that the case for dispersion is similar: dispersion between US and international stock was lower when both groups had negative returns than when they had positive returns. They point out that diversification benefits depend on both the correlations between the asset classes as well as on the standard deviations of these markets. Hence, they argue that dispersion is a suitable measure as it accounts for both. Second, they argue that investors tend to have poor intuition about correlation and its relationship to diversification benefits. Silvapulle and Granger (2001) is another study which finds that stock diversification benefits diminish with large negative movements in stock returns due to higher firm-level return dispersions when market returns are largely negative (Demirer and Lien 2004). According to Sancetta and Satchell (2007) the increasing importance of correlations during market downturns is related to the market's tail distribution.

5.2 The diversification puzzle

Campbell (2000), among others, notes that many investors appear to ignore the 'free lunch' offered by diversification. He states as examples that individuals overinvest their retirement savings in the company they work for, in one or two favored sectors such as technology, and in companies that are based in the region where they live. Even sophisticated investors concentrate their portfolios in stocks of their own country rather than diversifying internationally. Campbell highlights that in fact the benefits of diversification are substantial, and they have grown dramatically over time.

Especially in the case of individual investors it is plausible to assume that the seeming under-diversification has, at least partly, to do with high transaction costs due to high (in relative terms) brokerage fees and record-keeping of a high number of instruments. It might for example be the case that the implicit or indirect costs of monitoring and record-keeping may not be fully accounted for when deeming portfolios underdiversified. Even for institutional investors diversification is associated not only with potential benefits but also costs in the form of higher staff and other resource needs and transaction costs from rebalancing of the portfolio. For example, Shawky and Smith (2005) show that even the average number of stocks held in professionally managed equity mutual fund portfolios is about 90, with about third of the total value held in the top 10 stocks although the same authors find that the optimal number of stocks would be close to as many as 481.

Statman (2004), among others, examines the well-known finding that investors tend to hold fewer assets than suggested by diversification studies findings on sufficient diversification. Statman argues that the optimal level of diversification, measured by the rules of mean-variance portfolio theory exceeds as many as 300 stocks. Actual levels of equity diversification in earlier studies were, however, much lower. Statman's article examines the anomaly that the average investor holds only 3 – 4 stocks and argues that the so-called diversification puzzle can be solved in the context of behavioral theory. He argues that if investors construct their portfolios as layered pyramids where the bottom is for downside protection (risk aversion) and top for upside potential (risk seeking). Statman also maintains that the behavior of investors is motivated by their aspirations, not their risk attitudes. Some investors fill the upmost layer with only few stocks (an undiversified portfolio), some with lottery tickets. Although this behavior is not consistent with mean-variance portfolio theory, it is that with behavioral portfolio theory. Van Nieuwerburgh and Veldkamp (2005) argue that informational advantages explain the observation that investors tend to hold fewer assets than suggested by the literature on diversification benefits.

5.3 Diversification in a bond portfolio context

While the consensus among most financial professionals is that asset class allocation is one of the most important decisions that investors make, some investors are focused on purely or almost entirely on one asset class, such as a proprietary equity trader on stocks or a bond mutual fund on bonds. The reason for such concentration of investments may have to do with asset-liability management considerations or the level of desired riskiness of investments.

Fixed income style institutional investors such as many central banks or fixed income funds invest in a wide array of bonds and fixed income instruments, the former usually focusing on plain vanilla government, government-related and perhaps corporate bonds and the latter sometimes also on a wider scope of fixed income instruments such as bank loans or securitized debt instruments[4]. One of the best-known fixed income investors is Pacific Investment Management Company, LLC (PIMCO), a global investment management firm. Also central banks around the world are important fixed income investors. An example of an individual, as opposed to an institutional, fixed income investor could be a retired person or someone close to retirement who is looking for a constant and secure return on their investment.

In order to diversify their investments within the fixed income, or more precisely, bond universe, investors can diversify in a wide range of sectors, maturities or the rating. Varotto (2007) finds that interest rate factors followed by maturity diversification are important and credit rating and seniority diversification effects are unimportant causes of diversification of total portfolio returns in international corporate bond portfolios. He also finds that industry diversification, unlike interest rate and maturity diversifications, have little impact on the volatility of corporate bond portfolios. However, industry sector diversification still seems to remain as a commonplace industry practice, as for example bond mutual funds or sales and

[4] E.g. mortgage-backed bonds (MBS), asset-backed bonds (ABS), convertibles, trust preferred bonds, collateralized loan obligations (CLOs), and collateralized debt obligations (CDOs) (Fabozzi 2007)

trading teams at investment banks and brokerage companies are often set up by the sector. Furthermore, many fixed income indices are differentiated by the industry sector[5]

There seems to have been only few empirical studies examining how many bonds are required to attain a target level of diversification. The study by McEnally and Boardman (1979) investigates how many bonds are needed to achieve a certain level of diversification benefits in terms of volatility risk reduction of the portfolio. They find that eight to sixteen bonds notably reduce volatility when differentiating the investment opportunity sets by credit ratings. McEnally and Boardman also find that diversified high-yield bond portfolios possess lower systematic or non-diversifiable risk than investment grade bond portfolios. They hypothesize that the finding may be attributed to an industry effect as low-risk, defensive bonds are often utility bonds whereas bonds including more risk tend to be industrial bonds.

Dbouk and Kryzanowski, henceforth referred to as D&K (2009) challenge the applicability of McEnally and Boardman's results to more recent period based on several arguments. First, McEnally and Boardman only consider a random sample of 515 corporate straight bonds (i.e. bonds with no optionality features). Second, D&K argue that the findings are likely to be outdated as the bond market has expanded in size as well as in industry, country and credit quality scope and the market efficiency since the sample period of 1972 – 76. Third, as McEnally and Boardman themselves note as well, the sample period was characterized by extreme corporate bond market volatility and spreads. With the hindsight of the ongoing severe financial crisis, research concerning periods of financial crisis has become relevant again. Furthermore, diversification tends to be needed the most during periods of extreme markets. Fourth, D&K (2009) maintain that the only measure used by McEnally and Boardman, the unconditional variance is not enough to fully reflect the diversification benefits, as more recent diversification tests on equity data take advantage of a lot broader set of measures that illustrate return-to-risk characteristics, alternative definitions of risk or higher order moments. As measurement of diversification is closely related to performance measurement in general, this view is consistent with Le Sourd (2007) who notes that the increase in professional fund management and competition in the financial market has added to the need for clear, accurate and informative performance measurement.

[5] e.g. Merrill Lynch and maintains widely used bond indices such as an euro government bond indices, euro credit corporate bond indices.

Dbouk and Kryzanowski (2009) seek to contribute to the literature on bond portfolio diversification by examining the minimum portfolio size (henceforth PS) required to capture significant diversification benefits. They differentiate the investment opportunity sets (henceforth IO sets) by issuer type, credit rating, and time-to-maturity. Their results show that the differentiation characteristics of bonds as well as the metric used to measure the marginal benefits of extra diversification determine the minimum portfolio size. D&K find that the marginal benefits of further diversification are in most cases optimized with portfolio sizes of 25 to 40 investment grade bonds when using a criteria of small marginal benefit. More precisely, an optimal PS is found when the marginal benefit of additional diversification decreases below 1% in terms of reduction in the risk metric or improvement in the reward-to-risk metric examined. Their approach is consistent with the arguments by Statman (2004), who maintains that marginal, as opposed to overall, benefits are what matters when evaluating diversification benefits.

The metrics Dbouk and Kryzanowski (2009) use to measure diversification benefits include dispersion of returns, reward to risk and downside risk measures as well as the probability of underperforming a target rate of return among others. Dispersion metric improvement is examined in terms of standard excess deviation or mean derived deviation as well as cross-sectional dispersion, the latter of which seems to have been studied to a slightly lesser extent. When cross-sectional variations in returns are high, a fund manager is operating in a high-risk environment where the probabilities of market over- and under-performance are high (D&K 2009). Dbouk and Kryzanowski's (2009) analysis concerning diversification benefits in terms of reward-to-risk focuses on Sortino ratio but the authors also run the analysis for the more traditional Sharpe ratio. Downside risk and higher moments are examined by investigating cross-sectional skewness and kurtosis and the left and right tail weights. The probability of under- or overperforming a target rate of return, in turn, is examined by comparing the portfolio returns to the index constructed from all bonds.

Dbouk and Kryzanowski differentiate the IO sets by industry sector of the issuer, credit ratings and time to maturity of the bonds. The first contribution by D&K to literature is related to diversification for IO sets differentiated by not only rating but also industry sectors and maturities. The study also separates bonds issued by foreign entities as a separate sector. Second, they examine not only straight bonds but also ones with optionality features such as callability, puttability and convertability. Third, they use several metrics, some of which are rather modern and recently introduced into literature, to measure diversification benefits for bonds. Fourth, D&K state as one of their major insights that strictly speaking no minimum PS exists in the first place but that it depends on the risk, return and bond maturity objectives of the investor as well as on the issuer and bond features (issuer sector and rating). Fifth, D&K find that although the PS defined by the small marginal benefit criteria is usually achieved with 25 – 40, the PS leaves much potential diversification benefits unrealized for many IO sets. Dbouk and Kryzanowski maintain that this may explain the observation found earlier by Kahn and Rudd (1995) that individuals purchase bond mutual funds although studies find that bond mutual funds exhibit neutral and under performance when evaluated using gross and net-of-fees returns, respectively.

Similarly as most diversification studies focus on either pure stock or combined stock – bond portfolio, the focus of research has to a large extent been on the US rather the euro-denominated market. Also Dbouk and Kryzanowski run their analysis on US bond market data, which is not surprising as the US bond market was long the largest and the most liquid bond market in the world. The structure of the international bond market saw a considerable change with the start of the European Monetary Union (EMU). In fact, the Euro government bond market has surpassed the US market in terms of size and number of issues (Fabozzi 2007). Outstanding euro-denominated notes and bonds issued amounted to over 8,700 billion euros in September 2010 whereas the amount of US dollar denominated bonds and notes was around 7,300 billion euros (BIS Quarterly Review December 2010).

6. Data and methodology

This section presents the data and methodology as well as the motivation for the methodology chosen.

6.1 Data and sample

The bond sample data used in this study is extracted from the Bank of America Merrill Lynch EMU Euro Broad Market Index from January 2002 until December 2010. The sample consists bonds, their monthly returns and descriptive bond data. There are 11,504 unique bonds and 367,435 month-end bond returns when differentiating by issuer sector of the bonds.

According to the bond index information provided by the Bank of America Merrill Lynch, the Euro Broad Market Index tracks the performance of euro denominated investment grade debt publicly issued in the eurobond or Euro member domestic markets, including euro-sovereign, quasi-government, corporate, securitized and collateralized securities. The inception of the index took place on December 31, 1995.

Table 4. Sample sizes for the IO sets differentiated by issuer type

Characteristic	1. All	2. Government	3. Govt-rel.	4. Cov./Sec.	5. Financial	6. Industrial	7. Foreign	8. Utility
Panel A: Sample sizes for IO sets differentiated by issuer type and credit rating								
Unique bonds (#)	11,504	629	1,099	4,877	1,642	921	2,104	241
Bond prices (#)	367,615	25,656	36,344	134,965	51,182	33,279	77,085	9,284
Average coupon (%)	4.70	4.55	4.08	4.29	4.67	5.23	5.04	5.01
Average maturity (years)	5.51	6.18	5.21	3.51	4.75	5.67	5.82	7.42

Note: The table summarizes the sample sizes in terms of the total number (#) of unique bonds, the total number (#) of bond prices, the average coupon (%), and average maturity in years in each investment opportunity differentiated by issuer type. Note that the sum of sample sizes for the sectors does not sum up to the number of 'All' because nine (9) bonds belong to both foreign and domestic government over their life.

Qualifying securities must have an investment grade rating (based on an average of Moody's, Standard & Poor's and Fitch), an investment grade rated country of risk (also based on an average of Moody's, S&P and Fitch foreign currency long term sovereign debt ratings), at least one year remaining term to final maturity and a fixed coupon schedule. Callable perpetual securities qualify as long as they are at least one year from the first call date. Fixed to- floating rate securities also qualify provided they are callable within the fixed rate period and are at least one year from the last call prior to the date the bond transitions from a fixed to a floating rate security.

Merrill Lynch index qualifying issuers must be eurozone members and have at least one readily available, transparent price source for their securities as determined by Bank of

America Merrill Lynch. Qualification with respect to Euro membership is determined annually effective each December 31.[6] Qualifying sovereign securities must have a minimum amount outstanding of €1 billion. Bills, inflation-linked debt and strips are excluded from the index; however, original issue zero coupon bonds are included. Qualifying non-sovereign securities must have a minimum amount outstanding of €250 million. Euro legacy currency and defaulted securities are excluded from the index. Since the foreign bonds in the sample are issued in euro by non-eurozone entities from all sectors, all returns are based on euros. Clearly, the number of foreign bonds is much smaller than the number of domestic (eurozone) bonds, which means that the study has an investor home bias. Nevertheless, it can be thought to accurately reflect the home-bias selection followed by investors that are referred to in diversification literature, since international diversification not subject to this bias could achieve higher diversification benefits.

The index is rebalanced on the last calendar day of the month. Issues meeting the qualifying criteria are included in the Index for the following month. Issues no longer meeting the criteria (e.g. due to a rating downgrade or the time to maturity rolling below one year) during the course of the month remain in the Index until the next month-end rebalancing at which point they are removed from the index.

The extracted data contains monthly returns and descriptive bond information, such as industry, rating, monthly total percentage returns, bond coupons, maturities and day count methods. Monthly total monthly returns (obtained by dividing the total dollar return by the beginning of the period dirty price) were readily available in the Merrill Lynch Global Index system database.

[6] Slovenia joined the common currency as of 1 Jan 2007, Cyprus and Malta 1 Jan 2008, and Slovakia 1 Jan 2009 and are included in the index as of those dates. There are in total nine (9) bonds which are considered foreign at the time of their issuance and later change into domestic government bonds with the eurozone membership of the country.

6.2 Methodology

This section first presents the preparation of data which is needed as bonds mature over the course of the period under review. The following subsection after that presents the ex post simulation method.

6.2.1 Data preparation

Industry-wise the categorization of this study follows loosely that by Dbouk and Kryzanowski (2009) but is adjusted due to data availability constraints and author discretion. The sector categorization used in this study is the following: 1. All bonds, 2. Government bonds, 3. Government-related bonds, 4. Covered and securitized bonds, 5. Financial (unsecured) bonds, 6. Industrial bonds, 7. Foreign bonds.[7] In addition there were bonds issued by utilities, but their low number (243) prevents from examining portfolio sized beyond 39 bonds, which is the number of outstanding utility bonds in the beginning of 2002. Hence, utility bond results are left untabulated in this study. Unlike in D&K (2009), government bonds and supranational/government-related bonds are considered separate sectors in this study, as it is relatively commonplace for professional bond investors to consider these classes distinct from each other.

Categories 2. to 6. are bonds issued by entities domiciled in eurozone countries. Government bonds are bonds issued by euro country governments. Government-related are sub-sovereign entities like sub-sovereign agencies (e.g. the German special purpose bank KfW) or regional governments (e.g. German States). The category Covered and securitized bonds refers to bonds with a collateralization feature. There were only 31 domestic bonds categorized as securitized. Their risk profile fundamentally different as covered bonds are backed by assets in the issuer's balance sheet whereas securitized bonds are backed by assets that do not remain on the issuer's financials. Both are however, backed by a recourse to a pool of assets securing the bond, should the originator become insolvent. The category 'Financial bonds' refers to unsecured (non-covered and non-securitized) bonds issued by euro area financial institutions. Industrial bonds are corporate bonds issued by industrial companies. Perpetual

[7] The corresponding division of Dbouk and Kryzanowski (2009) comprised of 1. All bonds, 2. Treasury and agency bonds combined, 3. Industrial bonds, 4. Utility bonds, 5. Financial bonds, 6. Foreign bonds. Furthermore, the created IO sets differentiated by rating (1. All, 2. AAA, 3. AA+/AA/AA-; 4. A+/A/A-; 5. Sub-Investment grade) but such analysis was not repeated in this study.

bonds, i.e. bonds with no maturity date, were removed from the sample as they can be seen as non-representative of many traditional investors' investment universe.

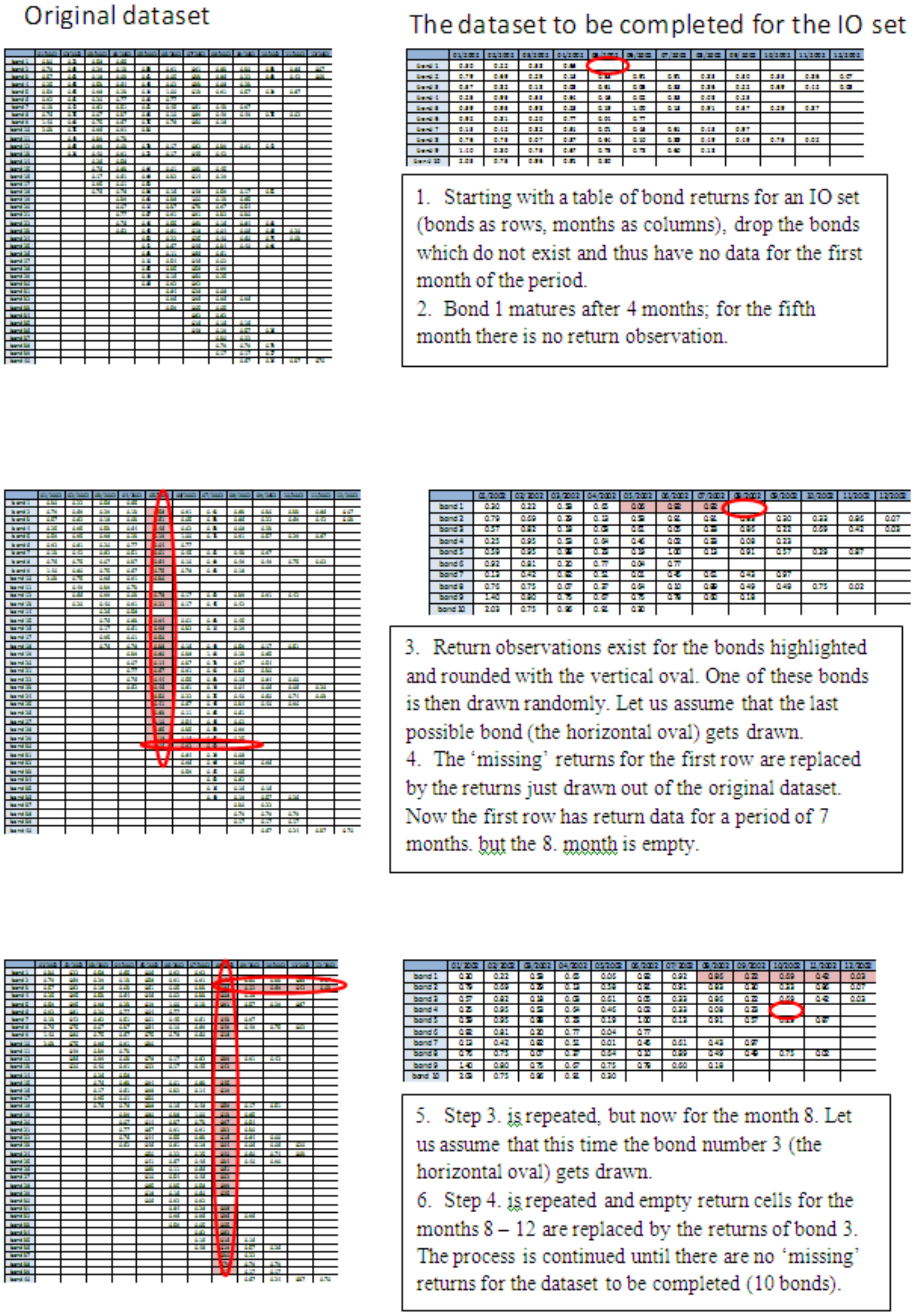
The database sample used in this study is roughly three times smaller than that used by D&K (2009) in their study[8] and some bond sector samples are rather small (e.g. 629 government and 921 industrial corporate bonds), which may pose challenges from the statistical inference point of view but the approach is more of a simulation than statistical hypothesis testing nature.

As bonds, unlike stocks, are instruments with a maturity date, the monthly bond return data had to be further prepared into a form allowing drawing random sample portfolios out of the data. To do this, the return time series of all bonds that mature during the period under review are continued by a return of a bond belonging to the same category (government, government-related, etc.). Conducting the procedure with a bond from the same or as similar an issuer as possible is impractical due to the sheer number of the bonds and not always possible. Furthermore, the choice of the replacement bond would in any case entail arbitrary features. Thus the continuing bond is drawn randomly (using random number generation in Stata) from the relevant category. The replacement process is continued until the dataset for each category is a panel data table without any blanks. For an illustration of the process, see Figure 2. The illustration is an imaginary example including 40 bonds and a period of one year only. A separate data table is compiled for the category 'All' as well as all the sector-specific IO sets (government bonds, government-related bonds, etc.), in the latter of which the matured bonds are replaced with another bond belonging to the same sector.

In order to examine the effect of time period choice on diversification benefits for the category 'All bonds', three tables of data are formed; first table consists of the first three years (36 months) of the period 2002 – 05. For the periods 2005 – 07 and 2008 – 10 a similar replacement process is repeated as before but using the bonds outstanding at the beginning of 2005 and 2008 as the bond universe starting point.

[8] The reference study by Dbouk and Kryzanowski (D&K, 2009) was conducted on bond sample extracted from the Lehman Brothers Fixed Income Database distributed by Warga (1998) and has been widely used in various bond studies (Elton et al. 2001; Liu et al. 2007). The database consists of 39,132 bonds and 1,289,010 monthly bond prices from January 1985 until December 1997.

Figure 2. A simplified illustration of the data preparation.



The way for preparing the data for the calculation was done in the aforementioned way because replacing the maturing bonds and drawing random portfolios simultaneously would not have been possible with the operating system and memory requirements of the computer and the statistical program Stata. The downside of the process is that is does not allow flexible examination of e.g. different time periods and that maturing bonds are not continued with a different random bond in each random portfolio. However, the large number of bonds ensures that the element of randomness is present at all times and constitutes a practically realizable process to enable the analysis. The data preparation method cannot be compared with that of the reference study by Dbouk and Kryzanowski (2009) as the authors do not address how the maturing bonds are handled and replaced.

6.2.2 Simulation analysis

In this study three different types of metrics are used to examine portfolio diversification benefits and to identify the optimal or minimum portfolio size (henceforth PS). The methodology follows closely the one used by Dbouk and Kryzanowski (2009) in their study. The ex-post simulation type calculations were run separately for each IO set, the preparation of which was described in the previous subsection.

The PS refers to the portfolio size at which marginal benefit in terms of diversifying away a specific percentage of nonsystematic risk or to capture a specific percentage of the reward from bearing risk becomes very small. In this study, as in the reference study by D&K (2009), that 'small marginal benefit' or SMB-criteria threshold is set to 1 percent. The actual identification of such a minimum portfolio size is done by selecting bonds randomly using a Monte Carlo approach in order to create 1000 portfolios for each investment opportunity (henceforth IO) differentiated by the bond industry sector set and each portfolio size and following how the diversification benefits develop when the portfolio size changes. The methodology of the study follows the D&K (2009) methodology but the analysis is restricted into three (3) metrics: mean-derived deviation, Sharpe ratio and kurtosis. The first measure represents a variation or volatility type of risk metric, the second a return-to-risk type of metric and the third measure takes into account the higher moments of distribution.

A PS ranging from 2 to 100 (at sizes 2, 5, 10, 15, …, 95, 100) is investigated as well as a portfolio 'All', where the latter is all the bonds in the index as a whole[9]. In this study equal weights are used in forming the portfolios in the same way as in Dbouk and Kryzanowski

[9] The figures for the totality of the bond universe is calculated from the self-completed data table in which return time series of maturing bonds are continued with another similar bond. Thus, the composition is not exactly the same as that of all the bonds outstanding but it should be rather similar in terms of the industry sector structure, as the replacement is done with similar but otherwise randomly drawn bonds.

(2009). The equity-related literature reports that no sample-based mean–variance portfolio formation strategy consistently performs better in terms of out-of-sample performance than an equal-weighted portfolio. Samuelson (1968) shows that if returns are identically distributed, an equally weighted portfolio can reach the optimum. D&K (2009) refer to the study by DeMiguel et al. (2007) where the authors find that none of the 14 models that they examine across empirical data sets is consistently better than the 1/N rule in terms of Sharpe ratio, certainty-equivalent return, or turnover. DeMiguel et al. (2007) hence conclude that estimation error more than offsets the out-of-sample gain achieved by optimizing diversification.

The overall benefits of diversification as PS increases can be estimated by examining the share in percentage terms of the potential benefits that are achieved, on average, for the specific PS vs. the potential benefits achievable from holding all the assets in the IO set. This is, indeed, also the most common method used to estimate the *overall* benefits of diversification (Dbouk and Kryzanowski, 2009). It should be noted that because the form of the metric distribution changes as the IO set, metric itself and PS change, the values of the dispersion metrics used in the determination of the minimum PS also will vary.

The most common method for estimating the *marginal* – contrary to the overall – benefits of diversification as the PS increases is to investigate the speed at which the value of the diversification metric changes (Campbell et al. 2001). Because the average correlation among security returns (i.e. market risk or non-diversifiable risk) limits the power of diversification to reduce risk, at a certain point an increase in PS results in only a small change in the metric measuring the marginal diversification benefits. As there are costs associated with diversification, rational investors will be reluctant to increase the PS when there are only 'small' benefits from increasing the PS to the next larger PS.

Like in the study by Dbouk and Kryzanowski (2009), a marginal change in the value of the diversification metric is taken here to be 'small' when it is 1% or less, as measured by a percentage improvement in the metric used. However, a 'minimum' PS based on this criterion of a small marginal benefit (SMB) may leave a considerable portion of the total diversification benefits unrealized, which is also shown in the results of Dbouk and Kryzanowski (2009). The methodology used by D&K and in this study is consistent with the argument by Statman (1987) who maintains that the level of diversification benefits in terms of reduction of unsystematic risk meaningless but that the marginal net benefit is what matters; diversification should thus be increased as long as the marginal benefits exceed the marginal costs.

The expected result for this study is that an increase in the PS reduces the overall risk of the portfolio with different measures of non-systematic risk or increase the reward from bearing the risk associated with the portfolio at different speeds and in different proportions when examining different bond sectors and measures. It is also expected that the identified minimum portfolio sizes should not consistently differ between those found by Dbouk and Kryzanowski (2009).

Furthermore, results for the periods 2002 – 04, 2005 – 07, and 2008 – 10 are compared against each other in order to see whether the diversification benefits have differed in different time periods that represent very different economic and financial market conditions. As the three subperiods represent extremely different economic environments, it may be that the results vary between the periods even though D&K (2009) do not find significant differences between different subperiods. The financial and economic crisis of the late-2000s, still ongoing as of this writing, can be expected to increase the minimum PS required, as in times of financial crisis the correlation of asset classes tends to increase, a phenomenon found in the literature and among practitioners (J.P. Morgan Asset Management, 2009). When correlation of asset classes is higher, more diversification than previously is required when targeting same diversification benefits than with lower return correlation. However, it is not entirely obvious whether increased correlation increases the minimum PS; it may hypothetically also be the case that the marginal diversification benefits do not get 'small' i.e. under 1 percent any slower in a high-correlation environment but that the unrealized overall diversification benefits become larger.

7. Results

The following metrics are used for estimating the benefits from diversification:

1. Dispersion of returns: mean derived deviation (MDD)
2. Composite return and risk metrics: Sharpe ratio
3. Metrics based on higher-order moments: kurtosis

7.1 Dispersion of bond returns: mean derived deviation (MDD)

The first metric examined in this section of the study is the excess standard or mean derived deviation (MDD) for a randomly selected portfolio. MDD is defined as the difference between the time-series standard deviations of the random portfolio and the whole investment opportunity set (henceforth IO set) to which that portfolio belongs. This metric, which is calculated for 1000 randomly selected portfolios for each (un)differentiated IO set, is given by

$$MDD_{j,s} = \bar{\sigma}_{j,s} - \sigma_J \tag{7}$$

where

$\bar{\sigma}_{j,s}$	is the mean of the standard deviations for the 1000 randomly selected portfolios with a PS or PS of s for un/differentiated IO set j
σ_J	σ_J is the average standard deviation of all the bonds in un/differentiated IO set j.

As expected, the MDD decreases with increasing portfolio size (henceforth PS), as can be seen from Figures 3. and 4. The minimum PS that satisfies the small marginal benefit (SMB) criterion ranges from 15 to 55 bonds for IO sets differentiated by issuer type. The overall diversification benefit at this minimum PS is quite substantial (MDD reductions ranging between 63% to 95%). It should be stressed that even though the statistical significance in the difference for the PS of 2 and All, the low number of bonds for some categories compels to reservations when interpreting the results. Hence, especially the results for government, government-related and industrial bond categories should be treated carefully. Large portfolio sizes in these categories constitute a substantial part of those IO sets, which may distort the results. Indeed, the results for government and government-related bonds seem to point to lower PS than the other IO sets.

Contrasting the findings with those by Dbouk and Kryzanowski (2009) who use dollar-denominated bond data from the period 1985 - 97, the results seem to point to similar portfolio sizes despite a slightly different IO set differentiation. The MDDs they find for IO sets differentiated by issuer type range between a PS of 35 for the Treasury/Agency sector to a PS of 45 for categories All, Utility, and Financial bonds. In D&K a PS of 40 is found for categories Industrial and Foreign. The category Covered/Securitized is not a category of its own in Dbouk and Kryzanowski's study but it can be said that the results here point to a risk level and behaviour closer to government- and government-related bonds than that of (unsecured) financial bonds. The overall MDD reduction ranges between 63 and 92 percent which is not very different from the findings by Dbouk and Kryzanowski (from 75 to 96 percent for IO sets differentiated by issuer sector and 72 to 93 for IO sets differentiated by rating). Furthermore, Dbouk and Kryzanowski find the MDD to decrease monotonically whereas the results of this study show MDD to first decrease monotonically but then fluctuating around a level seemingly the asymptote for each category. The slight instability or fluctuation around the asymptote may stem from a lower number of simulations of this study (1000) compared to the simulation size of 5000 in S&K (2009).

A comparison of the MDD for a specific PS for the time periods of 2002 – 04, 2005 – 07, 2008 – 10 for the category 'All' shows similar PS results than the sector comparison (for the sectors for which there is a large number of bonds and return observations available). The optimal PS satisfying the small marginal benefit criterion ranges between 40 and 55 and the respective MDD reductions vary within the very high range of between 87 and 92 percent. Dbouk and Kryzanowski do not observe significant PS differences during different time periods, but as their specific results in terms of the sensitivity of the results to the time period choice are untabulated a more detailed comparison with their results is not possible.

What is the most striking observation is, however, the enormous difference in the MDD level for the three periods especially with portfolio sizes clearly below 100. The average MDDs (multiplied by 100) for the periods 2002 – 04 and 2005 – 07 are below 10 (9.65 and 7.37, respectively) even with a PS of two, but for the latest period in the study, 2008 – 10, an MDD level below 10 is not reached before the PS of 40 (MDD 9.95). So even though in the period 2008 – 10 a PS of 50 is enough to bring the MDD metric down by 92 per cent, the absolute level of the remaining, undiversified risk is still practically as high as for a portfolio of just a couple of bonds in financially more stable times. When the PS approaches 100 the differences between periods become less pronounced but remain noticeable. Besides the severity of the financial crisis, the result also highlights that although the marginal reduction of dispersion is less than 1% beyond these optimal PSs, much potential diversification benefits may still remain unrealized.

Figure 3. Excess Standard Deviations (MDDs) for IO Sets Differentiated by issuer type

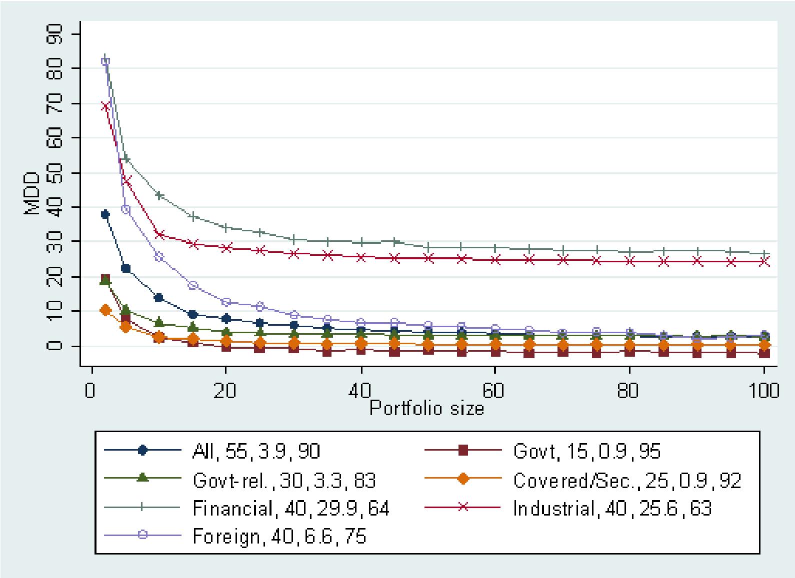


Figure 3. Excess Standard Deviations (MDDs) for IO Sets Differentiated by issuer type. The figure depicts the excess standard deviations or mean derived deviations (MDDs) of returns multiplied by 100. MDD is measured as the difference between the standard deviation of the 1000 random portfolios and an equally weighted index of all bonds in that IO set *j* for the whole period. The optimal PS is found when the reduction in the MDD is not more than 1% from incrementing the PS to the next larger PS provided that the difference in the mean MDDs for PS *s* of two and All are significantly different at the 0.05 level. The legend of each series lists the series name, optimal PS, the MDD value (multiplied by 100), and the percentage reduction in the MDD from a benchmark PS of two when the optimal PS is reached.

Figure 4. Excess Standard Deviations (MDDs) for IO Sets Differentiated by time period

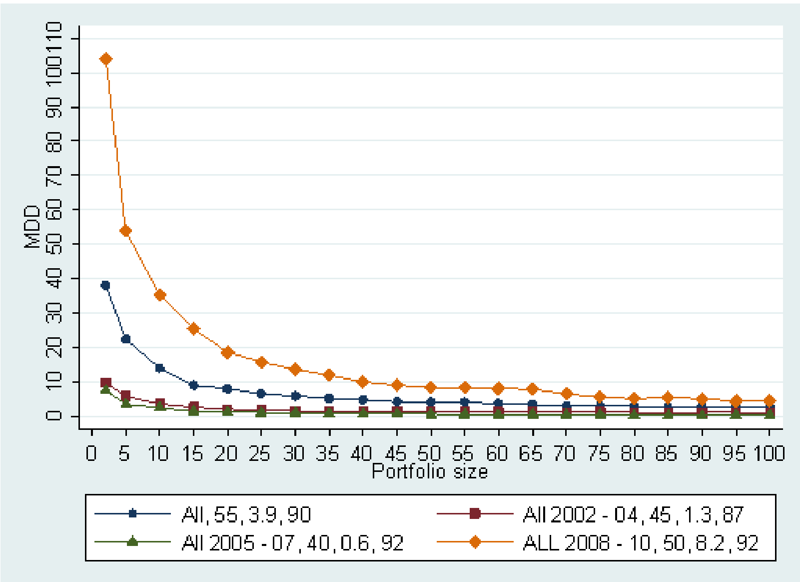


Figure 4. Excess Standard Deviations (MDDs) for IO Sets Differentiated by time period under review. The figure depicts the excess standard deviations or mean derived deviations (MDDs) of returns multiplied by 100. MDD is measured as the difference between the standard deviation of the 1000 random portfolios and an equally weighted index of all bonds in that IO set *j* for the period. The optimal PS is found when the reduction in the MDD is not more than 1% from incrementing the PS to the next larger PS provided that the difference in the mean MDDs for PS *s* of two and All are significantly different at the 0.05 level. The legend of each series lists the series name, optimal PS, the MDD value (multiplied by 100), and the percentage reduction in the MDD from a benchmark PS of two when the optimal PS is reached.

Dbouk and Kryzanowski examined also another dispersion metric, namely the average cross-sectional standard deviation, which sometimes is referred to as the mean realized dispersion or MRD (de Silva et al. 2001; Ankrim and Ding 2002). Risk-averse managers seek to reduce their exposure to higher MRDs. Dbouk and Kryzanowski find that, like the MDD, the MRD metric also points to a PS range of 35 – 45 both when differentiating IO sets by issuer sector and rating. With these portfolio sizes the overall MRD reduction is on average 76 – 80 %.

7.2 Composite return and risk metrics: Sharpe ratio

Investors are interested in not only the return or the risk of the portfolio separately, but finding the best return-risk tradeoff portfolios. Hence, a diversification strategy, such as increasing the portfolio size to decrease risk, should also improve the return-to-risk tradeoff. In this sub-section the Sharpe ratio metric is examined so as to see how it reacts to a changing PS for the IO sets differentiated by issuer sector and the period under review.

Metrics commonly used for assessing the return-to-risk relationship often normalize the excess return over the risk-free rate of the portfolio by the risk of that portfolio. The Sharpe ratio is defined as

$$S_P = \frac{\bar{R}_{j,s} - \bar{R}_F}{\bar{\sigma}(R_P)} \tag{8}$$

$\bar{R}_{j,s}$ denotes the mean return on the portfolios of size s for IO set j

$\bar{R}_F$ denotes the mean risk-free rate, here defined as the return of a 3-month German government bill index from Bank of America Merrill Lynch

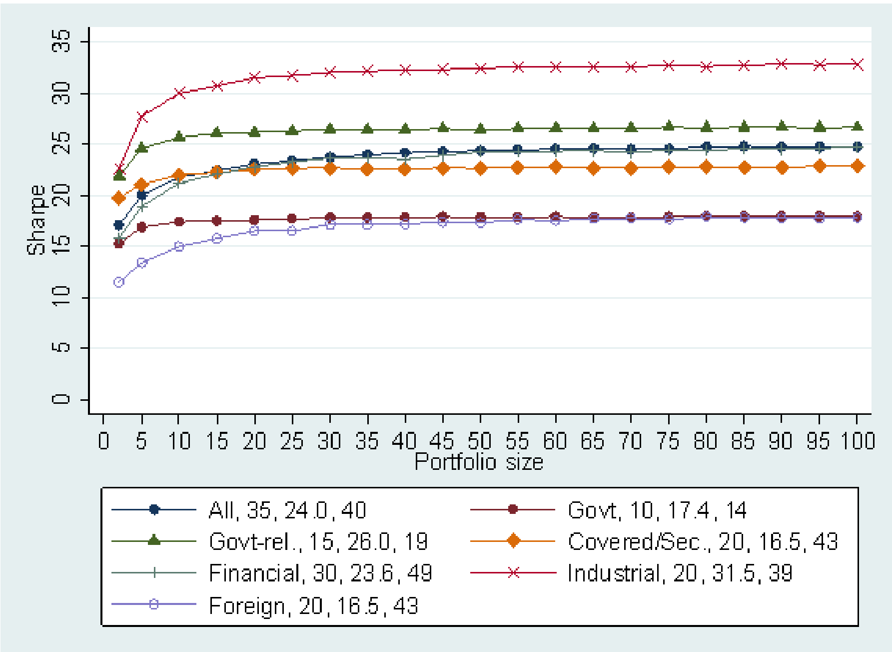
$\bar{\sigma}(R_P)$ denotes the average standard deviation of returns for portfolios of size s for IO set j

The results for the Sharpe ratio metric are shown in Figures 5 and 6. For IO sets differentiated by issuer type, the SMB-determined minimum PSs show a range between 10 (Government) and 35 (All) bonds for all IO sets, and the associated overall increases in their Sharpe ratios are in the range of 14 – 49 %. The results point to both lower PSs and smaller total improvements in the Sharpe ratios than those found in Dbouk and Kryzanowski (2009). They find that the SMB-determined minimum PSs for the Sharpe ratio range from 30 (Utility) to 50

(Industrial) bonds, and the associated improvement in the Sharpe ratios range between 80 percent for Treasury/Agency category and 95 percent for foreign bonds when differentiated by issuer type. For categories differentiated by rating the minimum PS varies between 25 (AA) to 60 (AAA) and the related Sharpe increase is from 67 to 92 percent.

Dbouk and Kryzanowski also examine diversification benefits using the Sortino ratio, in which the average standard deviation in the Sharpe ratio formula is replaced by the average semistandard deviation of returns for portfolios of each PS. When examining the Sortino ratio Dbouk and Kryzanowski observe optimal PSs in the range of 25 – 60 for categories differentiated by issuer type and portfolio sizes between 20 and 50 for IO sets grouped by rating. The results are thus very similar when using Sharpe and Sortino ratio.

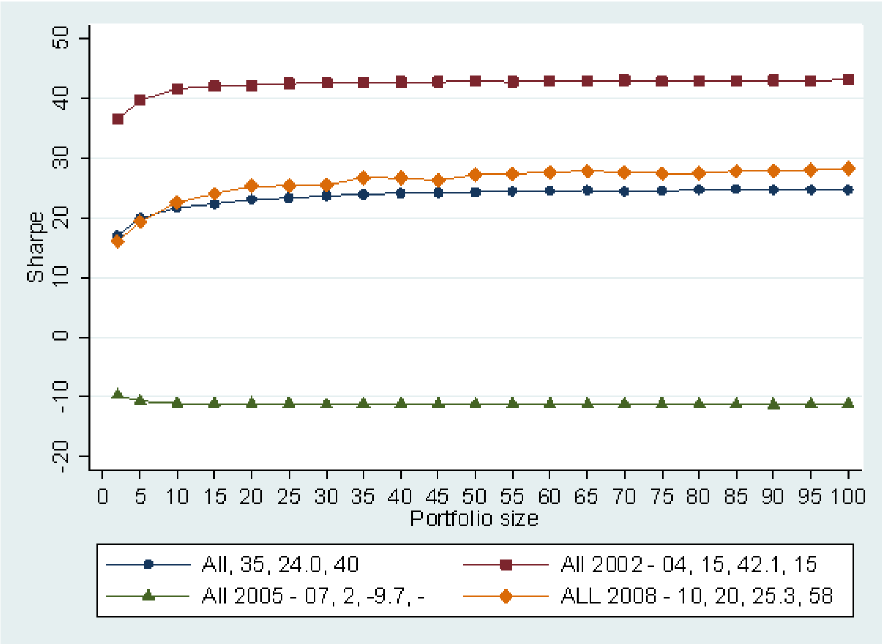
Figure 5. Sharpe ratios differentiated by PS and issuer type.



This figure depicts the Sharpe ratios multiplied by 100 of quoted returns differentiated by PS and issuer type. The optimal PS is reached when the reduction in the Sharpe is not more than 1% from incrementing the PS to the next larger PS provided that the difference in the mean Sharpe for PS *s* of two and All are significantly different at the 0.05 level. The series name, optimal PS, the Sharpe value, and the percentage reduction in the Sharpe from a benchmark PS of two when the optimal PS is reached are reported in that order for each series in the legend to the figure.

The results for IO sets differentiated by the issuer sector show rather low levels of overall diversification benefits. Government and government-related IO sets show particularly low levels of overall Sharpe ratio improvements (14 and 19 percent, respectively) with the optimal PSs. The SMB-determined optimal portfolio sizes are also smaller for these categories (10 and 15). The results may be related to the risk and return characteristics of government and government-related bonds in general which can be expected to show some resemblance. Other categories are showing higher PSs in the range of 20 – 35 at which the overall Sharpe ratio improvements are from 39 to 49 percent.

Figure 6. Sharpe ratios differentiated by PS and observation period

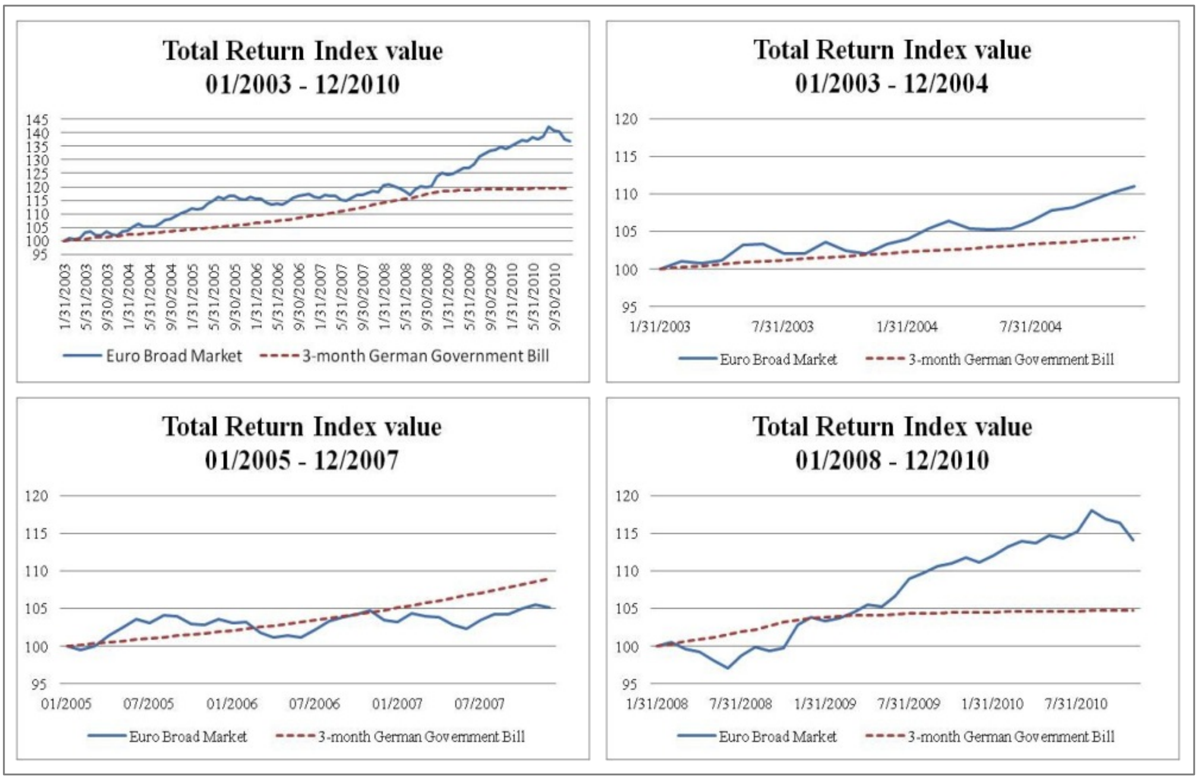


This figure depicts the Sharpe ratios multiplied by 100 of quoted returns differentiated by PS and time periods under observation. The optimal PS is reached when the reduction in the Sharpe is not more than 1% from incrementing the PS to the next larger PS provided that the difference in the mean Sharpe for PS *s* of two and All are significantly different at the 0.05 level. 'Tr./Ag.' refers to Treasury/Agency. The series name, optimal PS, the Sharpe value, and the percentage reduction in the Sharpe from a benchmark PS of two when the optimal PS is reached are reported in that order for each series in the legend to the figure.

The results for IO sets differentiated by the period under review are somewhat surprising and warrant further analysis. A clear outlier is the period 2005 to 2007 for which the Sharpe ratios for the IO set of 'All' are negative; in other words excess returns over the risk-free rate represented by the 3-month German government bills, have been negative in this period. Furthermore, the negativity seems quite notable. The observation is investigated further with comparing the total return index value development of the Merrill Lynch EMU Euro Broad Market Index and the 3-month German government bills for the three periods. Total return index is an index that calculates the performance of the groups of bonds (here the Euro broad market index bonds and 3-month German bills) assuming that all coupons and redemptions are reinvested. Stock indices (such as S&P500) tend to be better-known but they exist for bonds as well such as the ones investigated below.

The total return indices for the subperiods confirm the negative Sharpe ratio finding for the period 2005 to 2007. Total return index data for the year 2002 is not available from Merrill Lynch, but data from 2003 to 2010 support the results of the simulation analysis. For the periods 2003 to 2010 and 2008 to 2010 the excess returns are clearly positive. So is the case for the period 2003 to 2004 as well, but direct comparison to the Sharpe ratio results for 2002 to 2004 is not possible because of the time period discrepancy. For the outlier sub-period of 2005 to 2007 the broad market index has underperformed the 3-month German government bill, which explains the negative Sharpe ratios.

Figure 6. Total Return Index values for Euro Broad Market and 3-month German government bill index



Source: Bank of America Merrill Lynch Global Index System database, accessed 13 November 2011. Note that the total return indices begin at the beginning of 2003 (as opposed to 2002 like elsewhere in this study), because the total return index values are only reported as of beginning of 2003 in the Merrill Lynch database.

What is also notable in the 2005 to 2007 results is that in addition to negative excess returns the Sharpe ratio seems to decrease rather than increase when the portfolio size increases. The decrease is not, however, statistically significant at 5% confidence level when comparing PSs 2 and 100. The seeming decrease may be related to the observation by Israelsen (2005) that Sharpe ratio may be unreliable when risk premiums are negative. Israelsen notes that both the Sharpe and the information ratio may result in deceptive rankings of portfolios when risk premiums or excess returns are negative. In such cases a higher Sharpe (or information) ratio does not always point to the best portfolio and the reliability of the measure may be compromised.

7.3 Metrics based on higher-order moments: Kurtosis

The volatility and return-to-risk metrics examined thus far have the benefit of simplicity, robustness and independence of a reference benchmark market index[10], there are higher dimensions of risks that the metrics are not able to capture. They ignore the existence of third and fourth movements that may be unfavorable to the investor.

Higher moments beyond mean (first moment) and volatility (second moment) are skewness (third moment) and kurtosis (fourth moment). D&K (2009) note that lower second return moments may occur for PSs along with fatter tails. Moreover, Lo (2001), among others, argues the Sortino ratio can be manipulated by transferring part of the risk from the first and second-order moments to the third and fourth-order moments. Recent portfolio-choice, asset-pricing, value-at-risk, and option-valuation models highlight the importance of modeling the asymmetry and fat tails of returns (Jondeau and Rockinger, 2003). In a cross-section setting, they also document co-variability of moments beyond volatility, suggesting that extreme realizations tend to occur simultaneously on different markets.

Studies have documented that investors prefer to construct portfolios with positive skewness provided that the mean of returns normally falls above the median (Harvey and Siddique 2000; Premaratne and Tay 2002). Hence, an increase in PS that makes skewness more positive or less negative is considered valuable. Dbouk and Kryzanowski (2009) find that the the mean of the time-series of cross-sectional mean skewnesses is highly positive at a PS of two, and decreases monotonically as the PS increases from 2 to all bonds for all IO sets. Thus, the SMB-determined minimum PS of two bonds is preferred for skewness for all IO sets. There seems to be no reason to expect the study at hand to yield different results which is why this study focuses on the kurtosis metric instead.

Kurtosis is a measure of the peakedness and tail heaviness of the probability distribution of a random variable. Higher kurtosis means more of the variance is the result of infrequent

[10] The risk-free rate in the Sharpe ratio formula is one type of a reference index, but here a benchmark index is understood as a target rate of return or market portfolio type of reference index.

extreme deviations, as opposed to frequent, modestly sized deviations. According to Ruppert (1987), kurtosis measures both peakedness and tail weight, because if probability mass is moved from the flanks to the center of a distribution, then mass has to be moved from the flanks to the tail to keep the scale fixed. As a result, Brys et al. (2006) conclude that, since no agreement exists on what kurtosis really estimates, its use is often restricted to symmetric distributions. They also note that the kurtosis coefficient is very sensitive to outliers in the data.

The mean of the time-series of cross-sectional mean kurtosis is given by

$$\mu_{Kurt_{j,s}} = \frac{1}{N}\sum_{\tau=1}^{N} Kurt_{j,s,\tau} \tag{9}$$

where:

$Kurt_{j,s,\tau}$	is the cross-sectional kurtosis for the 1000 randomly selected portfolios for IO set j with a PS of s for month τ
N	is the number of cross-sections

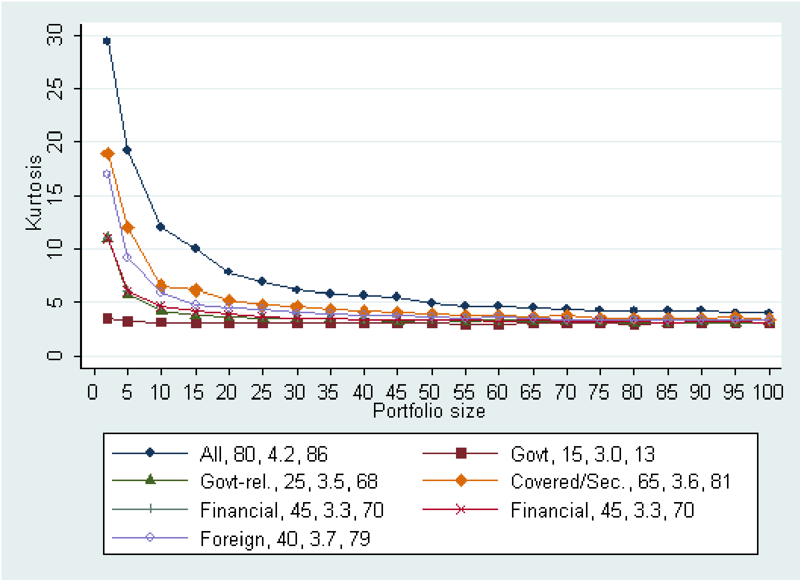
The kurtosis metric, as shown in Figures 7. and 8., improves (i.e. declines) with an increase in PS from 2–100 for all IO sets, which is plausible as in portfolios with more bonds the relative weight of one outlier observation is smaller than in a portfolio consisting of only few bonds. If risk-averse investors weigh potential downside returns more than potential upside returns as a typical investment grade bond investor probably does, then these investors will prefer a distribution with low kurtosis, since the tails are more likely to fall closer to the mean. Thus, the risk of an extreme loss decreases as PS increases from 2–100 bonds. Unlike the other metrics, diversification benefits are captured by the decrease in kurtosis between a PS of two, and the PS under investigation since measuring the potential diversification benefits as the difference in the kurtosises between a PS of two and PS of 'All' bonds is not applicable due to the very high kurtosis for a PS of 'All' bonds. Here the cross-sectional kurtosis for the category 'All' is 103.6 while the kurtosis for PS of 2 is 29.5.

Based on the results for IO sets differentiated by the issuer sector the SMB-determined minimum PSs for the Kurtosis vary in the rather wide range between 15 and 80 bonds. The diversification benefits for government bond category is a very low 13 percent, which probably has to do with eurozone government bonds having traded very close to German yields for the most of the period 2002 to 2010, although in especially 2010 the spreads of some eurozone governments blew up dramatically[11]. It should also be noted, however, that the IO sets pointing to the lowest PSs, namely government and government-related (PSs of 15 and 25, respectively) are sets with a relatively low number of bonds in the original sample (625 for the government bonds and 1099 for the government-related bonds). Furthermore, the small improvement in the Kurtosis metric for the government bonds is explained by the already very low starting level of 3.5 (the asymptote of the metric being 3.0). The governments the bond spreads of which compared to Germany widened the most in the late 2000s, most notably Greece, Portugal and Ireland, represent rather small portion of the overall government bond market. This is probably the reason why the government debt turmoil does not seem to be reflected in the results.

For other IO sets than government bonds diversification benefits are rather high (68 to 86 percent). Excluding the government and government-related bonds (showing PSs of 15 and 25, respectively), the PSs range from 40 to 80 and the associated kurtosis reduction from 70 to 86 per cent. Dbouk and Kryzanowski find a considerably lower range of 20 to 30 for the PS, but the Kurtosis metric improvements show a similar, if yet somewhat tighter a range of 80 to 86 percent for the overall kurtosis improvement. As to the kurtosis levels for portfolio size of 2, Dbouk and Kryzanowski find a kurtosis level of roughly 50 for categories All, Financial, and Industrial. The level found by them is clearly higher than in the results found in this study, for example 29.5 for the IO set All and kurtosis of below 20 for all other categories. This may stem from the most of the period under investigation being rather stable. Corresponding calculations for e.g. 2009 to 2011 would probably yield very different results. Here it should be highlighted that for government bonds the kurtoses of 2 and All are not statistically significantly different at the 95% confidence level.

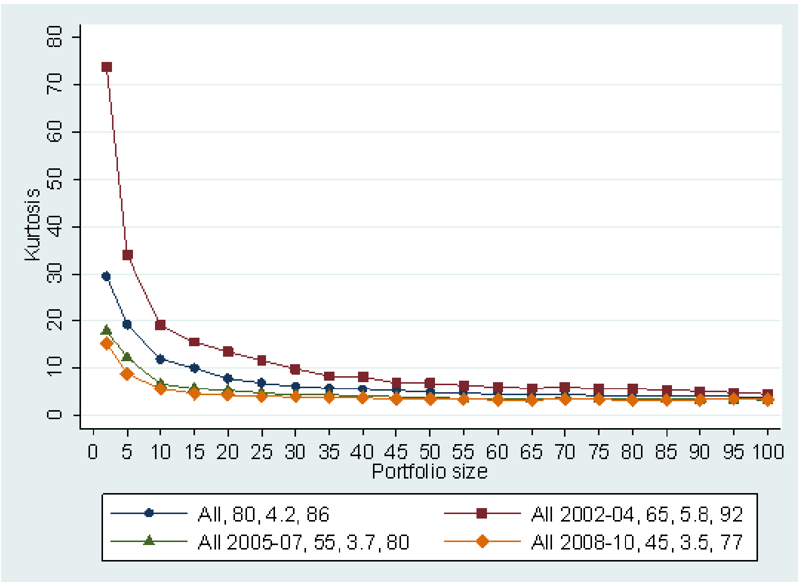
[11] The widening has continued and further exacerbated in 2011 but the period is beyond the scope of this study.

Figure 7. Kurtosises differentiated by PS and issuer type.



This figure depicts the kurtosises (Kurts) of quoted returns differentiated by PS and issuer type for all maturities. The optimal PS is reached when the reduction in the Kurt is not more than 1% from incrementing the PS to the next larger PS provided that the difference in the mean Kurts for PS *s* of two and All are significantly different at the 0.05 level. The series name, optimal PS, the Kurt value, and the percentage reduction in the Kurt from a benchmark PS of two when the optimal PS is reached are reported in that order for each series in the legend to the figure.

Figure 8. Kurtosises for differentiated by PS and period under observation.



This figure depicts the kurtosises (Kurts) of quoted returns differentiated by PS and issuer type for all maturities. The optimal PS is reached when the reduction in the Kurt is not more than 1% from incrementing the PS to the next larger PS provided that the difference in the mean Kurts for PS *s* of two and All are significantly different at the 0.05 level. The series name, optimal PS, the Kurt value, and the percentage reduction in the Kurt from a benchmark PS of two when the optimal PS is reached are reported in that order for each series in the legend to the figure.

When differentiated by the time period, a PS range from 45 to 80 is observed with an overall kurtosis metric reduction between 77 and 92 percent. The high cross-sectional kurtosis for small portfolio sized for the period 2002 to 2004 is notable. It is not entirely clear what causes the high kurtosis in the period. It may have to do with the early stage of development of the European, euro-denominated bond market. Similarly, the low level of the Kurtosis metric for the period 2008 – 10 is somewhat a surprising finding.

It is also notable that the relation between kurtosis and PS is convex, i.e. it first decreases as the portfolio size increases, but later as PS increases further, the kurtosis begins to increase as well. This results in the kurtosis for the portfolio of bonds in the index to be clearly higher than the corresponding kurtosis at a portfolio size of two (D&K, 2009). For example, the kurtosis for the whole bond universe for the whole time period of 2002 to 2010 is 103.6 which is considerably higher than the Kurt of 30 for PS of 2. The results of this study is consistent with the finding of D&K. As highlighted by them, the property shows a difficulty for interpreting kurtosis changes as an isolated issue. According to D&K this illustrates that kurtosis is measures both the tail heaviness of a distribution as well as the peakedness of the distribution. Hence, their relative influence on the kurtosis metric may vary as the portfolio size changes. More specifically, it happens because the value of the numerator of the Kurtosis formula first declines at a faster pace than the denominator (the standard deviation raised to the power of four), but later the numerator declines slower than the denominator. Skewness, in turn declines monotonically as both the numerator and the denominator are raised to the power of three.

In addition to skewness and kurtosis, Dbouk and Kryzanowski (2009) investigate the effect of diversification on the left tails of portfolio returns as well as on the probability of observing market underperformance. The left tail weight (LTW) measure shows minimum PSs in the range between 70 and 100 for market underperformance probability metric, in turn, points to very low PSs in the range 2 to 15.

8. Summary and conclusions

In this study, the minimum portfolio sizes (PSs) required to capture most of the diversification benefits from increasing PSfor various measures of diversification benefits are examined for investment opportunity sets (IO sets) differentiated by issuer type or time period under review. Diversification benefits are examined by investigating the minimum portfolio sizes beyond which the marginal benefits of increasing the PS are less than 1%.

Table 5. Summary of the minimum portfolio sizes beyond which the marginal benefits of increasing PS are less than 1%.

IO sets differentiated by issuer type				IO sets differentiated by subperiod			
IO set	MDD	Sharpe	Kurtosis	IO set	MDD	Sharpe	Kurtosis
All	55	35	80	All	55	35	80
Government	15	10	15	2002 - 2004	45	15	65
Govt-related	30	15	25	2005 - 2007	40	2	55
Covered/Sec.	25	20	65	2008 - 2010	50	20	45
Financial	40	30	45				
Industrial	40	20	45				
Foreign	40	20	40				

Note: This table reports the minimum PSs beyond which the marginal benefits of increasing the PS to the next PS are less than 1% for various diversification metrics. These include: MDD (mean excess standard deviation) , Sharpe ratio and kurtosis. Govt-related refers to government-related and Covered/Sec. to Covered and securitized bonds.

Based on the results summarized in Table 5. it is found that the the minimum PSs vary by the metric used to measure the marginal benefits of further diversification and the factor used to differentiate IO sets (here bond sector and the time period under review). The finding is consistent with the results by Dbouk and Kryzanowski who studied US dollar denominated bonds in 1985 – 97.

In the reference study by D&K(2009) the marginal benefits of further diversification are generally achieved with PSs of 25–40 bonds, but the unrealized advantages of full diversification (i.e., holding all bonds in the IO set) at these PSs are still notable. This study finds somewhat higher range of 40 to 55 bonds for most IO sets when using the MDD metric to measure the diversification benefits for both different issuer sectors as for different observation periods. The overall diversification benefits with the optimal PSs seem quite similar. Dbouk and Kryzanowski observe overall MDD improvements (reductions) in the range of 72 – 96 percent whereas this study finds a range from 63 to 92 percent.

Using the Sharpe ratio metric the PSs found in this study are between 10 and 35 for most categories and associated diversification metric benefits 14 – 58 percent compared to a benchmark portfolio of size two. The difference to the D&K findings is quite drastic. D&K observe a minimum PS between 30 and 50 for categories differentiated by issuer sector and between 25 and 60 for different rating categories and associated Sharpe metrics improvements of 80 – 95 percent. One hypothesis is that the market conditions and thus the period under review is a major driver for the Sharpe ratio. The period of 2002 – 04 when the bond market and credit spreads were recovering from the dot-com boom resulted in a favourable return-to-risk profile for bonds whereas the interest rate level increase in 2005 – 07 led to losses and poor risk-return characteristics for bond portfolios. These very different interest rate environments prevailing at different times may make the overall risk is more burdensome to diversify away and remains high even with large portfolio sizes.

Furthermore, as can be seen from the time period differentiated Sharpe results, the Sharpe ratio for the period 2002 – 10 hides considerable sub-period differences. Surprisingly, the results show very similar Sharpe ratios for the whole period of 2002 – 10 and the time of the current financial crisis in 2008 – 10. Interestingly, the euro broad market index in 2005 – 2007 were negative, thus resulting in negative Sharpe ratios. Furthermore, a larger PS during this period seems to result in a lower rather than higher Sharpe ratio, although the difference is not statistically different at 95% confidence level. The negative Sharpe ratios stem from 3-month German government bills (used as a proxy for the risk-free rate) having provided a higher return than the overall market index. The 3-month bills have not suffered as high capital losses as interest rates increased and they have allowed the investor to reinvest the proceeds at a higher rate. Furthermore, the Sharpe ratio seems to have declined rather than increased with higher PSs in 2005 – 2007.This observation is somewhat surprising, but it may be related to Israelsen’s (2005) argument that Sharpe ratio may be unreliable when risk premiums are negative and its use may lead to deceptive results.

Examination of the Kurtosis metrics in this study points to PSs within the wide range of 15 – 80 for the investment opportunity sets differentiated by issuer sector but excluding the categories of government and government-related bonds the range is between 45 and 80. The low PSs and rather small overall diversification benefits may stem from the low riskiness and that both categories were long considered quite homogenous. The divergence between governments has become notable in the past couple of years, but most of the period under review in this study more or less all investment grade rated governments were seen as relatively safe investments and were thus trading at levels close to each other and moving in unison. Furthermore, similar results for these categories as such are not very surprising, as both sectors contain government-specific risk.

The corresponding range found by Dbouk and Kryzanowski for different issuer sectors is 20 – 30 with the kurtosis metric. The PSs found by the authors are considerably lower than the ones found in this study. This may have to do with the higher overall Kurtosis levels found by D&K using US, dollar-denominated bond data. Excluding government and government-related categories of this study and the category 'Foreign' in D&K, the overall percentual improvements, in turn, look very similar (68 - 92 % here vs. 66 – 88% in D&K 2009).

When examining the three sub-periods the metric values and potential overall diversification benefits vary, a finding also observed by Dbouk and Kryzanowski. However, when looking at the periods 2002 – 2004, 2005 – 2007 and 2008 – 2010 the minimum PSs do not vary dramatically when using the MDD metric. The PS variation range for different periods with the MDD measure is 40 to 55. However, this study does find quite a drastic difference in optimal PSs for different periods when using the Sharpe ratio and cross-sectional kurtosis whereas D&K do not observe notable differences between different subperiods. They find that there are no significant changes in the SMB-determined PSs across the samples even though the metric values, potential diversification benefits and the form of distribution differ across these three time periods. In general, the optimal PS D&K (2009) find is about 40 bonds with associated diversification benefits of about 80%. In this study the results using Sharpe ratio point to PSs between 2 and 35 with associated overall benefits ranging from no benefit (or actual damage) to 58 percent. Examination of the cross sectional kurtosis metric points to PS between 45 and 80. The associated benefits in terms of kurtosis reduction are (77 to 92%) in line with D&K (2009) findings.

Perhaps the most notable limitation of this study is the data preparation method, which was resorted to due to operating system and computing capacity of the Stata and operating system. The downside of the process is that is does not allow flexible examination of e.g. different time periods and that maturing bonds are not continued with a different random bond in each random portfolio. There may be some bias caused by the data preparation, but its exact effect is hard to estimate. It can be assumed that the most important characteristics of the data persist despite the preparation, as shown by the consistency of Sharpe ratios for the three subperiods and the period of 2002 to 2010 compared to the total return indices of the whole Euro Broad Market Index and three-month German government bonds. Also, the large number of bonds in most IO sets ensures that the element of randomness is present at all times and constitutes a practically realizable process to enable the analysis. Furthermore, the choice of the replacement bond would in any case entail arbitrary features. For some IO sets the number of bonds and return observations is rather low for this type of an analysis, which may affect the reliability of the findings.

Bond diversification benefits are an interesting research topic and research on bonds seems to lag that on equity despite the importance of bond markets in the economy. Potential topics for future research might include examining the global bond market as opposed to a single market. Furthermore, modern measures utilizing the development of performance measurement should be used to examine diversification. As bonds are instruments with a fixed or limited upside but a potentially large downside, the use of measures focusing on the downside risk seems particularly interesting. Furthermore, a closer examination of the relationship between return time-series and volatility and cross-sectional volatility would be interesting. Dbouk and Kryzanowski (2009) find very similar PSs using both criteria, but whether the connection persists in other markets and other periods would be interesting to examine. Also the use of measures that account for the higher moments warrants more analysis. Another interesting topic will be to examine the impact of the current financial crisis on diversification benefits and how bonds are regarded as financial instruments. More specifically, it remains to be seen whether risks and returns of bonds containing credit risks is fundamentally changing.

A very natural topic for further research would be to combine the modern diversification benefit or performance measures and the small marginal diversification benefit approach and examine equity portfolios and other asset classes as well as multi-asset class portfolios.

To summarize, strictly speaking no definite minimum PS exists but it depends on the risk, return and bond maturity objectives of the investor as well as on the factor used to differentiate the IO set (here the issuer) and the period under investigation. Dbouk and Kryzanowski state the same insight. Furthermore, this study comes to the same conclusion that the small marginal benefit-determined portfolio sizes still leave much potential diversification benefits unrealized for many IO sets. As also noted by Dbouk and Kryzanowski (2009), this may explain the observation found earlier by Kahn and Rudd (1995) that individuals purchase bond mutual funds although studies find that bond mutual funds exhibit neutral and under performance when evaluated using gross and net-of-fees returns, respectively. This is caused by the difficulty and cost for retail investors of forming their own bond portfolios that capture a high percentage of the potential benefits of full diversification and to earn value-neutral benchmark- and risk-adjusted returns through self management. Diversification in a bond context is a research area where progress can made in future.

References

Abad, P., Chuliá, H., Gómez-Puig, M., 2009. EMU and European government bond market integration. Working paper series No. 1079. European Central Bank, Frankfurt.

Association for Financial Markets in Europe [Internet]. [updated 2011 Sep 15]. London, UK. [Accessed 2011 Sep 15]. Available at www.afme.eu

Bankers Almanac Credit Rating Definitions [Internet]. [updated 2011 Nov 4]. [retrieved 2011 Nov 4]. Available at http://www.bankersalmanac.com/addcon/infobank/definition-of-credit-ratings.aspx

Bank for International Settlements Quarterly Review September 2011 [Internet]. [updated 2011 Sep 15]. [retrieved 2011 Nov 4]. Available at http://www.bis.org/publ/qtrpdf/r_qa0812.pdf

Bank of America Merrill Lynch Global Index System [Accessed 2011 Jan 05 – 2011 Nov 10]

Bauböck, R., 2007. Why European Citizenship? Normative Approaches to Supranational Union. In: Theoretical Inquiries in Law. Berkeley Electronic Press, Berkeley.]. [updated 2007 Oct 22]. [retrieved 2011Sep 15] Available at http://www.bepress.com/til/default/vol8/iss2/art5

Bennett, J., Sias, R., 2006. Why company-specific risk changes over time. Financial Analysts Journal 62(5), 89–100.

Bhattacharya S., Pfleiderer P., 1983. A Note on Performance Evaluation. Technical Report 714. Graduate School of Business, Stanford University, Palo Alto, California, the US.

Bhanot, K., Guo, L., 2011. Negative Credit Spreads: Liquidity and Limits to Arbitrage. The Journal of Fixed Income 21(1), 32-41.

Biais, B., Declerck, F., Dow, J., Portes, R., von Thadden, E., 2006. European corporate bond markets: transparency, liquidity, efficiency. Discussion paper. Centre for Economic Policy Research, City of London.

Black F., 1972. Capital market equilibrium with restricted borrowing. Journal of Business 45, 444-455.

Blackstone, B., Lauricella, T., 2010 [Internet]. [updated 2010 Feb 4]. Global Markets Shudder – Doubts about US Economy and a Debt Crunch in Europe Jolt Hopes for a Recovery. The Wall Street Journal. [retrieved 2011 Aug 20]. Available at http://online.wsj.com/article/SB10001424052748704041504575045743430262982.html

Bloomberg L.P. [Accessed 2011 Jan 05 – 2011 Nov 15]

Brennan M., 1970. Taxes, Market Valuation and Corporate Financial Policy. National Tax Journal 25, 417-427.

Brown, K., Harlow, W., Tinic, S., 1993. The risk and required return of common stock following major price innovations. Journal of Financial and Quantitative Analysis 28(1), 101–16.

Brys, G., Hubert, M., Struyf, A., 2006. Robust Measures of Tail Weight. Computational Statistics and Data Analysis 50, 733 - 759.

Campbell, J., and L. Hentschel. 1992. An asymmetric model of changing volatility in stock returns. Journal of Financial Economics 31, 281–318.

Campbell, J., Lettau, M., Malkiel, B., Xu, Y., 2001. Have individual stocks become more volatile? An empirical exploration of idiosyncratic risk. Journal of Finance 56(1), 1–43.

Cantaluppi L. and Hug R., 2000. Efficiency Ratio: A New Methodology for Performance Measurement. Journal of Investing 9(2), 19 – 25.

Campisi S., 2004. The Case for Money-Weighted Performance Attribution. Journal of Performance Measurement 8(3).

Catts, T., 2011. Colgate-Palmolive Beats P&G With Low Coupons. Bloomberg [updated 2011 Nov 14]. [retrieved 2011 Nov 15]

Chollerton, K., Pieraerts, P., Solnik, B., 1986. Why invest in foreign currency bonds. Journal of Portfolio Management 12(4), 4–8.

Cvitanic J., Lazrak A., Wang T., 2007. Implications of Sharpe Ratio as a Performance Measure in a Multi-Period Setting. Working Paper. Available at SSRN: http://ssrn.com/abstract=962187.

Darling R. and MacDougall A., 2002. Using Performance Statistics: Have the Measurers Lost the Plot? Working Paper.

Dbouk, W., Kryzanowski, L., 2009. Diversification benefits for bond portfolios. The European Journal of Finance, 15(5), 533 – 553.

Demirer, R., Lien, D., 2004. Firm-level return dispersion and correlation asymmetry: Challenges for portfolio diversification. Applied Financial Economics 14(6), 447–56.

De Santis, R., Sarno, L., 2011. Assessing the benefits of international portfolio diversification in bonds and stocks. Working Paper number 883, European Central Bank. Available at SSRN: http://ssrn.com/abstract=1105383.

DiBartolomeo D., Witkowski E., 1997. Mutual Fund Misclassification: Evidence Based on Style Analysis. Financial Analysts Journal 53, 32-43.

Dimson E., Jackson A., 2001. High-Frequency Performance Monitoring. Journal of Portfolio Management 28(1), 33 – 43.

Dowd K., 2000. Adjusting for Risk: An Improved Sharpe Ratio. International Review of Economics and Finance 9(3), 209-222.

Elton, E., Gruber, M., 1977. Risk reduction and portfolio size: An analytical solution. Journal of Business 50(4), 415–37

Ericsson, J., Renault, O., 2001. Liquidity and Credit Risk. Working paper.

European Central Bank Statistics Pocket Book March 2011, 2011[Internet]. [updated 2011 Mar 9]. [retrieved 2011 Oct 16]. Available at http://www.ecb.int/pub/spb/html/index.en.html.

Evans, J., Archer, S., 1968. Diversification and the reduction of dispersion: An empirical analysis. Journal of Finance 23(5), 761–67.

Fabozzi, F., 2007. Bond markets, analysis and strategies. Prentice Hill, Upper Saddle River, New Jersey, USA.

Fisher, L., Lorie, J.H., 1970. Some Studies of Variability of Returns on Investments in Common Stocks. Journal of Business 43, 99-134

Ferson ,W.,Schadt R., 1996. Measuring Fund Strategy and Performance in Changing Economic Conditions. Journal of Finance 51(2), 425-461.

Fiori, A., Zenga, M., 2009. Karl Pearson and the Origin of Kurtosis. International Statistical Review, International Statistical Institute 77(1), 40-50.

French, K., Schwert, W., Stambaugh, R., 1987. Expected stock returns and volatility. Journal of Financial Economics 19, 3–29.

Goetzmann, W., Li, L., Rouwenhorst, K., 2005. Long-term global market correlations. Journal of Business 78(1), 1–38

Graham J., Harvey C., 1997. Grading the Performance Market-Timing Newsletters, Financial Analysts Journal 53(6), 54 – 66.

Grantier B., 1988. Convexity and Bond Performance: The Benter the Better. Financial Analysts journal 44(6), 79 -81.

Gray K. Jr., Dewar R., 1971. Axiomatic Characterization of the Time-Weighted Rate of Return. Management Science 18(2), B32 – B35.

Grinblatt M., Titman S., 1989. Portfolio Performance Evaluation: Old Issues and New Insights, Review of Financial Studies 2, 393-421.

Grinblatt M., Titman S., 1993. Performance Measurement without Benchmarks: an Examination of Mutual Fund Returns. Journal of Business 66(1), 47-68.

Groeneveld, R., Meeden, G., 1984. Measuring skewness and kurtosis. The Statistician 33(4), 391–399.

Harvey, C., Siddique, A., 2000. Conditional Skewness in Asset Pricing Tests. The Journal of Finance LV(3), 1263 – 1296.

Henriksson R., Merton, R., 1984. Market Timing and Mutual Fund Performance: an Empirical Investigation. Journal of Business 57(1), 73-96.

Houweling, P., Mentink, A., Vorst, T., 2003. Comparing possible proxies of corporate bond liquidity. Working Paper, Erasmus University Rotterdam.

Hosking, J., 1990. L-moments: analysis and estimation of distributions using linear combinations of order statistics. Journal of the Royal Statistical Society Series B, 52, 105-124.

Hunter, D., Simon, D. 2004. Benefits of international bond diversification. Journal of Fixed Income 13, 57–72.

Hunter, D., Simon, D., 2005. A conditional assessment of the relationships between the major world bond markets. European Financial Management 11(4), 463–82

Investing in Bonds Europe, 2011[Internet]. [updated 2011 Apr 14]. [retrieved 2011 Apr 14]. Available at http://www.investinginbondseurope.org .

Israelsen, C., 2005. A Refinement to the Sharpe Ratio and Information Ratio, Journal of Asset Management 5(6), 423 - 427.

Jacquier E., Kane A., Marcus A., 2003. Geometric or Arithmetic Mean: A Reconsideration. Financial Analysts Journal, vol. 59(6), 46 – 53.

Jensen M., 1968. The Performance of Mutual Funds in the Period 1945-1964. Journal of Finance 23(2), 389-419.

Jensen M., 1972. Optimal Utilization of Market Forecasts and the Evaluation of Investment Performance. In: Mathematical Methods in Investment and Finance, Szego G. and Shell K., Elsevier.

Johnson, S., 2007 [Internet]. [updated 2007 Jun 6]. Markets still underpricing risk. Reuters [retrieved 2011 Aug 20]. Available at http://www.reuters.com/article/2007/06/12/us-financial-summit-risk-idUSN1223893620070612

Jorion, P., 1987. Why buy international bonds? Investment Management Review (September/October): 19–28.

Jorion, P. 1989. Asset allocation with hedged and unhedged foreign stocks and bonds. The Journal of Portfolio Management 15(4), 49–54.

Jorion, P., 2006. Value at Risk: The New Benchmark for Managing Financial Risk. 3rd Edition, McGraw-Hill.

Kahn, R., Rudd, A., 1995. Does Historical Performance Predict Future Performance? Financial Analysts Journal 51(6), 43-52.

Kaplanis, E., Schaefer, S., 1991. Exchange risk and international diversification in bond and equity portfolios. Journal of Economics and Business 43(4), 287–307.

Lane, P., 2005. Global bond portfolios and EMU. European Central Bank Working Paper Series No. 553.

Latané, H., Young, W., 1969. Test for portfolio building rules. Journal of Finance 24, 595–612.

Le Sourd, V., 2007. Performance Measurement for Traditional Investment. EDHEC Risk and Asset Management Research Centre.

Lo, A., 2002. The Statistics of Sharpe ratios. Financial Analysts Journal 58(4), 36-52.

Lobosco A., 1999. Style/Risk-Adjusted Performance. Journal of Portfolio Management 25(3), 65-68.

Macauley, F., 1938. The Movement of Interest Rates, Bond Yields and Stock Prices in the United States Since 1856.

Malkiel, B., Xu, Y., 2006. Idiosyncratic risk and security returns. Working Paper, University of Texas.

Markowitz, H., 1952. Portfolio selection. Journal of Finance 7(1), 77–91.

Matlock G., 2010 [Internet]. [Updated 2010 Feb 16]. Peripheral euro zone government bond spreads widen. Reuters [Retrieved 2011 Aug 20]. Available at http://www.reuters.com/article/idUSLDE61F0W720100216

McDonald J., 1973. French Mutual Fund Performance: Evaluation of Internationally Diversified Portfolios. Journal of Finance. 28(5), 1161-1180.

McEnally, R., Boardman, C., 1979. Aspects of corporate bonds portfolio diversification. Journal of Financial Research 2, 27–36.

Melnikoff M., 1998. Investment Performance Analysis for Investors. Journal of Portfolio Management 25(1), 95-107.

Merton R., 1981. On Market Timing and Investment Performance I: an Equilibrium Theory of Value for Market Forecasts. Journal of Business 54(3).

Modigliani F. and Modigliani L., 1997. Risk-Adjusted Performance. Journal of Portfolio Management, 45-54.

Moody’s Corporation, 2011 [Internet]. [Updated 2011 Nov 4]. [Retrieved 2011 Nov 4]. Available at www.moodys.com.

Muralidhar A., 2001. Optimal Risk-Adjusted Portfolios with Multiple Managers. Journal of Portfolio Management 27(3), 97-104.

Muralidhar A. S., 2001. Skill, History and Risk-Adjusted Performance. Journal of Performance Measurement, Winter 2001-2002.

Odier, P., Solnik, B., Lessons for international asset allocation. Financial Analysts. Journal, March 1993, 63-77.

Pagano, M., von Thadden, E., 2004. The European Bond Markets under EMU. CSEF Centre for Studies in Economics and Finance Working Paper No. 126.

Pogue G., Solnik B., Rousselin A., 1974. International Diversification: A Study of the French Mutual Funds. MIT Working Paper.

Premaratne, G., Tay, A., 2002. How should we interpret evidence of time varying conditional skewness? Working Paper, University of Singapore.

Reilly, F., Brown, K., 2003. Investment Analysis and Portfolio Management, Seventh Edition, South-Western/Thomson Learning.

Roll R., 1977. A Critique of the Asset Pricing Theory’s Tests. Journal of Financial Economics,129-176.

Sancetta, A., Satchell, S.E., 2007. Changing correlation and equity portfolio diversification failure for linear factor models during market declines. Applied Mathematical Finance 14(3), 227–42.

Scholtz H., Wilkens M., 2004. Investor Specific Performance Measurement – A Justification of Sharpe Ratio and Treynor Ratio. Working Paper.

Scholtz H. and Wilkens M., 2005. A Jigsaw Puzzle of Basic Risk-adjusted Performance Measures. Journal of Performance Measurement.

Shadwick, W., Keating, C., 2002. A universal performance measure. Journal of Performance Measurement Spring 2002, 55 – 84.

Sharpe W. F., 1964. Capital Asset Prices: A Theory of Market Equilibrium under Conditions of Risk. Journal of Finance 19, 425-442.

Sharpe W. F., 1966. Mutual Fund Performance. Journal of Business, 119-138.

Sharpe W. F., 1994, The Sharpe Ratio. Journal of Portfolio Management.

Silvapulle, P., Granger, C., 2001. Large returns, conditional correlation and portfolio diversification: A value-at-risk approach. Quantitative Finance 1(2), 542–51.

Solnik, B., Boucelle, C., Le Fur, Y., 1996. International market correlation and volatility. Financial Analysts Journal 52(5), 17–33

Sortino F.A., Miller G., Messina J., 1997. Short Term Risk-Adjusted Performance: A Style Based Analysis. Journal of Investing, summer 1997.

Sortino F. A., van der Meer R., 1991. Downside Risk. Journal of Portfolio Management, 27-32.

Sortino F. A., van der Meer, R., Plantinga, A., 1999. The Dutch Triangle. Journal of Portfolio Management 18, summer 1999, 50-59.

Statman, M., 1987. How many stocks make a diversified portfolio. Journal of Financial and Quantitative Analysis 22(3), 353–63.

Statman, M., 2004. The diversification puzzle. Financial Analysts Journal 60(4), 44–53.

Statman, M., Scheid, J., 2005. Global diversification. Journal of Investment Management 3(2), 53–63.

Tett, G., 2011 [Internet]. [Updated 2011 Sep 1]. Get used to world without 'risk free' rate. Financial Times [Retrieved 2011 Oct 30]. Available at http://www.ft.com/intl/cms/s/0/52a9169e-d4b6-11e0-a7ac-00144feab49a.html#axzz1cLfggjlg

Thomas, L., 1989. The performance of currency-hedged foreign bonds. Financial Analysts Journal 45(3), 25–31.

Treynor, J., 1965. How to Rate Management of Investment Funds. Harvard Business Review 43(Jan/Feb), 63-75.

Treynor, J., Black F., 1973. How to Use Security Analysis to Improve Portfolio Selection. Journal of Business 46(1), 61-86.

Treynor, J., Mazuy K., 1966. Can Mutual Funds Outguess the Market? Harvard Business Review 44(Jul/Aug), 131-136.

Unicredit Bank, 2011 [Internet]. [Updated 2011 Nov 4]. Credit Risk [Retrieved 2011 Nov 4]. Available at http://www.unicreditgroup.eu/en/Investors/Credit_Risk.htm

Van Nieuwerburg, S., Veldkamp, L., 2005. Information acquisition and portfolio under-diversification. Working Paper. New York University.

Vinod, H., Morey, M. 2001. A 'Double' Sharpe Ratio. Working Paper.

Ziemba, W. T., The symmetric downside risk Sharpe ratio and the evaluation of great investors and speculators, Journal of Portfolio Management, 2005 (Fall): 108-122.