Investment Benchmarks for Alternative Asset Classes

PhD Thesis

Daniel P. Broby

Accounting and Finance Department
Strathclyde Business School
University of Strathclyde, Glasgow

March 7, 2021

Abstract

This thesis is a collection of essays that explore current theory and practice in respect of benchmarks for alternative asset classes. A benchmark is required in order to measure and attribute the performance of professionally managed investment funds. A principal component analysis (PCA) based index, based on factor weights determined by eigenvalues, is proposed to address identified weaknesses in current indices. The approach uses linear combinations of factor returns to construct alternative asset indices. These are statistically correlated with the principal components identified. The resultant indices provide an attributable benchmark, particularly for commodity futures. Collectively the essays identify and suggest enhancements to index construction methodology as applied to alternative assets. Increased investment in such assets has created a need for such new and innovative benchmarks. The essays focus on a variety of unique features present in alternative assets and the proxies used to invest in them. It was found that the lack of reporting on leverage and liquidity is the biggest impediment to index refinement in real estate and hedge fund indices. The gearing and lack of liquidity in the investment proxies make indices harder to replicate. Principal component derived factor weights can partially address this issue. The approach also proves useful to address identified problems in peer group benchmarking. It is concluded that PCA can be used to benchmark commodity futures and help in the classification of hedge fund strategies. The chapters herein have important implications for asset allocation, manager selection, index construction, portfolio risk assessment, alternative asset pricing, the testing of commodity market efficiency and the synthesis of hedge fund strategies. They point to the need for a more bespoke treatment of the benchmarking of investments in alternative assets.

Preface/Acknowledgements

"Today, almost every index is called investable, although little thought is given as to what that actually requires or how to achieve it. At the same time, investability is increasingly important to the way indices are used for investments and investment products. Liquidity and low tracking errors don't just happen, after all: They must be part of the index design."[1]

This thesis draws on my practical fund management experience in my previous roles as Chief Executive Officer, Chief Investment Officer and Chief Portfolio Manager. Combining this industry knowledge with my academic experience, I believe my findings make a contribution to bridging the practical with the academic in the field of benchmarking. They also fill a number of identified gaps in the academic literature.

Benchmarking and measuring financial returns is one of the most important topics in Quantitative Finance. The scientific management of investment funds, the measurement and attribution of returns and portfolio risk management, are all built upon benchmark choice and specification. In Broby (2010), I explained that the dynamics of asset prices can only really be understood with a clear and defined benchmark. The findings in this thesis corroborate that.

This thesis adds to my prior research in the field of benchmarks and indices, namely Broby, Bautleac, and McKenzie (2017), Broby and Lochhead (2002), and Broby et al. (2016). It was inspired by a book that I wrote in 2007, *A Guide to Equity Index Construction*, Broby (2007). It also draws on a CFA Institute reading that I co-wrote with Professor Mitchell Conover and Professor David Carino, *An Introduction to Investment Benchmarks*, Conover, Broby, and Cariño (2013), all leaders in their field.

In the writing of this thesis, I published a peer reviewed paper, *Investment benchmarks:*

[1]David Blitzer, Managing Director and Chairman of the Index Committee, Standard and Poor's

their ontological and epistemological roots, Broby (2017), which explains the philosophical background to this field and is a thought piece inspired by the research in this thesis.

I wish to thank my extremely supportive colleagues and mentors for their the assistance and guidance. In particular, my supervisors, Professor Krishna Paudyal and Professor Andrew Marshall. Thanks also go to those who have helped in the work towards publication of various aspects of the research contained herein, namely Professor Andrew Mackenzie and Olivier Bautheac, for their contribution towards Broby, McKenzie, and Bautleac (2017).

I would like to acknowledge those whom I would have liked to have spent more time with during the writing of this thesis, namely my children Leopold, Anastasia, Theodore, Elizabeth and Oskar.

Acronyms

The commonly used acronyms in this thesis are as follows:

APT: Arbitrage Pricing Theory

AUM: Assets under management

BCOM: Bloomberg Commodity Index

CAPM: Capital Asset Pricing Model

CRY: Thompson Reuters Commodity Index

DJCI: Dow Jones UBS Commodity Index

DJIA: Dow Jones Industrial Average

ETF: Exchange Traded Fund

FTSE: Financial Times Stock Exchange

GRS: Gibbson, Ross and Shanken

GSCI: Goldman Sachs Commodity Index

KMO: Kaiser Meyer Olkin Measure

MPT: Modern Portfolio Theory

MSCI: Morgan Stanley Capital Index

NPV: Net Present Value

PCA: Principal Component Analysis

REIT: Real Estate Investment Trust

S&P: Standard and Poors

TVA: Total Value Added

Contents

List of Figures

List of Tables

Chapter 1

Introduction to the essays

1.1 Introduction

This thesis comprises a collection of essays on investment benchmarks that demonstrate how the principal components of the underlying instruments can be used to create an index or help refine existing construction techniques. The chapters are ordered so as to collectively build on the existing literature and make the case that a more customized approach should be taken when benchmarking alternative investments. It provides an important foundation for theoretical work into benchmarks, defined below.

Definition 1.1.1. Benchmark. A benchmark is a standardized tool that that can be used as a point of reference - Stevenson (2011).

A benchmark can be tailored to measure and attribute the performance of professionally managed alternative investment assets. A good description of its uses can be found in Conover, Broby, and Cariño (2013). They point out that benchmarks are an important fund management and reporting tool. That said, there is no consensus on what benchmark is the most appropriate, particularly for alternative assets. The collective essays therefore close some of the gaps in the literature on benchmarks and provide an incremental contribution to both theory and practice.

An important distinction is that benchmarks can be in the form of an index that is constructed based on a pre-defined sampling and weighting methodology. An index is defined below.

Definition 1.1.2. Index. An index is a figure in a system or scale representing the average value of specified prices, shares, or other items as compared with some reference figure. Stevenson (2011)

The most significant contribution of this thesis is the refining of index construction method. A method is presented that captures those components of appropriateness that the literature suggests are not present in existing alternative asset indices. It does this through a novel principal component analysis (PCA) approach to index construction that represents an extension to existing practice. The essays support this by highlighting the gaps in existing methods that the PCA approach addresses (in each specific alternative asset class). It details where the current approach to index construction requires PCA to address appropriateness and how index design can be improved in order to achieve optimal benchmark outcomes for alternative assets.

Definition 1.1.3. PCA index. A PCA index is a principal components analysis derived index that is based on an optimally weighted linear combination of factor scores. Each asset instrument weight is derived from its factor loading. Each asset instrument contributing to the factor score depends on how strongly it relates to the factor.

The literature is extended in three ways by this thesis. (1) It is concluded that PCA derived indices can be used to benchmark commodity futures in an optimal mean variance way. (2) It is concluded that hedge funds strategies can be synthesised using PCA and be used to form strategy level indices. (3) It is further concluded that the use of PCA is limited in benchmarking real estate due to the illiquid nature of the asset.

PCA will be explained more fully in chapter four, *on the use of principal components in index construction*. It is a statistical tool for exploratory data analysis. One of the practical applications of PCA, as a multivariate approach, is its focus on the basic dimensions of variability of the underlying alternative assets. The advantage of PCA in benchmarking is that it can be used to assign index weights. This is done using outputs that specify the relative importance of the various dimensions identified by their factor loadings. These are based on the eigenvalues produced in the analysis of the individual asset instruments.

The investigatory focus of the thesis is the validity of PCA indices is alternative assets. These are increasingly being used by investors to achieve greater diversification and enhance investment returns. There is a strong theoretical basis for index construction. The use of PCA adds to this and addresses specific shortcomings related to the use of proxies and derivatives to gain exposure to alternative assets, defined below.

Definition 1.1.4. Alternative asset class. An alternative asset class is a common grouping of investment instruments that does not conform to traditional groupings. Their unconventional characteristics make them difficult to measure and evaluate.

At the same time as addressing the shortcomings of existing scholarly knowledge, the thesis provides useful insights into individual alternative asset sub-classes. These are supported by the theoretical and empirical literature. The PCA index construction approach is applied, and tested for appropriateness in commodities, hedge funds and real estate. The evidence presented in the empirical chapters suggests that a PCA index is particularly useful for commodity investments using managed futures. Its application was found to be less useful for benchmarking investment in hedge funds and real estate investment trusts. The collective contribution of the various findings is:

- in proposing an alternative way to construct indices from principal components.
- in providing a taxonomy of benchmarking.
- in delivering a well-founded critique of finance theory as applied to benchmarking.
- in illustrating that benchmarks should be market specific.
- in documenting the shortcomings of peer group indices.
- in demonstrating how established techniques and method can be applied in practice to benchmarking alternative asset classes.
- in highlighting how the nature of the time series of alternative assets can be addressed by different sampling methods.

Combined, the thesis provides a practical set of recommendations on current benchmark construction. It does this in a number of ways but foremost is the aforementioned method, an extension of a concept from principal components analysis. The potential of this approach is considered as an appropriate way, based on the literature in the field of benchmarking, to address the issues identified in the subsequent chapters.

An extended literature review and a chapter exploring the academic approaches to the construction of indices both contribute to the understanding of the issues and challenges involved. It is acknowledged that there are "simple" benchmarks that can be applied to alternative assets. The academic body of knowledge on the subject, however, suggests a preference for the use of an optimal set of assets, based on mean variance subsets of available investments. In this respect, the concept of risk and its relationship to return is key to the usefulness of an index and hence the focus of this thesis. PCA is consistent with this as the first principal component of a set of index constituents is the equivalent of the systemic market risk identified by capital asset pricing models.

As the essays demonstrate, there is a rich literature on benchmarking and sampling approaches. These approaches are general in nature but the choices are specific to alternative assets and indices that are used to benchmark them. The construction methods are therefore supportive of the literature that emanates from capital market, asset pricing and performance measurement finance research. The proposed PCA index builds on these and extends the literature in respect of what makes an appropriate benchmark choice.

In summary, the contribution of this thesis is both a theoretical and a practical one. The former, because benchmarks are required to test academic theories and the PCA index approach adds to the body of knowledge on them. The later, because benchmarking techniques are used in the professional management of alternative investment assets.

1.2 Background

An investment benchmark involves many dimensions and the literature tells us these are not all captured in existing alternative asset indices. A researcher has to deal with proxies for alternative assets and use those in deriving an index. With alternative assets such proxies are often illiquid and/or a derivative of the asset. As a result, a motivation for this thesis is that, the use of these proxies may well lead to unreliable and/or undesired investment indices. With PCA, a researcher can observe the characteristics of an alternative asset and assign coefficients, or weights, to those observed variables.

A benchmark that reflects academic precedents is a practical fund management tool. A PCA index does this by capturing the variance of the constituents. The theoretical backdrop is explained in the next chapters. The literature on benchmarks is used to further the insights that can be applied to alternative asset index construction, PCA indices, and hence alternative investments. It is assumed, throughout the thesis, that a good benchmark should exhibit a number of characteristics in order for it to be useful. It should be constructed in a disciplined and objective way, be formulated from accessible information, and be representative. As indices fulfill these criteria, Bailey (1992b) suggests that they can be used as benchmarks. His criteria are expanded on in the literature review and used as the basis for subjective evaluation throughout the thesis. He argues that the right indices should be useful to academics, aligned to capital asset pricing models, and be investable. His criteria are used to evaluate the proposed PCA derived contender indices, herein described.

In his paper on the design of better benchmarks, Martellini (2012) observed that theory and practice are not always aligned. In this respect, the problems academics identified in this thesis, as pertains to existing commercial alternative asset benchmarks, are that

- they are not representative.
- they have a sampling method designed for equities.
- they do not reflect the weighting preferences of investment managers.
- their return time series does not reflect the distribution of the returns of the assets accurately.

These deficiencies will be explained and explored further in the subsequent chapters. PCA indices can address and mitigate some of theses. The essays, meanwhile, provide a holistic view of the state of the body of knowledge in the field of benchmarking and provide suggested refinements to that knowledge.

Investigation into what makes an appropriate benchmark is relevant for two main reasons. Firstly, Kryzanowski and Rahman (2008) show that it is possible to exploit benchmark inefficiency to deliver active returns. Secondly, because as Jacobs, Müller, and Weber (2014) show, there has been increasing interest in diversifying into alternative asset classes. This is illustrated by tables 1.1, 1.2 and 1.3. These represent the growth in assets under management (AUM) of proxies for commodities, hedge funds and real estate investment trusts. Specifically; the money indexed against the *Bloomberg BAUMBCOM Commodity Index* saw AUM rise from USD 16.7bn on 31/2/2005 to USD 65.0bn on 31/12/2019; the AUM in the *HDGNAV hedge fund of fund Index* rose from USD 108bn on 29/1/1999 to USD 406bn on 28/2/2020; the market capitalization of the *Bloomberg REIT Index* rose from 106bn on 29/1/1999 to USD 406bn on 9/4/2020.

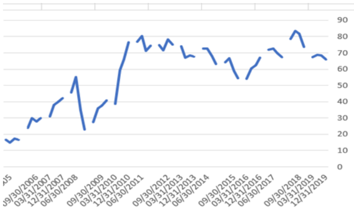


Figure 1.1: Growth in AUM in Commodity ETF's

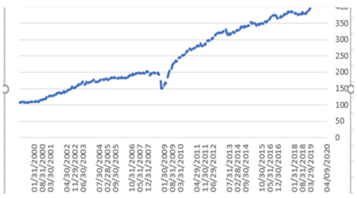


Figure 1.2: Growth in AUM in Hedge funds

Figure 1.3: Growth in AUM in REIT's

These three figures collectively show the representative growth in assets under management (AUM) in each of the alternative asset classes covered in the empirical chapters later in this thesis. Figure 1.1 represents the largest commodity Exchange Traded Fund 31/12/2005-12/31/2019. figure 1.2 represents total fund of fund assets 29/1/1999-28-2-2020. Figure 1.3 represents the market capitalization of REIT's in the US 21/9/1999-9/4/2020.

The observed increase in demand for alternative assets has been driven by the argued uncorrelated nature of their returns and the growth in derivatives as a way of gaining exposure to such investments, and a lower real rate of achieved investment return on traditional assets. This has resulted in the need to benchmark the many new and often more complex investment strategies that alternative asset classes facilitate. These include off-exchange investments and derivative based instruments, both of which present their own benchmarking issues.

There is a documented linkage between indices and pricing models in the literature. This is expanded on in the next chapter. Although there is consensus for equity benchmarks, as the subsequent chapters shall demonstrate, there has been less agreement on alternative assets. They are, by their nature, less easy to gain investment exposure to and as such indices tend to use proxies. This approach needs to be addressed in order for an index to prove useful and replicable.

The evolution of equity indices as benchmarks, upon which most alternative indices are based, has strong theoretical roots in traditional finance theory. Those designed for equity assets are typically mean-variance optimal portfolios. They derive from Markowitz (1952) and the *Capital Asset Pricing Models* that build on his work. This includes models pioneered by Sharpe (1964), Treynor (1965), Lintnert (1965) and Mossin (1966). Collectively, these provide a theoretical backbone to index construction. Academics, however, assume a number of prerequisites. One of these includes an assumption that there is a normal distribution of returns based on time series data. As shall be shown, this is not always the case for equities and often not the case for alternative assets. This warrants further investigation as it has implications for benchmarking. This is explored in the chapters herein and provides an additional motivation for this study. This is because grouping by principal components can identify common features which other techniques cannot.

Diverse alternative asset investment strategies and their benchmarks have been commented on by a number of academics. They include investment benchmark commentary into hedge funds by Atilgan, Bali, and Demirtas (2013), private equity by Cumming, Helge Hass, and Schweizer (2013a), managed futures by Szymanowska et al. (2014), real estate by Ross and Zisler (1991), and commodities by Fabozzi and Kaiser (2011). The wider research is reviewed in the second chapter. As noted by Cumming, Hassand, and Schweizer (2014) There is documentary evidence that investors are increasingly including alternative assets in multi-strategy portfolios. They attribute this to the less correlated nature of their returns to traditional assets. There is therefore a need for both appropriate aggregate strategic benchmarks and tactical ones at the sub-group level.

In an academic context the literature suggests that a benchmark should take risk into account. This should be both in the short and long run. Importantly, much of the literature fails to make a distinction between short term tactical and long term strategic benchmark construction. That said, regardless of what they are designed to measure, how they are defined and constructed is important. In an investment context Broby (2007) defines them as an independent rate of return (or hurdle rate) forming an objective test of the effective implementation of an investment strategy. This definition is useful from an academic perspective, particular when risk is added to the mix. It suggests that the construction method should be done according to transparent and repeatable rules.

In addition to providing an important methodological contribution, the research in this thesis is relevant because, as Amenc, Goltz, and Le Sourd (2009) suggest, a diversified portfolio of alternative assets can improve performance outcomes. Indeed, Ibbotson and Kaplan (2000) point out, most of the post market factor variation in investment returns are derived equally from asset allocation decisions made against a benchmark. Many asset allocations have been shown to be poorly constructed. Agnew and Szykman (2005) demonstrate that poor weighting decisions are the norm in retirement accounts. A proper benchmark would highlight, and possibly even help alleviate, such practices.

In summary, the enclosed chapters build on a rich strand of scholarly research focused on performance measurement, portfolio management, and index construction methods. The literature, expanded on in the second chapter, illustrates that benchmarks have a theoretical base in current benchmarking practice, grounded as it is in traditional asset classes. The essays investigate investment benchmarks used in the evaluation of alternative asset classes. They help identify where the state of the art is, where it cannot be improved on, and proposes areas where it can be improved. The thesis demonstrates that the examination and interpretation of the multivariate correlations of alternatives assets using PCA can result in the construction of meaningful indices that can be used for investment benchmarking.

1.3 The index as a benchmark

The index is the most common type of benchmark in daily use and has been since the advent of professional asset management. They are widely used in finance academia to test theory related to capital markets. As previously mentioned, a good index should incorporate risk and enable attribution. According to the *Index Industry Association* (IIA) globally in 2018 there were approximately 3.3 million commercial indices, the vast majority of them being traditional equity or fixed income benchmarks.

The Dow Jones Industrial Average (DJIA) was the first stock market index. It has a time series that started on May 26, 1896. Indices, however, have a longer academic pedigree. Indices were adopted from the discipline of economics, where they were first devised in order to measure price inflation. This evolution of indices is explained in chapter three, a taxonomy of benchmarking techniques.

A list of the common alternative asset indices used as benchmarks, as well as indices mentioned in this thesis, is provided in Appendix A. This includes commercial offerings from the key index providers, such as *Dow Jones, Standard and Poors, Morgan Stanley* and *Financial Times*. They all produce indices focused on alternative asset sub-classes. Amongst these, the *Standard and Poors GSCI Total Return Index* is the most widely used production based commodity index. It has an approach similar to equity market capitalization indices. Similarly, the *FTSE EPRA/NAREIT Global REIT Total Return Index* is the most widely used real estate index. It includes a more diverse set of global indices than its US based competitors. The *HFRX Global Hedge Fund Index*, meanwhile, is the most widely used hedge fund index. It is the most reflective of underlying strategies.

The chapters in this thesis provide proposed enhancements to these and other existing commercial offerings. There are also a number of indices that are used to benchmark indirect exposure to alternative assets. For example, *the Wealth Management Association's Private Investor Index Series* (WMA). This index was constructed to address the problem that direct exposure indices are extremely difficult to replicate, a problem common to many alternative assets. The PCA approach helps mitigate this issue.

In the literature on benchmarks, market capitalization indices are fairly recent innovations. They are now indispensable in the professional management of investment portfolios and are the most widely used benchmarks in practice. In the essays, these are explained along with the literature's critique that they are not appropriate in the context of alternative assets. The proposed PCA index adds to the literature. The approach provides a way in which a new generation of synthetic indices can be constructed. It is argued that this would provide an advantage over peer group fund indices, as it avoids the incorporation of the fees charged on the underlying funds.

In summary, this thesis extends and refines the existing benchmarking approaches. In this way, the essays provide a practical overview of the usage of benchmarks. The theory, which supports this, is covered in the review of the literature on investment benchmarks. It is also identified and critiqued in the subsequent chapters. A uniting theme through this thesis is that the index construction method, as it currently stands, does not address all the issues faced by investors in alternative assets. What currently exists has evolved from, and is built around, the concept of broad market proxy indices, typically justified by recourse to Modern Portfolio Theory (MPT), first proposed by Markowitz (1952). PCA does not contradict MPT. The first principal component is akin to the market factor. As such, the usage of PCA indices is incremental to existing knowledge.

1.4 Research questions

There are a number of research questions related to benchmarks specifically designed for alternative assets that arise from the literature. These relate to the appropriateness of existing commercial benchmarks. They include how an index handles investment proxies and how reflective it is of the asset class being measured. In addressing these, the thesis fills a research gap, builds on theory, and provides useful practical guidance to benchmark constructors. Funds that gain exposure to alternative assets, and hedge funds pursuing alternative strategies, have to address benchmark appropriateness in order to be managed optimally.

PCA is used to answer some outstanding questions that arise from the literature. It is a multivariate statistical methodology used in many scientific investigations. It is a dimension reducing technique that can extract the principal components of the intercorrelation of the assets instruments that can be used to constitute an index. The method is suitable for alternative asserts as the technique is most useful when high correlation is observed among the variables. Moreover, as PCA does not impose requirements for normality and homoscedasticity it can be used where returns are distorted by the use of proxies or in instances of illiquidity. Importantly, PCA can be used to develop a multivariate index, taking into account asset sub-groupings.

The areas that PCA indices address, those that relate to appropriateness, include the aforementioned use of derivatives, the presence of leverage and the lack of liquidity. These were identified by Amihud (2002). These all impact the distribution of the returns of asset class time series and similarly the cross section of returns. This in turn impacts the appropriateness of the commonly used benchmark proxies. There are, as a result, two clear research questions.

Research question 1. Can PCA be used to make an appropriate benchmark for alternative assets?

Research question 2. Does a PCA derived index for alternative assets provide useful insights and reflect the uniqueness of the distribution of the returns? [1]

The primary research question is unaddressed by the literature and is answered in the subsequent chapters by identifying differing variances and co-variance in the time series and the distribution of the specific sub-group returns. These are influenced by variation in underlying constituents and the groupings of their common principal components.

The secondary research question is answered at the asset class level and therefore aids scholarly investigation into asset pricing. A PCA benchmark consists of sub-sectors or individual alternative assets and as such facilitates attribution. Grouping such investments into a benchmark creates the homogeneity among the asset instruments. That said, it needs to be tested for its mean variance properties, the method for which is addressed next.

1.4.1 Establishing index superiority

In order to extend the literature, it is necessary to link the evaluated benchmarks to finance theory and the foundations of MPT. This is done though the first principal component representing systematic market risk. That said, there still needs to be a test for the optimal mean variance outcome.

According to academic precedent, the way to show that PCA is a superior index method are (1) to understand the existing competing index method and (2) show an improvement in risk adjusted performance and/or subjective criteria from identified sampling issues. This requires a common investigation in the empirical chapters.

In order to make a PCA contender index and test its appropriateness, the literature suggests an index should be mean variance optimal and deliver superior risk adjusted performance to a contender index.

There are a number of ways one can test portfolio superiority. That said, the GRS Test proposed by Gibbons, Ross, and Shanken (1989) is the most appropriate. Its advantage is that it uses multivariate regression analysis.[2] It is commonly used to test the equality of the Sharpe

[1] Impacted, for example, by portfolio leverage and/or any lack of liquidity in the underlying assets. These characteristics are derived from the way such assets are traded.

[2] To show that the test statistic formula will be distributed in finite samples.

ratios of two investment strategies. The Sharpe ratio is explained more fully in chapter two. The GRS Test can, by extension, be used to test the equality of the Sharpe ratio of two indices. The Gibbons, Ross and Shanken test statistic is stated thus:

$$GRST_t = \sum_{N}^{j=i} w_{\tilde{j,t}-1} - w_j, -13^{Rt^{b,jt-13}} \tag{1.1}$$

Where:

- $GRST_t$ = Gibbons, Ross and Shanken Test

Essentially, the ratio of the two indices is compared using their mean excess returns over a risk-free rate and the standard deviations of excess returns. The null hypothesis, shown below, is that a contender index should not be able to outperform an existing index on a risk adjusted basis, assuming the existing index is mean variance efficient [3]. The null hypothesis is rejected if the hypothetical value of zero is found to be outside the confidence range. This result would suggest that the contender index is superior to the existing index.

Null Hypothesis 1.

$$H_0 : GRST_t = \eta_i - \eta_j = 0$$

Where:

- $\eta_i = \mu_i/\sigma_i$ = Mean excess returns i (over a risk-free rate)/standard deviations of excess returns i
- $\eta_j = \mu_j/\sigma_j$ = Mean excess returns j (over a risk-free rate)/standard deviations of excess returns j
- $GRST_t$ = is made up of an estimator of the Sharpe ratio difference, the standard error of the difference estimator, the Sharpe ratios and the correlation between the excess returns of strategy i and j.

[3] The Matlab code used to calculate the test p-value in the empirical chapters is shown in Appendix B.

The conjecture is tested based on the null hypothesis using a one tail test (since a fund manager is only interested in out-performance) and can be assessed with the T-statistic. The formula shows the area outside of the confidence zone. Rejecting the null hypothesis with a weak p-value implies one of the indices out-performs the other on a risk-adjusted basis. The goal is to identify if it is possible to consistently derive better risk and return characteristics by using a PCA construction approach to an index.

1.4.2 Subjective evaluation

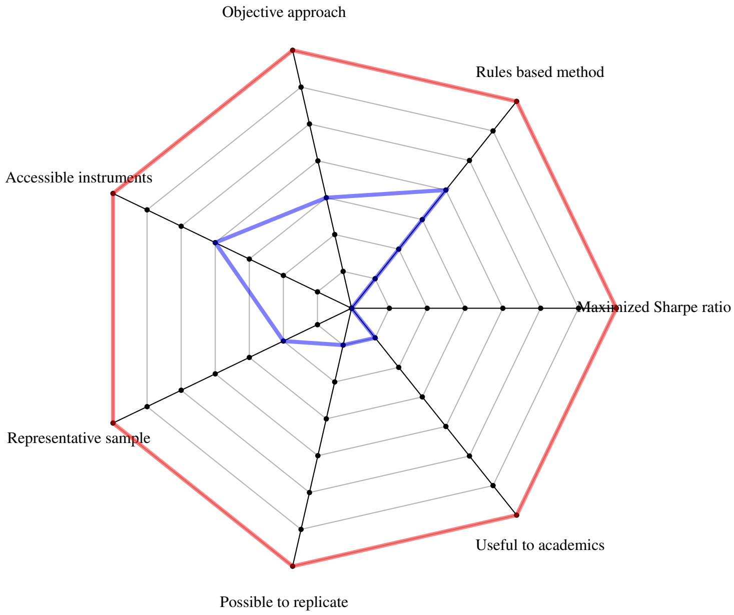
In order to bridge the gap in the literature on index appropriateness a number of subjective criteria are used to access the research questions. These are based on the observations from the literature that (a) using market capitalization is not appropriate for alternative investments, that (b) the sub-groupings of alternative asset classes are poorly reflected, that (c) the index has to be reflective of desired investment outcomes and (d) the index should incorporate risk as well as return factors.

The link between results and research questions will be demonstrated throughout the thesis using a variation of the spiderweb diagram in figure 1.4. This is done in order to makes it easier to visualise the subjective nuances in method and trade-offs. A Likert score is applied to a number of variables. This is an established psychometric tool in social sciences research. The seven dimensions are based on the finance literature on benchmarks, as introduced in chapter two. They include (1) Rules based method, (2) Objective approach, (3) Accessible instruments, (4) Representative sample, (5) Possible to replicate, (6) Useful to academics, and (7) Maximized Sharpe ratio. The nodes are determined by the relative importance of each of the specific dimensions.

Figure 1.4 reflects what a flawless *Optimal Index Spiderweb* would look like using these criteria. As the dimensions are individual index characteristics, the Likert numbers are ordinal variables. They cannot be subject to hypothesis testing as they were not collected from a sufficiently large sample. As such, they should be viewed as illustrative only.

In order to create each spiderweb a Likert scale is used. This assigns each dimension a point value, from 1 to 7. This creates subjective values for the index construction trade-offs as *extremely unimportant* at 1 point, *slightly unimportant* at 2 point, *unimportant* at 3 point

Figure 1.4: Optimal Index Spiderweb versus naive portfolio.



Optimal Index Spiderweb (the optimal dimensions are on the outer boundary) versus a naive portfolio (whose dimensions are shown on the inner plots). Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. The optimal index has 7 points on the Likert scale for each of the dimensions. The naive portfolio scores very poorly, with 4 points each for dimensions 1 and 3, 3 points for dimension 2, and one point each for dimensions 5 and 6.

neutral at 4 point, *slightly important* at 5 point, *important* at 6 point, and *extremely important* at 7 point. Anchoring the observations to Likert scale was conducted on an intuitive basis. There is much debate on the reliability of this method but it is generally accepted by social scientists that it is internally consistent for visual representation when the ranking is done by the same individual.

All the dimensions in figure 1.4 are considered extremely important. The diagram depicts a second plot which represents a naive portfolio. This illustrates that a naive portfolio is not a suitable benchmark by the criteria identified in the literature.

Likert assigned dimensions can be critiqued for being subjective in assigning numbers to each variable. The use of the output is, however, directed to meaningfully answering the research questions in this thesis. It is also used to illustrate the points rather than be used in quantitative evaluation of the scores.

1.5 Defining alternative assets

The definition used in this thesis, as outlined earlier, focuses on the uniqueness of the assets. Kritzman (1999) was one of the first to document alternative assets as an investment class in their own right. A clear definition of alternative assets is important because strategic allocation investment objectives are facilitated by asset class, and achieved through the time series of returns generated by them.

An alternative asset, however, is often defined by default. In that respect, in practical fund management it is simply an investment asset that is non-traditional. The definition of alternative assets is therefore a negative one, by exclusion. Assets that are not considered traditional are labeled as alternative. This presents a problem for those charged with identification and classification of such assets and by extension the benchmarking of them. It does, however, provide a justification for the PCA index approach. This is supported by the fact that PCA starts from the premise that there is no unique variance. The total variance is equal to the common variance, The first component will always have the highest total variance and in this respect is analogous to market risk.

PCA can be used to group assets. This in turn reflects their performance and can therefore be used as a benchmark. A good description of how asset classes determine investment performance over time is given by Brinson, Hood, and Beebower (1995). They show that the asset class decision dominates investment strategy. They conclude that market timing and security selection explain 94 percent of the variation of investment returns. As a result, it is not surprising that any asset class, be it traditional or alternative, requires a benchmark. It could even be argued that crypto-assets are an alternative asset sub-grouping.

In order to have a benchmark, the definition needs to encompass a universe of instruments that an investor can invest in. Kritzman (1999), amongst others, shows that is possible to identify, and hence define, four key differences between traditional and alternative assets. These are:

1. the ability to trade in alternative asset instruments. This is termed their liquidity and is often lower than traditional asset classes.

2. the leverage present in alternative assets. This is often a function of the instrument used

to gain the exposure and is generally larger than in traditional asset classes.

3. the venue of the trading of the assets. Traditional assets are traded in public markets. Alternative asset classes are, however, traded in different venues, often directly between buyers and sellers, or in the derivative markets. In this respect, they often utilize futures, options, over the counter products, or are instigated by private transactions.

4. the total return components of the underlying assets. Alternative assets do not generate income in the form of interest payments or dividend payments and as such their total return attribution is different from traditional assets. Real estate investments can generate rental income, but there is no yield on commodities.

The four criteria identified determine how a selected index is calculated and are taken into account in the Likert score for the dimensions in the *Index Spiderweb* diagrams produced for each type of index in chapter two. They also fit well with the concept of a PCA derived index. Amenc, Goltz, and Le Sourd (2009) addressed the methodological choices that this entailed. This is further explored in chapter three on benchmarking techniques. This in turn illustrates the need for separate academic investigation because it can be observed that alternative asset classes are distinct from equity, fixed income and cash equivalents. Such traditional asset classes are priced based on the discounted value of their underlying cash flows.

Alternative assets, however, are priced on supply and demand dynamics. This difference, it could be argued, has benchmarking implications. These are explored in the chapter two review of the literature. All the differences identified in the literature require specific treatment as far as benchmark selection goes. Whether distinct individual alternative sub-groups are an asset class has also attracted academic interest. Su, Lau, and Chau (2013), for example, touched on this from the perspective of commodity investment. They concluded that the robust inter-temporal hedging demand that they identified suggested a role for commodities as an asset class in strategic asset allocation. Jacobs, Muller, and Weber (2014), meanwhile, investigated why diversification into alternative asset classes is relevant.

As explained, there is a distinction between benchmarks and indices. The use of PCA indices as benchmarks, in this respect, is justified because individual alternative asset classes are comprised of similar instruments and/or types of investment. Indices can better reflect this.

These instruments, according to Agarwal and Naik (2015), are typically homogeneous as far as legal structure, trading environment and regulation are concerned. For example, investments in commodities futures are made using derivatives, instruments clearly structured in a different way from physical real estate investment.

It is suggested in this thesis that PCA indices help in addressing the uniqueness of the alternative assets return generating process. This should be taken into account in benchmarking such assets. The non-normal distribution of returns was a topic first addressed at the index level by Elton and Gruber (1992). Traditional finance theory assumes that a linear relationship exists between risk and return, and that returns are normally distributed. However, this assumption does not hold for many alternative asset class time series, which exhibit skewed distributions. By definition, principal components are jointly normally distributed. The Gaussian distribution assumption aligns PCA with traditional finance theory.

The identification of uncorrelated alternative assets is useful because there are clear diversification benefits from combining alternative and traditional asset classes. These are documented by Jacobs, Müller, and Weber (2014), among others. Diversification is important because failure to do so adequately will impact performance. Asset allocation is the main determinant of portfolio performance according to Brinson, Hood, and Beebower (1995). As such, the identification and classification of such assets detailed in this chapter has practical applications for the fund management industry.

Fund managers have to be able to benchmark the investment performance generated from an asset allocation decision to invest effectively. The identification of such assets can be made by applying a proactive quantitative approach. PCA is one way to do this, as it helps identify the sub-groupings. As explained, it can also form the basis of an index. This method is proposed as a compliment to other index construction methods based on sampling and top level classification. One justification of its use is that alternative assets are not homogeneous at the sub-grouping level. In commodities, for example, crude oil is very different from frozen orange juice. In hedge funds, a market neutral equity fund is very different from a directional fixed income fund. In real estate, a tower block is very different from a factory.

PCA asset classification is the process used for assigning assets to the various sub-groupings. This is done based on a number of common characteristics. Identification outcomes are based on the quantitative and qualitative rules that are used to define the asset and sub-asset group. These provide important cornerstones of what has become a large industry generating hundreds of thousands of investment indices. In addition to being of scholarly interest, PCA derived refinements in benchmarking techniques for alternative assets may result in improvement in the management of investment funds and thereby hopefully the performance of fund managers in the long run.

1.5.1 Identification of an appropriate universe set

In order to construct a PCA index, it is necessary to identify an appropriate universe set from which to sample. The literature introduced in the next chapter helps with this. Absolute benchmarks can be ruled out. The literature finds they use heuristics as performance indicators and therefore do not give a detailed breakdown at an instrument level. Peer groups can also be ruled out. The literature finds they do not incorporate risk, suffer from survivorship bias, and also from performance drag due to the fees applied.

The current sampling method to construct alternative assert indices has been adopted by default, an extension of that used in traditional asset classes. This method takes the time series back to first principals by transforming the return observations in order to identify patterns that can then be used to create optimal weights for an index. This is explained in more detail in chapter four.

The identification of an appropriate universe set can be done using PCA. Creating a PCA candidate index requires a goal-oriented approach. In this respect, a benchmark is central to getting exposure to the asset class and must therefore be appropriate. To do this, the design must be clear, precise and replicable. The resulting index should reflect the type of return distribution, variance, skew and kurtosis. With alternative assets, this means selecting the index universe from the respective cash market, the futures market or the peer group of collective investment vehicles. The index will then reflect investment objectives and improve diversification, thereby helping meet investment targets. Essentially, the goal of PCA is to divide the time series data into meaningful sub-groups. These sub-groups are then used to construct and replicate liquid and useful indices.

Once a universe of instruments is identified, the PCA method can be used to further identify sub-groups in classes of commodities, hedge funds and real estate. This helps with identification of an appropriate investment benchmark for alternative investments by investigating the co-movements of the asset class. The method to do this is to sample the available investment universe and suggest a weighting for the representative assets to be included in an index.

In summary, for traditional assets, index methods are well established and the literature supporting them are well defined (as is demonstrated in the next chapter). A number of issues with existing benchmarks exit but the identification of the sample universe allows for more representative outcomes.

1.5.2 Regression of excess returns

It is instructive to the construction of a PCA index to understand the nature of alternative assets. In this respect, alternative asset classes can be identified by doing a regression of the excess returns of a set on investment returns, taking into account aforementioned differences. The difference between a regression and PCA is the addition of a multiplication transformation and an inverse matrix. An asset class should exist in this regard if the constant term of such a regression deviates significantly from zero. This is the case, for example, in the time series of commodity futures fund, hedge funds, and real estate fund returns. It is not said to exist in the class of investments identified by Statman (2006), namely Socially Responsible Investments. The difference being that the later is the result of negative screening, whereby properties are screened out rather than identification of common properties.

Defining alternative asset classes through regression essentially focuses on the correlation of their returns. This has interesting implications. For example, Hafner and Wallmeier (2007) show that institutional investors could consider volatility as an asset class. This, they argue, is seen as attractive as volatility movements can be negatively correlated with stock index returns. Their conclusion can be critiqued, however, as volatility is time varying and driven by uncertainty.

Table 1.1 reports asset class ten year correlation matrix between traditional assets and alternative assets over the period 01/01/2000 and 01/01/2020. The results indicate that there is no significant correlation between the daily returns and the returns on other asset classes. Both fixed income and cash are negatively correlated over this period. There is a positive correlation to equities, possibly an indication that the market factor is important.

Table 1.2 reports the estimates of a regression where the dependent variable is the arithmetic daily return on the broad alternative asset class index from 01/01/2000 to 01/01.2020 and the independent variables are the corresponding daily returns on traditional asset classes. The lagged returns of both the response and the explanatory variables, as well as the lagged values of the average traded volume, are included as additional control variables. Applying the regression technique to alternative assets, the relationship between three alternative asset classes and the *Standard and Poor's 500* (as a proxy for traditional equity) was investigated. These regressions constitute initial analysis for the later chapters. The dates selected were 1/1/2000 though to 1/1/2018, capturing data points from both sides of the credit crisis.

The regression presented in table 1.2 has the *Bloomberg Commodity Index* as the dependent variable resulted in an R squared of 0.1, with a t-test of 10.764. The results demonstrate that commodities and equities are distinct asset classes. This can be used to pre-select the instrument set.

A second regression is presented on hedge funds in table 1.3. The *FRS Global Hedge Fund Return Index* was chosen as the dependent variable, resulted in an R squared of 0.4, with a T test of 24.641. This was not as strong a result as the commodity index but that is not surprising as hedge funds have exogenous underlying exposure to equities. The adjusted beta was 0.449. Once again, this is sufficient for hedge funds to be considered as distinct from equities as an asset class.

A third regression is presented on real estate in table 1.4. It shows the Real Estate Investment Trust regression. It shows the *DJ REIT Total Return Index* as the dependent variable. It had a much closer adjusted beta to equities of 0.926. This is not surprising either as REIT's are listed on the stock exchange. As a result, the R squared is a better measure and in this instance is 0.417 with a T test of 27.283.

1.5.3 Risk factors

An alternative asset class can also be identified quantitatively by risk factors. In this respect an asset class is a statistical interpretation of an instrument that earns a risk premium. Sharpe (1992) suggests a good way to do this is proportion of variance explained by the selected asset

Table 1.1: Asset class 10 yr Correlation matrix

	Equities	Fixed Income	Cash	Commodities	Hedge Funds	Real Estate	Private Equity
Equities	1						
Fixed Income	-0.218347	1					
Cash Proxy	-0.149930	0.285416	1				
Commodities	0.622228	-0.127008	0.001497	1			
Hedge Funds	0.638933	-0.150505	-0.057851	0.518732	1		
Real Estate	0.797844	-0.087296	-0.124367	0.422492	0.339589	1	
Private Equity	0.889389	-0.185871	-0.125064	0.588018	0.512362	0.782829	1

This table is a correlation matrix of Traditional and Alternative Asset Classes. It shows the correlations between equities, cash, fixed income, real estate, commodities and private equity. Fixed income and cash are negatively correlated to equities. The alternative asset classes are negatively correlated to fixed income. There is some correlation, however, between alternative asset classes and equities. The highest correlation is between equities and private equity. The high correlation between real estate and equities may be down to the use of Listed Real Estate Investment Trusts as proxies.

Table 1.2: Bloomberg Commodity Index versus Standard and Poors 500, 01/01/2000 - 01/01/2020, Regression Analysis

Raw BETA	0.29	BETA+(X>0)	0.194
Adjusted BETA	0.527	BETA-(X<0)	0.372
ALPHA (Intercept)	-0.017	Avg Slope	0.283
R^2 (Correlation^2)	0.1	Convexity	-0.089
R (Correlation)	0.317	ALPHA (Intercept)	0.133
Std Dev of Error	2.071	R^2 (Correlation^2)	0.105
Std Error of ALPHA	0.064	R (Correlation)	0.324
Std Error of BETA	0.027	Std Dev of Error	2.063
t-Test	10.764	Std Error of ALPHA	0.091
Significance	0	Std Error of BETA+	0.049
Last T-Value	0.501	Std Error of BETA-	0.044
Last P-Value	0.692	Number of Points	1042
Number of Points	1042	Last Spread	3158.67
Last Spread	3158.67	Last Ratio	0.025
Last Ratio	0.025		

This table shows that the Bloomberg Commodity Index has a significantly lower beta than the Standard and Poors with an adjusted beta of.0527. It is also poorly correlated to it with an R squared of only 0.1. This suggests that commodities are a separate asset class and distinct from equities.

Table 1.3: HFRS Global Hedge Fund Return Index versus Standard and Poors 500, 01/01/2000 - 01/01/2020, Regression Analysis

Raw BETA	0.176	BETA+(X>0)	0.091
Adjusted BETA	0.451	BETA-(X<0)	0.251
ALPHA (Intercept)	0.02	Avg Slope	0.171
R^2 (Correlation^2)	0.4	Convexity	-0.08
R (Correlation)	0.633	ALPHA (Intercept)	0.147
Std Dev of Error	0.488	R^2 (Correlation^2)	0.442
Std Error of ALPHA	0.016	R (Correlation)	0.665
Std Error of BETA	0.007	Std Dev of Error	0.471
t-Test	24.641	Std Error of ALPHA	0.022
Significance	0	Std Error of BETA+	0.012
Last T-Value	0.441	Std Error of BETA-	0.011
Last P-Value	0.67	Number of Points	912
Number of Points	912	Last Spread	1946.56
Last Spread	1946.56	Last Ratio	0.399
Last Ratio	0.399		

This table shows that the HFRS Global Hedge Fund Return Index has a lower beta than the Standard and Poors and is poorly correlated to it. This suggests that hedge funds are a separate asset class, distinct from equities.

Table 1.4: DJ REIT Total Return Index versus Standard and Poors 500, 01/01/2000 - 01/01/2020, Regression Analysis

Raw BETA	0.86	BETA+(X>0)	0.884
Adjusted BETA	0.907	BETA-(X<0)	0.84
ALPHA (Intercept)	0.076	Avg Slope	0.862
R^2 (Correlation^2)	0.417	Convexity	0.022
R (Correlation)	0.646	ALPHA (Intercept)	0.039
Std Dev of Error	2.424	R^2 (Correlation^2)	0.417
Std Error of ALPHA	0.075	R (Correlation)	0.646
Std Error of BETA	0.032	Std Dev of Error	2.422
t-Test	27.283	Std Error of ALPHA	0.106
Significance	0	Std Error of BETA+	0.057
Last T-Value	-0.086	Std Error of BETA-	0.052
Last P-Value	0.466	Number of Points	1042
Number of Points	1042	Last Spread	2836.55
Last Spread	2836.55	Last Ratio	0.125
Last Ratio	0.125		

This table shows that the DJ REIT Total Return Index has only a slightly lower beta than the Standard and Poors and is poorly correlated to it. This suggests that REIT'S can be used as a proxy for real estate as a separate asset class due to the lack of correlation, however the systematic element from the beta may add a market dimension to returns.

classes, depicted thus:

$$R^2 = 1 - \frac{Var(\tilde{e_i}))}{Var(\tilde{R_i})} \tag{1.2}$$

Where:

- R^2 proportion of the variance of Ri "explained" by the n asset class.

The incorporation of risk is why the GRS test is appropriate. By way of background, traditional asset classes, and indeed alternative asset classes, are said to have returns that are independent of one another. This property means that one can gain useful benchmark insight by describing them in terms of their probability distribution. After all, mean rates of return and standard deviation are important considerations in asset allocation and subsequent benchmark assignment. This is why the difference between normal and non-normal return time series is important. The rate of return of a portfolio of assets, which could also be a benchmark, is expressed as:

$$R_t = \sum x_{jt}, r_{jt} \tag{1.3}$$

where:

- x_{jt} = weight of assets j during holding period t
- r_{jt}= return on the asset j during period t

An asset class stated in this way is similar to the way it is expressed in *Arbitrage Pricing Theory* of Ross (1976). This is relevant as the PCA index approach is factor based. The systematic factors can be interpreted based on the systemic factors K thus:

$$t_j k = \sum_k (\lambda_j k_F)_k t(\epsilon_j t) \tag{1.4}$$

Where:

- $t_j k$ = Asset Class
- $\lambda_j k_F$ = Factor loading

- $\epsilon_j t$ = Idiosyncratic return

In a financial market context, alternative asset can be observed to exhibit unique risk factors. Such risk factors are clearly different from subjective factors, such as observing that art assets are different from commodity assets. In this respect, Idzorek and Kowara (2013) investigate whether risk factor based asset classes are superior to traditional asset class determination. They tried to identify if asset class returns are determined by the risk factors. They were not able to prove a conclusive answer with the historical data alternating between risk factors.

PCA has also been used in economics. It is observed that economic risk factors impact the time series of alternative assets. Different asset classes act in unique ways to economic news. Traditional assets are assumed to be impacted by systematic influences. The co-movement of alternative assets, however, suggests this is not the case. Real estate returns, for example, are linked to interest rate risk as well as long run demographics and GDP per capita. Such systemic factors would change the discount factor and the expected cash flows. As the cash flows are different for each of the alternative asset classes under investigation, the reaction to such economic news would be specific to that asset class.

Table 1.5 reports monthly data on the *Standard and Poors US REIT USD* (a real estate benchmark), the *Standard and Poors GSCI Commodity* (a commodity benchmark) and the *HFRX NA Index* (a hedge fund benchmark) taken from 31/12/12990 – 1/6/2018. These are compared with a number of time series consistent with those identified by Chen, Roll, and Ross (1986), namely the percent return on the US nominal Dollar major currency Index, the percent return on US Industrial Production, the percent return on the first difference of the *US CPI - All urban: All items SADJ*, the percent return on the US chain price index for personal consumption SADJ, and the percent return on the first difference *US PPI* and *crude petroleum NADJ Index*.

1.5.4 The impact of liquidity and leverage

It is relevant to the understanding of alternative assets to have some background on the impact of liquidity and leverage. As mentioned, the literature suggests the use of proxies complicates benchmark appropriateness. In this respect, the impact of liquidity and leverage on alternative asset returns is an important thing to understand. Liquidity and leverage shape the distribution

Table 1.5: Real Estate, Commodities and Hedge Funds correlation to Economic time series

Percent Return	S+P US REIT USD Percent Return	S+P GSCI Commodity Percent Return	HFRX NA	
US CPI - ALL ITEMS SADJ	First difference	0.19919	0.69267	0.42170
PRICE INDEX -PERSONAL CONS SADJ	Percent Change	0.19401	0.66452	0.45276
US PPI: CRUDE PETROLEUM NADJ	First difference	0.14149	0.58653	0.31356
	Average	0.17823	0.64791	0.39601

This table shows the correlation between alternative asset classes and economic time series. The decision to invest in alternative assets is sometimes made on the basis that they provide an inflation hedge. The first series is therefore the Consumer Price Index. The second series is the Personal Consumption. The third is crude oil.

of investment returns of alternative assets. These are both cross-sectional and longitudinally distinct. PCA derived indices partially takes this into account by grouping instruments impacted by these characteristics.

The definition of liquidity is that amount of an asset and/or instrument that can be bought or sold without affecting the publicly listed price. An actual lack of liquidity is one of the defining characteristics in certain of the alternative asset classes, and as such presents challenges for investment benchmarks. For example, if there is no change in a price of a constituent, because no trading has occurred, then the true nature of the value is not reflected in the index. PCA captures liquidity through common latent factors resulting in grouped variations in the returns.

The liquidity of alternative asset benchmark returns was investigated by Amihud and Mendelson (2015). In addition to the impact of liquidity on the nature of the returns, their research identified two priced liquidity risks that are relevant, namely shocks to the market and the overall market liquidity, termed the *market illiquidity return premium*. The former is not something an index aims to capture.

Closely allied to lack of liquidity is the *pricing* of that liquidity. Cao and Teiletche (2007) support this view arguing that it is in fact the key defining characteristic of alternative assets. Where there is no price discovery, it is difficult to benchmark an asset. The literature suggests this is particularly evident in real estate. This provides motivation for chapter seven. The lack of continuous data is an issue as alternative assets, like all asset classes, experience regime change over time. Commodities have price spikes, real estate has property bubbles and busts, and hedge funds have clear market related risk events. All these impact liquidity and all these require measurement, hence are relevant for benchmark enquiry.

The reason liquidity is relevant to index construction is that alternative assets trade less frequently than traditional assets. The lack of trading data, in turn, results in an inherent smoothing process in their investment returns. This represents a challenge for alternative asset benchmarks, as they can fail to reflect this. It can lead to an underestimation of the variance of returns and their correlation with other assets. It can also lead to misrepresentation of risk, as proxy-ed by variance. PCA is agnostic as to pricing, focusing instead on variance by scalar projection.

The other major variable is leverage. Leverage is the use of borrowed capital to generate returns on risk capital. It is a defining characteristic of many alternative assets, be it at the instrument level or capital structure level. In collective investments funds, such as hedge funds, leverage is used as part of an investment strategy using either financial instruments or borrowed capital. Leverage is present in many of the alternative assets. It increases the beta of the returns of alternative assets relative to the Securities Market Line derived from traditional assets. The collective essays explore the challenges that this presents in representing the nature of these returns from a benchmark context. The research questions that arise are discussed next.

1.5.5 Skewness and kurtosis in alternative assets

An additional element of understanding the way alternative assets act as a group or sub-group is to understand their return distributions. In this respect, alternative assets exhibit negative skew. PCA can still work when the skew is high. In practice, skewness and non-linearity can be reconciled. The transformation captures these if they are common. In a primer on alternative risk premia, Hamdan et al. (2016) show that alternative asset class returns are skewed. This is particularly the case in hedge fund returns as demonstrated by Eling et al. (2010). This is significant because there is a the need for a contender index to be representative. If this is not the case, benchmarks constructed using them will not reflect the time series generated by the assets.

The initial analysis of the time series is reported in tables 1.6 and 1.7. The ten year returns (01/06/2008 - 01/06/2018) in Table 1.6 , five year returns (01/06/2008 - 01/06/2013) in Table 1.7. The results demonstrate that there is a difference between traditional and alternative assets in respect of their skewness and kurtosis. [4]

[4]Skewness is a statistical term for symmetry and/or lack of distribution. A symmetric dataset has a slope of zero. The skewness essentially measures the relative size of the two tails which deviate therefrom, and is expressed as follows:

$$\gamma_1 = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^3\right] = \frac{\mu_3}{\sigma^3} = \frac{\mathrm{E}\left[(X-\mu)^3\right]}{\left(\mathrm{E}\left[(X-\mu)^2\right]\right)^{3/2}} = \frac{\kappa_3}{\kappa_2^{3/2}} \tag{1.5}$$

Where:

- γ_1 is the skew
- μ is the mean,
- σ is the standard deviation,
- E is the expected value expectation operator.

This is relevant because if a bespoke benchmark does not also exhibit similar return characteristics there will be return deviations from it. The return distributions of the alternative asset classes under investigation in chapters five, six and seven are presented in tables 1.6, and 1.7. The kurtosis of the ten year returns, from 01/06/2008 - 01/06/2018, is positive for each of the asset classes and especially noticeable in respect of both real estate and private equity. The skew is negative in each instance. Consequently, it is deduced that existing alternative asset benchmarks have a number of shortcomings that this thesis shall identify and address.

Kraus and Litzenberger (1976) were among the first to investigate skewness preference and the valuation of risky assets. They argued that such metrics have to be taken into account in order to apply the Capital Asset Pricing Model to performance measurement. Sometime after this, Harvey and Siddique (2000) developed the model to take account of skewness.

As far as specific alternative asset classes goes, Fernandez-Perez et al. (2015) investigated skewness from the perspective of commodity returns. They concluded it contributed to an eight per cent excess return in such assets. This is related to the nature of the futures contracts used to gain exposure to the asset class. In itself, there is a significant deviation from what would be expected and provides motivation for further investigation.

Alternative assets also exhibit different levels of kurtosis, the peak in the distribution of their returns. Kurtosis is shown with numbers in excess of zero and with alternative assets. What are termed leptokurtic distributions result from the use of leverage and derivatives. [5]

Campbell and Cochrane (2000) point out that if asset returns have systematic skewness, then expected returns should reflect the risk. The logical extension of this is that the benchmark should also reflect it. They did, however, further point out that a trade-offs had to be made between mean and skewness. This is depicted in Figure 1.5.

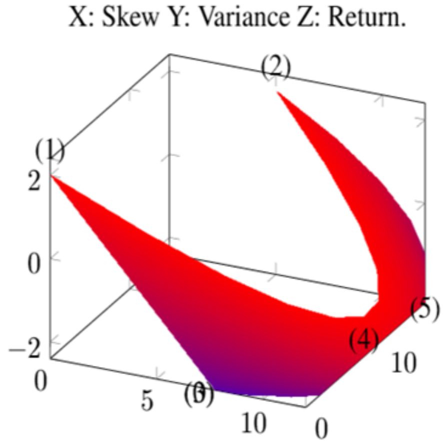
[5] The formula for leptokurtic distributions shown thus:

$$\mathrm{Kurt}[X] = \mathrm{E}\left[\left(\frac{X-\mu}{\sigma}\right)^4\right] = \frac{\mu_4}{\sigma^4} = \frac{\mathrm{E}[(X-\mu)^4]}{(\mathrm{E}[(X-\mu)^2])^2}, \tag{1.6}$$

Where:

- $\mathrm{Kurt}[X]$ is the kurtosis
- μ = the central moment
- σ = the standard deviation.
- E is the expected value expectation operator.

Figure 1.5: Diagram depicting how risk return index trade-offs are impacted by skewness



The negative trade-offs between mean and skewness. The X axis depicts the skew, whilst "Variance" is plotted on the Y axis and "Return" on the Z axis. The shape is generated through the positive trade-offs between variance and the mean. It has been shown that where conditional skewness is systematic, the returns should compensate for the additional risk. This would be the case in alternative assets which also exhibit skewness. Research shows that conditional skewness partly explains the cross-sectional variation of expected returns across asset classes and is significant even with controlled for size and book-to-market factors. The illustrative figure is based on data from Harvey and Siddique (2000), Panel A, page 1271.

Table 1.6: A comparison of different asset classes 01/06/2008 - 01/06/2018

Traditional and Alternative Asset class 10 Yr returns						
Asset	Cide	10Y Return	10Y SD	Skewewnes	Kurtosis	Characteristics
Equities (*)	S+P500	1.04401	0.05108	-0.74061	2.07418	Long only
Fixed Income (*)	US TREASURY CONST MAT 10 YR	0.25776	0.00057	0.48777	-0.66526	Long only
Cash proxy(*)	US TREASURY CONST MAT 1 MONTH	0.02873	0.00039	2.14879	3.64631	Long only
Commodities (**)	S+P GSCI	-0.84245	0.06940	-0.78845	2.15618	Futures
Hedge Funds (**)	HHFRX NORTH AMERICA	0.10810	0.01638	-0.57351	0.11270	Multi Strategy
Real Estate (**)	S+P500 REAL ESTATE	1.07273	0.08822	-0.37867	8.01564	Proxy
Private Equity (**)	SPLPEQTY	0.97676	0.08609	-0.20463	5.74851	IRR Peer Group

Annualized: 10 Yr returns						
Asset	Cide	10Y Return	10Y SD	Skewewnes	Kurtosis	Characteristics
Equities (*)	S+P500	0.07347	0.00495	-0.12526	0.11781	Long only
Fixed Income (*)	US TREASURY CONST MAT 10 YR	0.02300	0.00006	0.04019	-0.10285	Long only
Cash proxy(*)	US TREASURY CONST MAT 1 MONTH	0.00281	0.00004	0.12048	0.16455	Long only
Commodities (**)	S+P GSCI	-0.16746	0.00668	-0.14277	0.12074	Futures
Hedge Funds (**)	HHFRX NORTH AMERICA	0.01023	0.00161	-0.08104	0.01065	Multi Strategy
Real Estate (**)	S+P500 REAL ESTATE	0.07496	0.00842	-0.04610	0.24369	Proxy
Private Equity (**)	SPLPEQTY	0.06992	0.00822	-0.02245	0.20847	IRR Peer Group

*This table shows Traditional and Alternative Asset Class Returns over a 10Yr period and Annualized returns for the same assets. Traditional Assets are marked with an * and Alternative Assets are marked with ** Note the marked kurtosis on the real estate returns. The negative skew on the Standard and Poor's 500 is a function of the selection period. The columns on characteristics illustrates the challenge alternative asset indices face from proxies.*

Table 1.7: A comparison of different asset classes 01/06/2013 - 01/06/2018

Traditional and Alternative Asset class 10 Yr returns						
Asset	Cide	10Y Return	10Y SD	Skewewnes	Kurtosis	Characteristics
Equities (*)	S+P500	0.68867	0.03256	-0.45954	0.72885	Long only
Fixed Income (*)	US TREASURY CONST MAT 10 YR	0.11642	0.00031	-0.33018	-0.38219	Long only
Cash proxy(*)	US TREASURY CONST MAT 1 MONTH	0.01738	0.00040	1.50349	1.15100	Long only
Commodities (**)	S+P GSCI	-0.42277	0.05156	-0.69625	0.96770	Futures
Hedge Funds (**)	HHFRX NORTH AMERICA	0.03139	0.01716	-0.49628	-0.10809	Multi Strategy
Real Estate (**)	S+P500 REAL ESTATE	0.35667	0.03936	0.11086	-0.40373	Proxy
Private Equity (**)	SPLPEQTY	0.55788	0.03613	-0.34699	0.21631	IRR Peer Group

*This table shows Traditional and Alternative Asset Class Returns over a 5Yr period. Traditional Assets are marked with an * and Alternative Assets are marked with **. This table is presented to compare with the full 10 year dataset as markets went through regime change post the 2008 credit crisis. Note the positive skew in cash and the negative skew in all assets except real estate.*

1.6 Academic relevance

Academic research directed at identifying the right benchmark for alternative assets is of scholarly importance because indices facilitate the professional management of investments. An index enables the empirical testing of performance relative to a benchmark. A good benchmark permits the analysis of investment within the context of the trade-off between risk and return thereby allowing academic hypotheses to be tested. An inappropriate or poorly constructed benchmark makes academic tests and associated analysis invalid. PCA addresses this by supplying a dimensional insight to unstructured instrument data. It allows the individual alternative assets to be viewed from the perspective of how they relate to each other. The weights of a PCA index are based on statistical factors and as such the output is considered superior to a naive portfolio. The GRS test is applied to the PCA derived indices in the empirical chapters to test this.

The selection of the appropriate index for scholarly research is a data quality issue. A key motivation for the PCA approach is that the theoretical base, that works well for benchmarking traditional assets, is not as elegant when extrapolated to alternative assets. It is for this reason that further academic investigation into alternative asset benchmarks is warranted and relevant. There is, however, a consensus amongst finance academics that a good benchmark has to fit in with theory. This view persists despite the fact that there has been literature, similar in vein to Reinganum (1981), suggesting that some models are misspecified.

The nature and cross-section of index returns is also of academic interest. The CAPM predicts that benchmark investment returns from alternative assets be derived by market and instrument specific factors. Risk adjusted out-performance relative to benchmark returns should therefore be generated by skill or luck (α). It can be further observed that the returns of alternative investments are influenced by leverage and a number of risk factors not captured by the single index model (that influence β). As most academic alternative asset pricing models, however, start with the assumption of multivariate normality of the joint distribution of returns, the PCA approach, that captures these risk factors, is robust.

In summary, the taxonomy of benchmarking that the following essays deliver is important from an academic perspective. Benchmarks are used to convey information about asset classes and the investment set that fund managers benchmark themselves against. As investments in alternative asset classes increase, due to ongoing financialization, the need for appropriate alternative asset benchmarks also increases. This highlights the practical, as well as academic importance, of benchmark design.

1.7 A caveat: The Roll Critique

A word of caution needs to be raised in respect of testing how optimal indices are. Testing two indices against one another is similar to testing a two parameter asset pricing model. In this respect, there is a caveat answering the research questions by resort to the method proposed. This is whether the appropriate benchmark is observable. The *Roll Critique* shone light on this in respect of benchmark choice. In it, Roll (1977) showed that testing the Capital Asset Pricing Model (CAPM) equation is equivalent to testing the mean-variance efficiency of a portfolio, and hence a benchmark. As a result, the CAPM and what is the most appropriate benchmark derived from it, is tautological if the market fulfills the assumptions behind MPT.

The Roll Critique suggests that a valid test of asset pricing theory would require the benchmark to contain every single asset. It is not possible to do that with alternative assets. Commodities are not all extracted from the ground, hedge funds have restrictive covenants, and real estate is largely owned individually rather than institutionally. As such, the critique is equally if not more valid for alternative assets.

The Roll Critique further suggests that even if the PCA approach proves mean variance efficient, it may well be just because the existing asset price index is not efficient. Similarly, the first component factor weights will turn out well diversified portfolios but whether this is the optimal index would prove impossible to check. Furthermore, the existing indices and the PCA indices will be highly correlated with one another.

The Gibbons, Ross, and Shanken (1989) approach conveniently sidesteps the Roll critique. They point out that, as the CAPM has quite specific hypotheses one can accept that the market portfolio is mean variance efficient. The GRS test is therefore only focused on whether a particular portfolio (index) ex-ante is mean variance efficient and as such suitable for empirical use in later chapters.

1.8 Practical relevance

This thesis is practically relevant because benchmarks have a central role to play in fund management. The alternative asset data-set is often unrepresentative of the underlying asset class. As an exploratory data technique, PCA is a helpful tool to address these issues. As the essays herein illustrate, existing benchmarks are not appropriate if poorly constructed. Although existing offerings give the context for portfolio building and monitoring they can be sub-optimal if misspecified. Indeed, where misspecified benchmarks are used for manager compensation, Admati and Pfleiderer (1997) argue that their selection cannot be easily rationalised.

In the academic literature, the practical relevance of a suitable benchmark was underlined by Siegel (2003). He claims that they determine the performance of investment managers perhaps more than any other influence. The owners of assets, who appoint managers, also benefit from greater benchmark clarity. A good benchmark allows them to understand the *relative risk* that a fund manager is taking and compare it with his/her portfolio *expected return*. In a similar way in academia, research into benchmarks is also practically important because a valid benchmark provides interested parties with an objective means of evaluating skill. Investors looking for excess returns do not typically want to invest in risky assets unless they offer a risk-adjusted out-performance of a benchmark.

Another relevant practical issue is the large and growing indexation industry that mimics index returns. Academics Arnott, Hsu, and Moore (2005) question whether this industry is benchmarked to the optimal mean variance indices. Passive investment funds that follow such alternative asset indices need to be consistent and easily replicated. This is not the case in alternative asset classes as shall be explained in the chapters on commodities, hedge funds and real estate.

It is observed that a benchmark is an important communication tool for professional fund managers. It allows for better expression of expectations as to how a fund should be managed. It therefore fulfils an auditing and accountability function, the import of the term benchmark. A better understanding of costs, commissions and spreads can be had by attributing the return.

These practical issues are worthy of individual investigation but are not captured by academic databases. The role of these in index replication is addressed by Petajisto (2011) who label them *index turnover costs*. This is particularly the case for alternative assets where the transaction costs are typically higher than traditional assets. The benchmark is the first step in such an investigation.

In addition to indexation, an industry has grown around index measurement and attribution. A joint *STOXX/Deutsche Borse* White Paper, STOXX (2013), points out that innovations in the benchmark industry promote capital flows and helps the asset management industry by promoting niche benchmarking areas such as corporate social responsibility. The many investment strategies that address alternative assets, similarly, has helped evolve innovative benchmark methodologies and even driven financialization. A PCA derived index would add to the growing number of benchmarks.

The practical use of benchmarks includes the tactical as well as the strategic. Investors, such as commodity trade advisers, apply benchmarks tactically, according to Agarwal and Naik (2015). In this respect they give trading advice based on such benchmarks and therefore there is a high bar in their construction. They are also, like many others, using the benchmarks for marketing. As such, benchmark indices need to be clear and transparent.

An important practical issue from an individual fund manager's perspective is that many alternative investment funds use a performance fee structure based on a benchmark hurdle. This is designed to reward skill. According to Bacon (2015) benchmarks have a challenge in making these appropriate to the task. The returns of such hurdle rates ignore transaction costs and are better suited to determining timing skill as an added value contribution to a portfolio. Often, as investigated by Goetzmann, Ingersoll, and Ross (2003), these are combined with high water marks which further emphasise the importance of benchmark choice.

There is also a strategic element for practitioners to consider, as Anson (2004) explains. Long term institutions invest in alternative assets to match long term liabilities. They require benchmarks for strategic research and portfolio construction. In this respect, the benchmark is used for strategic rather than tactical decisions. The increase in portfolio allocation to alternative asset classes presents new challenges to both benchmark creation and their selection. Investigation of these is important because investment strategy targets cannot be built around a poorly defined portfolio. This subset of benchmarks is under-researched because institutional investment in alternative assets classes is a recent phenomena, driven by financial innovation and a desire for superior risk adjusted returns. As it is increasingly common for large institutions to include alternative assets in investment portfolios, it is clear that their benchmarks need to be better defined.

1.9 Thesis structure and the presented essays

This thesis consists of eight chapters, which are independent but related. The chapters are ordered in a way that allows the reader to get a critical insight into the specific issues that arise in applying PCA to the benchmarking of alternative assets. Individually, they address the research specific problems presented by benchmarking investment in commodities, hedge funds and real estate. The literature review in chapter two and benchmarking techniques in chapter three address how these have been approached by traditional finance academics. The empirical work suggests how PCA can be used to enhance the body of knowledge and add to the literature. As a broader taxonomy, the thesis fills gaps in the literature on theory of index construction and its uses in academic practice.

Chapter two, *a review of the literature on investment benchmarks* consolidates the diverse research that embodies investment benchmarks and indices in order to identify what gaps exist. In this context, it introduces PCA and suggest how it might be used to enhance existing benchmark approaches to alternative assets. As there is no distinct line of literature on benchmarking in its own right, it draws from papers and research which uses indices in academic areas such as performance persistence, market efficiency, and capital asset pricing. It illustrates what has been learnt from a benchmarking perspective, especially in respect of alternative assets and identifies the gaps that PCA indices might address.

Chapter three, *a taxonomy of benchmarking techniques*, provides a theoretical framework for index construction. It reviews the common approaches and challenges faced in the benchmarking of investment assets. There is no similar review in the literature. It provides a critical analysis of the nature of alternative asset returns and how these may present issues for the construction of indices designed to measure and attribute the performance of such assets. It identifies some of the problems with existing indices that PCA can help address. It further details each of the major contributions to index construction methodology, laying the foundation for chapter four on the refinement of the proposed PCA method.

Chapter four, *on the use of principal components in index construction* provides a theoretical framework for the use of PCA in index construction. It explains the method behind PCA derived factor index weights. It proposes PCA as a way to benchmark alternative asset classes in order to address some of the issues with existing indices identified in the literature. It introduces what has been published on the PCA as applies to individual sub-groupings. It details the mathematics behind PCA's use in index construction. The chapter proposes a refinement to the index sampling process and how it might be used to create optimal indices for alternative assets. The PCA method detailed in this chapter is applied in the subsequent asset specific empirical chapters to elicit insights into the underlying cross-section of returns and see if a PCA index is appropriate.

Chapter five, *commodity benchmarks, challenges and refinements*, reviews the key issues related to commodity investment. This is typically done through proxirs in the futures market, presenting issues of whether the universe is representative and whether the returns are influenced by factors like roll yield at the expiry of such futures contracts. It is shown that PCA indices can make a contribution to the optimal benchmarking of such assets.

Chapter six, *hedge fund benchmarks, challenges and refinements*, critiques the use of peer group indices, widely used to measure investment performance in hedge funds. It proposes a number of refinements and suggestions on how synthetic strategy returns might present a more replicable way to construct an investment benchmark for such assets. It contributes to the literature by proposing that synthetic hedge fund indicies can be created by PCA at the strategy level.

Chapter seven, *real estate benchmarks, challenges and refinements*, reviews the illiquidity and lack of homogeneity in this asset class. It does this through the lens of Real Estate Investment Trusts (REIT's), a liquid listed instrument. It investigates whether the principal component approach would add value in the weighting and construction of benchmarks focused on this subset of the asset class. The chapter concludes that their is little value in this approach using REIT's, thereby adding to the growing literature on these asset proxies.

Chapter eight is the concluding one that puts the findings in the context of theory and practice, detailing the contribution of the research and which research gaps have been filled.

1.10 Findings

In the researching of this thesis, it was found that index construction is a well defined science and has a strong grounding in financial theory. Whilst it may not be possible to improve on the mathematical method, it is found that the sampling of constituents can benefit from the use of alternative statistical method. PCA is proposed as a way to do this, grouping common components and weight constituents based on their factor weightings. In this respect, a major contribution of the empirical chapters is to demonstrate how the common PCA factors of the alternative asset sub-groupings can help capture homogeneous sets that can be weighted in indices.

The main contribution of the empirical chapters is to show how PCA can be applied to specific assets. Investigation of the PCA approach gives rise to a number of insights, all of which can help develop more appropriate indices for alternative assets. It is suggested that synthetic indices can be practically designed based on factor weights, thereby solving the investment constraints in existing passive indices. Specifically, the thesis finds that (1) commodity indices have an emphasis on production based weighting, which is sub-optimal, (2) hedge fund indices based on peer groupings suffer from tracking error as a consequence of the net of fee reporting and (3) real estate indices suffer from the lack of homogeneity in the physical property stock.

There are a number of important ancillary findings from the research presented in chapters 5, 6 and 7. These point to future areas of investigation. The commonality in benchmark deficiency proved to be their classification. It is found that the weighting criteria for alternative assets cannot follow the same protocol as that of traditional assets due to the presence of illiquidity and leverage. It is suggested that in the absence of an alternative asset pricing model that can capture these idiosyncrasies, this leaves open the possibility of generating alpha against the existing commercial indices. This line of thought is expanded upon in the later empirical chapters.

It is clear from the literature that time series distortions are not well catered for in current index design in respect of alternative assets. This is particularly poignant in hedge funds indices. This most likely reflects the lack of homogeneity in the hedge fund various sub-strategies. The nature of the returns of derivative instruments needs to be taken into account, by incorporating the roll yield or their defined payoffs. Improvement in benchamrks is therefore of scholarly importance.

It is found that improvement on current benchmarking practice may well prove difficult due to a lack of reporting about leverage. In addition, the illiquid nature of alternative assets and how this impacts the reputability of an index is explored. It is concluded that existing benchmarks are not always appropriate. It is found, as a result of this, that this is a deficiency in the capital markets and that index constructors should pursue synthetic replication more.

1.11 Conclusion

The chapter set the scene for this thesis on benchmarks for alternative assets. It defined what a benchmark is, namely a tool for the objective measurement of investment performance. It laid the foundation for the subsequent chapters which contribute to the literature by identifying where there are shortcomings in existing benchmarking practice for alternative assets and how PCA can be used to fill these. It set the criteria for the objective tests of the proposed PCA index method and illustrated how mean variance provides a theoretical justification of the role that indices fulfill.

Alternative assets were introduced in the context of what makes them a distinct asset class from traditional assets. In this respect, it was explained that as a consequence of the unique nature of the investment instruments, the distribution of the time series is non-normal and exhibits skewness and kurtosis. It was further explained that they are an asset class by exception and as such the sub-groupings have their own distinct issues. It explained that a focus on alternative asset benchmarks is important due to the increase in their importance in asset allocation and the dissatisfaction amongst professional fund managers that the return series of existing benchmarks does not correspond to the sort of assets and weights that they have gained exposure to.

This thesis addresses the gap in the literature on benchmarks that relates to the appropriateness of indices for alternative assets. It is shown that existing indices for alternative assets are not representative, have a sampling method that is designed for equities rather than proxies, do not reflect investment preferences or the nature of the returns of such assets. PCA is proposed as a way to address these. This chapter introduced the research problems that emanate from this, namely (1) whether PCA can be used as an appropriate benchmark for alternative assets and (2) what insights and/or improvements can be gained through PCA. It identified the problem of using proxies to measure alternative assets. It explained how they are used in the context of risk and preferences, particularly in traditional asset classes. The chapter explained how a PCA index focuses on correlated assets and can be used to quantitatively generate an index. This is expanded upon in chapter four.

The test of index superiority is common to the three empirical chapters. In the context of the different characteristics of alternative assets, this concept is that any challenger to existing benchmarks should be able to out-perform what exists on a risk adjusted basis.

The essays in this thesis present the concept of a PCA index as a benchmark for alternative assets. In this respect, an index has construction rules for both security selection and their weights and the proposed PCA method is consistent with this. That said, as alternative asset classes are used to diversify a portfolio, one has to have appropriate benchmarks for each and every asset class. The most common of these are commodities, hedge funds and real estate, the subject of the empirical testing of the PCA index approach. The investigation was shown to be relevant because the choice of an appropriate index has implications for the way assets are allocated. It also has implications for the way alternative asset managers are selected and the testing of market efficiency for such assets by academics.

Chapter 2

A review of the literature on investment benchmarks

This chapter is a literature review relating to investment benchmarks. It identifies gaps in the scholarly contributions and observations about indices. Its establishes the role of a benchmark in finance and what are the desirable properties of a suitable benchmark for alternative assets. In this respect, it provides a synthesis of finance related papers, drawing out their benchmark related contribution. It does this by examining research into time-series, cross-sectional variation in investment performance, persistence of returns and the measurement of skill. It reviews benchmark usage in investment style, investment risk factors, and risk exposures. It highlights how the science behind benchmarking has developed largely from practice and subsequently adopted by academia to test and validate theories. Its contribution is in showing how indispensable appropriate indices are to academic enquiry. It is argued that benchmarks have been on a development path, starting from a theoretical base in modern portfolio theory. This evolution has resulted in alternative asset investors adapting the index method from traditional asset classes. New evidence shows this to be sub-optimal. It is therefore concluded that evaluation of alternative asset can be done using bespoke benchmarks.

2.1 Introduction

This chapter critically reviews the academic literature that incorporates investment benchmarks. Its extended format creates a synthesis on the body of knowledge relating to benchmarks through the lens of papers that have incorporated indices in their approach and methodology. The aim of this review is to give an overview of index uses and present the results of earlier studies on the measurement of investments in both equities and alternative assets. The identification of gaps in the literature serves as a starting point for the PCA index analysis and application to alternative assets. In this respect, the review supports the investigation of the benchmarking of real estate, hedge funds and commodities.

To align with the research questions, the focus of this review is on the way academic benchmark usage has developed and what constitutes best practice. In other words, what makes an index appropriate. Bailey, Richards, and Tierney (1990) are a good starting point. They made an important contribution to this by exploring the appropriateness of benchmark concept and design. Bailey (1992c) extended this to an understanding of benchmark quality. Their insights, explained later, are used to evaluate the empirical and subjective quality of the proposed PCA derived indices in this thesis. They form the basis of the dimensions depicted in the *Optimal Index Spiderweb* analysis (introduced in chapter one) and to provide a subjective framework for evaluation of different index approaches.

In order to identify the deficiencies in the literature, this chapter drills down into the specific alternative asset classes and the unique challenges that these present in performance measurement, the primary role of benchmarks. As mentioned, the literature uses indices to test theory and measure outcomes. In this respect, a major theme in finance is the question of whether skill or chance delivers investment returns. This can only be deduced using an appropriate index. Although appropriateness is defined, Bailey, Richards, and Tierney (1998) suggests it is still possible to game benchmarks. As such, there is still a gap in the literature on what constitutes the appropriate index for each and every asset class.

In equities, the most appropriate index is often assumed to be a market proxy derived from Markowitz (1952). Major stock-market indices are used as substitutes by academics. That said, in a paper on the components of investment performance by Fama (1972), it is conceded that benchmarks other than the market proxy might be more appropriate. This is because, sometimes it is difficult to obtain reliable ex-post observations of risk and dispersion. He suggested this might be the case if returns were non-normal or had two parameter stable distributions. This observation applies to some alternative asset classes, as demonstrated in chapter one.

In the literature, there is less discussion on alternative asset benchmarks than there is on equity indices. That said, Cumming, Helge Hass, and Schweizer (2013b) corroborate the view that there may be more appropriate benchmarks in respect of alternative assets. They developed a benchmark that took account of the statistical parameters of the return distributions and best-fit distributions, showing that it produced better results than a Markowitz framework. This suggests there is merit in testing other benchmarks, such as the PCA approach proposed in this thesis.

Combined, the conclusions of Fama (1972) and Cumming, Helge Hass, and Schweizer (2013b) suggest that a PCA derived index has the potential to be more appropriate than existing commercial indices. Further, Kryzanowski and Rahman (2008) provide a motive for filling this research gap by showing that it is possible to exploit benchmark inefficiency.

This review draws conclusions from the prior research and gives perspective and insight into index construction method and appropriateness. It does this by answering eight academic questions. Its aim is to establish what the role of a benchmark is, and what are the desirable properties of a suitable benchmark for alternative assets. In this context it builds on the introductory chapter and critically assesses whether current benchmarks are appropriate for alternative asset classes. It illustrates where existing benchmarks are positioned using the *Optimal Index Spiderweb* framework as outlined in chapter one. This framework highlights the many dimensions that not captured in existing alternative asset indices and the failings of the proxies which are often illiquid and/or a derivative of the asset which can be captured by a PCA approach.

The aim of this review, therefore, is to illustrate what can be extrapolated on benchmark and indices from academic research in finance. Also, to identify the challenges found in benchmark construction method specifically related to alternative assets. Its focus is on the questions that need to be addressed in order to have an appropriate index for the asset class. The contribution that this chapter makes is, as a result, identifying additional subjective factors that index constructors should take into account in building appropriate benchmarks for alternative assets. The reviewed papers demonstrate the importance of benchmark choice and construction method. From them, lessons are learnt on which benchmark approaches are most appropriate. These lessons provide a useful contribution to the finance discipline in their own right.

2.2 Benchmarking as a research field

The first key observation from the literature is that benchmarking is not a *de jure* research field in its own right. Indeed, much of the literature resides outside of the highest ranked finance journals. Despite this, Conover, Broby, and Cariño (2013) document the many different types of benchmark and there is a rich source of authoritative practitioner literature that is produced by the commercial index providers; as exampled by MSCI Indices and Methodology (2015), Standard and Poors Dow Jones Indices. (2018), and FTSE Russell. (2017).

There has not been a literature review on benchmarks that has been published in a top academic journal. That said, a good literature review on portfolio performance evaluation can be found in Aragon and Ferson (2006). They divide their review along the lines of traditional, conditional and stochastic performance benchmark measures, thereby incorporating conditional alphas and relational betas. These same issues are covered in this chapter, albeit with a focus on what can be learnt from the method, rather than what is learnt from the measurements.

Equity benchmarks and indices have been developed since 1871 when Alfred Cowles and Associates created the first US index. Their use has since been widened and Morrison (2016) illustrates that benchmarking is not just used in portfolio management and investment assessment but also in regulations, laws, and standards. As such, benchmarks could be said to be a *de facto* research field. The first indices, developed by Dutot (1738) and Laspeyres (1871), are expanded upon in chapter three. Their introduction, however is limited in application and has evolved into their current state, particularly for equities.

Benchmarks feature heavily in finance literature. There is extensive use of benchmarks in academia to test capital asset pricing models and to identify the presence or absence of fund manager skill. The foundations of MPT are built on the concept of mean variance efficiency proposed by the aforementioned Markowitz (1952). He shows that optimal investor portfolio choice can be calculated. It can therefore be benchmarked and measured. As a result, index related lessons can be learnt from the literature on the application of mean variance approaches.

Academics, like Carhart (1997), use benchmarks to measure the persistence of returns. Others, like Amenc, Goltz, and Lodh (2012) use benchmarks to understand the various beta factor risks influencing returns in alternative assets. As such, where deemed appropriate, the prior literature is used to illustrate benchmark related issues to these two key areas of scholarly study. As a consequence of prior finance research, this review also covers other time series related topics such as factor models.

Fama and French (2010) highlight that much of the factor research done by finance academics is in fact simply data snooping against a benchmark. Within the field of finance, portfolios have also been designed to test concepts, such as market efficiency. These are also benchmarks. They include a whole genre of economic, statistical and fundamental risk factor portfolios. Similarly, these can be benchmarks in their own right, or lead to interesting insights for benchmarks in general. The research questions that dominate the literature, those that pertain to performance measurement relative to an index, conditional and stochastic risk, apply equally to alternative assets. As such, PCA indices could well have a similar if somewhat narrower role to play.

Alternative asset classes have a history of recorded price data points. These, however, have not been converted into a benchmark time series in the same way as equity time series have been. The lack of precision in older first generation alternative asset indices is because the asset class groups a number of instruments together, many of which are not homogeneous. A discussion of this phenomena can be found in Amenc and Martellini (2002) and adjustments can be made. In this respect, Abhyankar, Copeland, and Wong (1997) provide a methodology for adjusting non-linear returns in reconstructed time series. PCA derived indicies negate the need for such adjustments.

In addition to the measurement of return, a great deal of the literature on benchmarks is on risk. Elton and Gruber (1997) point out that increasing incremental diversification in a portfolio is not just about returns but also lowering risk. Meanwhile, the different nature of the time series of returns of the three alternative asset classes under investigation in this thesis, namely commodity futures, real estate investment and hedge fund strategies, require both a focus on diversification, risk reduction and the nature of their returns.

Another observation from academia is that benchmarks are also important in asset allocation. In this respect, Terhaar, Staub, and Singer (2003) investigate how to allocate between asset class benchmarks and illustrated some of the differences between them. Their work suggests that benchmarking is more than a portfolio level application, that benchmarking is applicable at the strategy level. This is potentially more relevant for alternative assets, which are initially selected from a strategic perspective.

In terms of benchmarking as a research field, much of the literature focuses on indices as benchmarks rather than hurdle rates as benchmarks (or other genre). In order to understand this distinction, the various different types are addressed next, as is how academics have used them in return analysis.

Within finance, there has been little discussion on the use of PCA indices for benchmarking. The use of PCA by researchers is not however new. PCA has been used to construct portfolios by Partovi and Caputo (2004) but it has not previously been applied to the construction of indices in finance. PCA indices have, however, been used is social economics, economic tourism, foreign direct investment, investment efficiency and even other disciplines like biology.

The literature on PCA suggests that its use in indices can help address the skewed and/or non-normal return time series of hedge funds, commodities funds and real estate funds identified in the literature. Non-normal distribution of returns present a challenge as to how universe samples are made when constructing an index. This line of literature is relevant because, as Young and Graff (1995) demonstrate, there is a general stochastic discount pricing framework that skews the expected excess returns of an index time series.

In summary, whilst benchmarking is sometimes overlooked as a scholarly research field, the choice of benchmark is of scholarly importance.

2.3 Tracking error

In the literature, tracking error is an important concept. It is used to measure portfolio performance against benchmarks. Several methods are reported in the literature on how to measure returns of a passive portfolio relative to an index. Tracking error is a measure of how the returns on a benchmark index differ from the returns on a passive portfolio. It is a measure of the unexplained element that arises in the return relative to a benchmark.

Tracking error is explained by Sharpe (1994). It is calculated by "taking the square root of the average of the squared deviations between the investment's returns and the benchmark's returns, then multiplying the result by the square root of the scale of the returns." Tracking error can be extended by using the value of R2 of CAPM. The formula is shown below.

$$TE_{a,T} = \sqrt{\sum_{t=1}^{T} \frac{(R_{a,t} - R_{b,t})^2}{len(R_a) \cdot \sqrt{scale}}}. \tag{2.1}$$

Where:

- $TE_{a,T}$: tracking error for asset *a* at time *T*
- $R_{a,t}, R_{b,t}$: return for asset *a*, return for benchmark *b* at time *t*
- $len(R_a)$: number of observations
- $scale$: number of observations per year.

There are numerous ways documented in the literature by which tracking error may be measured. The most common is to calculate the mean of the actual differences in returns between an index and the portfolio being measured. It is also possible to calculate the standard deviation of any return differences between the portfolio and index. Another approach is to do a regression and then calculate the standard error of fund returns over the benchmark returns. Pope and Yadav (1994) identify some failings in this approach, particularly in the impact of serial correlation on the result.

The existing research on indices documents a number of statistical challenges that impact tracking error. As indices are constructed by a process of sampling, the Standard Error of the statistical output is an important variable. Black, Jensen, and Scholes (1972) document where cross-sectional tests face similar measurement error bias. A low measurement on an index time series means there is less variability in the sampling distribution based on the standard deviation. [1] An index based on PCA, which minimizes the variance of the dimensions, provides an alternative statistical approach.

2.4 The different types of benchmark and their appropriateness dimensions.

There are many types of benchmark and the literature reflects this. These came into existence largely as the result of practical application. Much of the literature is therefore incidental. There is no common thread other than historical improvement based on finance theory. Broby (2007) documents this by recording the way index benchmarks developed from the price method. This is used in economics to calculate the retail price index. The documented history includes the subsequent development through to the current justification for the use of free float indices.

The literature also shows that the appropriateness of a benchmark should depend on its role. In this respect, a contribution is made by Siegel (2003) who suggests that benchmarks should be (1) useful for funds manged in an active way, but also should be (2) useful as a framework for passive funds. Hammond and Subramanian (2013) point out that this approach reduces their subjectivity and bias, especially in instances where they are designed for one set of users at the expense of another.

The literature, however, has no precise answer as to what the right benchmark proxy is. As previously mentioned in the introductory chapter, Roll (1977) famously critiqued tests of asset pricing models by questioning whether the benchmark chosen in such tests could actually be shown to be a true market proxy. Clearly, if the benchmark is miss-specified, the test is as well. The same would apply for alternative asset class. The GRS test proposed by Gibbons, Ross, and Shanken (1989), as explained in chapter one, sidesteps this issue.

The impact of various firm specific events on the prices of securities using indices as a benchmark has been the focus of a great deal of academic research. An example of this is presented by Edmister, Graham, and Pirie (1994). Indices help answer the research question as to what extent the price performance around an event has led to abnormal returns. There is a whole strand of literature that answers this using t-tests on market adjusted returns as measured

[1] The standard error of a regression can be used based on an estimate of the Capital Asset Pricing Model.

by a broad index. Researchers reject the null hypothesis to justify market efficiency. Such studies are not as common in alternative asset classes as they are in equities. They could be the subject of further scholarly investigation.

As explained, the literature refers to multiple types of benchmark to evidence expertise and evaluate performance. That said, in an academic context, a well specified benchmark should allow interested parties to identify and evaluate risk exposures. Along these lines, Bailey, Richards, and Tierney (1990) argue it should have a similar risk exposure to the underlying investment portfolio to which it is applied. They further go on to list several types of benchmarks that fulfill these requirements. These are

- absolute and/or target return benchmarks;
- manager universes and/or peer groups;
- broad market indices;
- style indices;
- returns based ratio benchmarks;
- factor model based benchmarks;
- custom security-based (strategy).

These types of benchmark are decomposed in the sub-section that follow, with the addition of "risk based ratio benchmarks". As can be seen from the classifications, the absence of PCA derived indices is a clear gap in the literature.

2.4.1 Absolute and/or target return benchmarks

In academic literature, the term absolute is not precisely defined. In current practice, absolute return benchmarks are typically applied to investment strategies that target either predictable or consistent returns, rather than the positive returns the term implies. Brown and Goetzmann (1995a) investigated this aspect of absolute benchmarks and found that the performance persistence of funds using such benchmarks decreased over time. In other words, from a benchmark perspective, absolute targets are unrealistic when risk is being taken. Although sometimes used in alternative asset classes the fact that such benchmarks do not account for the nature of risk, long term compounding of returns and the cyclical nature of markets make them unsuitable for academic review of this asset class.

The first usage of benchmarks in the literature, prior to the 1970's, was largely as a tool to measure absolute or target returns. Prior to that, the use of absolute benchmarks was common in the fund management industry. These are minimum target returns that are set for a fund manager to exceed. They are sometimes used in alternative asset performance evaluation, particularly with hedge funds.

Early research by Arrow (1964) showed how investor preferences shaped absolute versus relative risk taking in investment. In this respect, the evidence shows that some investors have preference for absolute benchmarks. Indeed, many alternative asset class investors have such a preference.

The properties of absolute return benchmarks are explained in the literature by Granger and Ding (1995). Initially, such benchmarks were designed to measure a positive rather than negative investment return over a specified time horizon. It was Fama (1972) who showed how selectivity can improve on such an approach and much of benchmark theory developed the concept from that time to the present. Now, the term absolute has its wider meaning, encompassing a more inspirational target to deliver positive returns. The preference in academia, as explained by Bacon (2011), is for benchmarks that can allow for more practical measurement and attribution.

As relates to investments into alternative asset classes, the academic case for using absolute benchmarks is made by Till and Gunzberg (2005). They suggest absolute benchmarks might be appropriate for commodity futures investment. The same arguments can be presented for other alternative assets, and indeed Edwards and Caglayan (2001) do this. They argue that there is more absolute downside protection in alternative asset classes and this would imply they exhibit returns that might benefit from benchmarking to absolute targets. Tudor and Cao (2012), found some evidence that hedge funds produced absolute returns, but caveat their observation by pointing out that this is not universal. Regardless, Admati and Pfleiderer (1997) suggested absolute hurdles do not fulfill all the appropriateness criteria for a useful benchmark, or in their words as "not adding up".

In summary, the consensus in the literature is that absolute benchmarks are inferior to well constructed indices. Bacon (2011) puts this down to the fact that they do not allow for attribution. They do offer a benchmark but they lack risk adjustment and instrument level returns.

2.4.2 Peer group benchmarks

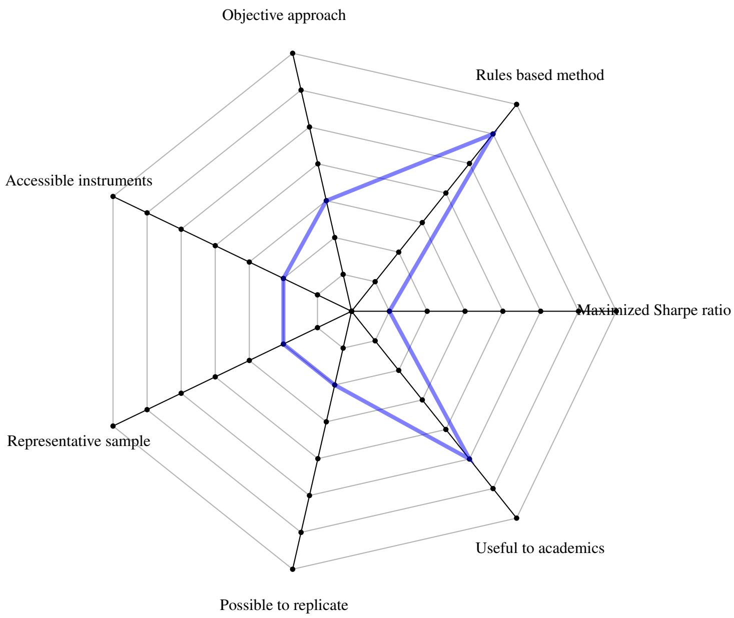
As the literature developed, in the 1980's, peer group comparisons became more common. This next level of complexity in benchmarking made the median manager returns the *de-facto* hurdle to beat. With alternative assets, this has become the modal benchmark. An explanation of the method can be found in Hunter, Kandel, and Wermers (2014). This approach uses manager universes and peer groups to compare fund managers with their competitors. Peer groups were developed by commercial providers of funds who compete for market share based on relative performance to competitor funds.

Peer group benchmarks have a number of shortcomings. PCA derived indices can address specific issues with indices based on them. These include (1) their being net of fees and (2) their subjective bias and (3) their poor replication ability. A discussion as to the suitability of peer group benchmarks can be found in Bailey (1992a). He observes that such peer group benchmarks are convenient and cheap. He claims that they have naive appeal but he is dismissive of their appropriateness due to the aforementioned issues. In figure 2.1, this is demonstrated by a *Peer Group Index Spiderweb*. Note the poor dimension scores for accessibility, representation and replication.

In papers that use peer group benchmarks, the weighting method is either equal or asset under management weighted. The choice impacts the cross section of returns. Investigating such weighting differences, Blake and Timmermann (1998) found big divergences in style and performance between the constituent funds of such benchmarks. This is concerning because the peer group approach is prevalent with the pension and mutual fund industry. Their contribution is therefore a critique of the use of such benchmarks. As far as alternative asset classes have implications for hedge fund indices.

The literature indicates that peer groups suffer from another weakness, namely that they influence fund managers asset allocation decisions. Blake, Lehmann, and Timmermann (2002),

Figure 2.1: Peer Group Index Spiderweb.



Peer Group Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. A peer group index scores a good six on dimension one, five on dimension six, but poorly at 3 on dimension two, with dimensions three four and five only scoring 2 and dimension seven a 1.

in further investigation of this issue, found that the use of such benchmarks resulted in performance clustering around the median fund, implying that the managers were minimizing risk against each other.

Peer group benchmarks are particularly prone to survivorship bias, a phenomenon first identified by Brown et al. (1992) and covered extensively in the literature. This is where benchmark returns can be overstated because under-performing funds are closed. This biases the sample towards those funds that perform well resulting in misleading returns. Grinblatt and Titman (1993) reported that the survivorship bias effect before transaction costs and fees is around 0.1 to 0.4 percent every year. Such bias manifests itself in commodity, hedge and real estate funds. It is therefore an issue when peer groups are used to evaluate them.

The literature further suggests that peer group benchmarks result in performance issues for the fund managers using them. Brown et al. (1992) and Fung and Hsieh (2002a) observe that peer group benchmarks have a tendency to under-perform passive index benchmarks. They suggest this is in part due to returns being stated net of costs. Admittedly, such a conclusion should be considered alongside the fact that active fund managers also tend to perform poorly against indices.

In terms of benchmark selections, the literature shows that there is also a problem with bias selection. Sensoy (2009), shows how such benchmarks can be gamed by the mismatching of peer groupings to fund managers style. The investment universe of peer groups is typically grouped by asset class and or the investment approach within that class. As such, they are often used for alternative investments where funds are developed to broaden exposure to such an asset class.

Despite the identified drawbacks, a number of investigations into specific alternative asset classes have used peer group benchmarking. Essentially, such papers take a collection of funds with the return series derived from similar investment disciplines. These include Brown and Goetzmann (2003a) for hedge funds, Geltner and Ling (2009) for commodities and Maxwell and Maxwell and Saint-Pierre (1998) for real estate. The papers that use this approach typically include rankings, scatter graphs and analysis for illustrative purposes.

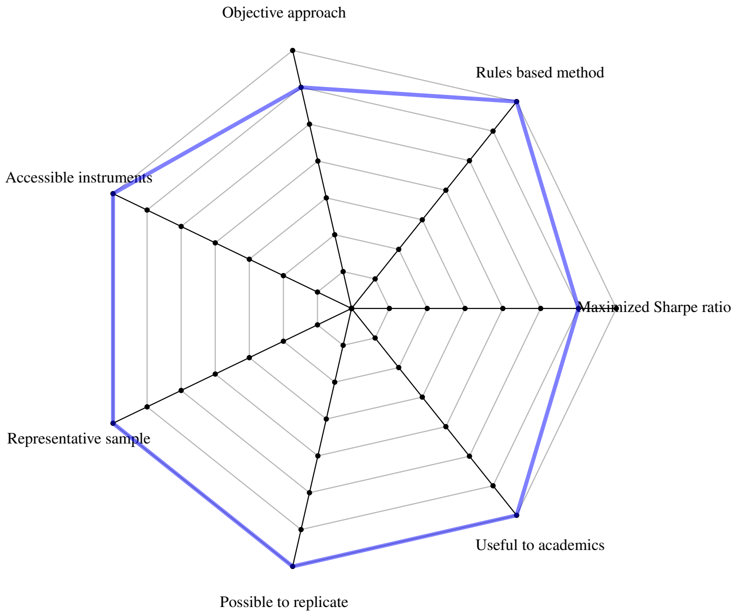
In summary, peer groups benchmarks suffer from a number of failings. This thesis proposes PCA indices as a way of addressing such issues and this is returned to in the essay on hedge funds in chapter six.

2.4.3 Broad market indices

The most widely cited benchmarks in the literature, and commonly used in practice, are the broad market indices. The construction of these is addressed in the discussion on market proxies in the first chapter. Enderle, Pope, and Siegel (2003b) describes the statistical and qualitative characteristics of such benchmarks. The *FTSE 100*, *Standard and Poors 500* and *MSCI* world are all examples of broad market indices. They represent an academically important construct due to the relationship between risk as proxied by standard deviation and expected return.

A good description of how broad market indices relate to the CAPM can be found in Broby (2011). In that paper, it is shown how indices allow the measurement of a whole asset class and how they are the *de facto* market proxy. Many professionally managed funds are benchmarked to the broad market indices. As can be seen in figure 2.2, they score highly on almost every dimension in the *Broad Market Index Spiderweb*.

Figure 2.2: Broad Market Index Spiderweb.



Broad Market Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. The broad market index scores highly on all dimensions, proving to be the closest to the Optimal index which would score seven on all dimensions. The index scores 7 on dimensions one, three, four, five and six, but a slightly lower 6 on dimensions two and seven.

In many instances, the literature shows that benchmark choice impacts manager fees. Heinkel and Stoughton (1994) demonstrate that fees are sometimes a linear function of a portfolio's performance in excess of a benchmark. Benchmark selection and construction are therefore important. They demonstrate that a fee arrangement based on this can overcome agency problems associated with adverse securities selection. With alternative assets however, such a relationship does not exist because the time series of returns are non-linear. That said, such funds are often sold with a performance fee relative to a benchmark, typically a broad one.

The literature on broad market indices builds on the theoretical framework of Markowitz (1952) and MPT. Indeed, broad market indices owe their lineage to him. In this respect he took benchmarks to the next level by describing the portfolio problem in the context of mean variance based on a quadratic optimization. Elton and Gruber (1997) point out the implications of such mean variance approach are intuitive. One of these is the importance of the passive index that represents the optimal portfolio. That said, not all indices are broad. Narrower market indices exist that represent investment styles within asset classes. Capital asset theory suggests the market portfolio is mean variance efficient, so such broader benchmarks are the most common variety in practical usage.

The literature highlights that the key advantage of broad market indices is that they can be used to compare relative returns across all spectra of investment approaches. In this respect, they are used by both academics and professional fund managers. This is unsurprising as Waring and Siegel (2006) point out that relative return benchmark investing against a broad market index "is the only means through which real value can be added to portfolios." The issue surrounding their usage, from an alternative asset class perspective, is that it is not always clear what the market universe is. This is important because return attribution has to be used to drill down and identify security selection, industry selection, or country selection in equity portfolios, or yield curve positioning in fixed-income portfolios. The method used in cross sectional empirical investigation in the literature can only be done if the universe is defined and clear.

In the finance literature, optimal portfolios are closely aligned to those broad market indices in as much as capitalization weighted portfolios are theoretically mean variant optimal. The weighting approach aligns the construction of the index with the CAPM which assumes that the optimal portfolio contains all tradable assets. For a discussion on this methodology see Haugen and Baker (1991). Such market capitalization weighting presents issues for alternative assets.

The equal weighted index, meanwhile, is an interesting side branch in the literature on broad market indices. These so called $1/N$ portfolios are described by DeMiguel, Garlappi, and Uppal (2009). There are two reasons for using equal weighted portfolios as benchmarks. Firstly, they are easy to implement, not relying on estimation, and secondly they are aligned with investors who use naive allocation rules. There is much discussion on whether the $1/N$ approach is as good as if not superior to an optimal mean variance one. There is broad consensus that it is a simpler approach that produces results comparable to mean variance portfolios. As a result of this, various hybrids have been proposed. An equal weighted index can be combined with, for example, a minimum variance index, as demonstrated by Maillard, Roncalli, and Teïletche (2010). For a survey of the various equal weighted combinations, see Chow et al. (2011) (2011). They document heuristic weighting strategies such as risk cluster equal weighting, diversity weighting, and fundamental weighting. It is possible to devise benchmarks based on this approach. The re-balancing period, however, has important return implications.

There has also been academic contribution on alternatives to capitalization and equal weighted indices. This line of research tends to focus on indices rather than benchmarks, so does not directly address this. Arnott (2011) questions it indirectly by investigating whether market capitalization is a reliable proxy for company, fundamental and/or economic factors.

A final strand of the literature on broad indices focuses on the impact of turnover on their time series of returns. In this respect, broad market indices need to be re-balanced and that generates turnover with its associated costs. This has a performance impact. Fong and Lee (2008) looked at this issue by comparing a sample of updating benchmark characteristics. These "updated' benchmarks gave a lower tracking error to the reference portfolios that had not been updated. The conclusion and contribution of this is that re-balancing make such benchmarks more efficient, despite the cost. That said, there is an optimal amount of re-balancing that can be done without impact, and that can be a subject of further academic research.

In summary, the literature provides a strong theoretical basis for the use of broad indicies but their appropriateness for alternative assets has not been documented.

2.4.4 Optimal indices

Simliar to broad market indices, the literature on optimal indices derives from the work of Markowitz (1952). His starting point is that investors select portfolios that take account of expected return and standard deviation. In his framework a portfolio is expressed in terms of maximizing return on investment risk. This significant insight is extended to the construction on mean variance optimal portfolios. It is these that can be used as applied benchmarks. In effect, an optimal index is a mean variant portfolio.

Theory suggests that the forecast rate of return on such a portfolio, and hence a benchmark, is therefore the weighted average of expected returns. The tangent point with the capital market line is the market proxy. This tangent point is what broad market indices aim to replicate.

$$\mathrm{E}(R_p) = \sum_i w_i \, \mathrm{E}(R_i) \tag{2.2}$$

MPT is further expanded in the literature to incorporate utility and the capital market line. Utility was explored at the same time as Markowitz was developing his theory by Stigler (1950). In order to optimize its utility, a benchmark portfolio should be mean variant efficient for it to be a useful portfolio management tool. In this context, Markowitz (1952) shows that the variance of the rate of return on this portfolio, and by extension a two stock benchmark, would be represented thus

$$\sigma_p^2 = w_A^2\sigma_A^2 + w_B^2\sigma_B^2 + w_C^2\sigma_C^2 + 2w_Aw_B\sigma_A\sigma_B\rho_{AB} + 2w_Aw_C\sigma_A\sigma_C\rho_{AC} + 2w_Bw_C\sigma_B\sigma_C\rho_{BC} \tag{2.3}$$

$$\sigma_p^2 = w_A^2\sigma_A^2 + w_B^2\sigma_B^2 + 2w_Aw_B\sigma_A\sigma_B\rho_{AB} \tag{2.4}$$

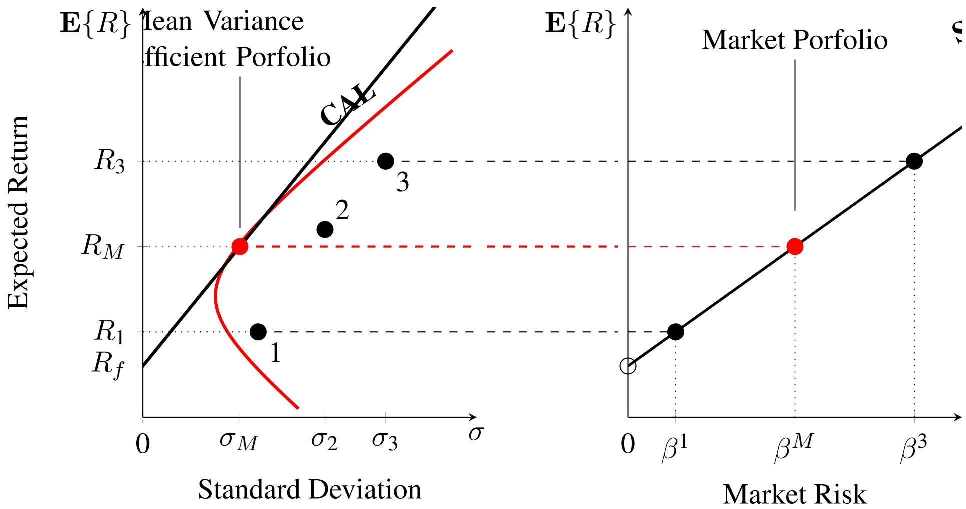
Where:

- $\sigma_p^2 = w_A^2\sigma_A^2 + w_B^2\sigma_B^2$ = The variance between the benchmark instrument A and benchmark instrument B.
- $\sigma_A\sigma_B\rho_{AB}$ = The co-ovariance between the returns benchmark instrument A and benchmark instrument B.

The portfolio and/or benchmark's standard deviation is, in a Markowitz (1952) context, a weighted average of the constituent standard deviations. This is why the constituent weights of an index are important, Any portfolio or benchmark delivers diversification benefits unless the assets are perfectly positively correlated. Mathematically, the constituent selection problem is actually a quadratic one. The weights should minimize risk. This requires that the assets must be positively weighted, hence why most early generation indices are long only.

The elegance of the mean variance approach is its introduction of risk to the literature, as explained by Sharpe (1964). This concept relates to the attributes of a good benchmark. With a risk proxy, it is possible to be more precise in respect of the trade-offs that portfolio construction requires. In effect the CAPM depicts a single asset class, that of all traded securities. MPT is thus used to show where benchmarks fall on the Securities Market Line. It gives rise to a world where one can diversify away unsystematic risk. In this respect, a well-constructed benchmark index can largely diversify risk away. The Capital Market Line is a graphic illustration of this and is reproduced as figure 2.3.

Figure 2.3: Market Proxy on the efficient frontier



This figure shows an efficient frontier with the transect with the Capital Asset Line (CAL), thereby illustrating the market proxy "M". The second diagram also depicts the relationship to the Securities Market Line (SML).

The literature suggests that the optimal portfolio, and hence benchmark, theoretically consist of all assets in all markets weighted by their capitalization. In this way, broad indices have a theoretical justification. That said, as Arrow (1964) explains in respect of commodity claims, in practice broad indices are confined by conditions of certainty. In that respect, indices cannot conceivably incorporate all assets. It is not certain what these are. As such a degree of sampling has to be employed, as shall be demonstrated in the next chapter. It is useful, however, in that a market capitalization index with a beta of one can be used in conjunction with a Capital Asset Pricing Model to allow for the calculation of a portfolio's alpha and beta relative to the benchmark. Not all benchmarks enjoy this advantage, hence the predominance of indices as benchmarks.

As the finance literature has developed, so too has the concept of an optimal portfolio. Research into this was conducted by Haugen and Baker (1991). The suggestion that the theoretical optimal tangent point is one that can be derived in an unconstrained fashion using both long and short positions. With alternatives, where leverage and derivatives are used, this represents an additional challenge to identification of the optimal mean variance efficient portfolio.

The market proxy portfolio is just one of many equilibrium portfolios found in the literature. This conclusion is very useful from a benchmark perspective as it can give a framework for the understanding of portfolios and hence their benchmarks that are driven by factors other than market risk. It is also less restrictive about the information inputs.

A whole genre of optimal indices are also derived from the *Arbitrage Pricing Model* based on risk factors, introduced as a concept by Ross (1976). Although related, this approach is quite distinct from CAPM. The later is a single factor model and as such cannot be used to develop bespoke benchmarks based on a risk decomposition. The main style based risk factors include volatility, capitalization, value and momentum.

2.4.5 Free float indices

Recent finance papers use benchmarks that incorporate a free float adjustment, an important index refinement. The concept is explained by Haugen and Baker (1991). It is now considered a *de facto* requirement for a market proxy. This is because not all shares are actively traded and as such replication becomes harder for large investors. The formula, detailed in Broby (2007) to do this is depicted below:

$$\text{Index level} = \frac{\sum_i \mathrm{P}_i \times \mathrm{N}_i \times \text{Free float adjustment factor}_i}{\text{Index divisor}} \tag{2.5}$$

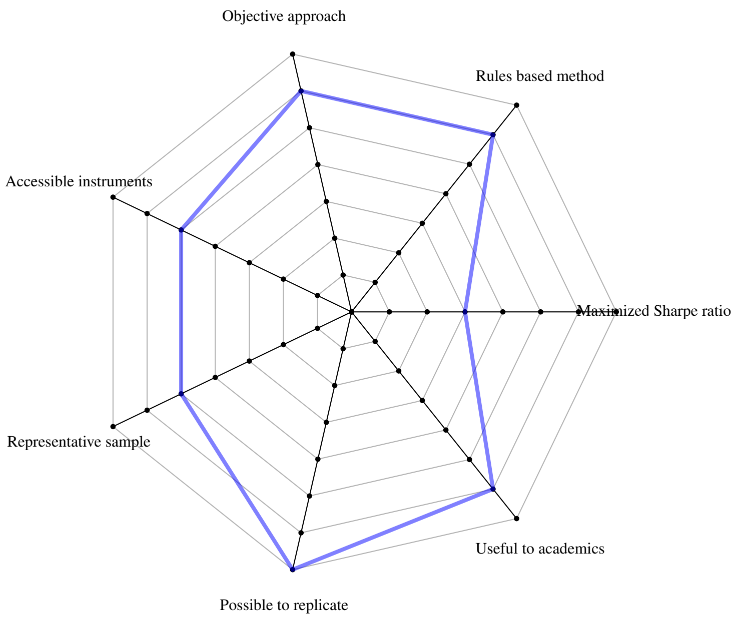
Where:

- P = Price of stock
- N = Number of shares

Extending the free float concept, Elton, Gruber, and Das (1993) pioneered the use of a benchmark focused on various market capitalization metrics. The approach is not easily transferable to alternative asset benchmarks as there are no insiders. That said, hedge fund owners have "skin in the game", owning units in their own funds. The use of free float is also not appropriate for alternative assets as they are not investments limited by issuance but rather derivatives of underlying instruments.

2.4.6 Factor indices

Figure 2.4: Factor Index Spiderweb.



Factor Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. A factor index scores well on the Likert scale. In dimension five if sores a seven, and dimensions one, two and six a 6. Dimensions two and three score a five. It scores poorly on dimension seven with a 3. The PCA index, introduced in chapter four, is based on factor exposure.

There exists a considerable body of literature on factor based benchmarks. These are very similar to a PCA derived index, as shall be explained in chapter five. That said PCA is distinct. Jolliffe (2002) states that they are "really quite distinct techniques". Factor benchmarks are used in active fund management. They build on the Arbitrage Pricing Theory of Ross (1976) to incorporate risk factor returns. A good explanation can be found in Daniel et al. (1997). They argue that empirical evidence suggests that factor model based benchmarks are better at delivering ex-ante forecasts of the future cross sectional returns.

Factor models, mentioned earlier, are particularly suited to alternative assets and are consistent with the common principal components approach introduced later in this thesis. Figure 2.4 shows the *Factor Index Spiderweb*. It compares well with other variants, suggesting the PCA approach has merits. Factors can be fundamental, economic and statistical in nature.

Fundamental factors include dividend characteristics, earnings per share, cash flow per share, revenue, the number of employees and book value. The literature illustrates that the weighting in a factor index is done as follows;

$$Wf_i = F_i / (\sum_{j=1}^{N} F_j) \tag{2.6}$$

Where:

- Wf_i = The weight of the factor index

There are a whole host of factor risk styles identified in the literature, Harvey, Zhu, and Liu (2016) identified 316 factor models. They believe this underestimates the true number. There is much debate on how many factors there are and indeed which factors are appropriate and/or statistically valid. Jaeger and Christian (2005), having examined factor models in hedge funds, find that factor based benchmarks can provide a valid, theoretically sound and cheaper alternative to market based indices.

The concept behind factor benchmarks is explained in the literature by Rosenberg and Marathe (1975). They suggest such benchmarks make it possible to optimize active return against active risk. They suggest the modeling of underlying constituents as a collection of factor exposures whilst incorporating unexplained risk. The factor sensitivities in such benchmarks can then used to predict portfolio returns given values of the systematic risk.

The most common factor model in the literature is the single factor market model. Its relationship to the Capital Asset Pricing Model and the Arbitrage Pricing Theory is explained by a single factor return. This can be further decomposed to identify other factors. In this respect, it can be extended to multiple factors by regressing a portfolio's return against those other factors that are believed to influence returns. Its general form is stated thus:

$$r_j = a_j + b_{j1}F_1 + b_{j2}F_2 + \cdots + b_{jn}F_n + \epsilon_j \tag{2.7}$$

where:

- r_j = Portfolio's return

- a_j = A constant for the asset
- $b_j 1 F_1$= A systematic factor
- ϵ_j = The risky asset's idiosyncratic random shock with mean zero.

Within the literature on alternative assets, Jaeger and Christian (2005) argue that factor based benchmarks were valid and theoretically sound. They benefit from being less expensive than peer group hedge fund indices. Rosenberg and Marathe (1975) develop a technique for the quantitative control of the active risk whilst still being useful to active managers. Rosenberg and Marathe (1975) were the first to investigate active return optimized against active risk. This is comparable to the optimization of the benchmark policy versus the benchmark policy risk. In this case, because the returns on securities are characterized by non-market co-variance, where the returns on securities correlate with factors other than the market factor, it is possible to model each security as a bundle of factor exposures plus an inexplicable risk measure.

An interesting allied question is whether bond investment can be measured by factor model benchmarks. This is addressed by Houweling and Zundert (2017). They find that investment grade bonds demonstrated positive alphas in respect of size, risk, value and momentum. For the high yield universe positive alphas are observed for size, low risk, value and momentum factors.

A further contribution to the literature is made by Fletcher (2017) who examined the diversification benefits of factor strategies in a benchmark universe using *Certainty Equivalent Returns*, obtaining posterior distributions to measure the effect of varying level of investor risk aversion. This is the level of return that an investor would accept now from a benchmark or other investment, rather than taking the uncertain possibility of a higher return in the future. He demonstrated that there are out-of-sample performance benefits by identifying common characteristics, a concept explored further in chapter four. Interestingly, from the perspective of long only benchmarks, he finds that imposing short selling restrictions on such characteristic portfolios results in more consistent performance.

A critique of the literature on factor models is that they are used by practitioners to claim out-performance through selective choice of style. The alpha generated from factor models comes largely from weighting methodology. This is potentially an area that PCA Eigenfactor weighted models can help with as weights are not dominated by the style tilt.

An important contribution in the literature to factor models and alternative assets is made by Fung and Hsieh (2002a). In this and later papers, they provide a link between alternative asset class risks and more traditional asset class risk. Roncalli (2017), meanwhile, coined the term *alternative risk premia* which incorporates this linkage. This concept encompasses two different types of systematic risk factor relevant to alternative asset classes, skewness risk premia and market anomalies.

Academics are somewhat divided on factor model indices. Ribeiro and Hodges (2004) signal a potential problem with the use of factor models when futures markets are used as proxies for underlying instruments. They explained that costs had to be taken into account and that this was typically a negative drag on the time series. The literature broadly confirms that factors, as variables, are independent of each other. They can also be used to determine weights, as explained in the fourth chapter where a methodology applying factor weights derived from common components is presented.

Overall factor models enhance the ability to explain time series. Indeed, Amihud and Goyenko (2013) argue that performance of a fund can be predicted using the correlation coefficient of the returns of such a benchmark. An example of this can be found in Gao and Rossi (2013). They developed a commodity pricing model based on three factors. This is illustrative of the sort of factor approach that can be applied to benchmarks. In it, they use the average commodity return, a basis spread factor, and a momentum factor. In order to adapt factors for use as benchmarks it is necessary to assign factor weights. The way this is done with PCA is illustrated in chapter four.

In summary, according to the literature, a factor model can help identify a portfolio and benchmark's sensitivities to risk factors. Valid benchmarks should have sensitivities that are similar over time. This means having similar metrics and being comparable in statistical profile. Although the risk sensitivities are not the same due to active positions, over time, a factor benchmark should reflect the risk of the portfolio. Without this, the benchmark would exhibit systematic bias.

2.4.7 Style and characteristic indices

As has been show, in the literature factor benchmarks are tailored and used in a broad context. They can also be used for specific style factors. In this way they can be used to measure different manager approaches. Style indices are a more recent development that assist with this. They owe their lineage to the research of Fama and French (1992). They show that its beta has only a small ability to explain the cross-section of market returns and proposed instead a three factor formula. This is detailed in their equation below.

$$Er_t = \alpha^c + \beta_{mkt_t} + \beta_{HML_t} + \beta_{SMB_t} \tag{2.8}$$

Where:

- MKT = Market Factor
- HML = High minus Low
- SMB = Small minus Big

The equation introduces three style factors into the literature. Through these Fama and French (1992) demonstrate that size and book-to-market are important additional style divisors. It is therefore possible to build tilted benchmarks biased to these styles that act as benchmarks. This led to the development of the style indices based on large-cap value, large-cap growth, small-cap value, and small-capitalization growth.

In terms of return distribution, Fama (1972) was the first academic to extend research on long only benchmarks beyond the traditional framework. He introduced both selectivity and timing as elements of return. Although his work initially focused on the valuation of mutual funds, it is relevant to alternative asset classes. Such collective investment vehicles differ from traditional funds due to market timing and the use of derivatives and leverage. These change the aforementioned linear relationship between risk and return. The decomposition requires a separate academic study to optimize the selection of the benchmark.

The *Style Index Spiderweb* is the same as that of the factor indices in figure 2.4. There has since been a large increase in the number of style tilts, the most famous of which was the Carhart (1997) four factor model, effectively an extension of the Fama and French model.

Similar to the literature on style, there are a few papers on characteristics, such as that produced by Daniel et al. (1997). They investigate characteristic based benchmarks, their name for style benchmarks, whilst looking at performance persistence. They constructed 125 passive portfolios based on market capitalization, book to market, and prior year characteristics, essentially each portfolio being a benchmark. Their contribution is to show the many ways a portfolio can be divided along style lines.

In the alternative asset class literature, Liang and McIntosh (1998) extended style analysis to Real Estate Investment Trusts (REIT's). Similarly, Fung and Hsieh (2001b), Agarwal and Naik (2000) and Dor and Jagannathan (2002) did the same for hedge funds. As styles are groupings, a PCA approach can be applied to the creation of similar benchmarks.

In summary, the nature of how the literature identifies the Fama and French Model styles limits the practical usage of this approach in benchmarks. That said, once one accepts that style can be decomposed along risk factor lines, a whole host of alternative benchmarks become available.

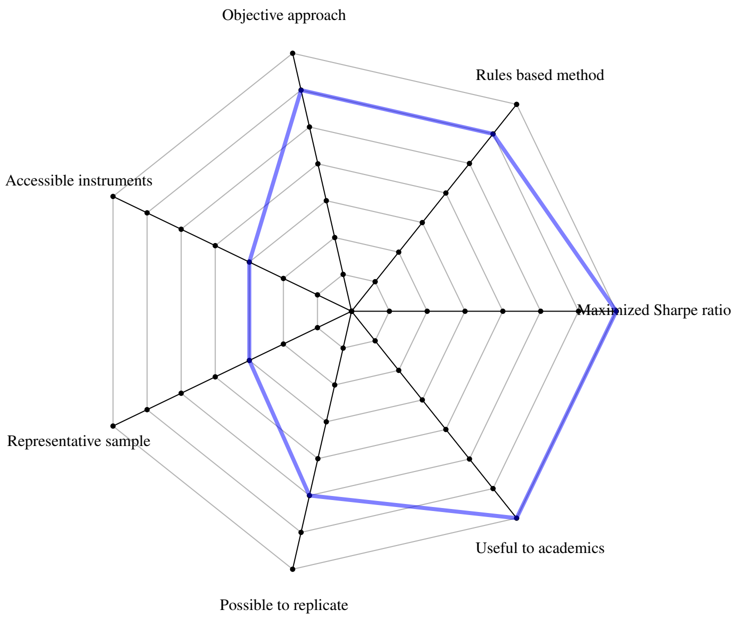
2.4.8 Returns based (ratio) benchmarks

There is a strand of literature focused on ratios as benchmarks. An academically important question associated with manager performance relates to returns and their rankings. How does one rank relative investment returns? In a significant branch of the literature, this is addressed by using return based ratios as benchmarks. They avoid database bias. The comparison of ratios removes any ranking issues.

The most popular of the returns based ratios is the *Sharpe ratio*. This was proposed by Sharpe (1994) and is based on Modern Portfolio Theory. Its appropriateness as a benchmark is depicted in the *Returns based (ratio) Spiderweb* in figure 2.5. The Sharpe ratio, and its derivatives, is based on average returns and standard deviation. It is used in the literature to compare traditional as well as alternative funds. It is particularly useful where there is not a full disclosure of underlying fund holdings. Similar performance indicators can be used that take into account the correlation between the fund and the benchmark. These are the *Treynor ratio* and the *Jensen ratio*.

As explained in the first chapter, if a contender index can deliver a higher Sharpe ratio than

Figure 2.5: Returns based (Ratio) Spiderweb.



Returns based (ratio) Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. This type of ratio performs badly in dimension three and four, with a score of 3. Dimension five score 5. Dimensions one and two score a size and dimension six and seven produces a 7.

an existing index then it may well be more appropriate. The ratio is given by the formula:

$$S_a = \frac{E[R_a - R_b]}{\sigma_a} = \frac{E[R_a - R_b]}{\sqrt{\mathrm{var}[R_a - R_b]}}, \tag{2.9}$$

Where:

- S_a = Sharpe ratio
- rp = The mean monthly return on the portfolio
- rf = The mean monthly risk-free return

As a benchmark, it is used extensively in manager selection and in the literature on manager evaluation. Despite its widespread use in long only fund evaluation, the literature suggests that it is not appropriate to use it where leverage is present. As such, its use in evaluating alternative asset classes may well prove more reliable in commodities than hedge funds or real estate investment trusts. Brooks and Kat (2002) point out, in this respect, that the earlier introduced Sharpe ratio substantially overestimates the true risk and return performance of hedge funds and presumably other alternative asset classes.

As has been observed, the ratio has been subject to critique. According to Sharpe (1994) "the literature surrounding the Sharpe ratio has, unfortunately, led to a certain amount of confusion. Whether measured ex-ante or ex-post, it is essential that the Sharpe ratio be computed using the mean and standard deviation of differential returns (or, more broadly, the return on what will be termed a zero investment strategy). Otherwise it will lose *raison d'etre*."

There are some papers on the Sharpe ratio and its use as a benchmark for alternative assets. According to Titman and Tiu (2011) "if a hedge fund is maximizing its Sharpe ratio, then the R squared of the regression of the hedge fund's excess returns on systematic factors is inversely related to the fund's Sharpe ratio and to its information ratio." As a result, the risk return relationship allows portfolios to be benchmarked based on returns adjusted by their alpha and/or beta and the relationship of those to the Capital Market Line.

Sharpe's critique of his own ratio was followed by a number of scholars proposing alternatives. These are still to become established for the mainstream of benchmarking, but are used in the hedge fund industry. They focus on the smaller partial moments and the negative relationships in return yields. They include the *Omega ratio* proposed by Keating and Shadwick (2002), the *Kappa 3 ratio* proposed by Kaplan and Knowles (2004) and the more widely used *Sortino ratio* proposed by Sortino and Van der Meer (1991). These alternative ratios are not dependent on a normal distribution of returns.

In academia, the Treynor ratio is also used as a benchmark. This is similar to that of the Sharpe ratio except that the denominator in this case is the beta of the measured portfolio. It is a risk weighted and adjusted return measure based on systematic risk. Graphically, it is the point at which the regression of the excess return intersects a portfolio to the benchmark. The Treynor ratio is given by the following formula:

$$T = \frac{r_i - r_f}{\beta_i} \tag{2.10}$$

Where:

- T = Treynor ratio
- β_i = The systematic risk of the portfolio
- rp = The average monthly return on the portfolio
- rf = The average monthly risk-free return

In summary, ratios can be used as benchmarks but the literature suggests they are not always appropriate.

2.4.9 Risk based ratio benchmarks (VaR and the Information ratio)

There is a rich stream of literature that focuses on alternative risk measures. An important risk-adjusted measure in the finance literature is proposed by Jensen (1968), namely *Jensen's Alpha*. It is the most commonly used academic skill measure, and hence a benchmark in its own right. It is a measure of whether the portfolio yield's extra return, justified the level of risk. The formula is given by:

$$\alpha_J = (R_i - R_f) - \beta_{iM} \cdot (R_M - R_f) \tag{2.11}$$

Where:

- α_j = The risk adjusted out-performance

There have also been other extensions of the aforementioned ratios in various papers, all of which can be used as a benchmark. The most relevant of these include, the *M-squared ratio* (a risk-adjusted performance ratio). This was developed and extended by Lobosco (1999). As a style and risk-adjusted performance measure, it can be used as a benchmark when there are strong style effects. This is likely to be the case in funds of alternatives.

Another ratio, the *Information ratio*, was added to the literature by Treynor and Black (1973). The Information ratio is a way of examining the performance of an investment by adjusting for its risk. It is a way of measuring skill by taking the expected active return and dividing it by the tracking error. The active return for an alternative asset is the return of the individual instrument minus the return of the chosen index. The tracking error is the standard deviation of this active return. It refers to the degree to which an investment outperformed the benchmark. In effect, it is the consistency with which the investment outperformed the benchmark.

Some practitioners use the Information ratio as a selection benchmark in preference to other academic measures. It is widely used and differs from the Sharpe ratio by virtue of measuring the excess return against a benchmark rather than a risk free rate.

The information ratio formula is:

$$IR_a = \frac{\alpha}{\omega}. \tag{2.12}$$

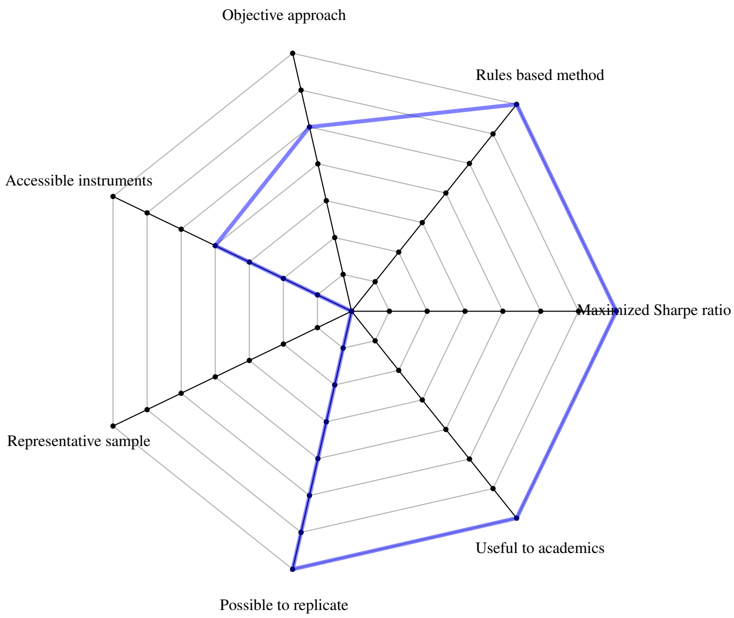
Where:

- IR_a = Information ratio of the asset *a*
- α = $R_a - R_b$. R_a = The return of asset *a* over the period considered
- R_b = The return of the user-specified benchmark over the period considered
- ω = $\sqrt{var\left(\alpha\right)}$.

The appropriateness of risk based ratios are depicted in the *Risk Based Benchmark Spider-web* in figure 2.6. There are a variety of such metrics. In addition to *semi-deviation* there is *value-at-risk*. Risk averse investors can use *expected shortfall* and others *tail risk*. The later is explained succinctly by Eling and Schuhmacher (2007). They further explain twelve such risk measures and risk based ratios. They examine whether they produced significant deviation from the normal distribution. They find practically no difference between the results that these produced in comparing hedge funds. However, the conclusion has not stopped academics using them in their analysis. Amenc, Goltz, and Lodh (2012), for example, adjusted the Sharpe ratio to measure maximum efficiency. Eling and Schuhmacher (2006) in a similar line of investigation found a high rank correlation between different performance measures using hedge fund indices data.

In summary, as detailed in Bodie, Kane, and Marcus (2014) the information ratio can be used to benchmark a fund's performance "provided it itself is based on the right benchmark".

Figure 2.6: Risk Based Benchmark Spiderweb.



Risk Based Benchmark Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. The returns based approach scores zero on dimension 4, limiting its use as an attribution method. It scores a 7 on dimensions one, five and seven, a 4 on dimension three, and a 5 on dimension 2.

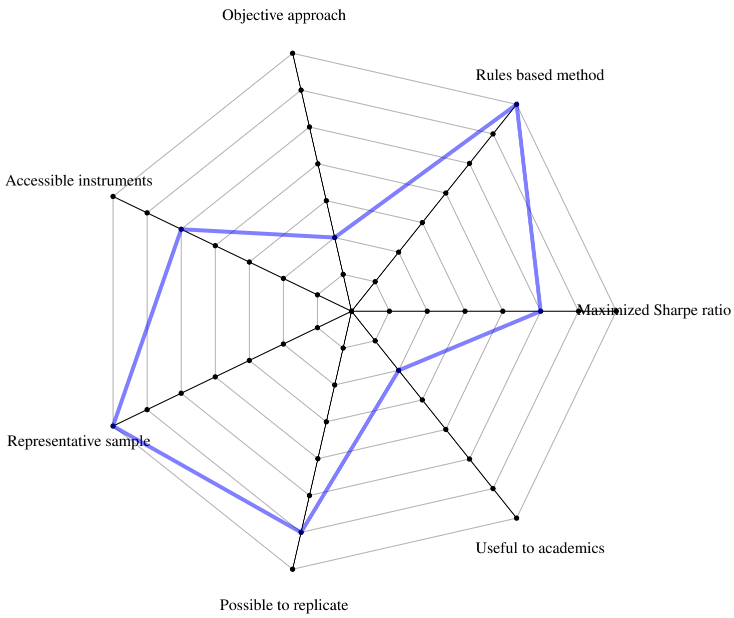
It is for this reason that the information ratio is used by investment consultants to rank funds. Other risk based measure are also used in this way. That said, their usage in alternative assets is restricted to funds rather than individual assets.

2.4.10 Custom and strategy benchmarks

The literature introduces custom benchmarks as a way to address specific asset classes. An example includes Campbell (2010) who investigated art as an asset class. Figure 2.7 illustrates the *t*he pros and cons of such benchmarks. Note how the appropriateness dimension is the most relevant.

When alternative assets are added to traditional assets in a combined portfolio, a strategy benchmark that encompasses the total assets is constructed, as well as a benchmark for the tactical exposure. There are some potentially open questions that arise from the use of such custom benchmarks. Their construction was investigated by Rennie and Cowhey (1990). They

Figure 2.7: Custom Index Spiderweb.



Custom Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. It is not surprising that custom indices score a 7 in dimensions one and four as they are bespoke. They do not perform so well on the broad usage and as such score only 2 on dimensions two and six, 5 on dimension three and 6 on dimension 5.

conclude such benchmarks could prove useful, particularly in the measurement of strategic investment decisions. Custom benchmarks are also termed strategy benchmarks.

According to Kuenzi (2003), strategy benchmarks are distinct from tactical benchmarks. In his work, strategy refers to the long-run mix of asset classes which explain an investor's objectives and constraints. Although not as common in the literature as broad indices, strategy benchmarks are often beneficial alternatives to published indices from the perspective of fund investors, consultants, and investment managers. Additional studies are needed to more fully understand this interaction. Anson (2004) points out that the drivers which provide added return beyond the return offered through passive exposure are tactical decisions against such strategy benchmarks.

Much of the literature on strategy benchmarks focuses on return and attribution. Hoernemann, Junkans, and Zarate (2005) show that strategic asset allocation against a benchmark explains most of a portfolio's return variability and is therefore the most critical decision in the investment process.

Many questions remain unanswered in the literature on the nature of long term rates of return of asset classes and the contribution to institutional diversification, and hence their benchmarks. Ibbotson and Kaplan (2000) illustrate how returns are affected by such strategy benchmarks. They used the adjusted value to determine how much of the total return variability could be described by asset allocation alone. Their equation took the form of

$$R_i t = (1 + PR_i t)(1 + AR_i t) - 1 \tag{2.13}$$

Where

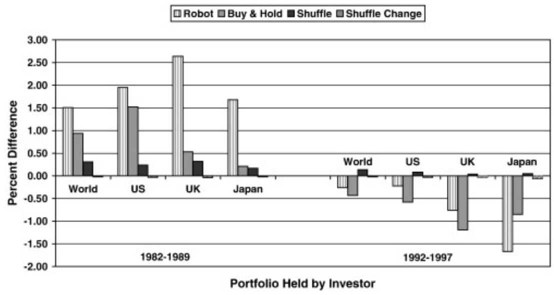
- $R_i t$ = The total return of fund i in period t
- $(1 + PR_i t)$ = The policy return of fund i in period t
- $(1 + AR_i t)$ = The active return of fund i in period t

A strategy benchmark is usually supported by a number of tactical benchmarks, often in the form of an investable index of the sort previously mentioned. Anson (2004) further expands the literature on this by explaining the difference between the strategic and tactical benchmark. This approach assumes a passive investment in the strategy benchmark is a neutral comparison point for measuring investment goals. Custom security based indices can then be used in strategy decisions to make asset allocation calls. Bange, Khang, and Miller (2008) found that fund managers typically under-performed their strategy benchmarks using a simulation exercise with 10,000 re-balanced portfolios. The impact of the weight changes of this exercise are shown in figure 2.8.

Although there are hundreds of thousands of benchmarks in use, the literature almost exclusively addresses performance measurement and investment returns from the perspective of single benchmarks rather than multiple ones. Multiple benchmarks tend to be referred to in the literature on strategy, such as that by Anson (2004). More research is therefore needed into such bespoke offerings.

In summary, where a benchmark is developed for a multi-asset class investor, the literature points out that the variance can be reduced by the choice of asset classes with low or negative

Figure 2.8: Impact of weight-change on strategy benchmark portfolio



The figure shows percent difference between a recommended portfolio standard deviation and a benchmark portfolio standard deviation, 1982 – 1989 and 1992 – 1997. The illustration highlights the impact that changes in constituent weights has on the risk of an index. Source: Bange, Khang, and Miller (2008), page 380.

co-variance. In other words, an asset benchmark is a risk reducing diversified portfolio. It is this property that underlies the investigation in this thesis. A benchmark can exist that is not a risk diversifier, such as cash. That said, if a benchmark lacks diversification it is unsophisticated. Even cash held on deposit should be diversified across different banks to reduce counter-party risk.

2.5 Question 1: Are benchmarks fit for purpose?

The first question to ask to gain insight into what gaps exist in the literature is whether benchmarks are fit for purpose. The literature indicates that finance academics tend to assume that professionally applied broad market proxy benchmarks, used alongside traditional asset classes, are mostly robust and fit for purpose. That said, Bailey and Chan (1993) show that it is possible to game benchmarks, suggesting that this is not the case. Despite this finding, there is little discussion in quantitative papers about the index selected for measurement. By way of illustration, Amenc, Goltz, and Lodh (2012) produce a survey on benchmark usage. They corroborate this gaming occurs in practice.

The *Index Spiderweb*, introduced in the first easy, is used in this review to address how the benchmark types can be subjectively measured as fit for purpose. The *Index Spiderweb* was built on the conclusions of Bailey (1992b) as to what are the fundamental properties of a benchmark. These are that they are (1) unambiguous as to weights of the instruments, (2) investable, to the extent that they can help passively, (3) measurable on a regular and frequent basis, (4) appropriate to a fund manager's style basis, (5) reflective of the investment objectives, (6) and specified in advance. In respect of alternative assets, which Bailey did not review, the question of investability comes up frequently. Likewise, what the manager and the investor are ultimately aiming to achieve through their investment, is relevant. The literature is explored further to propose ways to validate the significance of alternative asset indices.

As mentioned Bailey (1992c) further identifies the issues relating to benchmark quality in another paper. In this respect, any benchmark should have sufficient instruments to give good coverage of the investment universe. These instruments should not suffer from high turnover during re-balancing. An important contribution from this paper is the recommendation that a fund manager's active weight, the allocation to an instrument minus its weight in the benchmark, should not be negative in practice. This is relevant to alternative assets where instrument concentration might be present, for example the weight of crude oil futures in commodities indices.

Similar to Bailey (1992), Amenc, Goltz, and Lodh (2012) find that the important benchmark selection criteria are (1) objectivity, (2) liquidity, and (3) transparency, all common in accepted indices. The findings suggest that style and sector, which academic research has identified as important risk factors, play a less important role. This implies that an appropriate alternative asset index needs to focus on providing representation rather than simply exposure to the underlying asset class.

The literature unambiguously confirms that benchmark choice is important. Bacon (2015), in detailing the differences in benchmark methodology and construction, points out that the type of benchmark selected can lead to large differences in the time series of returns. In this respect, the differing approaches to benchmarking alternative assets suggest that there is further room for research. The need for this is particularly relevant with respect to benchmark choice and the chosen construction method. This is because, as Yasin (2002) identifies, there are no common grounds in theory for such a task.

Typically, a market proxy such as the *Standard and Poor's 500 index* is considered fit for purpose in finance studies. This particular index series has an extended track record all the way back to January 1871, as reconstituted by Wilson and Wilson and Jones (2002). It is the most often featured benchmark in the literature but similar national market proxy's such as the British *FTSE 100 Index*, the German *DAX 30 Index* and the French *CAC 40 Index* are abundant. This is because they are historic country based time series. That said, as equity indices they are not suited to benchmark alternative assets.

In summary, the literature does not find all benchmarks fit for purpose. That said, the PCA approach, focusing as it does on transforming the data with the greatest variance from the first principal component, has merit.

2.6 Question 2: How are benchmarks used for identifying skill?

The second question to ask to gain insight into what gaps exist in the literature is how benchmarks are used to identify skill. The literature clearly demonstrates that benchmarks facilitate the identification and evaluation of a portfolio's risk exposures and the interpretation of both past performance and its attribution. A strand of the literature focuses on separating skill from chance, testing portfolios against pricing models. In this respect, a benchmark index allows one to distinguish between active investment skill and luck, as explained by Fama and French (2010). Their work builds on a long tradition of using benchmarks by academics to investigate market efficiency and capital asset pricing. The selection of an appropriate benchmark is, therefore, scholarly significant.

Fund selection is not explicitly addressed in the majority of the literature. That said, at a fund manger level, Bailey and Arnott (1986) used cluster analysis to show how manager selection can be improved. This is a similar method to PCA. In this context, cluster analysis uses unsupervised learning to find groups of like managers based on similar attributes. Traditional manager indices simply use a process for sampling. PCA, meanwhile, reduces the set of attributes by selecting the attributes with the most variation. As a result, although cluster analysis can be used as a selection criteria for hedge fund indices, the PCA can also deliver weightings and as such, it is argued, is superior.

One way of identifying skill, as identified earlier, is to use ratios rather than indices as benchmarks. These have a linkage in theory to risk. In a benchmark context, the literature suggests that a Sharpe Ratio can be used to decide if a contender benchmark is superior to an existing benchmark. It is also possible to use it to measure the effect of adding new assets to a benchmark, based on an approach taken by Levy and Levy (2014). The literature suggests that a limitation of using the Sharpe Ratio to identify a superior benchmark is that it is only useful when returns exhibit a normal distribution. This can be criticised as it presents issues for its use in alternative asset evaluation where the returns are non-normal.

The literature further suggests that benchmarks have a policy role and the skill in allocation can be measured. That role was explained by Brinson, Hood, and Beebower (1995). Their framework helps users in the understanding of the consequences of a fund manager's investment decisions and the use of a strategic benchmark in the evaluation of their excess returns. Their paper illustrated the difference between a portfolio return and a benchmark return. This is reproduced in figure 2.9. They argue that the investment policy choice against a benchmark is the most important decision an investor makes. As such, the policy index is a tool for delineating responsibility and measuring the contribution of investment policy, market timing and security selection.

Figure 2.9: Policy returns versus a benchmark

	Helpful (to achieve the objective)	Harmful (to achieve the objective)
Internal origin (product/company attributes)	(IV) Actual Portfolio Return	(II) Policy and Timing Return
External origin (environment/market attributes)	(III) Policy and Security Selection Return	(I) Policy Return (Passive Portfolio Benchmark

This figure depicts the four policy return attributions that can be determined versus a benchmark. (Quadrant I represents policy). It shows that active returns are due to Timing, (II)-(I), Selection (III) - (I) and Interaction (IV)-(III)-(II) + (I). Source: Brinson, Hood, and Beebower (1995), page 134.

As in the practical world, in academia, identifying skill requires an understanding of investment objectives. In this respect, benchmarks communicate information about the set of assets in an investment strategy under investigation. They define what the relevant universe is and allow the returns to be set within the context of capital asset pricing and risk factor models. In this vein, both Fama (1972) and Jensen and Black Scholes (1972) could be said, with their research, to have facilitated the evaluation of portfolios based on a comparison of the ex-post returns of a benchmark index. A whole branch of the literature in finance builds on this, utilizing index returns as variables in statistical tests.

In summary, the role of benchmarks in identifying skill is crucial to the professional management of investments and their evaluation by academics. As such the question of what makes an appropriate benchmark for alternative assets is of scholarly importance.

2.7 Question 3: Is it possible to persistently out-perform a benchmark?

The third question to ask to gain insight into what gaps exist in the the literature is whether it is possible to persistently out-perform a benchmark. The literature on performance persistence in benchmarks is of scholarly importance. This strand is largely focused on either equity or fund universes, and measured against investment indices. In the alternative asset classes, there are persistence papers on the performance of commodity trading advisers and real estate investors, but these are largely on market timing and benchmark agnostic.

The large body of research that has focused on evidence of persistence in the returns of mutual funds can be used to obtain insights into benchmarks for alternative assets. This builds on the seminal work of Brown and Goetzmann (1995b), Grindblatt and Titman (1992), and Hendrciks, Jaydendu, and Richard (1993). These academics challenge the equity mutual fund literature with a view that superior performance against a benchmark is possible. The choice of benchmark used in their studies may contribute to their conclusion, hence why it is important to investigate the constituents of an appropriate benchmark and what it means for each alternative asset class.

Although rarely explicitly stated, performance persistence cannot be measured without a reliable benchmark. Poor performing funds tend to disappear from databases, either because they are closed down or merge. This can obscure any estimates of the degree of persistence.

This results in what is termed survivorship bias, and this in turn can lead to benchmark misspecification. Work on this has been done by Brown and Goetzmann (2003b) and Malkiel (1995). What can be concluded from this is that survivorship bias has important implications for index design. The literature is revisited later in relation to peer group indices.

Numerous studies have been conducted on insiders earning abnormal returns against an index. That said, the consensus view based on a large body of evidence is that the efficient market hypothesis holds for publicly available information. As such, it is assumed that investors in commodities, hedge funds and real estate investment trusts are informed.

Performance persistence in hedge funds is a growing field of research in alternative asset classes. It has been approached in the literature by investigators using a number of different methodologies. They can be broken down between the persistence of relative returns, such as the *Cross Product Ratio* of De Souza and Gokcan (2004); and the persistence of individual returns such as the *Hurst tests* used by Eling and Schuhmacher (2007). Other approaches include the regression models of Agarwal and Naik (2000) and the *Chi-square tests* of Carpenter and Lynch (1999).

In summary, the benchmarks chosen by these academics influences their results, once again showing the importance of optimal index construction and choosing the right benchmark for alternative asset. The question of whether it is possible to out-perform given the right appropriate benchmark is still outstanding.

2.8 Question 4: How does liquidity affect benchmark selection?

The fourth question to ask to gain insight into what gaps exist in the literature is how liquidity affects benchmark selection. This is important because benchmarking of alternative assets is complicated by liquidity and a whole genre of the literature addresses this. Alternative assets cannot always be sold at the prevailing price. Liquidity is a common factor in alternative assets as they are not listed instruments. In this respect, there is a big strand of literature devoted to investigating it. Tobin (1958) demonstrated that investors have different risk tolerances for different levels of liquidity. As such, a case could be made for a range of benchmarks.

Diesinger, Kraft, and Seifried (2010) demonstrate how liquidity forms a part in portfolio asset allocation decisions. The literature therefore suggests that different benchmarks may be appropriate for investors with different risk tolerances.

The impact of liquidity on the time series of returns has been researched within the context of Capital Asset Pricing Models. Amihud (2002), for example, demonstrate how the Capital Asset Pricing Model can be adjusted for liquidity in respect of the co-variances of its own return and liquidity, as well as that of the liquidity and return of the market. The lack of literature on such adjustments suggests that it is a challenging field in alternative assets. This may be because equity risk factor models may not apply. This suggests a focus on PCA may have merit. A caveat to this is that where stale pricing is present, some adjustment will have to be made to ensure that a benchmark can be replicated when fund flows occur.

The literature suggests that both market capitalization and the size of a market are variables that impact liquidity. The same volume of securities traded in a small company cannot be compared to the amount traded in a large company. In a similar way, silver trading cannot be compared to gold trading. In this respect, the literature identifies systematic cross sectional differences. This is attributed to transaction costs being higher for small companies and as a result of being traded less frequently estimates of systematic risk from their daily returns are biased downwards. With commodity futures, a corresponding effect could be seen based on contracts traded. With hedge funds, it is unlikely such an effect will manifest itself as the pricing is a function of net asset value and not price discovery. With real estate investment trusts, a similar effect would be likely as they are traded on a stock exchange.

In the literature, one of the most difficult aspects of liquidity, from a benchmark perspective, is its measurement. In this respect, Denvir and Hutson (2006) devised a unique methodology for calculating liquidity. They took the instruments in a traded index and compared it to the listed portfolio of those instruments (an Exchange Traded Fund). Whilst this produced interesting results, the method cannot be used in index construction as there are too many exogenous factors at work.

The literature documents that liquidity can be absent for a number of reasons, see Acharya and Pedersen (2005), Amihud, Mendelson, and Pedersen (2005) and Goyenko, Holden, and Trzcinka (2009). With real estate, Cheng and Roulac (2007) suggest it is because of the friction costs to trading and the physical use that property assets are put to. They suggest focusing on

Real Estate Investment Trusts as a way to overcome this. With hedge funds, Sadka (2010) points out that it is due to the complexity of their underlying investment approach and the imposition of investment gates and lock-ups on the investors in the funds. He investigated liquidity risk in the cross-section of hedge-fund returns. His results show that funds with low liquidity under perform those with good liquidity by approximately six percent annually. This would suggest that a good index would screen out the most illiquid funds. Commodities futures, meanwhile, do not suffer from liquidity issues, although certain contracts, such as milk futures, tend to have less trading than others.

The literature suggests that it would be useful if benchmarks for alternative assets could address the issue of liquidity. Pedersen, Page, and He (2013) suggest a solution in this respect. Their paper addresses liquidity and its impact on the measuring of market price risk on alternative investments. They estimate risk factors by smoothing the asset return series. Their approach can be extended to indices where there is a need to take account of illiquid or stale pricing.

In summary, there is little research in determining liquidity effect, suggesting current benchmarks are not useful at detecting this in alternative asset classes.

2.9 Question 5: How does leverage distort benchmark returns?

The fifth question to ask to gain insight into what gaps exist in the the literature is how leverage distorts returns. There have been no really successful efforts in the literature to design an empirical test consistent with combining the effects of beta and leverage. Instead the literature has focused on decomposing operating and financial leverage. This may be theoretically interesting but it does not advance the case of leverage impact on alternative asset returns.

There have been significant advances in capital theory in the literature in respect of the impact of leverage. This is documented by Ben-Zion and Shalit (1975). They investigated leverage from the perspective of equity risk, confirming that it is a distorting factor. Leverage skews the return distribution and as such is an issue for benchmarks. The presence of leverage is another defining characteristic in certain alternative asset classes but is poorly covered in the literature. This is a clear gap, as leverage distorts the return time series and therefore deserves further investigation. Leverage is present in hedge funds, private equity and structured finance. It is assumed the same would be the case in asset classes with more leverage.

Within the alternative asset class literature the focus has largely been on leverage in real estate investment. This is done through investigation of REIT structure, An example is provided by Giacomini, Ling, and Naranjo (2017). They find that the more debt there was in the capital structure, the greater the REIT'S under-performed their peers. This is surprising, as one would expect an equal propensity to out-perform. This would suggest an index based on REIT'S could be beaten by an investment strategy based on low leverage entities. As this is a market pricing inefficiency, the index construction method itself would not have to be adjusted for this anomaly.

Interestingly, Cheng and Roulac (2007) support the assumption in the literature that leverage detracts from performance, although they demonstrated it was only a weak negative relationship. It is unclear why leverage would be applied if it did not benefit the return, so more research on the return enhancing benefits of leverage would support the literature in this respect and presumably assist in active investment strategies.

According to Arnott (2011), the literature fails to address beta leverage in a satisfactory way. He shows that alternative beta strategies outperform their capitalization weighted counterparts, but suggests that it is driven by exposure to value and size factors. As far as benchmark construction is concerned, this would imply a style tilted benchmark would better capture this effect.

Research also demonstrates that leverage has the effect of increasing volatility. In this respect, reduction in volatility, the proxy for risk, was investigated in a benchmark context by Wang (1998). He shows that volatility can be an issue where benchmarks are built using futures contracts, as is the case with commodities.

In summary, the presence of leverage in collective funds and/or derivative exposure exaggerates investment returns. This has implications for fund of fund based indices and hedge funds in particular.

2.10 Question 6: What are the consequences of benchmark misspecification?

The sixth question to ask to gain insight into what gaps exist in the the literature is what the consequences are for misspecifying a benchmark. It can be deduced from the literature that the large universe of potential benchmarks not only presents a selection issue but also gives rise to the potential for the previously mentioned specification error. In this respect, there is a strand of the literature that addresses the potential for such errors. For example, there have been a number of studies that have applied different benchmarks to the same portfolios. These include Lehmann and Modest (1987) and Grinblatt and Titman (1994). They address the question of whether a benchmark has been correctly applied from the perspective of efficiency, timing and statistical power. The results show that, not surprisingly, inferences can vary using different benchmarks.

The literature that investigates the persistence of performance against alternative asset peer groups is particularly prone to the problem of benchmark misspecification. Agarwal and Naik (2000) claim there is no evidence of persistence in hedge fund performance. The benchmark they use is a peer group, so the alpha they refer to is versus other managers. The issue of what constitutes skill is closely tied to benchmark choice. Performance against a benchmark can be measured using the relative risk measure tracking error (the standard deviation of portfolio active return). The Jensen (1968) measure can be used as a benchmark test.

In terms of benchmark specification it should be pointed out that benchmarks in the literature are poor at attributing short term trading effects. Dybvig and Ross (1985) first identified this problem, finding that the portfolio returns of market timers are non-linearly related to underlying market.

An interesting approach to alternative asset benchmark misspecification in the literature was presented by Kaplan and Schoar (2005). They tested private equity returns for persistence, not against a peer group but against a stock market index traditionally used for equity funds. In this way, they were looking for alpha in the same way as previously described. They argue that the returns in alternative investing should be viewed relative to the alternative of investing in the public markets.

In summary, the literature shows that misspecifying a benchmark results in performance measurement errors and hence invalidation of results. This provides motive for the investigation in this thesis.

2.11 Question 7: How do you determine benchmark superiority?

The seventh question to ask to gain insight into what gaps exist in the literature is how performance superiority is measured. The most used method is *Jensen's alpha* α. This is derived from capital asset pricing models and is the most common benchmark mentioned in literature testing market dynamics. In a benchmark context it can be used to measure any abnormal return of a contender benchmark relative to an existing market index.[2] There has been some discussion in the literature on this approach. In their paper, Cremers and Petajisto (2009) ask whether benchmark indices should have alpha. They conclude that they should not and as such propose methodological changes to the Fama-French factor model in order to reduce the number of the non-zero alphas.

Bailey, Richards, and Tierney (1998) pointed out that many benchmarks involve an active return, but omit the risk element. In this context, it is suggested that when using a benchmark to evaluate a competitor benchmark, the risk can be assessed against a reward variability ratio Sharpe (1964). This supports the use of the Sharpe ratio in the evaluation of contender indices, as already discussed, and by extension, the GRS test.

Measurement of benchmark superiority is addressed by the GRS test of Gibson and Schwartz (1990). This was introduced in the first chapter. The GRS test of benchmark efficiency is effectively a statistical measure of how optimal a portfolio or index is. It is similar to the Sharpe ratio, but adapted for benchmarks. It has been used by some academics to evaluate whether the alphas of indices sum up to zero. It is a form of statistical F test. The residuals of this test, generated from a time-series regression, can be used to derive a multivariate normal distribution. The resultant distribution should have a mean of zero and can be presented in a constant co-variance matrix. In this respect, the GRS test is similar to a Wald test. In other words, it tests whether spurious correlations exist between multiple benchmarks.

[2] Jensen's alpha α is commonly used to test for skill in fund managers. It identifies that part of a fund managers returns that are not attributable to β. That is akin to the part attributable to a benchmark.

The question of how to benchmark without a benchmark, effectively what the GRS test aims to discover, has also been considered by academics. In this respect, Grinblatt and Titman (1993) also proposed the *GT measure*. They use the past portfolio positions as the benchmark, the co-variance between actual portfolio returns and lagged change in portfolio weightings. Martellini (2012) suggests that, if optimization procedures are used, better benchmarks can be designed. He suggests a focus on more robust estimates of moments and co-moments of stock return distributions.

The GRS test, according to Pesaran and Yamagata (2017) is not typically used to test alternative asset returns. They did however do this for hedge funds and the results are shown in Figure 2.4. The results suggest existing hedge fund indices are not optimal.

Figure 2.10: GRS test illustration

Hedge Fund Index relative to Standard and Poor's 500 returns and p-values of the GRS test based on CAPM regressions. The results suggest that existing hedge fund peer indices are not optimal as the p-values are not stable over time. Source: Pesaran and Yamagata (2017) Figure 2, page 35

In a similar line of academic inquiry, Gibbons, Ross, and Shanken (1989) developed the most widely cited test for ex-ante mean variance efficiency of a benchmark. The test is intuitive to implement because it uses realized returns rather than expected excess returns. It calculates the sample error, effectively, the difference between the efficiency in expected return and the efficiency in respect of the aforementioned realized returns. Put another way, the Sharpe ratio of the benchmark is compared with the information ratio of the portfolio being tested as a benchmark.

In summary, the GRS test can be used to test benchmark superiority.

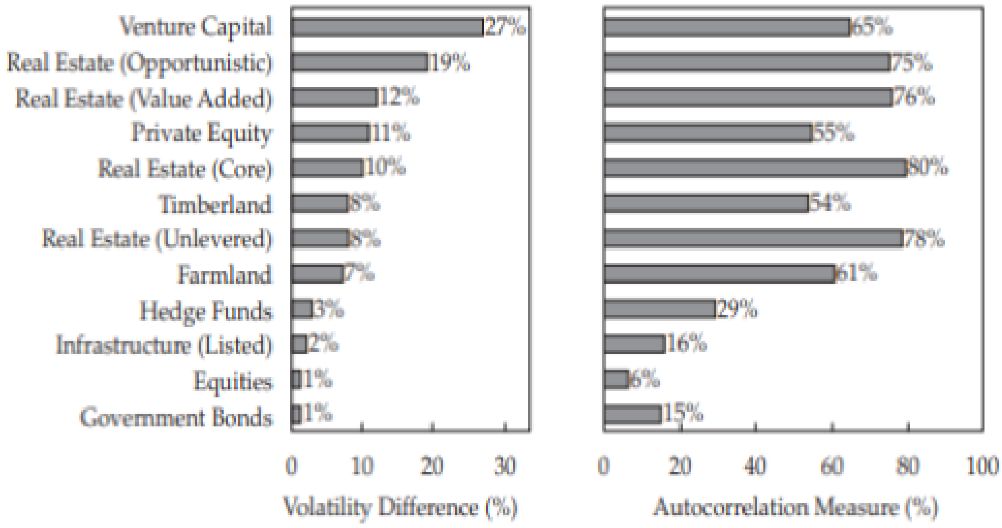
2.12 Question 8: What are the benchmarking challenges for alternative assets?

The eighth question to ask to gain insight into what gaps exist in the the literature is what the challenges are for alternative assets specifically. Alternative assets are not treated in the literature as an asset class in their own right but sub groupings are. Greer (2000) points out, in this respect, that an asset class is a grouping of investment instruments that bear some economic relationship to each other and Kritzman (1999) explains why it is relevant to define alternatives separately as an asset class. Groot and Swinkels (2008) show how it is possible to incorporate such asset classes into pension asset allocation taking into account the uncertainty in the risk premia.

The literature on long run benchmark returns is well established and these are recorded for each asset class in the Ibbotson (2012) yearbook. Traditional stock returns are higher than bonds and cash in the long run. Using the same method, extrapolating a linear relationship between risk and return, would suggest that alternative assets would benefit in the long run in the same way as equities. In other words, they should enjoy a risk premium for the lack of liquidity that would contribute to higher returns in the long run. That said, the nature of the way investors gain exposure to alternative assets would suggest that such an extrapolation is not appropriate.

Cumming, Hassand, and Schweizer (2014) contribute to the literature on liquidity. They point out that alternative asset classes experience data issues due to stale pricing and attempts to smooth the reported returns, both of which cause the previously identified skewness and kurtosis. They account for higher moments in the return distributions of alternative assets and find that replacing the traditional normal return distributions in their analysis yields better results. Pedersen, Page, and He (2013), as shown in Figure 2.12, identified a downward bias in the Standard Deviation in alternative asset classes based on smoothed data that took account of illiquidity.

Figure 2.11: Downward Bias of smoothed benchmark returns



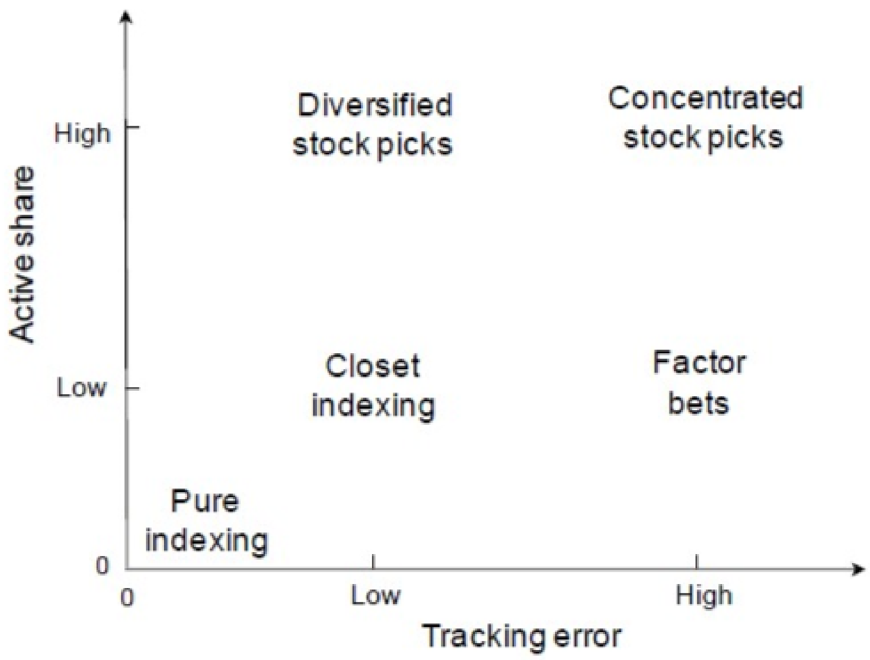
The diagram shows the difference between adjusted and reported index volatility for several alternative assets. It also shows auto-correlation, highlighting how return smoothing contributes to the poor estimation of volatility. This indicates that there is a downward bias to smoothed benchmark returns. Source: Pedersen, Page, and He (2013), Figure 1, page 35.

In summary, liquidity, or the lack of it, is a key feature in the literature on alternative assets. Ang, Papanikolaou, and Westerfield (2014) show that uncertainty about the length of the period of any illiquidity is the primary determinant of the cost of that illiquidity. As such, where illiquidity is a major asset class differentiator, such as in real estate, there has to be some approach for the benchmark to adjust for what are termed stale prices.

2.13 The role of active share

An important question linked issue associated with the appropriateness of indices is how much of the investment universe they represent. This is relevant to practitioners. Active fund managers use a benchmark with risk budgets and other management techniques. As a result, they use benchmark indices to justify their investment process. That said, in practice an index cannot conceivably incorporate all assets and a degree of sampling has to be employed. They find it useful, however, to know what the *active share* is. This is a relatively new concept in the literature on benchmarks. It is a measure of the exposure an active fund manager has to benchmark constituents. Active share was proposed by Cremers and Petajisto (2009) to measure what fraction of a portfolio is different from the benchmark index. In evaluating performance using this measure, they found that a higher active share resulted in higher performance, with the lowest percentile of active share under-performing benchmarks. The relationship between tracking error and active share that they identified is depicted in Figure 2.11.

Figure 2.12: The relationship of Tracking Error to Active Share



The diagram shows the relationship of tracking error to active share. The greater the active share the greater the tracking error. That is a consequence of concentrated stock choices. An index fund, by contrast, has low tracking error and active share. Source: Cremers and Petajisto (2009), page 3342.

By way of further academic investigation, Frazzini, Friedman, and Pomorski (2016) re-evaluate the Cremers and Petajisto (2009) paper on active share and argued that much of the out-performance of active managers identified was benchmark driven. This critique was strongly refuted but it could be said that the discussion only underlines the importance of benchmark constituents and their weights.

A small element of the active share can be attributed to specification error. As indices are constructed through sampling they face the problem that excluded instruments vary and co-vary differently from the constituents. In other words they suffer from specification error as Ramsey (1969) demonstrates. This condition is where it is not possible to know if over or under performance is due to a fund managers selection ability or to the existence of sample bias in the index. This problem is also present in the selection choices that have to be made within alternative asset classes.

2.14 Commodity benchmarks

There have been studies in the literature to understand more completely the nature of commodities as an investment sub-class. From a benchmarking perspective, Bessembinder (1992) notes that commodity markets differ from equities as an asset class. As a result, the question of the appropriateness of existing indices requires investigation.

Some academics argue that institutions invest in commodities as an inflation hedge and that therefore inflation could be a benchmark. In this respect, Erb and Harvey (2006) investigate changes in the annual rate of inflation. They found it explained forty three percent of the variation in the annual excess returns of the GSCI. This is shown in Figure 2.13. Although this is a high number, it is not high enough to be considered a benchmark by the authors.

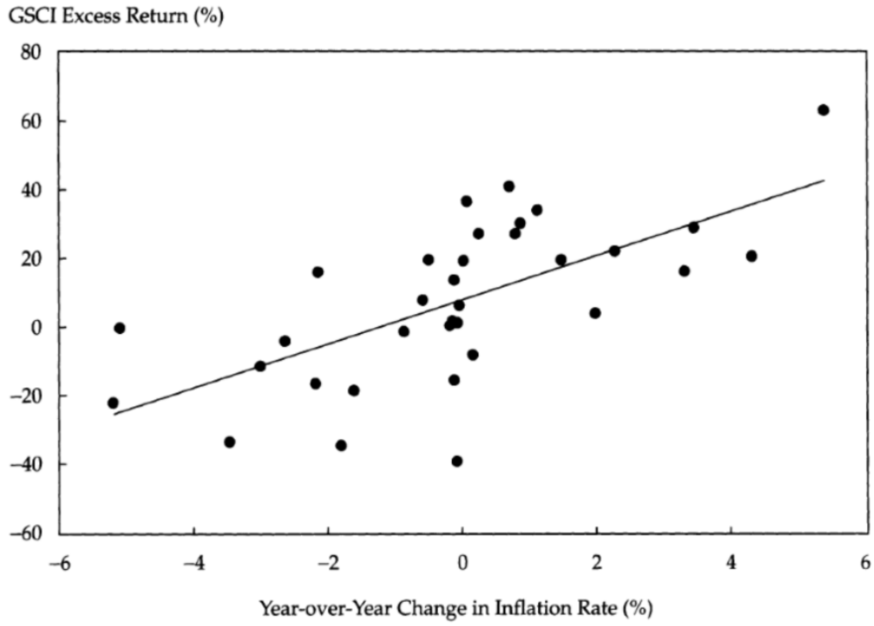
The commodity investment literature has come a long way since Cargill and Rausser (1975) first investigated temporal price behavior in futures markets. The concept of finacialization not only drives the need for benchmarks but also, according to Zaremba (2015), changes their very nature. He notes that financialization has caused a structural decline in roll yields which results in a revised efficient frontier. As already mentioned, the literature deals with their investments from the perspective of the futures market. This is due to the connection between the forward price of an asset and the cash price. This is represented in their paper thus:

$$F_{t,T} = S_t \cdot e^{r(T-t)} \tag{2.14}$$

Where:

- $F_{t,T}$ = i the forward price of an asset with initial price

Figure 2.13: GSCI Commodity Index Excess Return and Unexpected Inflation



This figure shows the GSCI Commodity Index Excess Return and Unexpected Inflation. It is emphasized that most commodities are priced in United States dollars. The GSCI excess return = 0.083 + 6.50 change in the inflation rate; the R2 is 0.43.Source: Erb and Harvey (2006), Figure 3, page 81.

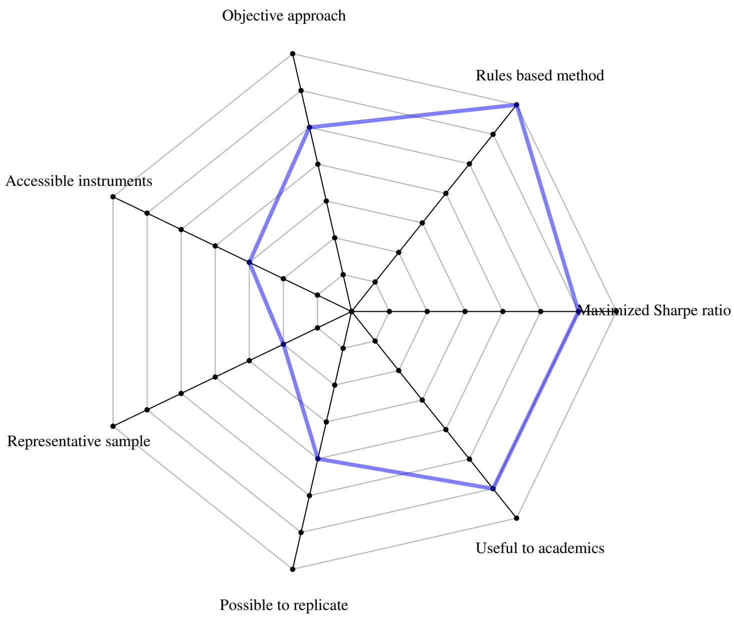
- St, S_t and maturity T, T

There is a good strand of literature on futures and their use as a commodity proxy. For example, Gorton, Hayashi, and Rouwenhorst (2013) detailed the fundamentals behind commodity futures risk premiums and how they vary across commodities and over time. This builds on the earlier work of Gorton and Geert (2006) who investigated the cross-section of returns. They established the precedent of using futures in indices from an academic perspective. Szymanowska et al. (2014) broke down the cross section of returns into two types of risk premia, namely a spot premia that is derived from the price of the underlying physical units and term premia that is based on movements in the yield (as expressed in basis points).

It is clear from the literature that there have been a number of advances in index design over time. The use of futures contracts as proxies and the subsequent roll yield was addressed by academics over time. Miffre (2012) classifies the existing commercial indices between what he termed first, second and third generation indices. The first generation commodity indices simply sampled the available futures and produced an index of those. The second generation indices adjust for contango, momentum and/or roll-yields. These are found to outperform the first generation indices. Third generation indices, it is argued, are more active. They have weighting strategies that go long backwardation and short contangos. That is, they weight instruments to reflect the rolling over of expiring futures contracts. These benefit from lower overall volatility than earlier generation indices. It is accepted that some of the criticisms of the applicability of early generation alternative asset indices equally applies to traditional assets.

As to indices as commodity investment benchmarks, in the literature Dunsby and Nelson (2010) documented the history of commodity futures indices. In addition, a good summary of commodity investing that builds on their classifications can be found in Miffre (2016). The *Commodity Index Spiderweb* in figure 2.14 sums this up.

Figure 2.14: Commodity Index Spiderweb.



Commodity Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. The dimension scores demonstrate particular weakness in accessibility of instruments and representative sample. The commodity index used is based on futures to gain exposure to the underlying asset.

The impact of index investing has been commented on in the literature on equities. The effect is also noted in commodities. Tang and Xiong (2012) found that there is a correlation between commodities included in an index than for commodities that are not included. Hamilton and Wu (2015) also observe similar benchmark effects. Elton, Gruber, and Rentzler (1987) explained how professionally managed commodity futures funds worked, prior to the development of commodity index funds. A reappraisal of such approaches was done by Sanders and Irwin (2012). They claim that index investment wasn't the major driver that others claimed.

On commodity strategies, a useful contribution was made by Fernandez-Perez et al. (2015). They illustrate the nature of the time-series and cross-sectional characteristics of commodity futures. They found that lower skewed commodity cross-sections had higher average returns, and vice versa. Their results showed that a test benchmark portfolio that buys low-skewed commodities and high-skewed short commodities earns 8 percent annually at a 4.08 t-statistic, an alpha of 6.6 percent and a t-statistic of 3.6 corresponds. This finding is significant in that it suggests that an index created without taking into account the skew will have the opportunity to undercut an index that does so.

A critical open question in the literature is the relationship between commodities and inflation. The literature has used indices to investigate this. As previously mentioned, commodity benchmarks are sometimes chosen for their perceived inflationary hedging properties. In this respect, the inflation link hasn't been directly demonstrated in the literature to date. That said, Gargano and Timmermann (2014) demonstrated a link between commodity price indices and macroeconomic variables, especially growth in industrial production and the investment capital ratio. Meanwhile, Arango, Arias, and Flórez (2012) demonstrated a negative relationship between interest rates and commodity prices.

There has been some prior investigation into common factors in the cross-section of commodity futures. The results are varied. Supporting the premise Jagannathan, Malakhov, and Novikov (2010) rejects the *Consumption Capital Asset Pricing Model*. On the opposite side, Miffre (2012) had results that indicated the volatility of commodity futures is unexplained by specific factors. The use of a PCA index methodology could provide interesting insights in resolving this open question.

As commodities are priced in the physical market and a benchmark proxy has to use the futures market, academic investigation has also addressed the pricing mechanism between these two mediums. Ng and Ruge-Murcia (2000) studied the determination of commodity prices at the cash level, assuming a rational profit-maximizing agent. At the economic level, they identified a range of trade and production influences. These include the nature of supply contracts, the seasonality in production, the heterogeneity of depreciation rates, the price elasticity of demand and the costs of spoilage.

2.14.1 Commodities and persistence of returns

Indices are prominent in the literature on persistence in commodity futures return time series. This was the focus of the research of Granger and Joyeux (1980). They identify very persistent but stationary processes, a conclusion relevant to benchmarks. Long term persistence in asset returns and asset volatility was meanwhile investigated by Pasquini and Serva (1999) and Connor and Rossiter (2005). This phenomena makes it easier to classify an asset class because the returns are highly correlated.

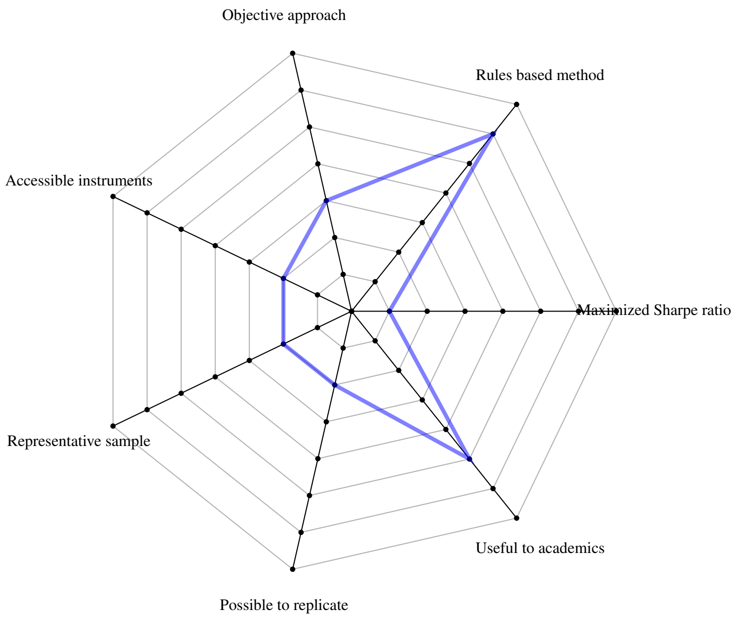
As with the broader literature on fund performance, survivorship bias is a big factor in determining the persistence in commodity fund returns. This was investigated by "schneeweis-spurgin - comparisons of commodity and managed futures benchmark indexes -1997" (1997), as well as Fung and Hsieh (1999) who looked into this. They illustrate that commodities trading advisors, as with hedge funds, see as much as 17 percent of funds drop out of databases every year.

In summary, the conclusions in the literature are relevant for the construction of commodity based PCA indices. A PCA index should take roll yield into account. It is also clear that there is a surviorship issue in peer group commodity funds which PCA can potentially address.

2.15 Hedge Fund benchmarks

There is extensive literature available on hedge funds despite their being relatively new financial products. This literature, however, does not develop a conceptual framework for their benchmarking and assessment. There is a good summary of such issues in the performance evaluation literature, produced by Snoussi and Hellara (2004). Figure 2.15 summarises these conclusions for hedge funds in the *Hedge Fund Peer Group Index Spiderweb*. As can be seen, the dimensions are mostly week, meaning existing indices are not appropriate.

Figure 2.15: Hedge Fund Peer Group Index Spiderweb.



Hedge Fund Peer Group Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. Only the rules approach to inclusion scores highly on the dimensions. A peer group index suffers from being stated net of fees.

A key academic issue behind any index construction is that of data quality. Within benchmarks this was addressed by Liang (2003). He argues that the poorly defined investment universe impacts hedge fund indices, with issues largely stemming from survivorship bias. He further suggests such issues are exaggerated when peer groups are used as the benchmark for performance evaluation. A key conclusion from this is that an appropriate hedge fund index should avoid the use of peer groups.

The typical non-linear time series of returns in hedge funds described in the literature results in differing risk profiles and benchmarks should reflect this. As mentioned, academics address this through the use of the reward to variability ratio. This is adapted for hedge funds by Shlens (2005). Various extensions of this include measures of downside risk. The use of ratios has benchmark choice implications. They can be used to measure risk but they cannot be used to attribute risk.

The risk profile of different hedge funds is impacted by their particular style. How to benchmark the distinct strategies remains an open question in the literature. This was investigated by Olmo and Sanso-Navarro (2012) who analysed the relative performance of returns of the different styles. In testing time-varying nature, they conclude that certain factors have predictive power in respect of the clustering of hedge fund returns. This is supportive of the PCA approach, as the method is similar in nature to cluster analysis.

The literature supports the view that benchmark should not measure alpha generation from style shifts, as that is a function of manager skill. Style based factor benchmarks that measure a single style risk are, however, appropriate in certain circumstances.

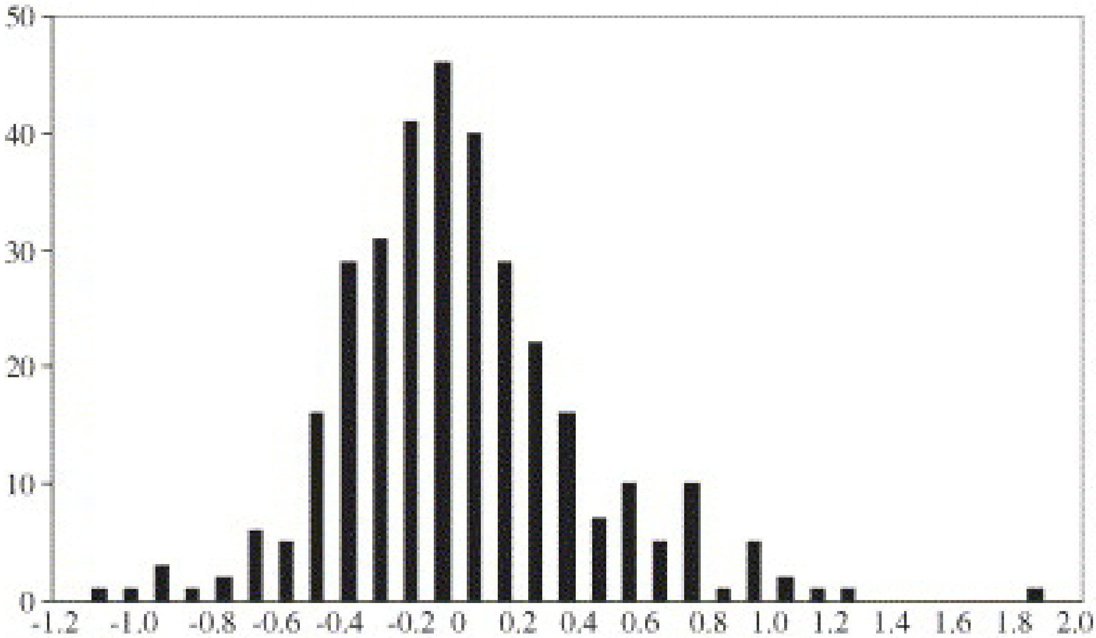
The drawbacks of the most commonly used hedge fund benchmarks in practice, the peer grouping, were highlighted in the literature by Fung and Hsieh (2004b). They suggest peer groups present challenges from survivorship bias. Not all hedge funds survive the test of time. They also include funds that are gated, so investment cannot be made. It also includes funds with a minimum investment amount, which is sometimes substantial, thereby marking precision weights difficult to achieve.

Despite the identified drawbacks of comparison to peer groups, much research has been done using such benchmarks. The returns of hedge fund time series against them was investigated by McCarthy and Spurgin (1998). They felt that the goal of benchmarking should be dual focused. Firstly, to be reflective of the asset class return and secondly to permit asset allocation between asset classes. One could add to these a third, namely their use in evaluating performance. These three roles are all addressed in separate strands of the literature, albeit in an indirect manner. Existing benchmarks fall short in satisfying these criteria in full.

The peer group critique by academia highlights that hedge fund benchmarks have to address the dual challenges presented by underlying and fund investability. Amenc, Goltz, and Lodh (2012) show that these factors result in marked heterogeneity between various competing hedge fund indices. They test the widely used commercial indices, drilling down into three main peer group indices, namely *Evaluation Associates Capital Market Indices, Hedge Fund Research Indices, and Credit Swiss First Boston/Tremont Indices*. They further investigate nine less widely used ones. They demonstrate manager style can be broken down into sub-indices. What this illustrates is that such benchmarks can be either (1) narrow or broad based, or (2) bottom up or top down.

On a similar tack, Denvir and Hutson (2006) investigate the skewness of a hedge fund return time series that manifests itself in the peer group data in the Fund of Hedge Fund universe. They find that underlying funds under-perform the hedge fund index on a risk-adjusted basis but did so with a negative skew. Their results are shown in figure 2.16. It is concluded, that a fund of peers should reflect the same characteristics as the underlying funds in aggregate, but this is not the case with peer group benchmarks.

Figure 2.16: Fund of Hedge Funds mean monthly excess returns



This figure shows the mean monthly excess returns of a hedge fund "fund of fund" relative to the hedge fund index. This diagram shows the negative skewed nature of hedge fund returns. Note the long bias in the distribution. Source: Denvir and Hutson (2006), Figure 1, page 12.

The literature on peer group benchmarks was enriched by Jagannathan, Malakhov, and Novikov (2010). They argue the limitation of a peer group approach is that risk is understated or misreported. Bollen and Pool (2009) find evidence of “temporarily overstated returns”, which further complicates benchmark attribution, if not making it impossible. He further finds that returns are incentivised in an asymmetric way by the performance fee. The later is only taken when returns actually out-perform the benchmark. It should be noted, in this respect, that peer group benchmarks are by definition net of fees. Mutual funds that compare to benchmarks are typically gross of fees.

The literature has investigated the many facets of hedge fund pricing. In doing this, academics treat hedge funds as a *de facto*. This is illustrated by Bussiere, Hoerova, and Klaus (2014). The literature also treats hedge funds as homogeneous in their returns, despite the presence of financial complexity in the form of derivatives or leverage. In practice, a multi asset benchmark is often applied. In the context of Modern Portfolio Theory, this can be considered a multi asset extrapolation of the single benchmark model.

Benchmark selection by academics was raised as an issue by Brooks and Kat (2002). In hedge fund performance studies, the benchmark used in evaluation has a direct impact on the outcome of the results. They suggest, therefore, that using traditional linear models of hedge funds and a mean-variance approach is compromised by benchmark misspecification. This is exaggerated when skewness, kurtosis and auto-correlation are taken into account. As such, the implication is that new benchmarks that address such issues may well be superior to existing ones.

The literature investigation into fund performance under non-normal or skewed returns in hedge funds is typified by Fung and Hsieh (2001a) as well as Agarwal and Naik (2004b). They investigate the attribution of returns by separating the beta risk premiums from the alpha that hedge fund managers generate and conclude that hedge fund strategies generate returns that are like option payouts. They conclude from this that "linear factor models using benchmark asset indices have difficulty explaining them."

2.15.1 Hedge funds and persistence of returns

The literature on persistence of returns against an index is one of the largest contributions to hedge funds. Carhart (1997) led the way in this line of investigation into persistence of skill. This is relevant because, as stated, performance measurement in hedge funds have relied on peer groupings as their benchmark. Indices based on hedge fund strategy's underlying exposures would be more appropriate. The issue is that the fee is incorporated in the net return and the net asset value itself is only calculated monthly. Many funds also have fund specific heterogeneous issues, such a a high minimum investment criteria, gates and lock-ups.

The literature's failure to identify a benchmarking approach for non-linear hedge fund returns suggests that further research is warranted. The closest approach would be that of identifying commonality in hedge fund returns. This was done by Busse, Brümmer, and Ihle (2012).

They find that the exposure to emerging market equities was the most important such factor. This probably reflects the widespread tilting of portfolios to the world's fastest growing regions. It is not clear if a hedge fund benchmark that incorporates a bias towards emerging markets would be appropriate as this could well be a temporal bias. That said, such a tilted benchmark would help with hedge fund performance attribution.

The literature is largely directed at performance and factor model attribution of returns rather than their construction method. An example of such an approach is given by Agarwal and Naik (2000). They investigated the time varying nature of returns to the underlying factors. An attribution approach, however, was conducted by Fung and Hsieh (2002b) who decomposed the returns of the underlying assets in the hedge funds themselves. Meanwhile, Mitchell and Pulvino (2001b) developed factors for risk arbitrage and Fung and Hsieh (2001a) trend mimicking strategies which could be used as benchmarks.

The literature on hedge fund of funds investigates the under-performance of such funds relative to their benchmark. An example of this is the research by Ang, Rhodes-Kropf, and Zhao (2008) who observe that the under-performance that they identify in this respect may be down to their being improperly benchmarked. The reasons, they speculate, could be because such funds incorporate two sets of fees, those of the underlying fund and those of the fund investing in them. Hedge fund of funds exhibit tracking error relative to the benchmark as a result of the aforementioned fees on fees. A better way, as they point out, is to benchmark against the underlying fund exposures. One way to benchmark against underlying exposure is to use synthetic replication. Research on this has been done by Gregoriou and Kooli (2011), a topic that will be returned to later in the thesis. They look into both the non-directional and directional hedge funds, using fund of funds as benchmarks. A fund of funds, they point out, is not replicable and therefore not an appropriate benchmark for academics.

A similar strand of research was pursued by Bhardwaj, Gorton, and Rouwenhorst (2014). They used a replication strategy to synthesise returns. They constructed benchmarks for commodity trading advisors, a sub-set of commodity focused hedge funds. This was built upon a trend mimicking strategy. Their benchmark portfolio was therefore made up of currency, commodity, and equity momentum factors. Their methodology is very dynamic and therefore hard to implement in practical use where replication is important.

Another strand of research perused the route of creating bespoke benchmarks. Kat and Brooks (2001), for example, develop a benchmark using dynamic trading strategies that involved futures contracts. This approach is also too complex for practical use. Having a derivative benchmark requires continuous updating and re-balancing. One way to overcome this is to use Static rather than dynamic re-balancing. The lack of trading is more cost effective.

In a sub-set of hedge fund strategies, fixed income arbitrage, there has also been some research into replication. This was conducted by Duarte, Jefferson, and Yu (2007). They used swap spreads, the yield curve, mortgages, volatility, and capital structure to replicate the factors driving hedge fund returns. In a similar fashion, fixed income Agarwal and Naik (2004b) analysed convertible arbitrage funds. They used data convertible bonds and underlying stocks constructing a simulated buy and hedge factor benchmark.

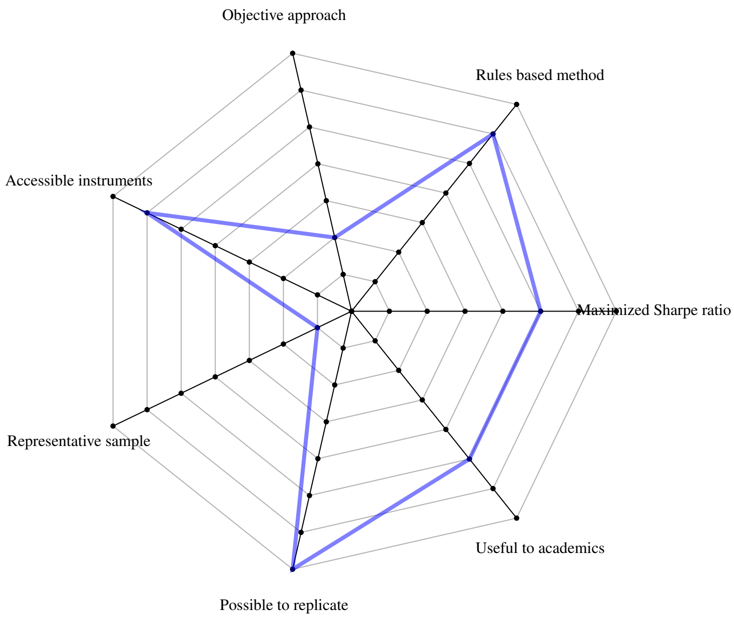
In summary, the majority of the literature on hedge funds focuses on building suitable benchmarks from constituent portfolios. A smaller subset focuses on ratios. Where the benchmark is broken down into risk factors, such as in the literature on style difference, the focus is on explaining the very broad cross-sectional and time-series variation of returns.

2.16 Real Estate benchmarks

The literature on the application of real estate benchmarks is extensive. That said, it is weaker on the construction method. The investigation of it as a separate institutional asset class is fairly recent. A paper that specifically delves into the design and construction of real estate benchmarks, is produced by Geltner (2001). He argues that the investment universe and objectives of real estate investors are too diverse for a single benchmark. This makes sense in a world where frequent transactions are recorded but it still leaves the question about the operation of such indices in the absence of a healthy market. They also argue that academics are better served by transaction rather than appraisal based indices. The variables he identifies as relevant are the reporting frequency, the revaluation data, the universe of underlying properties and the method used in the underlying valuation. These are summarised in the *Real Estate Investment Trust Index Spiderweb* in figure 2.17

Many of the concepts in traditional finance were slow to migrate to research into real estate

Figure 2.17: Real Estate Investment Trust Index Spiderweb.



Real Estate Investment Trust Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. Dimension four a reflecting representative sample only score 1. The strongest dimension is five, the ability to replicate. This is because the indices use REIT's that are listed on stock exchanges.

pricing. Mei and Lee (1994) were two of the first to look into whether there is a real estates factor (which could be applied as a benchmark). They find that there is and argued that this means real estate should be treated as an asset class.

As is explained, in the literature Real Estate Investment Trusts (REIT's) are the primary liquid investment vehicle for collective investment in this asset class. Most of the literature on performance in the public markets focuses on this as it is easy to measure. One such example is Kuhle (1987) who investigated portfolio diversification and return benefits of real estate investment trusts. This started a strand of research based on the time series of investment trusts and mutual funds investing in real estate, rather than whether real estate exhibit factor returns in their own right.

In both practice and the literature, there is a very real need to benchmark real estate investment. That said, there are a number of issues with pricing data. Price movements crystalize on a property by property basis, essentially each time a sale is made. This can be overcome by estimating the time series by regression, as explained by Case and Quigley (1991). This approach was first proposed by Bailey, Muth, and Nourse (1963a). As a result, many real estate benchmarks are not based on actual prices. Despite this, the cash-on-cash return on real estate is calculated thus:

$$\text{cash-on-cash return} = \frac{\text{annual before-tax cash flow}}{\text{total cash invested}} \tag{2.15}$$

The other approach to valuation in the literature is appraisal. This is an estimate. The valuation appraisal issues were investigated by Geltner (1993). There are several problems in measuring real estate returns based on appraisal data.The first of these is that not all properties are valued at regular intervals. This is due to the expense associated with producing these. The valuation appraisals are also very subjective as a consequence of the infrequent transactions in them.

Some of the identified issues were subsequently addressed by Geltner, MacGregor, and Schwann (2003). They point out that traditional appraisals suffer from lagging and smoothing biases. In some instances, reported real estate valuations in key indices do not change for up to six months. As a result, volatility is downward biased. They concluded that real estate benchmarks are too subjective in nature.

In the real estate literature, the impact of leverage, serial correlation and smoothing also impact benchmark time series. The literature on property returns has therefore tended to focus on tests for normality and the way that returns are distributed. A paper that focused on normality of returns was produced by Lu and Mei (1999). They reject normality based on a sample of emerging markets using the Anderson Darling test. Newell and MacFarlane (1996), similarly, investigated the smoothing of real estate pricing. They also test whether the time series of property unit trusts followed a chaotic pattern, but found that they did not.

A good review of the various appraisal methods can be found in Pagourtzi et al. (2003). They identify seven traditional methods, namely comparisons, investment, profit, development, cost, multiple regression and stepwise regression. They further identify five advanced valuation methods, namely artificial neural networks, hedonic neural method, spatial analysis, fuzzy logic and autoregressive integrated moving averages. All these valuation methods, when combined in an index, generate a return time series.

Property returns have been identified in the literature as non-linear. This is partly due to the liquidity and infrequent valuations. Hoag (1980) developed a fundamental approach to real estate indices. Meanwhile, Bailey, Muth, and Nourse (1963a) overcame the fact that properties are sold at different periods in time by using a regression approach to real estate index construction.

Fisher, Fisher, Geltner, and Webb (1994) compared a variety of different methods including benchmarks that apply un-smoothing approaches, such as the appraisal based Russell-NCREIF Index. They suggest using trace average ex-post transaction prices.

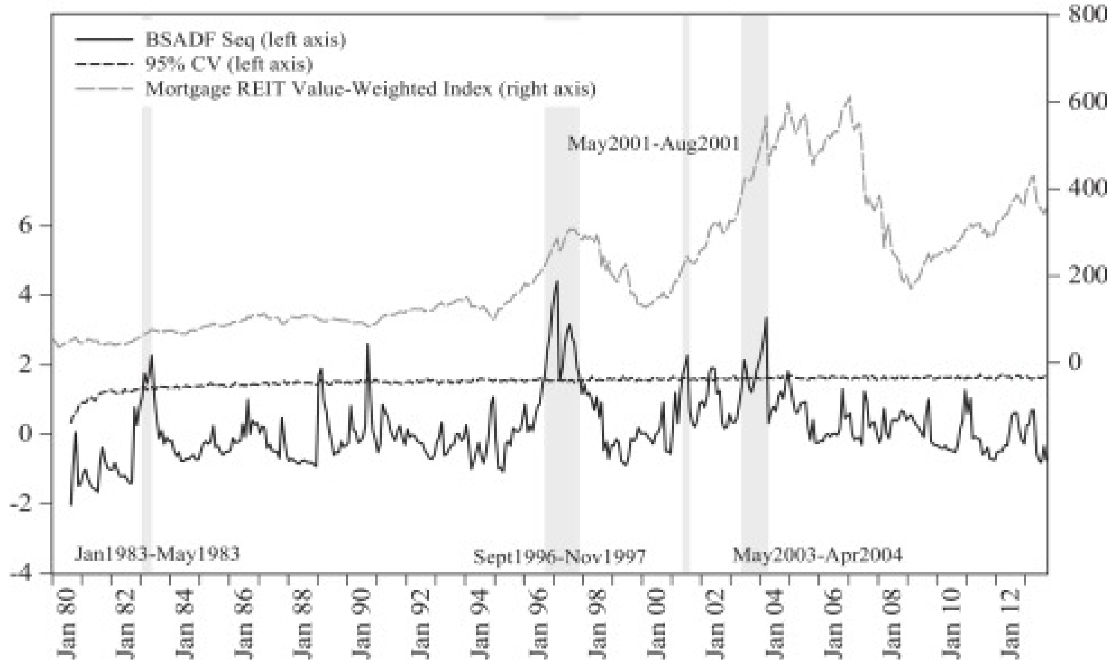
2.16.1 Real Estate and persistence of returns

There is a sub-branch of the real estate literature that focuses on performance persistence relative to a benchmark. Escobari and Jafarinejad (2016) investigate real estate bubbles using the REIT market. They clearly identify four statistically significant bubble periods, as can be seem in Figure 2.18.

Clearly, the real estate market is impacted by interest rates. Most studies, however view the interest rate as commonly treated as exogenous. Chang (2011) however investigated the nonlinear effects of expected and unexpected components of monetary policy on them. Unsurprisingly, he finds that tightening interest rate policy keeps returns down.

In summary, the literature in finance has tended to focus on REIT benchmarks. This is because of their liquidity. The literature, however, finds that this is not reflective of the underlying physical assets and that it is more highly correlated to the equity market than the property market.

Figure 2.18: Pricing in the Real Mortgage REIT Value-Weighted Index



This figure shows the various bubbles that REIT's have experienced in the United States. The real Mortgage REIT index is obtained by dividing the Mortgage REIT monthly value-weighted index by the Consumer Price Index. These periods are distinct from stockmarket related excesses, confirming that REIT's are a sub-grouping in their own right. Source: Escobari and Jafarinejad (2016), Figure 2, page 227.

2.17 Other alternative asset class benchmarks

In addition to the three main sub-groups of alternative asset classes, there are a few others. These include quite different assets such as weather derivatives Lennep et al. (2004). Although fixed income is not part of this investigation, the same challenges that this traditional asset class presents to index constructors also impacts hedge funds. As a result, some of these are presented from the literature. The first word of caution is on terminology. As a result of liquidity being concentrated it is difficult to construct reliable indices and as such the industry often prefers a single asset to take the role as benchmark. The attributes of such a single asset benchmark bond are described by Dunne, Moore, and Portes (2007). These are those with the security that has the lowest yield at a given maturity. There are, however, many multi-asset fixed income benchmarks.

Art is another example. It is investigated by Campbell (2010). He compares and documents the uncorrelated nature of benchmark returns from investment in the art market as relates to traditional asset classes. It is clear from his results that art, as in all alternative asset classes, requires a unique benchmark construction method. In this way, specification error can be minimized. The same applies to all alternative assets. Art was also investigated by Jianping and Moses (2002)

Infrastructure is increasingly being viewed as an asset class in the literature. Inderst (2011) sets out the reasons. He suggests such investments have useful characteristics for pension and insurance institutions. These include low correlation long duration predictable income relative to other assets and a favourable default and recovery profile. He argues, however, that financial theory is not enough to support the concept of infrastructure as a separate asset class. The assets themselves are very heterogeneous, and empirical evidence suggests that they should be treated as a sub-class

2.17.1 High yield debt

Within the literature, the multi-asset fixed income Benchmarks are often typically separated into credit buckets. In this way, similar securities are grouped together by issuer or by credit risk. This is explained by Barnhill and Maxwell (2002). Credit benchmarks can be further grouped by other features, such as whether they are fixed versus floating coupon rates, whether they are callable prior to maturity, and/or whether they are linked to inflation. In addition to companies, governments and government entities are also issuers of bonds and most issuers have several bond issues of different maturity, seniority, and structure. Bond issues tend to have less active secondary markets compared with equities as many investors like to hold them to maturity. Trades are also done through telephone markets, not through a centralized exchange.

The lack of pricing data makes bond benchmarks problematic. The literature suggests that one way to overcome this is to estimate them. In this respect, the values of bond prices can sometimes be estimated using the inferred current market value of similar bonds. When standard deviations and correlations are used the results can potentially be biased downward. Bond prices can also be impacted by credit spread changes. An explanation can be found in Collin-Dufresn and Goldstein (2001).

Fixed income pricing is impacted by the stale, late or unrecorded data on trades. This is particularly true of corporate bonds as explained by Gruber, Agrawal, and Mann (2004). As mentioned, the lack of liquidity is one of the biggest issues with alternative assets. Liquidity and lack of price data is also a problem for fixed income. The typical criteria used to construct a benchmark include country, credit risk, liquidity, maturity, currency, and sector classification.

Bond indices risk characteristics rarely match an investors desired risk profile. This is because a measured bond index faces the problem, according to Granito (1987), that it is a randomly changing reflection of the true market. All bonds have different credit ratings, duration, and prepayment risk reflecting the bond issuers' preferences. Meanwhile, outstanding bonds are continually changing maturities and new issuance is ongoing whereas bond indices are usually recreated monthly.

Siegel (2003) referred to another problem. This arises because capitalization-weighted bond indices give more weight to issuers that borrowers. The poor credits in an index may be more likely to be downgraded in the future and experience lower returns. With global bond indices, the countries that go the most into debt have the most weight. This means a portfolio tracking such an index will have exposure to the biggest borrowers, typically the worst credits. One solution to the problem is to use GDP weighted indices, proposed by Borensztein and Mauro (2004). That said, this approach can be critiqued as such weighting schemes may be less liquid and overly subjective.

2.17.2 Private Equity

Private equity is considered an alternative asset class in the literature by virtue of its lack of liquidity and is also particularly difficult to benchmark. That said, as it is traditional equity, there is much evidence that it can be effectively valued using Capital Asset Pricing Models and applying a risk premium. As a result, benchmarking private equity is either a case of creating a hurdle rate of return or a peer group of like investments, typically grouped by investment year as illustrated in Lin and Phalippou (2007).

One of the interesting findings in the literature about private equity relates to performance persistence. Kaplan and Schoar (2005) document substantial persistence. They identify that General Partners whose funds outperformed in their first fund are more likely to outperform in the next. These findings are clearly different from the literature on traditional asset classes such as equity funds.

The most widely cited listed proxy is the *Standard and Poor's Listed Private Equity Index*. This uses the modified market capitalization weighting approach combined with a divisor. As a listed proxy this is not necessarily representative of the achieved returns of General Partners. As a result of this, in addition to listed private equity Indices, Cumming, Helge Hass, and Schweizer (2013a) opined that institutional investors use an additional two key methods to constructing PE benchmarks:

- transaction-based PE indices.
- appraisal value-based PE indices.

These all indirectly link to the Internal Rate of Return (IRR) achieved on the portfolio of investments.

The most common type of return measurement, and hence *defacto* benchmark, in the private equity literature is the *Internal Rate of Return* (IRR). This widely used percentage is effectively the interest rate at which the *Net Present Value* (NPV) of positive and negative cash flows of an investment is equal to zero. The discounted cash flows that form the basis of this are expressed as follows.

$$\text{NPV} = CF + \frac{CF}{(1+r)^1} + \frac{CF}{(1+r)^2} + \frac{CF}{(1+r)^3} = 0. \tag{2.16}$$

This can be adjusted for comparison between funds in the following way.

$$AV_{PME} = \sum_{s}^{T} C_s \times \frac{I_T}{I_s} \tag{2.17}$$

Where :

- NPV = Net Present Value
- C_s = the cash flow from the investment at time *s*, positive for a contribution, negative for a distribution.
- I_s = the value of the index at date s.

2.18 Conclusion

In conclusion, there is a wealth of research that uses indices but there are also clear gaps in the literature. The largest of these is what constitutes an appropriate benchmark for alternative assets, the focus of this thesis. There is a consensus that index construction approach is important and that academics should have an index that is constructed out of a representative sample that reflects the asset class. As a result, there is clear merit in investigating what constitutes an appropriate index for alternative assets.

The literature clearly demonstrates that a benchmark should be appropriate for the asset class. In this respect, an appropriate benchmark should allow for accurate measurement of security, asset allocation and interaction effects. To do this it has to be constructed with a clear set of rules and be investable. It should be able to be replicated for the purpose of indexation, or be able to be sufficiently representative to facilitate active management. The use of proxies and peer groupings as benchmarks for alternative assets was identified in the literature as not being optimal. This leaves room for a method that addresses these, such as PCA indices.

In order to construct an appropriate index for these, and indeed other indices, the literature showed that the creation of an index involves three steps. The first step is the selection of the asset instruments. The second step is weighting items as they arise in the index. The final step involves the mathematical method to calculate the index. The next chapter details the theoretical backdrop for such methods.

The literature further highlighted that (1) indices facilitate attribution as an important function of their appropriateness and that (2) creating an index is essentially a sampling exercise. In this respect, the literature suggests there are issues in alternative assets benchmarking that need to be addressed. PCA is proposed as a way to do this and chapter four sets out the theoretical framework for the method.

It was also shown that the literature on appropriateness draws heavily on the market portfolio, as derived from Modern Portfolio Theory. In this respect, the literature indicates that commonly used capitalization weighted indices may not be optimal for alternative asset classes. Other benchmark solutions do not allow for attribution. As a result, the literature suggests that there is room for PCA derived indices to fill that gap. As a multivariate technique, this approach has been documented in many texts. That said, it has largely been used for interpretation in the finance literature.

The literature suggests PCA can be extended from an examination of a sub-grouping of alternative assets in n-dimensional space to a determination of objective weights for the construction of indices. It is suggested, based on the literature, that the cost of construction and re-balancing of such indices would be lower than traditional indices.

The chapter framed the research question by detailing the literature on the use of benchmarks and how indices are addressed by academics. Its contribution is as a review defining the body of knowledge behind index usage. The chapter posed eight questions to extract insights from the literature.

- Question one asked whether benchmarks are fit for purpose? It was concluded that indices have a good grounding in theory but can be improved.
- Question two asked how benchmarks are used for identifying skill. It was concluded that the CAPM provides a strong support for mean variance indices.
- Question three asked if it is possible to persistently out-perform a benchmark. The answer is that it is extremely difficult. This suggests that if a contender index is found to have superior risk adjusted characteristics to an incumbent, then it is likely a be a better index.
- Question four asked how liquidity affects benchmark selection?. The literature shows that liquidity commands a premium and that illiquid instruments make for poor index constituents.
- Question five asked how leverage distorts benchmark returns. The literature suggests that it skews the returns. As such, indices that incorporate leverage will not have the normal distribution that the CAPM assumes.
- Question six asked what the consequences of benchmark misspecification are. The literature suggests that this results in unreliable research and attribution outcomes. This provides a justification for benchmark enhancement.

- Question seven asked how benchmark superiority is determined. The answer ties back to chapter one and suggests a good way to do this is to use the GRS test. The various tests to measure benchmarks were presented in order to demonstrate how remedy of these deficiencies can be achieved using both quantitative and subjective approaches.

- Question eight asked what benchmarking challenges are faced in alternative asset classes. The literature suggests that addressing liquidity and proxies are the priorities.

It is clear from the literature that investment returns and risk, are all effected by benchmark choices. The Index Spiderwebs used throughout the chapter show that there are deficiencies in current alternative asset indices. They illustrate that:

- futures based commodity benchmarks were shown to be reliant on production based weighting methodology. This thesis addresses this through a focus on determining the weights by the co-movement of sub-groups through their principal components

- hedge fund peer group benchmarks were shown to have replicability issues. They were also critiqued due to their measurement of net of fees returns. This thesis addresses replicability through a focus on the underlying assets. It addresses the net of fee issue by a focus on gross returns.

- real estate benchmarks were shown to have investability issues. They were critiqued due to the illiquid nature of the underlying assets. This thesis addresses investibility through a focus on Real Estate Investment Trusts, a listed alternative to direct property investment.

In summing up, the literature suggests that the two key attributes of a good benchmark for alternative assets are (1) clear rules and (2) defined structure. The detail behind these attributes are addressed in the next chapter which delivers a taxonomy of benchmarking techniques. This includes a multitude of different indices all of which can be used to make academic inferences, all with their own idiosyncrasies. The lack of clarity and consensus on what is most appropriate suggests that there is a need for greater clarity. There is also a need for a taxonomy of index construction technique. Both are provided in the next chapter.

Chapter 3

A taxonomy of benchmarking techniques

This essay provides a theoretical framework for index construction. It also offers a critique of the common approaches and challenges faced in benchmarking investment assets. It provides an analysis of the nature of alternative asset returns and how these may present issues for the construction of indices designed to measure and attribute the performance of alternative assets. It investigates the intricacies of construction method. It shows how such index benchmarks have been adopted and adapted. This critique highlights that benchmark and index construction is a well developed science. Despite this, it raises the question of whether market capitalization weightings are appropriate for alternative asset classes. Its contribution is in illustrating the extent of the definition and sophistication of benchmarking technique and where refinements can be made to the theory. It concludes that benchmarking choices are largely down to subjective trade-offs based on constituents, investibility, breadth, re-balancing, and mathematical approach. These all support the overall conclusion that a more bespoke approach needs to be taken in the benchmarking of alternative assets.

3.1 Introduction

This chapter introduces the theory behind benchmarks and index construction. It details the methodological framework extracted from the literature that is used to construct composite indices and by extension PCA indices. The various approaches are natural extensions of the use of indices in academic research. It shows that choices and trade offs have to be made. The review of the literature, in the last chapter, clearly showed that there is no common benchmark used by finance academics but that there is agreed method. That said, where indices are used, there is a multitude of different underlying methodologies. This chapter therefore lays the ground for the PCA index approach which adds a construction method to what currently exists and therefore addresses a gap identified in the literature.

Index construction theory was not well captured in the literature on existing alternative asset indices. That said the last chapter highlighted that the creation of an index involves three steps, the selection of the asset instruments, the weighting decision and the mathematical method to calculate the index. This chapter investigates these steps from the perspective of alternative assets and critiques the various different approaches to the sampling, weighting and mathematical method. This chapter further provides insight into how benchmarks need to take account of proxies that are often illiquid and/or a derivative of the asset and how PCA can be used with existing methods to assign coefficients, or weights, to those observed variables. Its contribution is in (1) explaining the theoretical background to benchmarking alternative assets and (2) in identifying the limitations therein.

The literature shows that academics use bespoke, style, factor and ratio benchmarks in addition to broad market indices. These use benchmarking techniques adopted from professional asset management. As identified in the introductory chapter, not all are appropriate for alternative assets. The commercial indices that are considered appropriate deliver quite different return time series for what purport to be the same universe of assets. The differences arise as a result of either the calculation method or the sampling method.

Regardless of whether it is for academics or practitioners, the starting point of all benchmark techniques is the universe of assets. This is the same for traditional or alternative assets, As such, a sampling approach has to be applied to the identified assets in order to narrow down that universe. Only then is it possible to apply the weightings and decide on which mathematical technique should be used to combine the returns of those assets into a meaningful time series. As a prelude to detailing how this can be done using PCA, this chapter therefore investigates the different ways that this can be achieved. This provides the reader with a better understanding of how to address alternative asset benchmarks.

It has been shown in the preceding chapters that the more sophisticated benchmarking techniques incorporate risk and allow for instrument level attribution and these are based on equity market risk factors. Practitioners, however, require this feature to professionally manage alternative asset funds. Passive fund managers require indices that can be replicated. As such, there is a clear need for clarity and risk decomposition in alternative assets that this chapter will address.

The focus on asset pricing and performance has resulted in the widespread use of existing commercial indices in academia. This has been regardless of whether they fulfill the appropriateness criteria of Bailey, Richards, and Tierney (1990). Very few papers justify their choice of index in a detailed way. This is particularly the case for non-traditional assets. The contribution of this chapter is therefore to provide a greater clarity on such methodological choices.

Benchmarks and their practical use in the investment profession

The reason it is important to understand and critique benchmark techniques is that, as a practical tool, a benchmark is a measurement proxy that is central to professional fund management. In this respect, professionals differentiate between a benchmark and an index, as in this context the terms are often confused. An index or a ratio are both benchmarks that can be used in evaluation by both academics and practitioners. They prove useful when presented as a time series. Using a ratio as a benchmark, as explained in the previous chapter, they can link the returns to that of the market risk factor. PCA can also do this and in an index form facilitates numerical understanding and allows for possible multivariate analysis and instrument level attribution.

Alternative asset managers are focused on the granularity of performance and as such benchmarks that are not portfolios of securities, funds and/or instruments are not useful. This is why practitioners prefer indices over other sorts of benchmark. As previously mentioned, the simplest of benchmarks is a hurdle rate but this does not provide the aforementioned precision. They are also not useful as a tool for the professional management of funds. In contrast, indices are often interlinked with the investment process, as demonstrated by Brinson, Hood, and Beebower (1995). This is because they can, as Fama and French (2010) point out, differentiate between skill and luck. They can provide the basis for portfolio construction, be it active or passive in nature.

According to Conover, Broby, and Cariño (2013) an index is a benchmark with a very specific construction method. It has to be created in a step-wise fashion to reflect the performance of an asset class or a sub-segment of such a traditional asset class or an alternative asset class. At its broadest level, that of the asset class, it can be used to measure policy returns. At its granular level, that of the instrument level, it is a portfolio of investments created with a set of predefined rules. These rules form the core of its construction method. A benchmark, to be clear, need not go into the same level of detail.

A benchmark can be as simple as an investment objective stated in a quantitative way. It is therefore the case that an index is a benchmark, but not all benchmarks are indices. As a result of this nomenclature, and because benchmarks are used for investment measurement, the quantitative method should be precise. In not taking into account the different nature of the return distributions, existing benchmarks for alternative assets do not meet this criteria. The weighting of instruments in an index impacts the way the returns are delivered. In PCA indices, a weight may be considered as a coefficient of an instrument, exhibiting its importance relative to the rest of the asset class.

The literature showed that in the context of alternative assets, it is only possible to determine skill if the investment universe is defined clearly. In practice, commodity, hedge funds and real estate indices have wildly different constituents and have a large return difference between each other, let alone a fund mangers performance.

It is not uncommon for fund managers that prefer absolute returns to suggest that benchmarks are a distraction. This does not mean, however, that benchmarking techniques are not important to both active and passive managers. Passive investment strategies require precision because they replicate an index, a process termed indexation. This approach is designed to achieve the investment returns of the market proxy by exposure to its underlying constituents.

Active investment strategies also need precision in index method because they aim to deliver investment returns in excess of the benchmark, often on a risk adjusted basis. As a result, active asset managers need to know what the constituents of that benchmark are.[1]

The literature review showed that many managers under-perform their chosen benchmark. Indeed, Fama and French (2010) point out that active asset managers have *zero* α relative to passive benchmark alternatives. Academics have flagged the impact of performance of costs and market impact in this respect. These drags on performance result in marked tracking error, as explained by Frino, Gallagher, and Oetomo (2005). This concern is magnifies if there is benchmark mis-specification.

It is clear therefore, in the investment profession, that the more quantitative the fund management process, the more granular the benchmark needs to be. This is also the case for alternative assets. There are many strategies that a manager can apply to such assets and as such there are a variety of different benchmarks to choose from with which to measure investment performance. There are also numerous ways in which to present return, illustrated next.

3.1.1 The need for indices to be replicated

The use of indices in passive asset management is widespread. As a result, there is a need for them to be able to be reproduced as a portfolio by a fund manager wishing to track their return. In constructing indices in a predefined and quantifiable way, it is possible to do this. This is not the case with using, for example, the consumer price index as a benchmark for anything other than inflation. In this way, indices can be used as passive alternatives to active fund investment strategies focused on alternative assets. This is important because a passive index based approach to investment has a number of benefits. These include, a lower maintenance cost, less transaction costs, representative returns and a strong theoretical grounding. The problem with other forms of benchmarks is that they cannot be replicated at the instrument level.

Supporting the need for clarity, in a *Guide to Index Construction*, Broby (2007) noted that

[1]The returns in excess of an index are referred to as a manager's active returns, and the variability of the active returns is termed Alpha α. The later is a proxy for active risk. The difference between these two approaches, the passive versus active divide, reflects whether a portfolio is managed to match the return of a benchmark or to exceed it.

there is no *de facto* statistical method for index construction but there are *de jure* approaches. The granularity of such approaches is better suited to the measurement of the rates of return of investment assets and is especially important for index tracking strategies. An index facilitates a fund manager measuring the metrics of the benchmark with the rest of his portfolio and/or other portfolios. This is useful from the perspective of the policy decisions that are made when investing in alternative assets. Most importantly, an index allows a fund manager to measure the out or under performance, the correlation, and the volatility. These characteristics differ between asset classes and quite markedly in alternative assets.

A passive investor has to have a strategy to minimize tracking error as characteristics change due to constituent replacements, roll-over or removal. Unlike in traditional asset classes, the ability to hedge exposure is not always possible in alternatives, resulting in higher tracking error for such strategies.

Meanwhile, an index can be flawed and/or sub-optimal. In other words, not mean variance optimal. It can also be miss-aligned to the underlying universe of assets. Whilst the sampling method is the biggest reason for this, the various calibration methods detailed in this chapter also play a part. They are critiqued from the perspective that method used by practitioners, designed essentially for equity investment, are not suited for all asset classes. That said, it is argued they must be constructed using rules that make them desirable to users. Those users, however, vary, but the key variables and sub-components are all shaped by investment professionals. Their inputs, in turn, are determined by investment preferences.

3.1.2 The need to reflect investment preferences

Benchmarking techniques should reflect the investment preferences of their users. In this respect, preferences matter in so far as benchmark appropriateness is concerned. They allow individuals and institutions to understand their risk appetite.

It is assumed that an investor, in selecting a benchmark, would want to maximize utility relative to the available asset classes and the investment universe. Stigler (1950) illustrated that it is rational for an investor to prefer an investment universe that optimizes utility from which to select his/her benchmark. As a result, in addition to diversification, the benchmark selection problem becomes one of choosing the collection of assets (be they traditional or alternative) that meets this criteria. This is the utility approach to benchmarking.

The utility approach to benchmarking preference is interesting from the perspective of alternative assets. Investors that chose to gain exposure to them do so with different objectives in mind from those who select traditional assets. Investors in commodities might, for example, be seeking an inflation hedge; investors in hedge funds may be seeking absolute returns; investors in real estate may be seeking long term capital appreciation.

The concept of utility in investment and hence benchmarking was illustrated by Sharpe (1964). He linked it to what he termed the *Investor Preference Function*. This sits on the Securities Market Line proposed by Tobin (1958). It describes a function as a mathematical relationship that essentially assigns a numerical value to each possible theoretical benchmark. The formula for this benchmark utility function is as follows:

$$U = U(Ew, \sigma w) \tag{3.1}$$

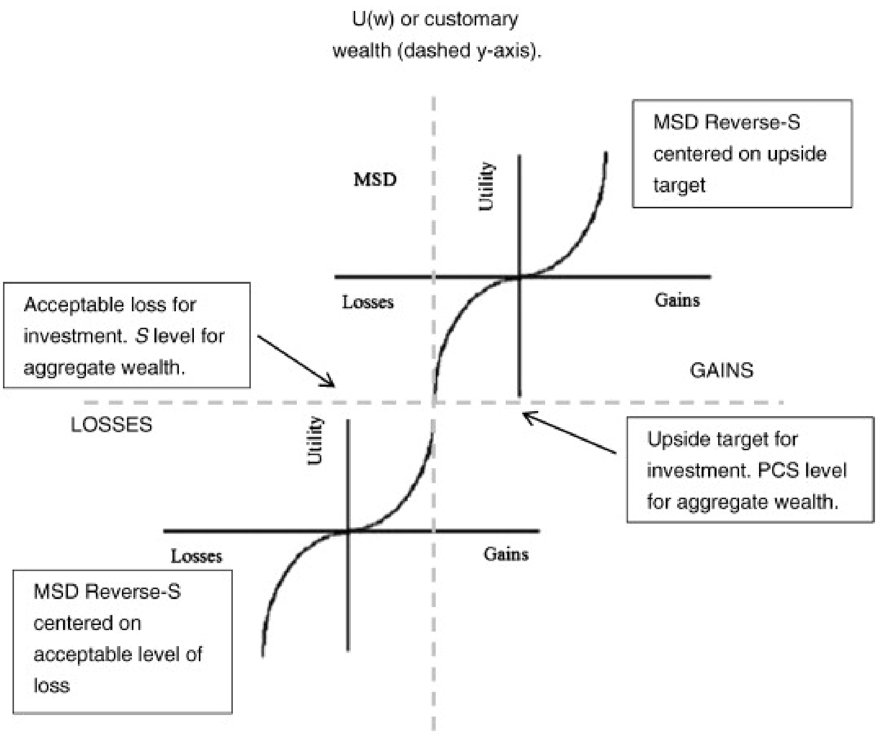
where:

- U = function that assigns a number or utility to any given set of values.
- Ew = expected future wealth
- σ = the standard deviation of the probabilistic divergence from that wealth.

The utility function is relevant to benchmark inquiry because, if it holds. the relationship between risk and return likewise holds. As a result, it is now widely accepted that academically useful benchmarks should measure risk as well as return. In this way, standard deviation is an accepted proxy for this. It follows that an optimal benchmark would be found by relating it to individual investor utility functions and how they relate to the efficient frontier. This relationship between utility and the efficient set applies to individual instruments as well as alternative asset classes. It is the building block of the market proxy concept, as demonstrated in Broby (2011) and, as such, relevant to benchmark selection.

Viole and Nawrocki (2013) extended utility functions to the field of benchmarking by linking the concept to investor risk tolerance. This is done through the use of lower partial moments. In this way it is possible to describe the utility of losses and the resultant upper partial moments that explain the appetite for gains or losses. The resulting Markowitz stochastic dominance wealth utility function is shown in the figure 3.1. This depicts acceptable downside loss and upside preferences and how they interact with wealth utility. Investors in alternative assets exhibit different upside and downside risk preferences. A benchmark that addresses these preferences should reflect this.

Figure 3.1: Markowitz stochastic dominance - The wealth utility function as a benchmark



The figure shows the total Markowitz stochastic dominance wealth function. The benchmark is the personal consumption satiation and the lower benchmark is safety first level of subsistence. Using this, mathematical statements about wealth preference and risk aversion can be used to make optimal decision rules for selecting investments. Source: Viole and Nawrocki (2013), Figure 2, page 193.

With traditional asset classes, investment benchmarks are commonly applied to reflect this utility preference approach to risk. Because the time series of alternative asset classes have a high degree of non-linearity due to the time value of money, leverage and the use of derivatives, the nature of the benefit is influenced by the different risk. By way of critique, it can be observed that these distinctive characteristics are not easily translated to alternative asset benchmarks. The potential usage, however, is relevant. Investors in alternative assets have different risk preferences than those who invest in traditional assets.

3.1.3 Which approach to use

The many benchmarking and index techniques require an impartial test to determine which one to choose. In this respect, Eichhorn (1976) proposed a number of tests that he suggested should be applied to index numbers. Adapting these to alternative assets, one should therefore take account of the:

1. proportionality as to investment objectives. This is required so that the index correctly reflects the plan sponsors reason for investing in an asset class.

2. proportionality as to cross-section of returns. This is required so that the index distribution is similar to the desired outcome.

3. the instrument exposure. This is required as a function of how exposure is gained to an asset class. For example, using REIT's as a proxy for exposure to real estate.

4. the constituents. This is required as a sample must be taken. Not all constituents can be used to make the index.

5. the entry and withdrawal of index constituents. The index must reflect changes in the universe. For example, futures contracts rolling over, hedge funds starting and closing, or new REIT's being floated on the market.

6. an understanding of the changing base. The ratio between the index should be unaffected by changing the base year.

7. the changing unit of measurement. This is where the ratio between the various alternative asset indices should not be affected by changing any common measure (eg: security instrument splits)

These points are taken into consideration in the development of the PCA index described in detail in chapter four,

3.2 Investment performance terminology

As stated, the primary use of benchmarks is measuring investment performance and in this respect one has to understand that there are different terminologies and approaches used to describe investment performance. This includes the concepts of *rate of return, expected return* and *active returns*.

Indices are constructed after observation and utilizing sampling through statistical method. They are mathematical formula that are more sophisticated than a hurdle rate. As such, it is not just the terminology of the sampling method that can vary but also the terminology of the calculation method.

In a mean variance index construction context, expected returns are also subject to what is termed *estimate uncertainty*, a statistical perspective addressed by the likes of Jobson and Korkie (1980). This is considered a major weakness in the theory behind optimal portfolios as the market proxy may well not be observable. Estimate uncertainty is present in all the index construction methods and would only be removed if all available assets were part of the benchmark in weights that reflects their size.

In summary, the aim of any benchmark is to deliver a time series of the rate of returns. Indeed, a benchmark should only be critiqued in the context of an understanding of such returns. The following sections therefore explain the different types of returns.

3.2.1 Rates of return

Benchmark rates of return provide a cross-sectional time series. This consists of the constituents of an asset class as recorded over time. A great deal of research since Fama and French (1992) has focused on the variables that explain such cross sectional returns. There has been less academic attention on the nature of the calculation. This is an omission as the different return methods can be quite marked.

As index construction is based on empirical method, the pricing data of such returns are said to be one way and indexed by time. In this respect, time is algebraically stated as t and returns as R. This is collectively referred to as a *rate of return* Rt, which in turn is broken down over time as follows:

$$R_t = B_t = P_t - P(t-1)/P(t-1)* \tag{3.2}$$

Where:

- R_t= Rate of total return (*inclusive dividends)
- B_t= Benchmark total return (*inclusive dividends)
- $P(t) - P(t-1)$= The return from t-1 to t.

Benchmark rates of return on a price of an alternative asset class or individual asset class are typically presented over daily, weekly, monthly or annual time horizons. The choice of which, in academic terms, is important when conducting statistical analysis. With alternative assets, the challenge in calculation is the frequency of the instrument pricing. Hedge fund pricing, for example, is often monthly. Real estate pricing is less frequent, dependent on actual transactions.

Rates of return can also be presented as multiple time series or as a fixed time horizon t. This is why the rate of return in such a case is presented as above. Benchmark returns are, however, typically stated as a set of prices in cross section, as explained by Mundlak (1978). They are stated thus , $I, 2, \ldots, m$. Cross-section data is thus captured at a single point in time rather than related periods.

As far as continuous benchmark time series go, rates of return can be stated in an *inter-period*, *inter-day* and/or *intra-day* fashion. A continuous benchmark is the standard and indeed simplest form. it is stated as follows:

$$\bar{R}_t annual = (1 + \bar{R} periodic)^m - 1 \tag{3.3}$$

Where:

- $R_t annual$ = Mean annual rate of return
- $Rperiodic$ = Mean period return

This equation gives a number for the rate of return which can be compared with a benchmark. That said, aside from the approach taken by Grinblatt and Titman (1993) who base their approach on a sum of time series co-variances, it is widely accepted that investment benchmarks require more than this to be useful. This is where indices prove a better alternative to simple hurdle rates, remembering that there is a distinction between a benchmark and an index. The index rate of return, in this context, is continuous which practitioners consider superior to a static hurdle rate.

The advantages of a time series from a benchmark perspective is that it is additive. That is, it can be split up into its constituent parts, such as security, sector, and country. Alternative asset classes can also be decomposed, as explained by Meric, Ratner, and Meric (2008). For hedge fund benchmarks, co-movement can be divided into funds, strategy and jurisdiction. In the case of commodity benchmarks, it can be broken down into commodity type, futures expiry, and market traded. In the case of real estate benchmarks, it can be broken down into property type, usage, and location.

As investment benchmarks tend to be compared on time periods of equal, rather than continuous lengths, rates of return and benchmark returns are typically expressed in the following way:

$$\bar{R} = \frac{1}{n}\sum_{i=1}^{n} r_i = \frac{1}{n}(r_1 + \cdots + r_n) \tag{3.4}$$

Where:

- $\bar{R}$ = Mean rate of return
- n = The number of observations in equally spaced periods

In finance theory, as explained by Kaplan and Ruback (1995), the total rate of return of a traditional asset is equal to the discounted cash flow that an investment generates over its life. As alternative investments have a discrete rate of return, the same definition does not apply.

In summary, in critiquing current benchmarking technique, it is necessary to introduce the concept of returns relative to a benchmark. This ties the theory with practical usage.

3.2.2 Relative returns

An alternative asset benchmark should provide a reference point from which relative performance can be evaluated. Once a benchmark is constructed, be it a hurdle rate, index or ratio, it should also be possible to decompose its returns and state the time series in a relative fashion. Bailey (1992c) used algebra to demonstrate how this is done. He suggests one first identify a portfolio to be measured, which in this instance is a portfolio of alternative assets. Then one has to identify where its return is equal to itself for the purpose of demonstrating the role of the benchmark algebraically, beginning with return over time R_t. One then inserts the appropriate benchmark time series. At the same time, both adding and subtracting the returns from the aforementioned equation:

$$R_t = B_t + (R_t - B_t) \tag{3.5}$$

Where:

- R_t = The return of the portfolio being measured over time
- B_t = Benchmark return over time

It is then possible to add to this the alternative asset manager's active management choices, producing either positive or negative α (alpha). Alpha in this context is defined as risk adjusted performance relative to a benchmark, sometimes referred to as *Jensen's Alpha* after Jensen and Black Scholes (1972). This can be stated algebraically, the portfolio minus the benchmark returns. It is therefore possible to say that:

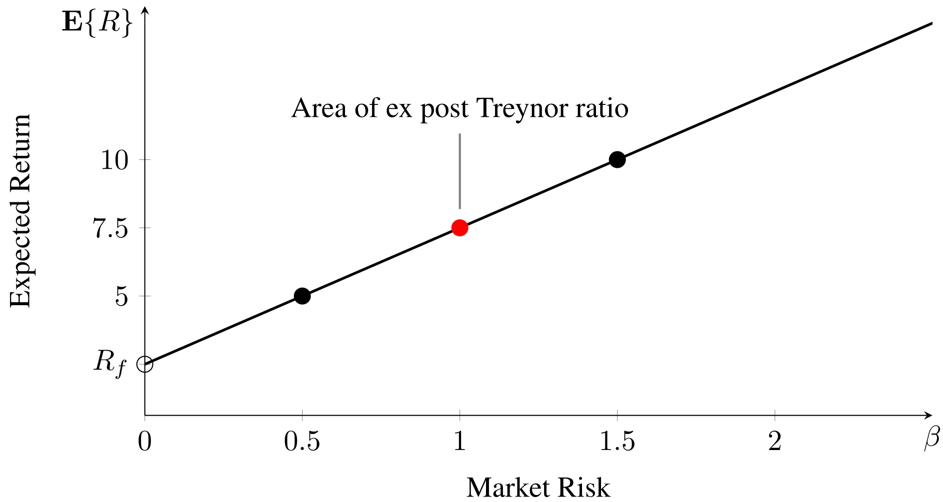
$$R_t = B_t + \alpha \tag{3.6}$$

Where:

- R_t = The return of the portfolio being measured over time
- B_t = Benchmark return over time
- α = Jensen's alpha

This can be graphically depicted as shown in the figure 3.2. It is any positive excess risk adjusted return.

Figure 3.2: Jensens Alpha as depicted by ex-post Treynor ratio



This figure shows the ex post positioning of the region where the ex post Treynor ratio is positive against the securities market line. The X axis depicts expected return and the Y axis market beta. The area above the line has positive Jensen's Alpha and that below the line has lower.

An alternative asset manager's return, therefore, incorporates an element of the benchmark return and an element of active management decisions. This is the relative return. For example, within alternative assets a commodities fund manager may decide to invest in wheat rather than pork bellies. Adding and/or subtracting the market index return from the right side of the equation results in the return relative to the index:

$$R_t = RI_t + (B_t - RI_t) + \alpha \tag{3.7}$$

Where:

- R_t= Return over time
- RI_t = Return of the index
- α= Alpha

With this equation, one can define the investment approach of a manager $'S'$. This is his style and it manifests itself in the difference between the benchmark return and the market index. For example, a real estate manager tilting his portfolio to commercial properties with guaranteed yields and good tenants. This can be shown as:

$$R_t = RI_t + S + \alpha \tag{3.8}$$

Where:

- RI_t = Return of the index
- α= Alpha

This equation merely shows that the return of a portfolio manager is related to the market index return, his style *S*, and the active management return α. A poorly constructed alternative asset benchmark would blur these differences. If a fund manager's portfolio is a broad market index, *S* and α are nil. In this instance, the portfolio manager generates just the market return. The definition of what constitutes the market is important. On the other hand, if the benchmark used is a broad market index, then *S* is assumed to be zero. As a result, a portfolio manager in effect generates a market return element from his active decisions. With alternative investments, the market return element is harder to identify.

In summary, if the market proxy is used as a benchmark, and a fund manager's style differs from it, the return time series can be very different. An alternative asset manager would have a very different style and risk profile from a broad market index and requires an appropriate benchmark. Without this the manager is incorrectly benchmarked. In such a case, the manager's style returns will be reflected in the measured active management component and as such an analysis of a manager's true value added will not be possible.

3.2.3 Active returns

It is also necessary to introduce the concept of active returns in order to understand how an alternative asset benchmark index can be practically useful to fund managers. Active fund management incorporates forecasting and benchmarks can be used to contextualize these into expected returns. In this respect, expected returns *(Er)* are an estimate of future rate of return and active expected returns differ from consensus expected returns. They are estimated bottom up using either *ex-ante* (forward looking) or *ex-post* (historical) time series data, or top down based on economic series. If the manager can see more expected rate of return than required, he would then look to make an investment. The expected return of a benchmark becomes the weighted sum of the individual returns from the securities making up the benchmark, annotated thus:

$$E[R] = \sum_{i=1}^{n} R_i P_i \tag{3.9}$$

Similarly, where a benchmark is an index or a collection of securities, its arithmetic rate of expected return is therefore stated thus.

$$E(R_b) = W_1 R_1 + W_2 R_2 + ... + W_n R_1 n \tag{3.10}$$

Where:

- $E(R_b)$ = Expected return of the benchmark
- W = Index weight
- R = Return of index constituent

This formula can be used as the basis of a simple index, namely a collection of weighted instruments. More technical construction issues are covered next.

3.3 The differing approaches to index construction and their relevance for alternative assets

An index is the most widely used benchmark employed by professional asset managers. Index construction is essentially a sampling, weighting and optimization problem. It has to be done from a mean variance perspective in order for it to prove useful from both an academic and a practitioner perspective. That said, these elements require trade-offs to be made, particularly in alternative assets.

The method behind index construction has evolved over time. Within equities, free float market capitalization, equal and/or value weighted indices, are now well established and considered the norm. New variants like fundamental and factor indices are also gaining in popularity. The indices in the alternative asset classes have similarly evolved through various stages. This was documented by Miffre (2012). The commercial versions detailed in chapter six represent more advanced versions of their predecessors.

A critique of any index is that it is dynamic, changing when new instruments either enter or exit the market, or as prices rise and fall. To preserve an index's representativeness, constituents need to be refreshed over time in the same way that goods in the inflation basket change. A robust index for alternative assets has to have a replacement rule.

Index construction is a process that requires first establishing the statistical method, then ensuring it is representative of the target assets. The collection and cleaning of the data is also important, due to the aforementioned stale pricing and survivorship bias, a concept explained by Brown et al. (1992). In the practical world, the maintenance of indices is an on-going process and requires frequent revision and re-balancing. A robust index for alternative assets has to have a re-balancing rule.

The methods applied to construction of indices were adopted from economics and as a result the techniques have subsequently been adapted for investment. Originally a number of indices were created to measure price inflation. These were summarized by Eichhorn (1976). Their origins are summed up by Chance (1966).

The index method themselves go right back to Dutot (1738). His ratio of average prices was devised to measure the price movement over time of a whole basket of goods. In investment indices, the basket of goods is replaced by a portfolio of securities and in alternatives with instruments.

3.3.1 The arithmetic versus geometric approach

All the aforementioned index approaches rely on different mathematical approaches. In this respect, the biggest differentiator is that index returns can be stated in either arithmetic or geometric form. The arithmetic form has already been explained and this is the form that finance practitioners typically use in their ex-post expected return calculations, expressed as:

$$\bar{r} = \frac{1}{n}\sum_{i=1}^{n} r_i = \frac{1}{n}(r_1 + \cdots + r_n) \tag{3.11}$$

Where:

- r = Arithmetic mean return

In contrast to the arithmetic yield, geometric returns yield an average equal to the product of one plus the total return for each period. The geometric mean reflects the composite return that an investor would have earned, if that investor had bought the index weight at the beginning of the valuation period and then reinvested all dividend income and subsequently re-balanced according to the index re-composition. The geometric return uses the n_th root and then subtracts one. The geometric average is expressed as:

$$\bar{r}_{\text{geometric}} = (\prod_{i=1}^{n}(1+r_i))^{\frac{1}{n}} - 1 = \sqrt[n]{\prod_{i=1}^{n}(1+r_i)} - 1 \tag{3.12}$$

Where:

- $\bar{r}_{\text{geometric}}$ = Geometric mean return

Note that in fund management practice, it is assumed that rates of return are expressed arithmetically unless stated otherwise. The caveat in using arithmetic means that they are not so good at multi-period analysis. They generate residuals over time. Geometric returns are assumed to be normally distributed and arithmetic returns are log. In this respect, arithmetic returns are useful in alternative classes due to the skew and kurtosis present.

Jevons (1865) defined the geometric index equation as "a harmonic average of price relatives". The geometric index approach has different properties. At any point in time, it is equal to or less than the arithmetic mean return for the same index. When the variability of the return ends, the arithmetic average benchmark return is always greater than the geometric mean benchmark return. The differential grows with increasing return variability. A good rule of thumb is that the geometric mean return is approximately equal to the arithmetic mean return minus half the variance.

As explained by Jorion (2007), There are usually few differences between arithmetic and geometric approaches when returns are low. The difference between the two methods was summarized by Walsh (1901) at the turn of the last century. That said, there is a difference. The extent of the difference was investigated by Blume (1974). He measured the distortion and showed that the use of an arithmetic approach results in overestimation of benchmark returns. The use of a geometric approach underestimated them. Significant differences can occur between the returns when volatility increases and the time period of measurement decreases, as might be the case in some of the more volatile asset classes like commodities but less so in real estate.

Only where there is certainty will the geometric mean index match the arithmetic mean index. Another issue is that the use of the arithmetic approach using annual data results in a lower mean than one calculated using monthly data. This bias has to be adjusted for. A good discussion on this subject can be found in Brennan and Schwartz (1985). They point out that as share prices are uncertain, a bias factor exists, as shown below.

$$B > 0, \sigma_i^2 + \sigma_j^2 > 2\sigma_i 2\sigma_j. \tag{3.13}$$

Where:

- B = Bias factor
- i = Index i
- j = Index j

To have two major methodologies is clearly not optimal. Broby (2007) illustrated the methodology selection between the two by considering "the same scenario with each index methodology. Where an index starts at 100, falls to 80 and then increases again to 100, the arithmetic mean return would be 5 per cent. This represents the average of the initial 20 per cent decline and the 25 per cent rise in the next period. The geometric mean return, however, would be zero because the value of the index at the end of the two periods is the same as at the beginning."

As a result of these shortfalls, a geometric index needs to apply a *divisor* to overcome the faults in a geometric index. In a price weighted index, the divisor is typically created by taking the price of each constituent and then adding all the constituents together to create the base value of the index. The different geometric and arithmetic weighting approaches and index adjustments are explained in more detail next.

3.3.2 The Dutot formula

The first indices introduced into the literature were little less than naive groupings. An index formula designed to measure inflation was proposed by Dutot (1738), the *Dutot formula*. He developed a ratio of the average aggregate prices of stocks. It is based on quarterly data, divided by the average aggregate price of stocks in the base period. Originally it was the price of stocks of goods but can by extension be equity stocks and financial instruments. The Dutot formula is sometimes called either the arithmetic formula or the relative of means formula. It is stated thus:

$$P_D = \frac{\frac{1}{n} \cdot \sum p_t}{\frac{1}{n} \cdot \sum p_0} = \frac{\sum p_t}{\sum p_0} \tag{3.14}$$

The Dutot formula designed as an inflation index and as such is part of the literature in economics rather than finance. It was created because of the aforementioned issues with geometric averages, the arithmetic average producing a higher average price change.

3.3.3 Lespeyres Index

An improvement on Dutot (1738) was developed by Laspeyres (1871). The *Laspeyres Index*, named after its inventor, is derived from a price index. It is the favoured index in the investment profession, endorsed by the *Global Investment Performance Standards*. It is also the most commonly used in later generation alternative asset indices. It is defined by Laspeyres as "a weighted arithmetic average of the price relatives using the values of the earlier period as weights". It was designed to measure a basket of goods and services. It essentially answers the question of how much a consumer basket bought in a base period would cost in another period. In this respect, it has fixed weights. As a result, it is termed a *base-weighted index*. It is represented thus:

$$P_L = \frac{\sum (p_t \cdot q_0)}{\sum (p_0 \cdot q_0)} \tag{3.15}$$

Where:

- P_L = is the price of asset L (L = 1,. . . ,m) at time t

- Pio = Price of asset i (i = 1,. . . ,m) in start time period

The Lespeyres index is illustrated in its simplest form, as a price of an asset. Within alternative assets, it is particularly well suited for physical commodity indices and has therefore replaced Jevons in standard usage in that asset class. Another advantage is that the Laspeyres index does not require up to the minute constituent information. Therefore in practice the Laspeyres formula is usually preferred where constituents are not continually updated, as is the case with alternative investments. Real estate pricing, for example, is based on observed prices of similar but distinct properties. With alternative assets, for non physical investments with multiple constituents, the Lespreyres formula can be extended. Its formula is as follows:

$$Index_t = \frac{\sum_{i=1}^{n}(p_{it} \cdot s_{it} \cdot ff_{it} \cdot cf_{it} \cdot x_{it})}{D_t} = \frac{M_t}{D_t} \tag{3.16}$$

Where:

- t = The time the index is computed
- n = The number of instruments in the index
- p = Price of the instrument (i) at time (t)
- s = Number of instrument (i) at time (t)
- ff = Free float factor of instrument (i) at time (t)
- cf = This weighting cap factor of instrument (i) at time (t)
- x = Exchange rate from local currency into index currency for instrument (i) at time (t)
- M = Free-float market capitalization of the index at time (t)
- D = Divisor of the index at time (t)

Capitalization weighted indices are the most recent iteration of such indices, replacing the first generation of arbitrary instrument selection and gaining more support by practitioners than value weighted indices. Method on this approach can be found in Enderle, Pope, and Siegel (2003a)

A "free float" capitalisation weighted index captures as precise a measurement as possible and is calculating the total of float outstanding for each share class. The total is then divided by the sum in what is termed a *divisor*. This then reflects the cumulative structural index change over time. Such adjustments present serious issues for alternative asset classes. These include identification, calculation and appropriateness.

3.3.4 The Jevons formula

The *Jevons formula*, proposed by Jevons (1865), is a price relative index based on a geometric mean. This is typically used in some of the older real asset commodity and physical real estate indices. The Jevons formula is sometimes referred to as the *geometric mean of relatives formula*. It is shown below.

$$P_J = \left(\prod \frac{p_t}{p_0} \right)^{1/n} \tag{3.17}$$

Where:

- P_t= Price of stock (n = 1,. . . ,N) in period 1
- P_0 = Price of stock (n= 1,. . . ,N) in the base period (period 0)

This formula was used historically to measure cash commodity returns.

3.3.5 Hermann Paasche Index

The other well known index method, the *Hermann Paasche Index* is described in Diewert (1998). Once again, it is derived from the measurement of price levels relative to a base period. In this respect it provides a comparison between the cost of constituents, as valued at current prices, with the value of those same constituents at the base period. This can also be used for physical real estate indices.

The Hermann Paasche price index dampens price increases, since it incorporates an element of the changes that have already occurred. In this respect, increased consumption reflects reduced relative prices. This observation presents an interesting problem in its use in real estate price movement measurement, where transactions are not as frequent as in other asset classes.

From an alternative investment perspective, by replacing *volume of goods* sold with *volume traded* the index can be adapted to provide a liquidity weighted index. This is useful for adoption of hedge fund indices. It can also be useful for physical real estate indices.

$$P_t = \frac{\sum (p_t \cdot w_t)}{\sum (p_0 \cdot w_t)} \tag{3.18}$$

Where:

- P_t = Price i $(i = 1, \ldots, m)$ in period t
- P_o = Price i $(i = 1, \ldots, m)$ in the base period
- w_t = Volume purchased in period t

3.3.6 Fisher Index

More complex versions of indices can incorporate different elements of these methodologies. The Fisher index, for example, is a geometric mean of the combined Laspeyres and Paasche indices.

$$P_F = \sqrt{P_L \cdot P_P} \tag{3.19}$$

Where:

- P_F = Fisher index
- P_L = Lespeyres index
- P_P = Paasche index

The *Fisher formula* is widely used in inflation measurement economics. It is often called the *Ideal Price Index*. The Fisher formula is more difficult to calculate because more data is required and both other indices need to be worked out separately. This increases the cost and time for the index provider. It can also result in a longer calculation time.

The Fisher index provides a measure of the *preference function* as introduced earlier. The Herman Paasche index acknowledges that choice of methodology implies that bias may creep in but that statistical method for indices should be unbiased. It is not, however a popular method. In this respect, Blume (1974) suggested that it should not be used for alternative assets but instead a suitable estimator could be calculated that had less bias. This could be achieved through a weighted average of the arithmetic and geometric means. In alternatives, this could be where an index constructor chain links older geometric indices with more recent arithmetic ones. This approach is seen as a compromise only suitable for academic studies that require very long time series.

3.3.7 Marshall-Edgeworth Index

When investability is a prime concern, for example in private equity, more attention has to be spent on the liquidity side of the equation. There is no really well-defined liquidity methodology but once again investment professionals look to inflation methodology. In this regard, the *Marshall-Edgeworth Index* is a price index that measures the weighted arithmetic average of current base price relatives. It uses volume to determine weights. It is not practical to serve as a timely indicator of price changes, as volumes are required in the current period. It is therefore delayed but of interest to investors in illiquid alternative asset classes.

$$P_{ME} = \frac{\sum \left[p_t \cdot \frac{1}{2}\left(q_0 + q_t\right)\right]}{\sum \left[p_0 \cdot \frac{1}{2}(q_0 + q_t)\right]} = \frac{\sum \left[p_t \cdot (q_0 + q_t)\right]}{\sum \left[p_0 \cdot (q_0 + q_t)\right]} \tag{3.20}$$

Where:

- P_t = Price of stock i $(i = 1, \ldots, m)$ in period t
- P_o = Price of i $(i = 1, \ldots, m)$ in the base period
- Q_t = Volume of i purchased in period t
- Q_o = Volume of i purchased in the base period

The quantities of the weighted volume are arithmetic mean values of the quantities of the base period and the current period. It is a symmetric and pseudo-superlative index. The Qo volumes are to be considered as volume-weighted trading prices on the underlying assets of the alternative assets. Qt volumes should be weighted projections on the previous month's or three months' average.

3.3.8 Tornqvist Index

The *Tornqvist Index*, like the Fisher index, uses inputs from two time periods to determine the weights. It attaches the same emphasis on the values in both periods. As a result, its output is similar to creating an average of the Laspeyres and Paasche indices, especially if the index spread is not very large. The above comments on combining long time periods with alternative asset time series also apply.

$$P_T = \prod_{i=1}^{n} \left(\frac{p_{ti}}{p_{0i}} \right)^{\frac{1}{2}\left[\frac{p_{0i} \cdot q_{0i}}{\sum(p_0 \cdot q_0)} + \frac{p_{ti} \cdot q_{ti}}{\sum(p_t \cdot q_t)}\right]} \tag{3.21}$$

Where:

- p_{ti}= Price of stock i (i = 1,. . . ,m) in period t
- P_i = Price of stock i (i = 1,. . . ,m) in the base period
- Q_t = Volume of stock i purchased in period t
- Q_o = Volume of stock i purchased in base period

The difference between the Tornqvist and Fisher indices is typically small relative to the difference between the Laspeyres or Paasche index.

3.3.9 Carli Elementary Index

There are many other less used indices. The *Carli Elementary Index*, for example, uses an average of price relatives. It can be used where there are multiple exchanges setting prices for the same instrument. An example for alternatives is the cryptocurrency market where no central marketplace exists. A price relative is the stock in one exchange divided by the stock at that exchange in the base period.

$$P_C = \frac{1}{n} \cdot \sum \frac{p_t}{p_0} \tag{3.22}$$

Where:

- P_t= Price of stock (n= 1,. . . ,N) in period 1

- P_0 = Price of stock (n= 1,. . . ,N) in the base period (period 0)

The above formula is sometimes referred to as the *price relatives formula* or *mean of relatives formula*. As an index it it is not reversible in time. The Carli index between periods 0 and t, exceeds one calculated backwards. In this respect, back testing is not possible making it not appropriate for alternative asset contender indices such as a PCA derived index.

$$P_C = \frac{1}{n} \cdot \sum \frac{p_t}{p_0} \tag{3.23}$$

Where:

- p_t= Price of stock (n = 1,. . . ,N) in period 1
- p_0 = Price of stock (i = 1,. . . ,M) in the base period (period 0)

3.3.10 Chain linking

All the aforementioned approaches and indeed any index and/or benchmark can be chain linked, as mentioned under the discussion on the Fisher index. This is the statistical method that index constructors employ to combine each of the different methods into a long-term index series. It works by calculating overlapping indices and restating them. To do this, an index with an old base is divided by an index with a new base, thereby generating a new one. A linking factor is applied to the old index series. This is sometimes done in academic research, for example when the time series that the researcher wishes to apply does not go back far enough.

Chain linking equates two index series and/or benchmarks that could be significantly dissimilar, such as an equity and a bond index. This could be used to benchmark multi-strategy hedge funds. The advantage is that balanced funds can be measured but the chain linking can cause problems with the long-term series.

The approach is used in linking short-term indices. This is because, the long-term series movement is not a pure price index. Very long-term comparisons using indices that are rebalanced can be meaningless. This approach was used by Wilson and Jones (2002). Like is not being compared with like with respect to the underlying stocks. An example of this is the number of new technology stocks included in the indices since the 1990s.

As an illustration, a commodity index from 1970 would be very different from a commodity index from 2000, as a result of financialization. As a result of this, chain linking can contribute to bias in the long-term series. Because of the nature of chain linking, any bias existing in the old index series is preserved in the long-term series. The impact of new items is not measured. Another example of where chain linking is relevant is in the historic conversion of the European currencies to the euro. From an alternative asset perspective, chain linking is most relevant in cash commodity and long real estate time series.

Alternatively weighted or dynamic benchmarks that are not based on optimal portfolios can be used in combination with existing benchmarks. Christoffersen, Olesen, and Christoffersen (2014) advises that investors should still consider a capitalization-weighted total market index to evaluate how much return enhancement and risk reduction they obtain whatever their choice of alternative benchmark.

3.4 Weighting and index adjustments

The weighting approach is equally, if not more, important than the calculation method. An index can be constructed using price, market capitalization, equal or factor weightings. It can be designed to measure pure price appreciation, either arithmetic or geometric, or total return including dividends and distributions. The different weighting constructs result in a different time series of returns.

In some alternative asset classes, such as commodities, there is no income element. In others, such as real estate, there is. In a discrete setting such as index construction, the weight function is a positive function defined on a discrete investment instrument universe. In this instance it is defined as:

$$U = \sum_{a \in A} f(a) \tag{3.24}$$

Where:

- U = Investment Universe
- $f(a)$ = Function of the assets

3.4.1 Price weighting

Price weighting determines the weight of each constituent by dividing the price by the sum of all constituent values. As a result, a change in the instrument will result in a change in the weights of all the securities in the index. A mechanism has to be put in place to avoid this. With a geometric mean, this is done through the use of a divisor.

Price weighting is depicted thus:

$$W_i^M = W_i P_i / \sum_{j=1}^{N} P_i \tag{3.25}$$

Where:

- W = The weight

When weighting by market capitalization, the weight of each index constituent is determined by dividing the market capitalization of all securities in the index. Market-capitalization weighting is sometimes called value weighting. It is calculated by taking the number of shares issued multiplied by the stock-market price per share.

$$W_i^M = W_i P_i / \sum_{j=1}^{N} W_i P_i \tag{3.26}$$

In alternatives. market capitalization weighting presents issues. What figure does one take, for example, for an open futures contract when constructing a commodities index? The answer is not given in the literature. The most common such constructs are expressed as:

$$\frac{P_2}{P_1} = \frac{SP_2/N}{SP_1/N} \tag{3.27}$$

Where:

$$\frac{P_2}{P_1} = S(P_2/p_1)/N \tag{3.28}$$

- summed from n=1 to N

Where:

$$\frac{P_2}{P_1} = S(P_2x/p_1x) \tag{3.29}$$

- summed from over prices n=1 to N

Where:

$$\frac{P_2}{P_1} = P(P_2/P_1)/(1/N) \tag{3.30}$$

- summed over prices n=1 to n, where the x's are a quantity vector from 1 to n
- P = price

3.4.2 Equal weighting

Another approach is equal weighting. Equally weighted the weights in an equal-weighted index are assigned as follows.

$$w_i^E = 1/N \tag{3.31}$$

Where:

- w_i^E = Equal weighted index
- N = the number of holdings in the index

There has been much academic research into the equal weighted portfolio, referred to as the 1/N portfolio, as discussed in the review of the literature in chapter two.

3.4.3 Adjustment divisors

Divisors are a way of adjusting a geometric index for corporate changes. In order to take account of the dynamic nature of companies, that are constantly issuing or buying-in stock, issuing dividends, splitting or consolidating, a divisor has to be used. The approach is illustrated in the formula below:

$$Index_t = \frac{\sum_{i=1}^{n}(p_{it} \cdot s_{it} \cdot wf_{it} \cdot cf_{it} \cdot x_{it})}{D_t} = \frac{M_t}{D_t} \tag{3.32}$$

Where:

- t = Time the index is calculated

- n = Number of instruments in the index
- p = Price of instrument (i) at time (t)
- s = Number of shares of instrument (i) at time (t)
- wf = Free float factor of instrument (i) at time (t)
- cf= The weight factor of instrument (i) at time (t)
- x = Exchange rate from local currency into index currency for company (i) at time (t)
- M = Free-float market capitalization of the index at time (t)
- D = Divisor of the index at time (t)

A good description of the various approached to divisors is given by MSCI Inc. (2018).

3.4.4 Market capitalization weighting

Instruments in an index with restricted capital need to adjust for the market capitalization of the free float. This should be calculated using adjusted closing prices and the new free float-based number of shares This is calculated thus:

$$D_{t+1} = D_t \cdot \frac{\sum_{i=1}^{n}(p_{it} \cdot s_{it} \cdot ff_{it} \cdot cf_{it} \cdot x_{it}) \pm \Delta MC_{t+1}}{\sum_{i=1}^{n}(p_{it} \cdot s_{it} \cdot ff_{it} \cdot cf_{it} \cdot x_{it})} \tag{3.33}$$

Where:

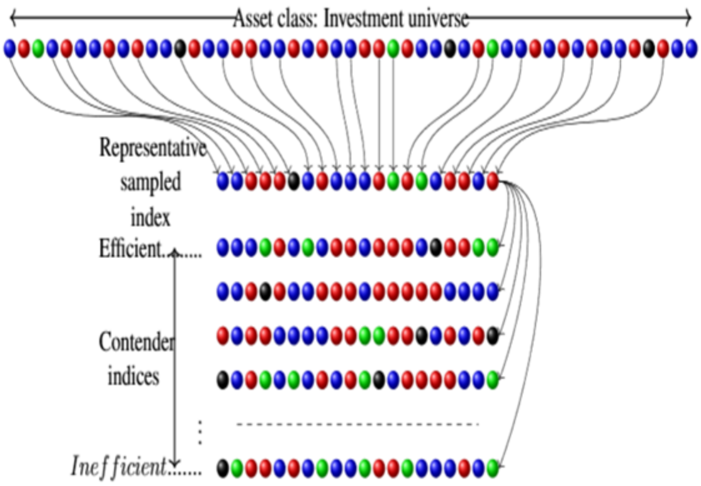
- $\Delta MCt + 1$ = The market closing capitalization of the index minus the adjusted market closing market capitalization of the index. MSCI (2014)

In the same way as adjusting for corporate actions, it is also possible to adjust for derivatives using a divisor. As explained, some alternative asset classes use the futures market to gain exposure. This is the case in commodity indices. The two different items are therefore the long and the short positions. The "spot–forward parity" links the spot market and the forward market.

3.5 Sampling the universe

As shown in the prior chapters, once the specific alternative asset class has been identified as an investment universe, it is necessary to select a sample in order to generate a representative benchmark. Marshall (1996) describes the various approaches to sampling. In the context of alternative assets, it involves dividing the asset class population into groups, termed *clusters*. This is depicted in figure 3.3, where the sampling is shown to impact the efficiency of the index. This is because sampling focuses on the relationship between the properties of these and gauging the precision of the outcome. The objective of creating a sample of instruments in this way is to draw inference concerning the whole asset class. This is an inductive method. That said, it is first necessary to apply a deductive approach and as previously observed investigate the behaviour of the time series.

Figure 3.3: Sampling the investment universe.



This figure is a stylized diagram showing how sampling the investment universe reduces the investment universe but retains some of the unique characteristics of the sub-components. The contender indices range from Mean variance efficient to inefficient.

In alternative assets that sampling to create existing indices has been pragmatic, and therefor is sub-optimal. This is because, in financial markets, the norm is first to determine the investment objectives and investment process, rather than the risk at the instrument level. The focus of the method has to be on both the outcomes and risk. These are necessary for the identified appropriateness criteria, that of being able to replicate the index. Existing alternative asset benchmarks aim to capture the above results but not the risk factors. They are therefore sub-optimal.

3.5.1 Top down versus bottom up sampling

As the construction of an alternative asset index is a sampling exercise, a decision has to be made between top down or bottom up approaches.

The top down approach to benchmarks is typical of Bali, Brown, and Caglayan (2011). They argue that anyone indirectly investigating hedge fund exposures must take account of macroeconomic risk. As such, they use a factor approach that emphasizes the cross-sectional relationship between alternative measures of systematic risk. Bali, Brown, and Caglayan (2014) built on this by including variables that measured macroeconomic uncertainty. The top down approach may well be more relevant for alternative assets and is consistent with Principal Component Analysis.

The bottom up approach to benchmarks, the sort that incorporates peer group evaluation, was addressed by Elton, Gruber, and Blake (1999). They created an index to identify systemic influences affecting funds and used this to compare with the peer groups. This, they suggested, had fewer cross-sectional observations. In creating this, they highlighted that systemic style could be used to classify and group funds and/or alternative asset class instruments.

3.5.2 Reducing sample bias

The literature showed that it is important to reduce sample bias. The sampling, therefore, has to be based on common characteristics. That said, when constructing a benchmark, the out of sample forecasting bias is a problem. Index constituents are subject to these during identification of the investment set. The ideal investment universe needs to be clearly defined, accessible, reliable, and designed to achieve the outcome of interest.

Sample bias is also an issue for alternative assets. The problem, as relates to the bottom up approach, was addressed by Elton, Gruber, and Das (1993). They re-examined a study by Ippolito (1989) by changing the benchmark used. Their results demonstrated that when a more appropriate multi-index model was used the alpha of the earlier study disappeared. The contribution of their work was to show the importance of benchmark selection in performance evaluation and as a result is the reason why a better solution is required to benchmark the alternative asset classes.

An interesting way to think of the sampling issue is to consider active share, introduced and proposed by Cremers and Petajisto (2009). This approach takes into account size, expenses and turnover characteristics of a fund relative to a benchmark. Sampling tends to work best with normal distributions, it should be pointed out, however, that Abhyankar, Copeland, and Wong (1997) showed that even broad market proxies such as the *Standard and Poors 500 Index*, the *DAX 30 Index*, the *Nikkei 225 Index*, and the *FTSE-100 Index* showed signs of non-linearity. The problem is magnified when alternative assets are considered.

3.6 Trade-offs in index construction

Constructing indices for alternative assets is essentially a sampling exercise, various trade-off decisions have to be made. Figure 3.4 is a ven diagram that shows how these decisions relate to one another. It is these decisions that make the benchmark useful in the practical sense. The were identified and documented by Siegel (2003). The choices made in these trade-offs is reflected in the tracking error between various sorts of index and has to be taken into account in benchmark returns.

The first step that an index constructor has to take is data collection and to decide the size of sample. The compiler must collect the yield data of the securities it has determined to be included in the index. Which instruments should be included and which should be excluded will affect the replicability and / or utility of the index.

For alternative investments, benchmark returns are very sensitive to the data inputs. There is a great deal of survivorship bias, particularly in hedge funds and other collective investment funds used to gain exposure to alternative assets. Some companies get taken over and some go bankrupt. The same is the case in, for example, hedge funds. Any bias in the historical data, such as stale pricing, unmatched strategies, or misreported net asset values will impact the total returns. As a result, clean data and random sampling is used. This approach allows the index constructor to drill down into the sub components and allows the aforementioned trade-offs to be made.

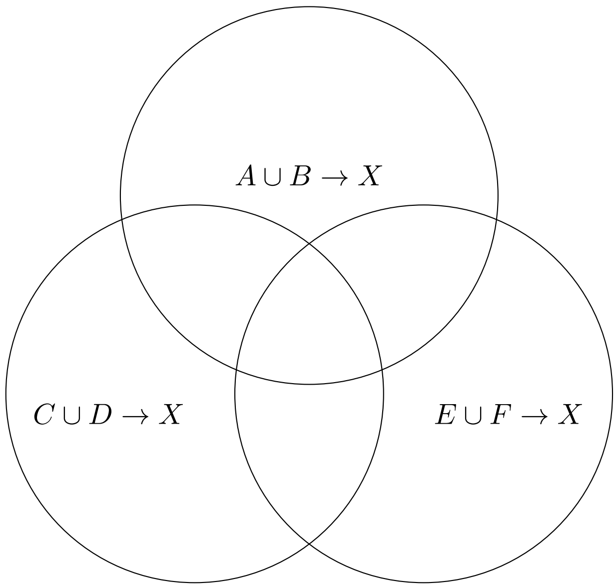
Boolean algebra is presented in order to explain the trade-offs and the outcome of the trade-offs. Qualitative choices have to be made and in this respect is expressed thus.

$$A \cup B \to X, C \cup D \to X, E \cup F \to X \tag{3.34}$$

Where:

- A or B Implies X,C or D Implies X,E or F Implies X
- $A \cup B \to X$ = Breadth or Investability implies an appropriate benchmark
- $C \cup D \to X$ Rebalancing or turnover costs implies appropriate benchmark
- $E \cup F \to X$ Rules or sampling value judgements implies appropriate benchmark

Figure 3.4: How trade offs interact to determine a benchmark's universe.



This figure shows a Venn diagram with the various intersections between the trade-offs that have to be made in index construction. $A \cup B \to X$ *= Breadth or Investability implies an appropriate benchmark,* $C \cup D \to X$ *Re-balancing or turnover costs implies appropriate benchmark,* $E \cup F \to X$ *Rules or sampling value judgements implies appropriate benchmark.*

3.6.1 Breadth or investability

The first major trade-off is breadth versus investability. The breadth of a benchmark has to be decided at the outset. Breadth is important for diversification. This refers to how many instruments there are in the sample that represents the whole asset class universe. A benchmark that contains too many instruments is not useful to fund managers as the costs of replication become prohibitive. In order to achieve investability, there should be a degree of cross coverage with the selection choices of an active manager's portfolio and his benchmark. It is for this reason that a broad equity index is not a good proxy for an alternative asset class.

One aspect of breadth versus investabitity is the frequency of trading. When financial instruments do not trade, they are not investable. Hedge funds, for example, mostly accept the creation or redemption of units on a monthly basis. Physical real estate is even less frequent in its trading. Dimson (1979) investigated the phenomena of infrequent trading and developed a method for estimating beta when shares did not trade, thereby facilitating their inclusion in an index even without a full history.

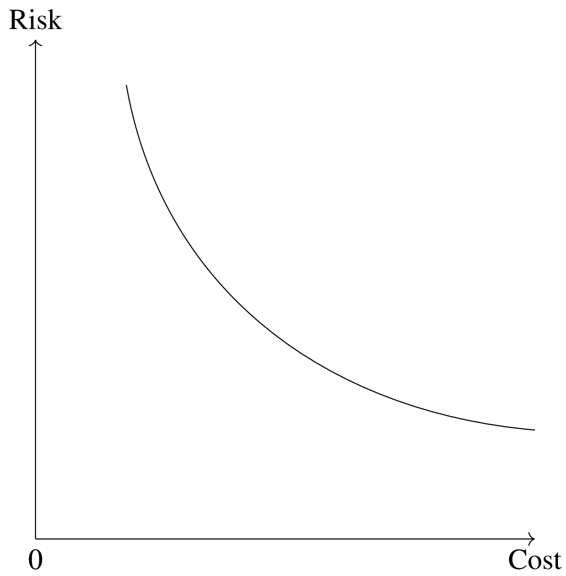
A full market capitalization index, one that has full breadth, would including every possible security in the investment universe available to investors. This would include many companies too small or too illiquid to realistically invest in.

3.6.2 Re-balancing frequency

The second major trade-off is the frequency of re-balancing. Part of investability is the question of benchmark turnover. If constituents change too frequently, there is a cost in replication. Bailey (1992a) estimated that turnover of as high as 20 percent can be tolerated on a quarterly basis.The term re-balancing refers to a readjustment in the weights of existing securities. Re-balancing occurs when changes to the instrument occur.

The need for re-balancing can be due to expiry, removal and replacement. Re-balancing is important for maintaining an index on the efficient frontier, introduced in the prior chapter. Re-balancing frequency was addressed by Arnott and Lovell (1993). The two key benchmark approaches they suggested are (1) Calendar re-balancing based on a fixed schedule, such as every week, month or year and (2) Range re-balancing, where a predetermined band is set and the returns reinvested into it. The former is preferred from the perspective of precision.

Figure 3.5: Re-balancing versus buy-and-hold strategy



This figure has a diagram that shows the trade-offs between a buy and hold strategy and a more active strategy. The more active a strategy, the higher the costs born by the benchmark. The risk shown is from the perspective of tracking error. Source Lam (2014), page 40.

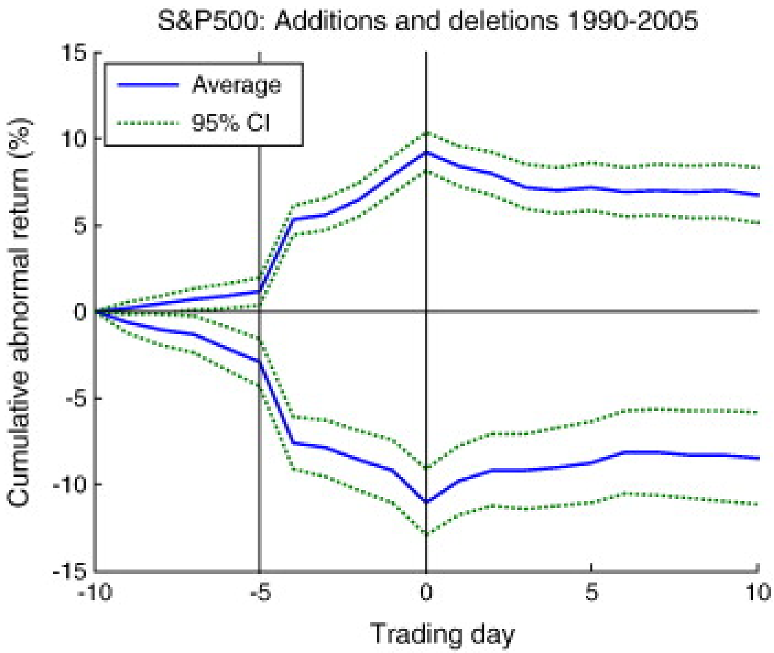
Each time a re-balancing trade is made, a cost is incurred. This is sub-optimal. An index is reconstituted and re-balanced to keep it close to its intended investment universe and weighting criteria. The term reconstitution refers to the addition and subtraction of securities from an index. The trade off is shown diagrammatically in Figure 3.5, taken from Lam (2014)

Frequent changes make the benchmark index more representative but the cost of replicating it increases due to the costs of turnover. In their analysis of the re–balancing frequency Kuhn and Luenberger (2010) conclude that a high re-balancing frequency reduces the portfolio performance due to transaction costs. Fletcher and Marshall (2005) employed re-balancing in their method to evaluate strategies based on re sampled portfolio efficiency. Re-balancing is also importance in academic portfolios designed to test market efficiency.

Edmister, Graham, and Pirie (1994) observe that excess returns of index replacement stocks are attributed to price pressure. In this respect, they point out that announcement date and information that impact returns. Graham and Pirie (1994), in a separate piece of stock specific research did not find evidence of abnormal returns in a widely flagged change despite the presence of abnormal volume. Dhillon and Johnson (1991) found abnormal returns in the period before such changes occurred. All underlie the importance of index construction choice during re-balancing.

Petajisto (2011) explained that index turnover costs represents a hidden cost that has to be borne by indicies due to the what is termed the index premium. This occurs due to market impact. Buying and selling large amounts of a security moves the underlying price. The impact is illustrated diagrammatically in figure 3.6.

Figure 3.6: The index premium and its hidden cost



The figure shows a chart with the cumulative average market-adjusted returns for shares added or removed from the index and the confidence intervals of 95 percent. It shows that the index premium, derived from constituent turnover, has a hidden cost. Source: Petajisto (2011), Figure 1, page 276.

Investability is not, however, the same as liquidity. A security may be liquid in one market but have restrictions on investment. Investability is an important considerations for futures based funds, which are used, for example, in commodity investment. Consider also real estate. Not everyone can afford the unit cost of an office block or shopping center.

3.6.3 Rules versus value judgments

The third major trade-off is between rules and value judgements, A benchmark can be constructed using either or both of a set of rules or a set of value judgments. The later is not an optimal portfolio. It is also not replicable, as the rules ensure the integrity of the index.

Einhorn and Hogarth (1981) investigated rules versus value judgments from the perspective of behavioral decision theory. They suggest these have to be treated within the context or rationality and optimally.

Potential bias tends to crop up everywhere in statistical method within index construction. These include, stock substitution, pricing, new issue impact, candidate selection and quality method bias. An index constructor should do all they can to ensure bias is avoided, so all these need to be addressed. Examples of ways to do this can be found in the Russell Index Guide Indexes (2008).

Table 3.1: Trade-offs in index construction expressed in Boolean logic.

	$A \cup B \rightarrow X$	$C \cup D \Rightarrow$	$E \cup F \rightarrow X$
DNF	$(\neg A \cup \neg B) \cup X$	$(\neg C \cup \neg D) \cup X$	$(\neg E \cup \neg F) \cup X$
CNF	$(\neg A \cup B) \cap (\neg B \cup X)$	$(\neg C \cup \neg D) \cap (\neg D \cup X)$	$(\neg E \cup \neg X) \cap (\neg F \cup X)$
ANF	$\neg((A \cup B) \forall (A \cup X) \forall (B \cup X) \forall (A \cup B \cup X) \forall A \forall B)$	$\neg((C \cap D) \forall (C \cap X) \forall (D \cap X) \forall (C \cap D \cap X) \forall C \forall D)$	$\neg((E \cap F) \forall (E \cap X) \forall (F \cap X) \forall (E \cap F \cap X) \forall E \forall F)$
NOR	$(\neg A \bigtriangledown \neg B) NOR (\neg B \bigtriangledown X)$	$(\neg C \cup \bigtriangledown D) NOR (\neg D \bigtriangledown X)$	$(\neg E \bigtriangledown \neg X) NOR (\neg F \bigtriangledown X)$
NAND	$(\neg A NAND \neg B) NAND \neg X$	$(\neg C NAND \neg D) NAND \neg X$	$(\neg E NAND \neg F) NAND \neg X$
AND	$\neg(A \cap \neg X) \cup \neg(\neg B \cap \neg X)$	$\neg(C \cap \neg X) \cap \neg(\neg D \cap \neg X)$	$\neg(E \cap \neg X) \cap \neg(\neg F \cap \neg X)$
OR	$\neg(A \cup B) \cup X$	$\neg(C \cup D) \cup X$	$\neg(E \cup F) \cup X$

This table illustrates the various trade offs that have to be made in index construction. The algebra is presented in Boolean logic form where all values are either True or False. These values are index choices that result in stock selection and iteration . In index construction, equivalent function may be achieved with clear choices. The resultant index will therefore have increased reliability and transparency.

3.7 Incorporating risk to mathematical method

The Capital Asset Pricing Model (CAPM) was introduced in the literature review. It is used by academics and practitioners alike to benchmark traditional investment returns in relation to their risk. In its basic form, it is given by the following equation:

$$E(R_i) = R_f + \beta_i[E(R_m) - R_f] \tag{3.35}$$

Where:

- $E(R_i)$ = The expected return on asset i
- R_f = The risk-free rate
- β_i = The beta of asset i,
- $E(R_m)$ = The expected return of the market

The CAPM forms the basis on broad market indices. This linear relationship is to justify the use of this *market proxy* as the *de-facto* benchmark that represents a sample of all traded assets. An alternative equilibrium asset pricing model, the Arbitrage Pricing Theory (APT) developed by Ross (1976), is also appropriate for alternative assets as it can break down returns into unique risk factors.

The market portfolio itself is the point where the Capital Market Line, which represents the trade-off between these variables, is tangent to the efficient frontier, a representation of the set of optimal mean variance portfolios. It forms the justification for the use of free float market capitalization indices.

The market proxy in equities markets is an approximation. This is also the case with alternative assets. Its use, therefore, does not capture all the dynamics present. These dynamics in alternative assets are further complicated by either the extent of the illiquidity, leverage and/or infrequent or inaccurate pricing that are present is such assets.

Alternative assets have idiosyncratic consumption and production functions. As far as alternative asset indices go, such as those for commodities, real estate, and private equity, the primary focus has been on measuring return rather than risk. As a result, there has not been the same effort made in the literature to justify them. The return, risk and correlations are, meanwhile, quite different from traditional assets. The constant price of alternative asset risk, and the specific dynamics of that risk, do not always hold. This dynamic attribute is something that PCA indices address over time.

Benchmarks for alternative asset classes are still dependent on the CAPM or APT. Black, Jensen, and Scholes (1972) observed that all derivations of these are built upon a number of simple assumptions that can be challenged. That said, the fact that assumptions are unrealistic does not necessarily invalidate the theories. Moreover, some of the criticisms of their applicability to alternative assets equally applies to traditional assets.

The first of these assumptions is that investors are risk averse and want compensation for risk. They are assumed to operate in a single period environment, which does not fit well with investors in hedge funds, a major alternative asset class. The CAPM and APT further suggests investors are wealth maximizers and can select portfolios on the basis of mean and variance. This contrasts with some real world situations where investors select alternative assets in order to hedge risk and achieve absolute returns.

The next assumption is that there are no taxes or transactions costs, a much criticized simplification and clearly not the case in for example, real estate investment. In reality, taxes and costs are the norm.

The CAPM and APT further assumes that all investors focus on the joint probability distribution of investment returns, which is not the case in alternative assets where, for example, commodity investments utilize managed futures. Finally there is an assumption that investors can borrow and lend at a risk free rate of interest. In reality, this is not the case.

The assumption of a normal return distribution is relevant for alternative assets and therefore their benchmarks. This is because a major difference between traditional and alternative investments is that the former are assumed to have a normal distribution of investment returns. In practice, as demonstrated in chapter one, these two distributions are distinct and therefore require different benchmarks. They are depicted algebraically as:

$$x = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}; y \neg \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \tag{3.36}$$

Where:

- x = The distribution of traditional returns.
- y = The distribution of alternative returns.

In practice, even traditional assets do not have a perfect normal distribution. Lanne and Pentti (2007) demonstrated the presence of conditional skewness in US stock returns. The fact that the returns of alternative assets do not have a normal distribution is key to understanding how benchmarks might be adapted for alternative asset classes. Having provided the context, the rest of this chapter goes on to describe the nature of benchmarks and the trade-offs that have to be made in order to make a benchmark index appropriate for use.

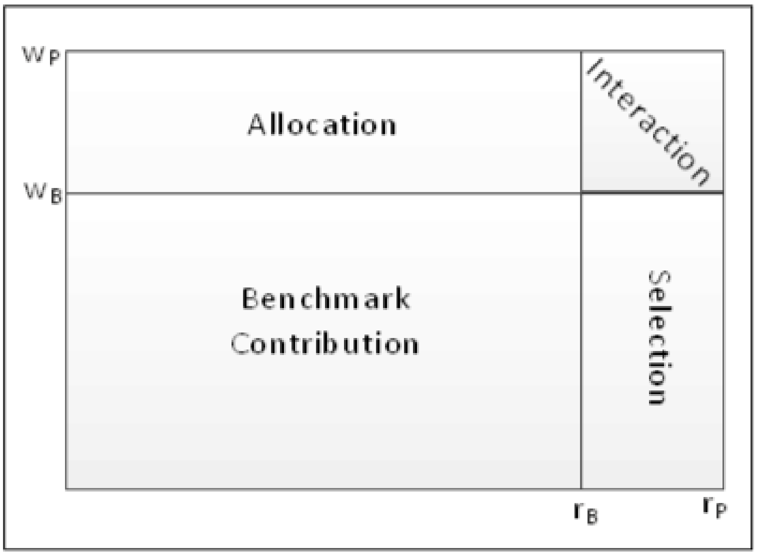
3.8 Allocation, selection and interaction effects

Benchmarks need to be able to attribute, Allocation, selection and interaction effects are the by-product of active asset management. Benchmark indices are used for attribution and this requires a clear break down of the sub-groupings in alternative assets. The most widely accepted approach to this kind of investment return attribution is based on the two factor model that was proposed by Brinson, Hood, and Beebower (1995). This is adaptable to alternative asset attribution, as it is based on top level interactions. Their method allows one to calculate segmented returns and drill down to the weights in each asset class. This is now considered the core of performance reporting. The various interactions are depicted in Figure 3.7.

As can be seen, the key contributions are from the benchmark returns. The alternative asset active manager adds value through an allocation, selection or interaction contribution to the returns. This can be either positive or negative.

An index return using the Brinson approach is thus assigned to each benchmark asset class and thereby makes it possible to further divide returns based on allocation to each asset class. This is the case, be it traditional or alternative, or where the returns are generated from policy decisions or timing decisions. It is also possible to account for the interaction effect between policy and security selection returns at the benchmark level.

Figure 3.7: Benchmark allocation, selection and interaction



This figure shows both the benchmark, allocation, selection and interaction effects interacting in diagrammatic fashion. The importance of the contribution are highlighted by the size of the box. It was based on a study about asset allocation of 91 pension funds from 1974 to 1983. Brinson, Hood, and Beebower (1995), Table 1, page 40.

3.8.1 Security selection

The first level of the Brinson approach is the security level. Indices are constructed from a portfolio of securities or in the case of alternative asset classes, instruments. As such, it is possible to identify the security selection contribution. This is calculated thus:

$$SE = w(b,i) * r(p,i) - r(b,i) \tag{3.37}$$

Where:

- SE = the Selection Effect
- $w(b,i)$ = The benchmark weights
- $r(b,i)$ = The benchmark returns
- $w(p,i)$ = The portfolio weights
- $r(b)$ = The total benchmark return

- $r(bp)$ = The total portfolio return

The implication of this for benchmarks is that there has to be a decision on the normal, or long-term, weights for each of alternative asset class. An illustration of the selection effect in alternative asset classes is where a Real Estate Investment Trust is chosen above another, or one commodity futures fund is chosen over another. This is an instrument specific impact.

3.8.2 Asset selection

The second level of the Brinson approach is the asset selection. The choice of benchmark weight gives rise to an asset allocation effect. This is calculated thus:

$$AE = w(p,i) - w(b,i) * r(b,i) - r(b) \tag{3.38}$$

Where:

- AE = The allocation effect
- $w(b,i)$ = The benchmark weights
- $r(b,i)$ = The benchmark returns
- $w(p,i)$ = The portfolio weights
- $r(b)$ = The total benchmark return
- $r(bp)$ = The total portfolio return

An illustration of the asset selection effect in alternative asset classes is where real estate is chosen above a commodity futures fund allocation. This is an asset allocation decision. The implication of this is that the benchmark decision has to incorporate which asset classes to include and which to exclude.

3.8.3 Interaction effect

The third level of the Brinson approach is the interaction effect. The interaction effect is, as its name implies, the interaction between the aforementioned effects. It is depicted thus:

$$IE = w(p,i) - w(b,i) * r(p,i) - r(b,i) \tag{3.39}$$

Where:

- IE = The Interaction Effect
- $w(b,i)$ = The benchmark weights
- $r(b,i)$ = The benchmark returns
- $w(p,i)$ = The portfolio weights
- $r(b)$ = The total benchmark return
- $r(bp)$ = The total portfolio return

The implication of the interaction effect is that the benchmark method need to be robust enough to avoid changing the weightings strategically away from the strategic view in order to try to obtain excess returns by exploiting fluctuations in asset prices. In the context of alternative investments, what constitutes a strategic weight is clearly important. Crude oil, for example, has small variances in production but big variances in price returns.

An illustration of the interaction effect in alternative asset classes is where both a Real Estate Investment Trust and also a commodity futures fund are chosen for a portfolio but in different weights from the benchmark. There is therefore an element of the return that differs from the specific asset chosen or the specific weight of that asset in the benchmark.

3.8.4 Total Value Added (TVA)

The *Total Value Added* (TVA), in the Brinson approach, is the combined decisions on the asset selection, the allocation and the interaction. In this instance the decision to invest in the Real Estate Investment Trust and the commodity futures fund in the weights chosen.

$$TVA = w(p,i) * r(p,i) - w(b,i) * r(b,i) + SE_((i)) + AE_((i)) + IE_((i)) \tag{3.40}$$

Where:

- TVA= The Total Value Added
- $w(b, i)$ = The benchmark weights
- $r(b, i)$ = The benchmark returns
- $w(p, i)$ = The portfolio weights
- $r(b)$ = The total benchmark return
- $r(bp)$ = The total portfolio return

In focusing on TVA it is possible to see that the way alternative assets are defined and subdivided has implications for their benchmark returns. For example, in real estate, between the different types of property and the way income is realized, either through rental yield and price appreciation. The next step is to explain the different ways the returns of these asset classes can be expressed and calculated.

3.9 Conclusion

This chapter fills a gap in the literature.by providing a taxonomy of benchmarking techniques. In critiquing the various benchmark return measures and index methodology this chapter illustrates the importance of construction technique. Its contribution is in clarifying and in simplification of the choices that an index constructor has to make. In that respect, choices have to be made. Those choices have an impact on the nature of the benchmark and its returns. These choices are either mathematical or subjective trade-offs. This is also the case for alternative asset benchmarks and applications of PCA.

The importance of attribution was introduced. This is essentially what separates the simple benchmark from an index. As investment in alternative asset classes becomes more institutional in nature. Attribution is one of the determinants of whether a benchmark is appropriate. The common index approaches were detailed, both from a theoretical and a practical perspective. The Lespreyres approach was identified as the most appropriate for alternative asset class indexes. The chapter also illustrated that the weightings can be applied using a PCA factor approach, an insight that supports the current phase of innovation in index construction.

The contribution of this critique is in identifying the trade-offs that have to be made in index construction. These were presented through the use of Bodleian logic. These are relevant from the perspective of the question of what makes an index appropriate. In choosing an appropriate index for alternative asset classes, one has to balance breadth and investability, the frequency of re-balancing and rules versus value judgments. Further investigation into the impact of these challenges on alternative asset benchmark construction is therefore warranted.

In conclusion, the various methods presented illustrate the precision and granularity of indices. The choice of method impacts both measurement and attribution. The precision and this ability to attribute and decompose the time series are equally important for the PCA approach, introduced in the next chapter.

Chapter 4

On the use of principal components in index construction

This essay introduces principal component analysis (PCA) as a tool to better understand how alternative asset sub-groupings interact with each other and what lessons can be learnt from the perspective of index construction. It presents an established method for the construction of indices using PCA, but a new application in finance. It is postulated that it can be useful to address a number of issues with alternative asset indices. These include the ability to replicate and the repressiveness of the sample. The method creates a factor model index derived from PCA that delivers identifiable sub-sectors and weightings. The resultant portfolio recasts the efficient frontier and the weights can then be used to construct an index. This can potentially be used for a number of alternative asset sub-groupings. A key finding is that using this approach makes it possible to facilitate synthetic replication of fund risk factors. This has advantages over peer group indices in a fund context. The essays contribution is in giving an holistic overview of how PCA can define and classify alternative asset classes. In this way, it provides useful practical insights for portfolio managers and an avenue of investigation for future academic research.

4.1 Introduction

This chapter builds on the earlier chapters and fills a gap in the literature relating to the application of principal component analysis (PCA) indices. It contributes to that literature by providing a theoretical justification for their use. Finance academics have PCA used as a reduction dimension technique but have not used it to construct indicies for alternative assets. As the literature shows, factor indices exit and these are similar in approach. That said, PCA is distinct from factor analysis. Not only does it explain the diagonal terms of a co-variance matrix or correlation matrix but also the off-diagonal terms. This essay therefore shows how characteristics of an alternative asset and its assigned coefficients can be used in this context to determine instrument weights and construct an index.

This essay explains and expands on the PCA method for the selection of alternative asset sub-groupings and how this can be used to group asset proxies. It builds on the dimension reduction concept proposed by Pearson (1901). He documents a multitude of uses of PCA. This chapter will review its use from the perspective of index construction. It demonstrates the determination of weights in indices through the eigenvector of the first principal component in an alternative asset sub-class. It lays the ground for the method used in the empirical chapters on commodities, hedge funds and real estate investment trusts.

The use of PCA is well documented in disciplines other than finance. It is explained by Jolliffe (2002) in his textbook on the method. As a statistical tool it is used in a number of fields where data is investigated in an exploratory manner. It is used in time series to seasonally adjust data, for example in the analysis of whether PCA can be extended for use in financial time series. From the perspective of this thesis, it is proposed as a method to create uncorrelated indices or components for alternative assets. Specifically, it can be applied where each alternative asset instrument is a linear weighted combination of the set of available instruments. In other words, it as an index.

PCA is an established procedure in academic investigation but has only recently started to be used as a method in finance as a response to over-fitting in traditional multivariate regressions. In economics, it is used to show correlated response and predictor variables and is used in their statistical analysis. PCA has not been used previously to construct indices for alternative assets. That said, it was used in a price context in commodities by Barlett (1948). He applied PCA to the time series of cotton over the period 1924 - 1938 in order to understand the nature of their returns.

PCA is also used in a portfolio context. Partovi and Caputo (2004) demonstrate how it is possible to recast the efficient frontier using PCA from a set of uncorrelated assets. It should be noted, however, that most assets have some degree of correlation. That said, they illustrate how PCA simplifies portfolio structure and delivers a conceptually more transparent solution. Meanwhile, PCA is also applied to the analysis of equity portfolios by Pasini (2017). He proposes that the method can be used to determine an index to test how much a time series departs from being a sequence of independent and identically distributed random observations with finite mean and variance. This is a different application from the one proposed in this thesis, but delivers an interesting insight. He identifies that the first principal component is typically equivalent to the market factor, and that the second principal component is typically sufficient to represent the remaining risk.

4.2 PCA explained

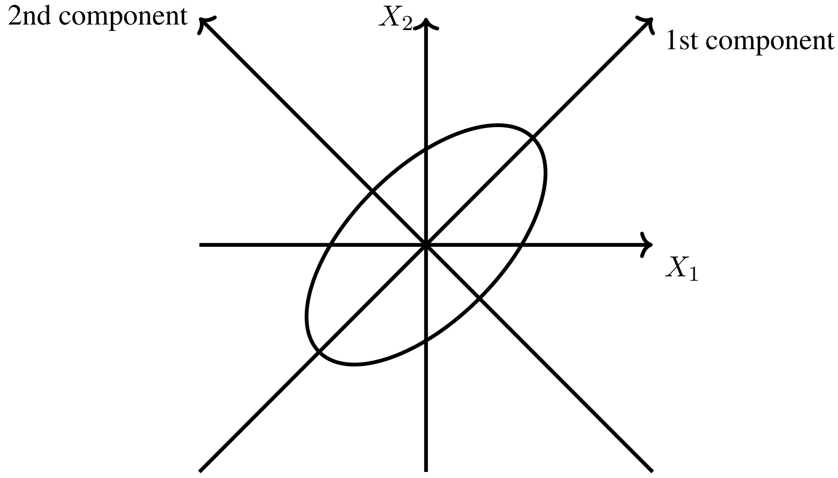
It is easier to understand PCA visually. Figure 4.1 presents a geometric representation based on two variables, X_1 and X_2. These are centered on their respective means. The ellipse illustrates the scatter of sample points. The line that transects the first principal component is derived from the widest point. The second component is the line which is at right angles to this first principal component. The initial reference point is used and a rigid transformation is applied around the origin. This results in a new set of axes. The origin is given by the sample mean average on the two X_1 and X_2 variables.

Meanwhile, figure 4.2 shows the transformed axis. The components in it can be explained algebraically based on the two variables, X_1 and X_2, with the following variance-covariance matrix

$$\boldsymbol{\Sigma X_1, X_2} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{pmatrix}.$$

a_{11} and a_{21} denote the weights from the first eigenvector of $\boldsymbol{\Sigma}$; a_{12} and a_{22} are the weights from the second eigenvector. If can be represented by a 2×2 orthogonal (or rotation) matrix

Figure 4.1: A depiction of the first and second components of an asset class set.



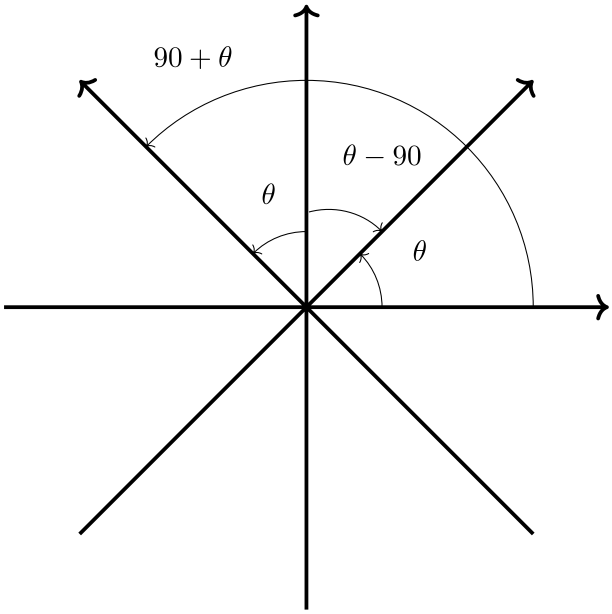
A geometric representation based on two asset variables, X_1 and X_2, showing the first component and second component rotations. In the case of alternative assets these could be the first component in the direction along which the asset instruments have the largest variance. The second principal component is the direction which maximizes variance in those instruments from all directions orthogonal to the first component.

$\mathbf{T}$, with the first column containing the first eigenvector weights and the second column the second eigenvector weights. This then allows the calculation of the direction cosines of the new axes based on the following:

$$\mathbf{T} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \cos(\theta) & \cos(90+\theta) \\ \cos(\theta-90) & \cos(\theta) \end{pmatrix} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

The cosines of the angles are based on the positive (horizontal and vertical) axes. The orientation of the transformed axis can therefore be found by multiplication of the relevant eigenvector values by -1, as depicted in figure 4.1.

Figure 4.2: A PCA transformed axis.



A transformed axis showing the cosines of the angles on the horizontal and vertical axes. The case of two-dimensional rotations can be extended to three or more dimensions by using the appropriate matrix of the direction cosines. In this way, one can build multi-factor models from which to build indices. The axis shows the direction of maximum spread. This is the principal axis. With this it is possible to subtract the variance to obtain the remaining variance. The same procedure is applied to find the next principal axis from the residual variance. The principal axis must be orthogonal to any other principal axes. The transformed data become the principal components.

4.3 PCA usage in alternative assets

The advantage of PCA usage in alternative asset class evaluation is that clusters are easily identified. It overcomes the problems with peer indices identified by Bailey (1992a) and those constructed without attention to correlation, co-variance skew and kurtosis. It also explains variability, and when adapted using factor analysis, correlation. The theorems behind PCA, matrix algebra and multivariate analysis are explained well by Rao (1979), amongst others. It can be used on investment proxies, thereby filling an identified need in the literature as relates to alternative assets.

PCA can help define a more appropriate index for alternative assets, instrument weightings are based on the output. This is also a function of its robustness. A good explanation of how it can be applied is given by Jolliffe (2002). He points out it is a particularly useful method if there is a large amount of data and one wants to view the various sub-groups visually in two dimensional space. As there are a number of sub-groups in alternative asset classes that are very different, this is deemed appropriate. For example, in commodities, metals are very different from agricultural commodities. Similarly, in hedge funds, directional strategies are very different from market neutral.

The PCA approach is useful in alternative assets where sub-groupings may not be immediately obvious. In order to identify if this is the case in alternative assets, data on the sub-groupings of commodities, real estate and hedge funds from 1/12/1990 - 1/6/2018 is analysed. Figure 4.3 illustrates the results. It demonstrates that commodities, hedge funds and real estate have little relationship to each other and can therefore be assumed to be separate asset classes.

The PCA has similarities to a regression model. In this respect, it creates an orthogonal transformation of the individual instruments, thereby better explaining the way they group together. In technical terms it results in a linear transformation of the data at the same time as preserving the statistical symmetry. It does this by taking data representing the first principal component and data representing the second greatest variance from that, and then regroups them. As a result, it can be used on the time series of alternative assets to re-evaluate the variances, co-variances and correlations.

This PCA approach is different from traditional index instrument sampling. That is done

Figure 4.3: No evidence of common components between commodities, real estate and hedge funds 1/12/1990 - 1/6/2018

This figure shows the eigenvalues between the three alternative asset classes, commodities, hedge funds and real REIT's. The lack of common components supports the case that they should be treated as separate asset classes.

through the first components followed by the assigning of weights based on an optimization of the eigenvalue results. As such, it represents an extension of the currently applied construction approaches, done typically on an optimization of the market capitalization as explained in the third chapter.

In order to use the PCA approach as an alternative asset classification method, the common components should be initially identified. An index is then constructed based on the weights of the factors present. This index can be tested to see if the weights are optimal. The construction of indices in this way builds on prior research into was applicable for testing co-integration, demonstrating this in respect of equity indices. As such, it is considered valid to extend it to alternative assets.

In summary, the use of PCA in the time series of returns ensures the resultant indices are robust. PCA is essentially a data reduction technique. It allows the index constructor to create factors than can be meaningfully interpreted and upon which weights can be assigned. The next section explains how the orthogonal transformation occurs.

4.4 Prior PCA usage in indices

There have been precedents in the use of PCA methods to define index constituents by their common attributes. Daniel et al. (1997), for example, argue that such characteristics provide a better ex-ante forecast of the cross-sectional future returns. As such, they argue characteristic identification is a superior way of matching the likely realized returns of an asset class against a benchmark.

The proposed PCA method for alternative assets, as detailed in this chapter, adds to the number of different approaches to construction method found in the literature. In order to be appropriate for a stated objective, the PCA index construction should result in a combination of co-varying securities and/or asset classes as well as being able to be grouped into appropriate sub-sectors. The various additional ways that this can be done is described in Meade and Salkin (1989). An example of one common method can be found in the MSCI Methodology Booklet (2018) MSCI Inc. (2018).

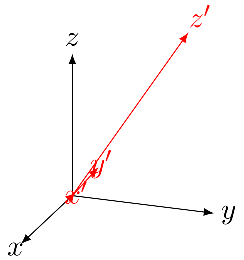
In the literature there are several hierarchical models similar to PCA that are used to create optimal weights, as described by Polson and Tew (2000). They show how they can be used to construct portfolios that can in turn be used as benchmark indices. They detail how Bayesian methods can be incorporated to treat parameter uncertainty, such as missing return data. This approach could be useful in infrequently priced asset classes, such as real estate. That said, most current methods, as explained, rely on representation rather than replication. Amenc, Goltz, and Lodh (2012) explain that in the index replication stage that one should have two steps in the construction process, these being constituent and weighting scheme selection.

A critique of PCA indices is that correlations vary over time. This was addressed by Brown and Warner (1980) who showed that when events are not clustered in time the differences between the various methodologies are quite small. As a result, there is no evidence that alternative methodologies convey any benefit over and above a one factor model. It is therefore proposed that the PCA method is equally as valid a method in the index construction process as any other.

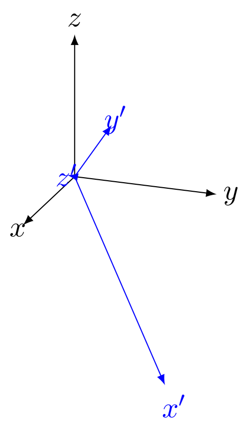
4.5 Orthogonal transformation

In order to understand how PCA can be used as a sampling method to construct an index it is necessary to specify the process. The technique is primarily a data analytic technique, so its use in indices is not widely appreciated. A tutorial is given by Shlens (2005). It uses linear algebra to obtain transformations of the data. These are orthogonal in nature and help with identifying how the data is grouped. The non-orthogonal vectors are depicted in figure 4.4. In index construction, this results in a linear transformation that preserves the integrity of the relationships between the various asset instruments. This allows for weights to be assigned. As explained in the previous chapter, this traditionally is done in index construction through sampling rather than statistical technique.

Figure 4.4: Non-orthogonal 3D coordinate systems.



(a) Vectors in 3d and non–orthogonal basis vectors.



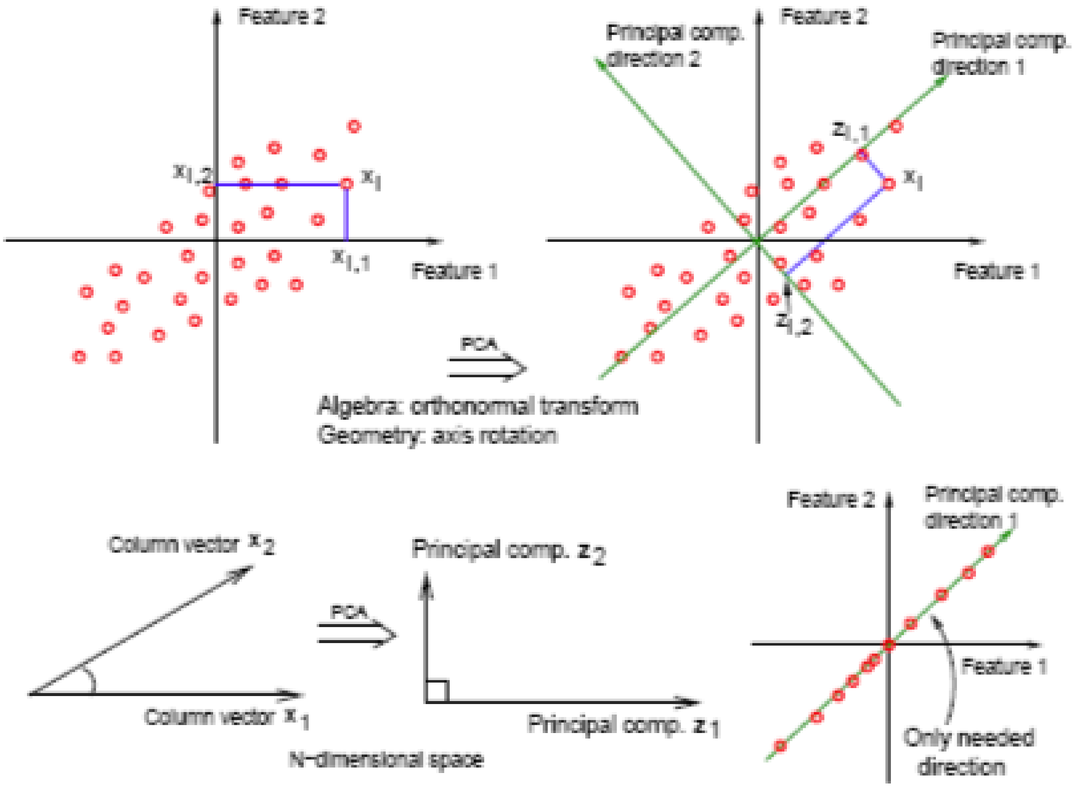
(b) Vectors in 3d and non–orthogonal basis vectors.

Non–orthogonal 3D coordinate systems. Orthogonal is a term used to mean normal . In Euclidean space, two vectors are orthogonal if they make an angle of 90 degrees, or one of the vectors is zero. This figure represents the transformation that a set of asset instruments would go through when PCA is applied.

The orthogonal transformation that alternative asset time series undergo using PCA can be expressed mathematically as a set of p, a dimensional vector of instrument weights. These are called loadings $\mathbf{w}_{(k)} = (w_1, \dots, w_p)_{(k)}$. $t_{k(i)} = \mathbf{x}_{(i)} \cdot \mathbf{w}_{(k)} \qquad \text{for} \qquad i = 1, \dots, n \qquad k = 1, \dots, l$. It is these loadings that form the basis of an index. Creating an index is not a common usage of PCA, but it is a secondary usage as documented by Jeffers (1967) and as used in practice by Chao and Wu (1996).

As part of the index creation, the individual instruments $t_1, \ldots, t_l <$ of t, the alternative asset class constituents, are calculated. This is where the dimension represents the maximum possible deviation from x, with each loading vector representing an index weight w. This orthogonal transformation can be shown diagrammatically, as in Figure 4.5.

Figure 4.5: Representation of an orthogonal transformation



This figure shows how the various principal component transformations look in a graphical two dimensional presentation. Note how the light coloured intersect is through the points of least variability in each of the dimensions. This is an orthogonal transformation and in an index construction context allows for the identification of single instrument weights. Source: STAT 508: Applied Data Mining and Statistical Learning. On-line resource.

As a result of transforming the first loading vector in the way depicted in the diagram, the variance of the individual asset instruments is maximized. The total variance remains the same. It results in a redistribution of the new alternative asset instruments on a different dimension. The outcome is determined by the most "unequal" result. In this way, the first alternative asset not only explains the most variance among the new assets, but the largest variance of any single instrument. This is illustrated mathematically where w equates to:

$$\mathbf{w}_{(1)} = \underset{\|\mathbf{w}\|=1}{\arg\max} \left\{ \sum_i (t_1)^2_{(i)} \right\} = \underset{\|\mathbf{w}\|=1}{\arg\max} \left\{ \sum_i \left(\mathbf{x}_{(i)} \cdot \mathbf{w} \right)^2 \right\} \tag{4.1}$$

Where:

- $w_{(1)}$ = Weighting load factor one.

This is represented in matrix form as:

$$\mathbf{w}_{(1)} = \underset{\|\mathbf{w}\|=1}{\arg\max} \{ \|\mathbf{X}\mathbf{w}\|^2 \} = \underset{\|\mathbf{w}\|=1}{\arg\max} \left\{ \mathbf{w}^T \mathbf{X}^T \mathbf{X} \mathbf{w} \right\} \tag{4.2}$$

Where:

- $w_{(1)}$ = Weighting load factor one.

When the transformation has been made, the next step is to extend the statistical input by the calculation of an additional component. This kth component is found by subtracting the result from the first component. In effect, another rotation is made. With commodities, hedge funds or real estate, this has the effect of splitting out different types of asset groupings. Think of it as potentially isolating different investment styles in hedge funds, different commodities, or different types of real estate investment trust strategies. The equation below shows how this is presented algebraically, highlighting the weighting of the respective identified factor.

$$\hat{\mathbf{X}}_k = \mathbf{X} - \sum_{s=1}^{k-1} \mathbf{X} \mathbf{w}_{(s)} \mathbf{w}_{(s)}^{\mathrm{T}} \tag{4.3}$$

Where:

- $w_{(k)}$ = Weighting load of the Kth factor.

Once the weighting has been identified, the loading factor vector should then be calculated. This is the point of the maximum variance from the new data matrix. It is shown algebraically thus:

$$\mathbf{w}_{(k)} = \underset{\|\mathbf{w}\|=1}{\arg\max} \left\{ \|\hat{\mathbf{X}}_k \mathbf{w}\|^2 \right\} = \arg\max \left\{ \frac{\mathbf{w}^T \hat{\mathbf{X}}_k^T \hat{\mathbf{X}}_k \mathbf{w}}{\mathbf{w}^T \mathbf{w}} \right\} \tag{4.4}$$

The results can then be presented as a set of weights which can be used in an index. These are mathematically expressed as P dimensions. In the construction of an index using this method, the random vector of returns is found from the universe of the relevant alternative assets. This is done with a mean vector where the vector is the *common asset factors* and the matrix of factor loadings are the *specific factors*. Note that this is similar to the output to the market model, which has a common market factor and various stock specific factors.

The creation of a common asset factor mean that the PCA approach has a theoretical link to the market proxy, as found in finance theory. That proxy is derived from the market model and mean variance portfolio theory. As shown in the literature review in chapter two, it can be used to justify broad market indices. The output shows that the variance for the asset equals the sum of the squared outputs for that alternative asset.

Using this approach, the yield structure of alternative assets generates an estimate of the relevant factors from their eigenvectors. That is, it identifies those factors associated with the largest eigenvalues of the matrix output. It is these that form the basis of the weight of the contender alternative asset index, as shall be further explained.

4.5.1 Interpreting the co-variance matrix

In order to create a mean variance optimal index it is necessary to interpret the co-variance matrix in the context of systemic risk. The application of this method represents a contribution to the literature on alternative asset benchmarks. This is because it allows for benchmarks to be constructed without knowing anything other than the return time series. PCA has been used in finance before, so the contribution is incremental and a refinement.

The most important principal component is a proxy for systemic market risk, and the subsequent components contain useful information about financial time series (specific risk factors). Yang, Rea, and Rea (2015) used PCA to interpret a co-variance matrix of asset returns. Interestingly, the last few common components are found to be meaningful, as they identify instruments with near linear correlations. This is relevant because there is an on-going debate in finance about the number of factors that explain return. It implies that with alternative assets one does not have to resort to Capital Asset Pricing Models to identify common factors.

As far as index construction goes, the second eigenvector is a combination of asset weights orthogonal to the first eigenvector and so on. In this way, the factors identify the variance not explained by the first eigenvector. This can be critiqued as difficult to use to identify a specific asset class group, as it means there is no real way of determining the number of eigenvectors without knowing the original number of sub-groupings that the alternative asset class exhibits. Financial industry experience, however, can be used to manually identify these but for the purpose of index creation the first eigenvector is sufficient.

4.6 Deriving factors from principal components

The properties of the PCA output means that it is possible to further derive investment factors. This is done in the same way that one can deduce that systemic risk is an important factor through principal components. One does this by starting with a matrix of the alternative asset class opportunity set. This can be expressed mathematically, for example on a stock matrix with five factor loadings, the model used in subsequent chapters. This as illustrated in the equation below:

$$\mathbf{X} = \mu + \mathbf{LF} + \epsilon \tag{4.5}$$

Where:

- X: vector of the alternative asset class returns. $X = \begin{pmatrix} \mathbf{X_1} \\ \mathbf{X_2} \\ ... \\ \mathbf{X_{100}} \end{pmatrix}$

- μ: X is drawn from a population with mean vector $\mu = \begin{pmatrix} \mu_{\mathbf{1}} \\ \mu_{\mathbf{2}} \\ ... \\ \mu_{\mathbf{100}} \end{pmatrix}$

- L: 100×5 matrix of factor loadings. $F = \begin{pmatrix} l_{1,1} & l_{1,2} & ... & l_{1,5} \\ l_{2,1} & l_{2,2} & ... & l_{2,5} \\ ... & ... & ... & ... \\ l_{100,1} & l_{100,2} & ... & l_{100,5} \end{pmatrix}$

- F: vector of common factors. $F = \begin{pmatrix} \mathbf{f_1} & \mathbf{f_2} & ... & \mathbf{f_5} \end{pmatrix}$

- ϵ: vector of errors (specific factors). $\epsilon = \begin{pmatrix} \varepsilon_{\mathbf{1}} \\ \varepsilon_{\mathbf{2}} \\ ... \\ \varepsilon_{\mathbf{100}} \end{pmatrix}$

Using the PCA approach, the variance for the alternative asset class *i*th is going to be equal to the total of the squared loadings and the variances of the instruments:

$$var(\mathbf{X_i}) = \sum_{\mathbf{j=1}}^{\mathbf{5}} \mathbf{l_{i,j}^2} + \psi_{\mathbf{i}} \tag{4.6}$$

Where:

- $\sum_{j=1}^{5} l_{i,j}^2$: communality of the alternative asset class *i*

- ψ_i: specific factor for the alternative asset class *i*.

The PCA approach is effectively an estimation using the the maximum likelihood statistical method. Although the original loadings are ambiguous, and therefore the instrument factors are not easy to determine, the rotations give a better understanding of how the asset class is broken down. The model that shows this is as follows:

$$\mathbf{X} = \mu + \mathbf{LTT'F} + \epsilon = \mu + \mathbf{L^*F^*} + \epsilon \tag{4.7}$$

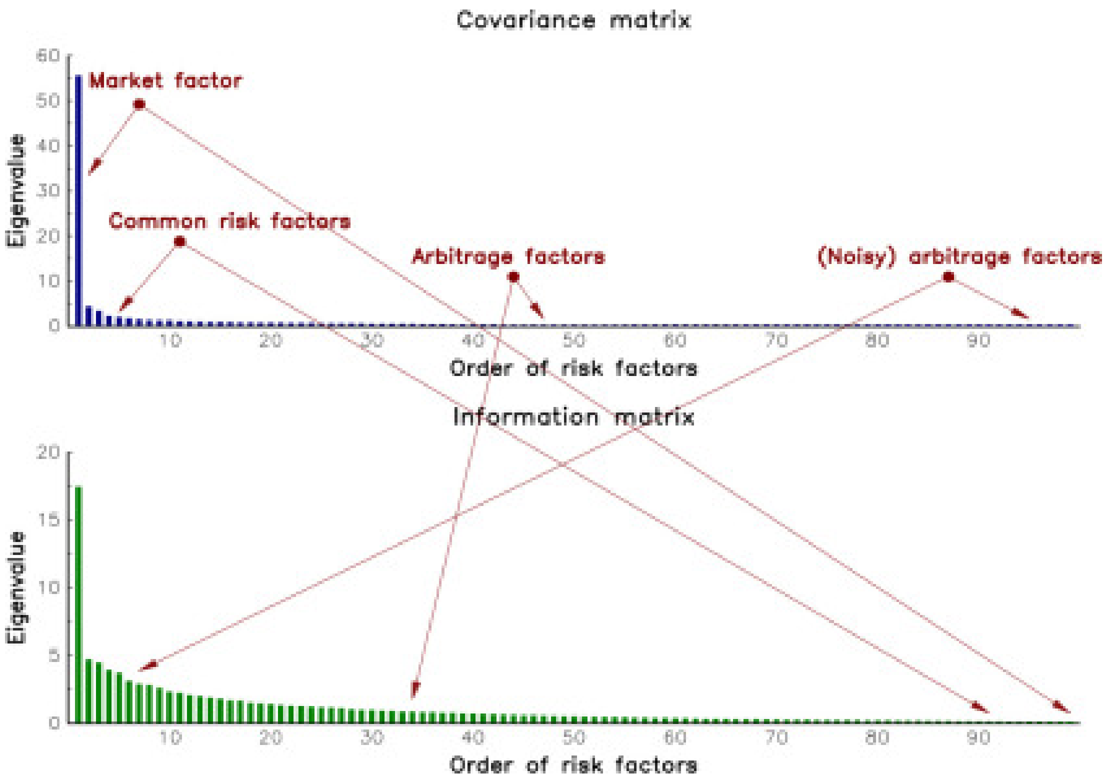
for the orthogonal matrix $\mathbf{T}$.

This approach was first presented in a generalised way by Hendrickson and White (1964). They argue their method should be used in preference to the varimax orthogonal method, suggested by Kaiser (1958).[1] In this way, the oblique solution is effectively obtained by trial and error, increasing the larger loads and reducing the smaller ones. As suggested earlier, a good knowledge of the time series of the asset in question is helpful to an index constructor.

The link with factors and eignevalues, which determines how they can be used in index weights, was explained by Roncalli (2017). The method is applied to the identified index universe and results in a risk factor being generated in the co-variance matrix in the first instance. This is the same as the market risk factor. The eigenvectors that follow are, in this universe of assets, common risk factors. Roncali further demonstrates that the last eigenvectors are arbitrage factors and the link he established is reproduced in the diagram in Figure 4.6.

[1] For the purpose of the subsequent chapters on commodities, hedge funds and real estate the SPSS software was set for promax oblique rotations.

Figure 4.6: Empirical eigenvalues of covariance matrix - index returns



This figure shows eigenvalues of covariance and information matrices of index returns. Note the dominance of the highlighted market factor, consistent with the Capital Market Pricing Theory. Note also the number of common risk factors. In practice, the first five of these are sufficient for a factor model. Source: Roncalli (2017)Figure 1, page 4.

4.7 Re-grouping PCA results into index components

To create an index, the underlying instruments must be regrouped from their raw form. This is done through a variance reduction method. The first step is the application of PCA. The next step is to analyze the data using the variance-covariance matrix for alternative asset returns. The aforementioned literature suggest the results, if applied to alternative asset classes, should have a number of factors that determine the variation of the data. The matrix has been created, it is then necessary to use the output for subsequent index construction by using factor analysis. In this way, it is possible to create an index. A yearly or monthly ranking can be calculated and a re-balancing holding period can be applied.

Once the common components have been established, it is possible to determine the factors present using the associated dimension reduction technique. This is a method for modeling observed variables and their co-variance structure for a small number of underlying un-observable latent factors. It can be considered as an inversion of the PCA. The next step is to create linear combinations of the observed variables. To do this the proposed contender index weights are derived from a factor analysis implemented through a variance-covariance matrix of the returns of alternative asset instrument sets. This is repeated on each date of the new reconciliation.

The results deliver a variance fraction for each of the identified factors. With these results, for each identified factor, the formation of a sub-portfolio is possible. This is based on only instruments with a significant loading to the identified factor. A loading of greater than 0.3 is recommended.

It is suggested that the proposed PCA contender index weights be derived from data observed over annual observation periods. This is for ease of computation. That said, the method can be used to construct alternative asset portfolios held over a shorter re-balancing period. Each portfolio that is created in this way is essentially the index at this time. At the end of this period, any proposed contender index weights are updated and the portfolio re-balanced using the same procedure.

In the next stage of the PCA approach, each alternative investment vehicle receives a weighting equal to the *nth* ratio of its load relative to the sum of the loads of the commodities contained in the sub-portfolio. The resulting group of sub-portfolios is then aggregated into an overall portfolio in which each sub-portfolio receives a weight equal to the ratio of the variance component. This is explained by the factor resulting from the total variance explained by the factors determined.

The experimental factor results can then be put into an oblique rotation. This allows for some correlation between the underlying factors and provides a clearer picture of the variance decomposition. It should show groups of alternative asset instruments as single factors. It stresses the underlying interactions between the factors and should be thought of as similar to creating sectors in traditional index construction.

4.7.1 Deriving index weights

To recap, the proposed PCA contender index is derived through the constituent weights for each period using factor analysis implemented using oblique rotations. The proposed contender index return is therefore a weighted average of the returns of the derived alternative asset portfolio constituents. The resultant factor model can be described as such:

$$PCA_r i, t = \beta 1 i F v_t + \beta 2 F v_t + ...\beta_n F v_t \tag{4.8}$$

Where:

- $PCA_r i, t$ = Proposed contender index, the excess return of portfolio i in month t,
- Fv = Factor identified by eigenvalues

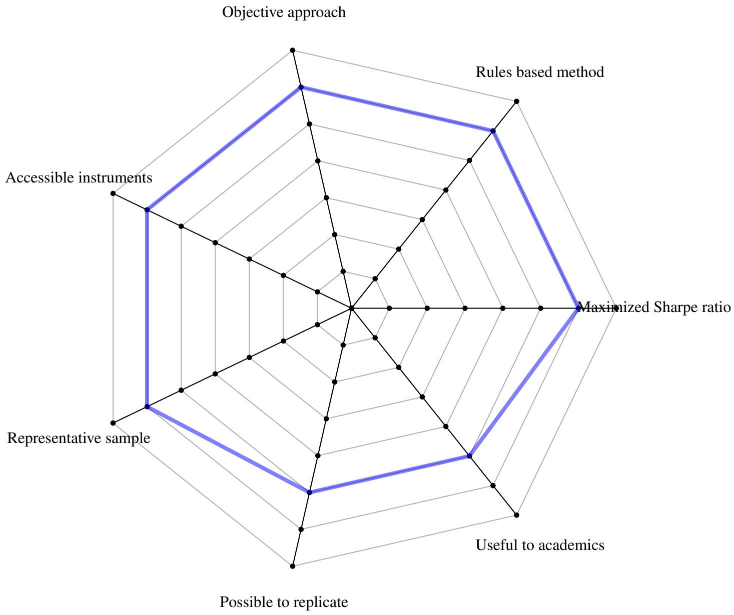
This factor model will be re-introduced in the empirical chapters. It can be used to construct an index in a stepwise fashion with a three-year observation period along with a six-month rebalancing frequency conducted at each rebalancing date.

4.8 Contribution and limitations

A key contribution of the PCA approach is that it explicitly addresses the time series of alternative assets as being non-normal in their distribution. This fills a gap in the literature as this was not previously addressed by scholars. Hubert, Rousseeuw, and Verdonck (2009) demonstrated how robust PCA can be in the face of skewed data. The method is justified in this context because, in order to represent the data in lower dimensions, PCA assumes the joint distribution of data follows a non multivariate normal distribution. Indeed, the only distribution that can represent the time series in a compact form is Gaussian distribution. As a result, the PCA makes an implicit assumption that data should follows Gaussian distribution.

The contribution made by the application of the PCA method to index construction is its ability to allow replication without necessarily requiring direct investment in the underlying instruments. Appendix F includes a diagram that depicts the holistic contribution to the body of knowledge. The method therefore allows for synthetic reproduction of alternative asset class risk factors. The method proposed gives rise to endogenous weights. These can be contrasted with the traditional security weights for traditional assets. The *PCA Derived Index Spiderweb* depicting its dimensions is shown in figure 4.7.

Figure 4.7: PCA Derived Index Spiderweb.



This figure shows a PCA Derived Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. Note that the dimensions are more fully filled out than in other Index Spiderwebs, with five 6 dimension and two 5 dimension scores. Whilst this is still not as good as the Optimal Index Spiderweb introduced in chapter one, it is subjectively better than many existing commercial offerings.

The use of PCA has shortcomings. The biggest limitation, that of requiring expert knowledge of the asset class to identify the factors, has already been flagged. The other limitations identified in the literature are domain shape dependence, the lack of stability, and the presence of sampling errors. As far as alternative assets are concerned, also by the fact that asset prices go up as well as down. In this respect, Wold (1978) points out that as the number of factors approaches the smaller of the dimensions, spurious correlations may occur. This may mean the smaller alternative asset class instruments might get miss-classified.

A final limitation of the PCA approach, identified by Fraley and Raftery (n.d.), relates to computing requirements that grow at a nonlinear rate relative to the size of the groupings. This can limit the size of the data set being analysed when the researcher does not have adequate computing power. As alternative assets have a large number of instruments, this is relevant. The proposed index construction method cannot realistically be done without the relevant software.

An insight gathered from the application of the PCA method is that when testing for the appropriateness of an alternative asset benchmark, a dialectic approach is best. This avoids accepting statistical output at face value. For example, as the real estate chapter illustrates, it is possible to use the PCA approach with REIT's but the resultant index in not practically useful. Its use in commodities, however, proves to be more intuitive, as shown in the chapter on its application in that asset class.

4.9 Conclusion

This chapter fills a gap in the literature on benchmarks. It presented a PCA derived approach to alternative assets and how a universe of alternative assets can be grouped together as an index. Its contribution is in illustrating how alternative assets can be regrouped without proxies. The use of PCA derived indices allows for synthetic replication of factor risk exposures in alternative asset classes, and therefor synthetic index construction.

The chapter showed how PCA derived indicies can be constructed using eigenvector based weighting combined with rules that allow for changes in continuity, context, causality, and consistency. The subsequent chapters present empirical evidence of the appropriateness of such indices for commodities, hedge funds and real estate,

The introduction of the PCA approach extends the literature behind benchmarks, both direct and indirect. The subsequent chapters further extend that literature contribution to alternative assets. The PCA approach is a departure from common practice but represents a contribution to benchmark theory. It is an enhancement of the methodology and an important contribution to knowledge that supports investment benchmarks. The next three chapters apply the method empirically on commodities, hedge funds and real estate proxies.

Chapter 5

Commodity benchmarks, challenges and refinements

This essay investigates the challenges faced in creating a benchmark to measure investment in commodities employing principal component analysis (PCA). There are a number of problems associated with constructing a representative index for commodities as an asset class. Commodities are physical in nature, bulky, and sometimes perishable. As a result, commodity futures are often used as a liquid proxy for direct investment. This derivative exposure results in an additional element to return, namely what is termed the roll yield. As a result, the attribution of returns is not the same as the cash market. Compounding this, the weighting of various commodities in a representative index is currently based on a production based approach. This essay compares the existing indices, which are not representative of consumption based demand, with a PCA contender index. The approach suggested is found to be theoretically appealing and based on its superior performance is also a practical alternative to existing commercial indices. The essay's contribution is in its innovative approach to construction of a replicable index that can be used by fund managers, gaining exposure to commodity futures.

5.1 Introduction

This chapter investigates benchmarks from the perspective of investment in commodities as an asset class. The main aim of this chapter is to test the appropriateness of the principal component analysis (PCA) index approach as applied to commodity futures. The essay builds on the insights extracted from the earlier chapters and investigates whether PCA derived indices can address some of the literature's identified shortcomings in the way existing commodity indices determine weights.

Definition 5.1.1. A commodity is a raw material or primary agricultural product that can be bought and sold, such as copper or coffee - Stevenson (2011).

Commodities can be either perishable or non-perishable, mined or grown. Their importance to the economy means that they are traded heavily. This represents a challenge for determining index weights because production is different from traded volume. Their inclusion in indices is also problematic due to their bulky physical form. As a result of this, in contrast to industrial users, investors often gain exposure through the futures markets. Demand for investment is driven in the long run by the finite supply of commodities and the relationship to consumption. In the short term, demand for commodity futures is driven by institutional asset allocation.

The literature suggests that the characteristics of futures based proxies, used in existing indices, are not optimal for assigning weights. The growing interest in commodity investment necessitates that this be addressed. This increased interest is termed *financialization*, whereby institutional investors are diversifying their portfolios into the asset class. The literature on the phenomena is covered by Cheng and Xiong (2014). Benchmark indices are used by scholars who investigate this in order to measure the impact of financialization on returns. In this respect, benchmark constituents and their weights impact research results.

Index appropriateness is compromised because current futures based commodity benchmarks are constructed using a production based weighting method. These are neither reflective of consumption patterns or investment demand. It is suggested that PCA can address this issue by focusing on the correlation of the individual commodity proxies.

As the second chapter showed, there is a rich literature that uses commodity benchmarks to perform academic investigation. Miffre (2012) identifies three generations of those based on commodity futures. These indices range from the initial poorly conceived long only first generation, through the second generation offering which take into account roll yields, to the third generation indices which take account of backwardization and contangos. Existing commercial indices still, however, have documented failings, thereby justifying this investigation.

The appropriateness of an index is partially shaped by how it reflects investment objectives. In this respect, the determinants of commodity prices were investigated by Arango, Arias, and Flórez (2012). The asset class has attracted the attention of professional fund managers, all of whom require a benchmark. To get an understanding of the nature of institutional commodity investing one can refer to Fabozzi and Kaiser (2011). Also, Yin and Han (2015) provide a useful study in commodity price co-movements.

Increasing investment in commodities has made investigation into appropriate commodity benchmarks more important. There are a number of commodity exchange traded funds that facilitate investment into commodities as an alternative asset class and all these replicate an index. The biggest is the USD 2bn *PowerShares DB Commodity Index Tracking Fund*. This tracks the non Optimum Yield version of the *Deutsche Bank Liquid Commodity Index*. As passive investing replicates an index, the way an index is constructed has implications for investors.

Irwin and Sanders (2011) discussed the role of indexation in the pricing of commodities. They found that although there was empirical evidence on price discovery in the literature, the data and methods used to gauge the the process was open to critique. In this respect, the challenges faced in benchmarking commodity investment is to:

- gain liquid exposure without physical delivery of the underlying commodity.
- reflect what investors want to get exposure to, be it an inflation hedge or a business cycle.
- group commodities into tradable units.
- weight the relative importance of commodities.

It was shown in the earlier chapters that existing indices can be refined to determine their weighting method, as determined by market capitalization. This is done by reference to capital asset pricing theory. Additional refinements can include trading and production volumes.

These are particularly useful adjustments to make to price commodity futures derived indices. Similarly, a commodity index should have a division between hard and soft commodities, as these have distinct influences on co-variance.

In summary, the challenges faced in benchmarking commodities are directly related to their nature and to their being physical assets. Agricultural commodities deteriorate. Mined commodities are bulky and heavy. As a result of this, professional investment is done through a proxy, namely the futures market. Futures do not require storage and delivery, both of which are expensive and require physical space. They are, however, an artificial construct and as such do not accurately reflect the economic significance of the underlying assets. PCA derived indices based on commodity futures can address this in respect of the weighting as they can group the instruments based on how they co-vary. As such, this essay investigates whether this is appropriate for investors and whether it can be done in an optimal way,

5.2 Commodities as an asset class

Although commodities are often assumed to be an asset class, it should be emphasized that in the traditional sense they are not. This is because they are not capital assets in the same way as equities or bonds. Capital assets are financial instruments that facilitate investment in commercial projects. Commodities, however, are used in the production cycle. They are linked to the real economy.

There is, however, a consensus in the investment industry that commodities are indeed a separate asset class. This is supported by the findings of Mongras and Marchal-Domrat (2006). They link this view to their observation that commodities outperformed risk-free returns and exhibited low or negative correlation to other more traditional asset classes. As such, they suggested commodity returns could not be replicated by a simple linear combination of assets. As a consequence, if they are a separate asset class, they need to be benchmarked separately.

Returning to the purpose of a benchmark, investors sometimes invest in commodities as an inflation hedge. This would suggest the *Consumer Price Index* (CPI) could be used as a benchmark. As chapter three explained, this would not provide attribution or be replicable. As such, an index is the best way to measure returns in this alternative asset class.

In addition to changes in the CPI and interest rates, cash commodity prices are driven by physical supply and demand. In some instances cartels can manipulate prices, such as the OPEC cartel, as discussed in Irwin, Sanders, and Merrin (2009). This phenomena was investigated by Markham (1991). They show futures prices are prone to manipulation due to market power, market rigging and/or market rumors. That said, in the majority of commodity time series, prices are determined largely by timing uncertainly in production and market clearing; as well as a variety of exogenous functions, such as the weather.

The nature of supply and demand is not just relevant to price discovery. It is also relevant for commodity index construction. The traditional capital asset pricing framework is a consumption, not a production based model. As such, it is not an ideal theoretical base from which to build a commodity benchmark. Production does not alter proportionally to supply in response to economic cycles or changes in demand.

As argued in the last chapter, a PCA factor approach gives a better insight into price movements than a production based approach. This view is supported by Christoffersen, Olesen, and Christoffersen (2014) who demonstrate the structure of daily commodity futures. They found that the factor structure was relevant to their volatility, meaning that both a factor approach and commodity futures market could be used to index returns.

The existence of economic factors that impact commodities is relevant for investors, and by extension the benchmarks they use. This was investigated by Arango, Arias, and Flórez (2012). They showed that economic metrics and, in particular, real interest rates are the main driving factors behind commodity returns. As a result of this linkage to the economy, there has been a growth in commodity investing by investors with demographic based liabilities. This includes pension and insurance life funds, all of which require suitable benchmarks.

5.2.1 Financialization

Institutional access to futures markets has enabled investment in commodities, and indirectly led to their financialization. The weight of money that this creates has had an impact on market returns, driven by investment in commodity indices. Domanski and Heath (2007) describe the impact of new investors to this asset class. They found that the weight of money was driving up prices during their sample period. They point out that the share of institutional traders rose from around seventeen percent in a decade to around 25 percent. This increase is mainly due to an upward trend in the long positions of so-called non-commercial investors. Their work preceded a commodity market down cycle. It can be observed that the inverse was also true during the bear phase. The weight of institutional money exaggerated subsequent downside market movements.

The process by which financialization impacts markets is relevant to index provision. This is because institutional investors access commodities through index related products. This happens because the futures markets are more liquid than the cash commodity markets. Basak and Pavlova (2016) developed a model to explain this, shown below.

$$dD_kt = D_kt\left[\mu k_dt + \sigma_k d\omega_k t\right] \tag{5.1}$$

Where:

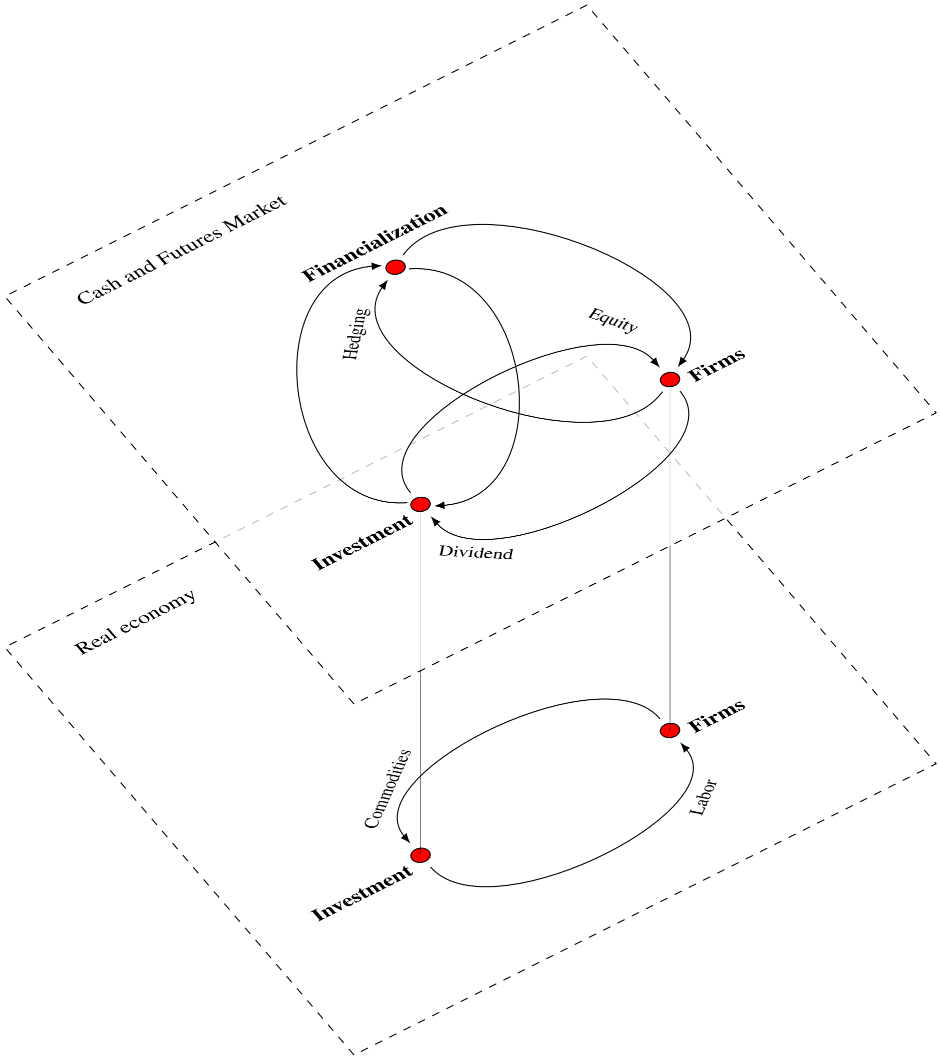
- dD_kt = the arrival of commodity supply news.
- μk_dt = is the terminal value of the process
- σ and ω > 0 are constant

The process of financialization, they argue, is due to a classic demand-side effect of the benchmark. This is depicted graphically in figure 5.1. It shows the linkage between the real economy and the cash and futures markets. It also shows an indirect linkage between the latter and commodities and labour.

The distortions in the return time series by the financialization effect is hard to capture in a benchmark context. That said, benchmarks can be used to evaluate the success of such a process. Institutional investment in the cash and futures markets needs to be benchmarked, but just what element of the return comes from new institutional asset allocations and what comes from more traditional sources cannot be measured by an index.

In summary, Hamilton and Wu (2015) show that investment in commodity futures markets has a price impact due to financialization. This is because the futures market have facilitated the aforementioned institutional investment in commodities and it is now an accepted alternative asset class. Many investors, new to commodities, gain their exposure through a passive index exchange traded fund that invests in commodity futures.

Figure 5.1: Commodity financialization



This figure shows the relationship between commodities, the real economy and the cash and futures markets. The linkage is important because commodity futures are used to benchmark the underlying real commodity market. The cash and futures markets can be used to benchmark investment, measure financialization and show impact.

5.3 Commodity futures

As explained, investment in commodities is typically done through the futures markets. Futures contracts are not homogeneous. Their use has benchmark construction considerations that are distinct from cash commodity benchmarks. In the futures markets, for example, contracts are traded in exchanges that are often in different time zones. On these exchanges, energy and industrial metals have liquid futures contracts that expire every month and in different months. Others, such as agricultural and livestock futures, have only a few contract months per year and are less liquid. Milk, for example, only occasionally trades and has a very stable time series. Crude oil is the most heavily traded, dominating market volume. Crude oil futures are a very different investment proposition from milk futures. Sampling a representative universe for an index therefore presents challenges.

In addition to the variety of traded contracts, indices based on futures are not always reflective of the underlying physical commodity prices. Index futures are sometimes mispriced relative to the underlying cash market. This was investigated by, among others, Yadav, Pope, and Paudyal (1994). The use of futures to construct an index gives rise to expectations based variance in the returns relative to the cash market, as explained by Greer (2000). This is where, in an efficient market, the future range of expected returns can be both positive and negative. This phenomena creates an element of uncertainly in returns that a benchmark cannot capture.

Creating a benchmark based on futures contracts has to take account of the type of investor and their motive. It is typically long duration investors that are interesting in commodities, due to their perceived inflation hedge. In this respect, Gorton and Geert (2006) offer one of the most comprehensive long-term studies on commodities as an asset class. They built a benchmark from the data of the *Commodities Research Bureau*. This used an equally weighted index of 34 commodity futures markets. The data allowed them to measure their index against properties found in traditional benchmarks. More recently, Levine et al. (2016) took a commodities return series as far back as 1844. This was based on the cash market. As such, there is reasonable historical data, but the mix between cash and futures contracts in their construction means that the returns are more indicative than representations of achievable investment returns over that period.

Using futures as a proxy for investment is now fairly well established and it is recognized that there is an impact on investment return. In this respect, the use of commodity futures to construct an index presents unique challenges for their construction and the attribution. The issues in managing funds, that utilize futures against an index benchmark, was first addressed by Elton, Gruber, and Rentzler (1987). They highlight the surprisingly high transaction costs that such funds accrue. This acts as a drag on investment returns, making replicability of returns difficult. This is because the index has to be continually re-balanced to reflect the expiry of contracts.

The expectational variance in the returns of commodity futures indices is a direct result of what is termed the carry or roll. According to Levine et al. (2016), these vary mostly due to moves in the underlying spot price. Although it is possible to devise a benchmark that utilizes futures, it will not be representative of the underlying cash market to which investors want exposure. That said, it may well have a high correlation. As such, indices are useful but have to take account of the roll yield.

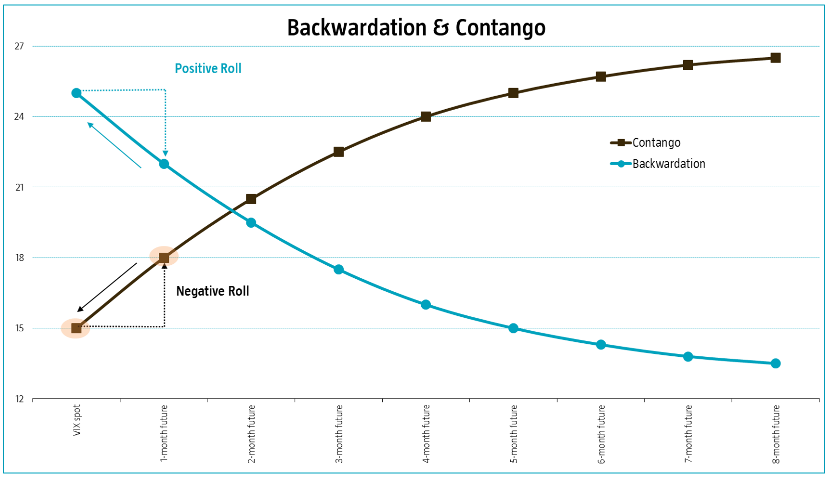
5.3.1 Roll yield

The roll yield is the carry element that is present in the returns generated in futures markets. Investment benchmarking of commodity returns that use futures as a proxy exposes investors to this. The roll yield is affected by what is happening to the supply and demand. Inventory adjustments, the costs of storage and the premium a physical buyer is prepared to pay, all affect the pricing of futures. As a result, commodity futures can trade at a premium (contango) or a discount (backwardization) to spot prices. This is depicted in figure 5.2. This difference impacts the return from selling expiring contracts and buying later dated contracts. The roll yield can and does vary through time.

The presence of the roll yield requires an adjustment in order for a futures based commodity index to have a continuous time series. A protocol to replace expired contracts is required. To do this, commodity futures need to be rolled before their delivery dates. This has a cost, which impacts return. This return is net of the cash collateral. The excess returns has two components, spot and roll returns.

In commodities, the literature on roll yields suggests that the phenomena is part of the term structure of contracts and therefore a function of momentum. This is not something that can be captured in a passive index, although it can be measured by it. Chaves, Kalesnik, and Little (2011) suggest that the term structure of commodities futures has informational content and hence informed investors can analyze them and make predictions. As a result of this, it is self evident that the roll yield needs to be adjusted for in the creation of an index, or else the index itself would contain alpha delivering properties.

Figure 5.2: Contango versus backwardization, a futures derived benchmark consideration



This figure shows the positive roll from backwardization and the negative roll from contango. Roll yield is the return generated in the futures market after a contract is rolled into a longer-term contract. The investor profits from the convergence of the futures price toward a higher spot or cash price.The roll yield has to be calculated and adjusted for in the futures based indices. Source: Google Images.

In constructing a commodity index to reflect the roll yield using PCA, it is possible to build on the work done by Tang and Xiong (2012), among others. They constructed a hypothetical investment position based on taking positions in the first monthly futures contract in a variety of commodities on a fully collateralized basis. This would equate to the base index. The method they employ is to hold futures contracts until the seventh calendar day of their month of maturity before being accepted into the next contract. This is synonymous with the re-balancing of the base index. The excess return of this approach represents the return on surplus futures on initial capital (since capital is still accrued with interest).

5.4 Existing benchmarks

Existing commodity benchmarks include a number of commercial indices based on futures. The *Standard and Poors Goldman Sachs Commodity Index* (GSCI)[1], the *Bloomberg Commodity Index* (BCOM) [2] and the *Dow Jones UBS Commodity Index* (DJCI) are the most used indices by institutional investors. However, these commercial indices partially base their weights on

[1]Standard and Poors Dow Jones Indices. (2018)
[2]Bloomberg Indexes (2018)

global commodity production. This is done to reflect the overall economic importance of the underlying constituents. Some also partially base their weighting on traded volume. Many combine both approaches. All these approaches produce differing return time series and standard deviations, and as such may not reflect the investment goals of investors.

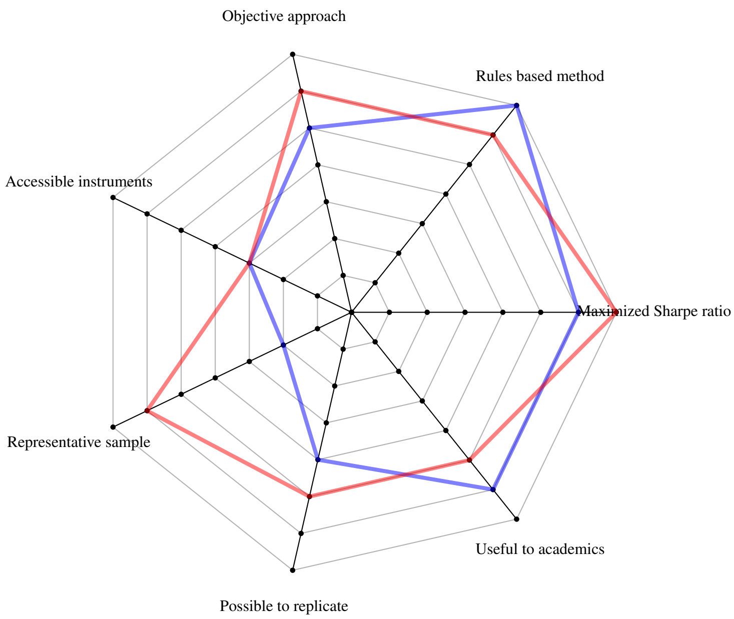
The most widely used index is the *Standard and Poors Goldman Sachs Commodity Index* (GSCI). This index uses an average production metric for its groupings over a five year period. The next most popular is the *Bloomberg Commodity Index* (BCOM), largely because it is available for free from the Bloomberg terminal. It is also based on production data, although annual in this instance. It incorporates a traded volume weighting adjustment which is a refinement that is useful for institutional investors. The third, the *Dow Jones UBS Commodity Index* (DJCI), also takes account of liquidity in a two for one ratio to production. This is excluded from the analysis as it combines both production and volume weighting. The PCA approach does not take account of liquidity and as such this would present an exogenous factor.

There are a number of investment strategies indexed to current commercial indices, including exchange traded funds, various notes and return swaps. Collective funds similarly require proper benchmarking. Such investment funds and products are linked with the aforementioned commercial indices. Schneeweis and Spurgin (1997) provide a useful comparison of the various indices and other commodity future benchmarks. They observe that the risk return characteristics of the various indices differ even though they gain exposure to the same underlying asset class. They conclude that scholars should expect more commodity futures indices to be created, further justifying the investigation into PCA indices for this asset class.

The problem with existing benchmarks is captured by Domanski and Heath (2007). They make the observation that "in the longer run, production can be changed and the elasticity of commodity supply depends on the marginal cost of production. In the short run, supply from production is relatively inelastic and depends more on above ground stocks. With the exception of gold, above ground commodity stocks are small relative to demand." As most common commodity indices are based largely on such consumption metrics, they could be deemed to be less than optimally weighted. Figure 5.3 shows the existing commercial indices on a *Commodity Index Spiderweb*. It contrasts these with the PCA factor derived index used in the empirical evaluation later in this essay.

In the *Commodity Index Spiderweb*, explained in more detail in chapter one, the PCA index and the commercial offerings both achieve the same dimension in instrument accessibility. The biggest difference is in the representative sample. The groupings on the PCA analysis score more highly.

Figure 5.3: PCA versus a commercial Commodity Index Spiderweb.



This figure shows a PCA derived index (red) versus a commercial Commodity Index Spiderweb (blue). Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. The commercial commodity indices perform better on the Likert scale in dimensions one and four. Dimension three is the same as both indices use the same commodities futures contracts. The PCA derived index performs better in the remaining dimensions.

As commercial financial products invest in fully collateralized commodity futures, any index has to be based on these instruments. The index constituents and futures contracts selected are therefore critical to relative performance.

5.5 Data

Data was selected in order to construct a contender index by deriving factors using PCA. Twenty four commodity time series were identified using Bloomberg. These formed the basis of analysis. The data extracted from Bloomberg comprised daily futures prices of commodities based on the same ones that form the *Standard and Poors GCSI Index*. This set includes a range of both soft and hard commodities such as wheat and corn, copper and crude oil. It should be noted that some commodities, like wheat and oil, trade in different forms in different venues.

Additional data from Bloomberg included daily prices of a comparison set of major commodity indices. The data used incorporated daily observations from 18/1/2008 to 12/2/2016, with a ranking period of 756 days (i.e. three years before first observation date). The return data, as part of the next stage, was selected, and adjustments were made for the roll yield (see section 5.6). The three comparison indices were:

1. The *Standard and Poor's GSCI* (GCSI). The main index for the quantitative analysis. This index assigns liquidity and production weights to its constituents. It is a long and unleveraged index. As a broad commodity portfolio, it is also well diversified. It is fully collateralized and designed so that as each commodity future rolls over and it is fully reinvested.

2. The *Bloomberg Commodity Index* (BCOM). This index has the easiest access. This index weights liquidity and production. No commodity accounts for more than 33 percent of the index, and no single commodity makes up more than 15 percent. During the rolling period defined in the methodology, the index rolls from current to subsequent contracts, which are normally within the 6th to 10th business days.

3. The *Thomson Reuters Commodity Index* (CRY). This index is the oldest. It was founded in 1958 and is the oldest index made up of 19 commodities. These are divided into four groups, including crude oil products, various raw materials and so-called liquid and highly liquid raw materials.

The GSCI component subsections were selected based on the popularity of this index. These were used to determine the PCA weights using factor analysis (based on the work of Lawley and Maxwell (1971)). Thereafter, common component weights were applied, as derived from the PCA. The outputs of the analysis was used to create a factor model time series.

Initial investigatory analysis on the data was conducted in order to address the Domanski and Heath (2007) concerns about gold not being correlated to commodities in general. As part of the investigatory analysis, a regression was made on the gold price in USD as the dependent

Table 5.1: Gold price (USD) versus Bloomberg Commodity Index, 1/1/2000 - 1/1/2020

Raw BETA	0.443
Adjusted BETA	0.629
ALPHA (Intercept)	0.182
R^2 (Correlation^2)	0.174
R (Correlation)	0.417
Std Dev of Error	2.111
Std Error of ALPHA	0.065
Std Error of BETA	0.03
t-Test	14.782
Significance	0
Last T-Value	0.7
Last P-Value	0.758 ***
Number of Points	1042
Last Spread	-1429.21
Last Ratio	18.568

*This figure shows the regression: Dependent variable: Gold price (USD), Independent variable: Bloomberg Commodity Index, 1/1/2000 - 1/1/2018. The correlation R shows the absolute value of the correlation coefficient between these variables. *** P value shows a probability that the coefficient of the independent variable is not reliable.*

variable versus *Bloomberg Commodity Index* as the independent variable using the time series 1/1/2000 - 1/1/2020. As can be seen, the R squared correlation is 0.174. The p value of 0.757 shows this is statistically relevant. This lack of correlation suggests that gold prices co-varies differently from the prices of other commodities, as per the Domanski and Heath (2007) findings. The results are shown graphically in figure 5.2. It is suggested that the way this sub grouping co-varies further supports the PCA approach. This is because PCA will group precious metals together, thereby determining an optimal weight for their inclusion in an index.

The same exercise was undertaken for crude oil, as it is the most important commodity by production volume. a regression was made on the crude oil price is USD as the dependent variable versus *Bloomberg Commodity Index* as the independent variable using the time series 1/1/2000 - 1/1/2020. The R squared in this case is 0.521, with a p-value of 0.758. This also suggests a separate sub-grouping will arise using PCA. The beta is 1.635, which as the largest commodity by production suggests its weight in indices will vary considerably over time. The results are shown graphically in table 5.2.

The initial analysis confirms that an index built to reflect production has different charac-

Table 5.2: Oil price (USD) versus Bloomberg Commodity Index, 1/1/2000 - 1/1/2020

Raw BETA	1.635
Adjusted BETA	1.423
ALPHA (Intercept)	0.192
R^2 (Correlation^2)	0.521
R (Correlation)	0.722
Std Dev of Error	3.424
Std Error of ALPHA	0.106
Std Error of BETA	0.049
t-Test	33.608
Significance	0
Last T-Value	-0.006
Last P-Value	0.498***
Number of Points	1042
Last Spread	19.63
Last Ratio	0.759

*This figure shows the regression: Dependent variable: Oil, Independent variable: Bloomberg Commodity Index, 1/1/2000 - 1/1/2020. The correlation R shows the absolute value of the correlation coefficient between these variables. *** P value shows a probability that the coefficient of the independent variable is not reliable.*

teristics from one designed to reflect consumption. The different drivers of supply and demand impact the price discovery. This is compounded by the dominance of crude oil in the weighting of raw materials in production, which in turn reflects the world's economic dependence on fossil fuels. Meanwhile, gold, though heavily traded, has only a small weighting, as production is only a small part of the supply. The *World Gold Council* estimates that total above ground stocks of gold are approximately 190,040 tonnes, with below ground reserves of only 54,000 tonnes.

5.6 Method

A PCA index was constructed using the method introduced in chapter four. The underlying observed random variables were taken from the sourced data based on the futures contracts. These were then expressed as linear functions of the common factors present. In order to do this, the data was subject to dimension reduction, so as to derive factors based on the eigenvalues. The output was then regrouped in order to determine weights for an index.

It is observed that the PCA and factor model approaches have different ways of treating multivariate data. As such, the method used to derive the factors is based on Flury (1984). The analysis of the variance-covariance matrix of the commodity futures returns is then implemented over the whole observation period (2005-2016). The results using this method are not dependent on normality assumptions.

After completing the dimension reduction, the portfolio weights were derived. For each identified factor, commodities with significant common components were considered (loading $\geq |0.3|$). These were then included in a sub-portfolio with the other commodities exhibiting the same factor. This was identified with its weight. For each factor a sub-portfolio of only commodities with a significant contribution (loadings $\geq |0.4|$) were included with weight using the following formula:

$$w_i = \frac{\hat{l}_{i,j}}{\sum_{i=1}^{p} \hat{l}_{i,j}}, \tag{5.2}$$

Where:

$\hat{l}_{i,j} \geq |0.3|$; i: 1, 2, ..., p; j: 1, 2, ..., 5

p: number of commodities with loading on factor *j* $\geq |0.4|$

w_i: weight for commodity *i* in sub-portfolio *j*

$\hat{l}_{i,j}$: estimated loading for commodity *i* on factor *j*.

The resulting five sub-portfolios were then combined into a commodity super-portfolio with individual weights:

$$w_j = \frac{\hat{\lambda}_j}{\sum_{j=1}^{5} \hat{\lambda}_j}, \tag{5.3}$$

Where:

w_j: weight for sub-portfolio *j* in commodity super-portfolio

$\hat{\lambda}_j$ = eigenvalue for factor *j*

$\sum_{j=1}^{5} \hat{\lambda}_j$ = total communality.

In order to create a base date, the PCA index at time t = 0 was scaled to equal 100. The performance series for each commodity reflected a constant investment, so that the use of futures contracts did not require either the supply or the withdrawal of funds.

The commodity portfolio was held over the holding period. At the end of this period, the weightings of the constituents was updated and the portfolio was re-balanced using the same procedure. The commodity proposed contender index return was then calculated as the weighted average of the portfolio constituents' return. The output can be expressed as a parsimonious factor model as follows:

$$\zeta_r i, t = \beta_t^{'} \cdot F_t, \tag{5.4}$$

Where:

$\zeta_r i, t$: factor model index return for period *t*

β_t = vector of factor sub-portfolio weights for period

F_t = vector of factor sub-portfolio returns for period *t*.

As per the earlier discussion on the roll yield (section 5.3.1), the time series generated for each commodity represents a constant investment, meaning that the rolling of futures contracts require neither adding or withdrawing investment. In this way, the time series of the futures contracts were adjusted for the roll yield. To do this the front month and expired month prices was derived from the settlement price based on the roll date of both contracts. This ratio was then multiplied by the entire historical pricing so each month the ticker rolls, the entire historical pricing adjusts. Additionally, a new ratio was calculated and multiplied every time a new contract rolled (to readjust the entire time series for new data).

5.7 Results

The results identified the presence of five main principal components that drive the variation in the data. These five factors were then used for the subsequent construction of the contender index using factor analysis.

The return of the competitor index is the weighted average of the returns of the portfolio constituents. This results in a five-factor vector model that is described as such:

$$\zeta_r i, t = \beta 1 \mathbf{v} \eta_t + \beta 2 \mathbf{v} \theta_t + \beta 3 \mathbf{v} \chi_t + \beta 4 \mathbf{v} \tau_t + \beta 5 \mathbf{v} \mu_t \tag{5.5}$$

Where:

- $\zeta_r i, t$ = Proposed contender index, the excess return of sector commodity portfolio i in month t, item **v** = an eigenvector
- η = Grain complex
- θ = Industrial metals complex
- χ = Livestock complex
- τ = Energy complex
- μ = Precious metals complex

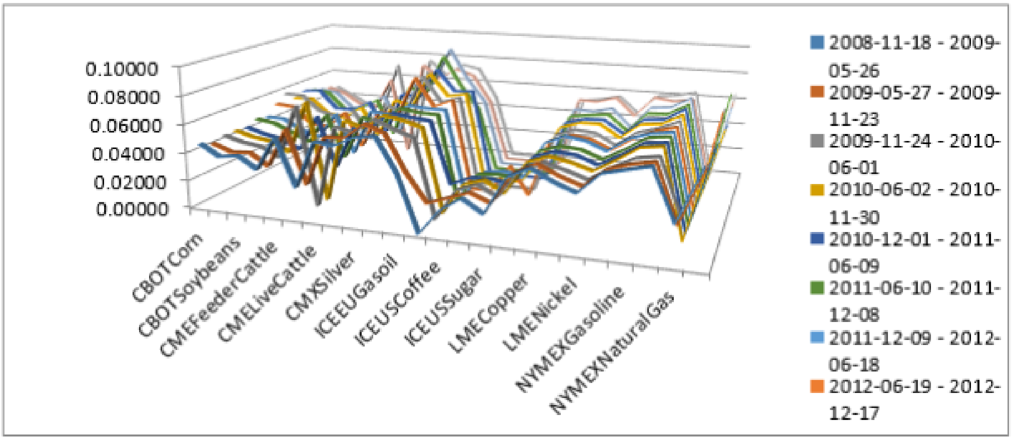
The model represents a factor mimicking portfolio as calculated in month t, and is *i*, *t* the residual return of portfolio i at time t. It captures the aforementioned way energy (including crude oil) and precious metals (including gold) co-vary. The results confirm that using principal components eigenvectors to determine the weights of an index is a valid one. The time series returns differ depending on the underlying commodity contracts. The volatility also differs dramatically.

5.7.1 Results of the PCA index weighting exercise

The results of sector loadings based on the model are shown at the end of this chapter in table 5.3. As can be seen, and as surmised by the high beta of crude oil, there are bug variations in the mean weights over time. Table 5.4 illustrates the resultant PCA vector factor weights in each of the discrete observation periods.

The results of the PCA index weighting exercise are shown diagrammatically in figure 5.4.

Figure 5.4: Dynamic PCA vector factor weights



This table illustrates the changing weights of the PCA contender index over the sample period. Note that although distinct, the time series are directionally related. Oil and natural gas consistently have the highest weights.

If a sub selection of the 2016 data is taken, as shown in table 5.3, mean returns ranged from +13.26 per cent for *NYMEX Gasoline* to – 33.05 per cent for *NYMEX Natural Gas*. This highlights the earlier point (section 5.5) about the beta of crude oil relative to broader commodities.

Table 5.4 (at the end of this chapter) highlights the descriptive statistics for the factor model, as introduced in the previous section. These are compared with the index returns of existing indices, as detailed in the data section.

The weights changed over the sampled time horizon. Table 5.5 (at the end of this chapter) shows the impact on the non zero weights. The PCA index had a minimum of 18 commodities and a maximum of 19, with a standard deviation of 22.36. Table 5.6 (at the end of this chapter) presents the calendar returns.

The index return time series obtained were compared to a number of important commodity indices using the information ratio This is shown in table 5.7 (at the end of this chapter). This is often referred to as a variation or generalized version of the Sharpe Ratio. It developed when users of the Sharpe ratio replaced the risk-free rate with passive benchmarks. The information ratio was also calculated. In this case, this benchmark was superior to any of the three aforementioned commercial indices in terms of having a positive alpha.

Once this tabulation was done, the results were compared to the three indices most commonly used by commodity investors (see again Table 5.4). A sample of 33 commercial indices was used. This was done so as to not to miss any of the existing dynamics.

The weights generated by this method of analysis are dynamics. They do not change dramatically over time. This means that the index method is investable as it can be replicated, one of the criteria for a solid index. An analysis of the weights of the proposed competitor index showed that gold is a particularly interesting observation, as it has one of the lowest volatilities but has one of the highest open interests. It is grouped with metals, but not unique enough to form its own group. The weight of the metals in the proposed competitor index did not fluctuate as much as anticipated.

The weights structure of the PCA contender index remained fairly stable throughout the period of observation (see figure 5.4). The results deliver narrow weight ranges and low weight volatilities at the individual commodity and sector levels (refer to table 5.5). The eigenvalue weights suggest a range widths from 0.38 to 2.04 percentage points for LME copper and CME live cattle respectively while volatilities range from 0.13% for LME copper to 0.51% for CME feeder cattle with the notable exceptions of CBOT soybeans and NYMEX natural gas that bounce in and out of the index[3]. Individual commodity weights are also fairly homogeneous and thus rather small in levels with no weight or average weight ever getting close to the 10% threshold. Individual commodity average weights range from 0.27% for NYMEX natural gas to 7.03% for CME live cattle. This range narrows down when only accounting for the permanent index constituents with minimum and maximum weights of 4.11% and 7.03% for ICE EU gasoil and CME live cattle respectively. CME lean hogs and the softs get no weight at all in the index over the period of interest.

The range of individual commodity weights is much wider for the indices of comparison with the BCOM weights ranging from 1.38% (ICE US cotton) to 12.6% (CMX gold), those for CRY from 1% (CBOT SR Wheat, CME lean hogs, CMX silver, ICE US orange juice and LME nickel) to 23% (NYMEX WTI) and those for the GSCI from 0.41% (CMX silver) to 23.04% (NYMEX WTI).

[3]This weight ranges widths and volatilities for CBOT soybeans and NYMEX natural gas are 3.48 & 2.77 percentage points and 0.74% & 0.84% respectively.

Active sector weights in the proposed PCA contender index range from 1.22 for livestock to 3.21 percentage points for industrial metals with volatilities ranging from 0.29% for livestock to 1.02% for industrial metals. The energy sector shows an average weight of 29.1%, followed by industrial metals (25.09%), grains, (20.8%), livestock (12.96%) and precious metals (12.05%). For sectors again, the indices of comparison show wider ranges weights. For the BCOM they range from 5.7% (softs) to 26.7% (energy), while for the CRY they range from 7% (livestock and precious metals) to 44% (energy) and for the GSCI from 3.65% (precious metals) to 63.05% (energy).

The active key-period returns, tracking error, and information ratios are calculated (table 5.7). Throughout the sample period 2008-2016, the information ratio of the proposed competitor index against the active positive risk-adjusted GCSI is 3.34 percent, with an active return of 3.98 percent. In this regard, the proposed competing index achieves an active, risk-adjusted return compared to the GCSI. The results for BCOM and CRY are impressive, resulting in an active positive risk-adjusted return of 6.65 percent and 5.72 percent, respectively.

The active risk-adjusted return is positive versus all three indices in six out of nine discrete year periods. The proposed competitor index did not fare so well in 2009 and 2011 against the GCSI and achieved an active negative risk-adjusted return of -20.19 percent and -7.08 percent. In both years, however, it achieves a positive active risk-adjusted return over the other two indices.

In the next step, the tracking error between the proposed competitor index and the three indices is compared. There is a tracking error between each index created by another method. The tracking error measures the spread of portfolio returns versus benchmark returns. This is best when a benchmark is considered a market index. The tracking error is simply the standard deviation of the active return, where the active return is defined as the difference between the competitor index return examined in this case and the return of the benchmark index. To regard the proposed contention index as being different from existing models, the time series of the return should produce a quite different tracking error.

The proposed contender index produces a tracking error of 11.91 percent over GCSI Index throughout the 2008-2016 sample period. A high ex-ante tracking error such as this means that the proposed contention index will be more likely to deviate from the GSCI return. It reflects the risk situation of the proposed competitor index compared to the GSCI. The ex ante tracking error of the proposed competitor index is essentially a function of its weightings and the correlation between the time series of the commodity futures. The complete scan tracking error of the BCOM is 5.55 percent and the CRY is 0.06261. This implies that there is more difference between this research method than between the GSCI, but there is still a significant difference between the two other indices.

The tracking error results are surprisingly volatile. GCSI has a tracking error of 3.19 percent in 2008 and only 0.08910 in 2012. The other two indices are less extreme, with BCOM and CRY peaking at 11.63 percent and 14.22 percent in 2008, respectively. It's interesting to see the difference between the proposed rival index and the GCSI, as the next step was to use the same constituent commodity futures as this index. This supports the earlier research on the use of the production-based method as the GCSI which is the only one of the three that relies solely on this weighting.

Throughout the sampling period, information ratios in the larger sample ranged between -0.16940 versus the *SDCITR Index* and an impressive 1.06420 against the *DCIBGLTR Index*. Nine of the information ratios are positive and only one is negative. The strength of the positive results with an information ratio above 0.75 is considered good.

It should be noted that benchmark indices should by definition have an active return of zero. The presence of such an alpha means that a superior return is possible or that sampling is not comprehensive enough. In the next stage, it should be noted that the GCSI, covers only 18 out of the total of 32 traded commodities contracts

As with Rallis, Miffre, and Fuertes (2013) the next step was to show a significant active positive risk-adjusted return compared to the GSCI and the UBSCI. This means that they are not optimally weighted, rather than that this research index provides alpha.

Examining the absolute performance of the commodity proposed contender index, BCOM, CRY, GSCI and an equal-weighted portfolio of the twenty-four commodities under study using standard performance measures, including annualized return, annualized volatility and corresponding Sharpe ratio as well as maximum draw-down.

In order to evaluate the diversification and inflation hedging features of the proposed compound commodity index, the pairwise correlations with the *Standard and Poors 500* had to be calculated, the BUSY index and unexpected inflation measured as the difference between realized and expected inflation. As in Rallis, Miffre, and Fuertes (2013) who model the later using two methods. First, an *ARMA(1, 1)* model is used to account for quarterly inflation (change in the CPI in the US) and make its forecast expected inflation. Second, as in Erb and Harvey (2006), a random inflation number is assumed, and taking into account the seasonality the expectation for inflation at quarter t, the inflation level at $t - 4$ (same quarter of the previous year), the results are again calculated and proved robust in terms of alpha generation.

5.7.2 Results of the test for index superiority

The PCA derived index was tested for superiority based on the null hypothesis introduced in chapter one. This suggests that that it should not be possible to out-perform the existing commercial alternative on a risk adjusted basis.

Null Hypothesis 2.

$$H_0 : GRST_t = \eta_i - \eta_j = 0$$

Where:

- $\eta_i = \mu_i/\sigma_i$ = Mean excess returns i (over a risk-free rate)/standard deviations of excess returns i
- $\eta_j = \mu_j/\sigma_j$ = Mean excess returns j (over a risk-free rate)/standard deviations of excess returns j
- $GRST_t$ = is made up of an estimator of the Sharpe ratio difference, the standard error of the difference estimator, the Sharpe ratios and the correlation between the excess returns of strategy i and j.

Table 5.7 shows the GRS test rejects H_0 hypothesis of zero risk-adjusted returns with a p-value of close to 0.000 for the Bloomberg commodity index and the Thomson Reuters/Core index. This suggests the PCA derived index is a valid contender.

There are many avenues for further investigation stemming from this research. For example, one could explore the nature of benchmarks for CTA funds which have both long and short futures exposure. This is illustrated by the low correlation between the *CTA Index* and the *Bloomberg Commodities Index*. A closer investigation of volume and open contracts is merited, as evidenced by the impact these two additional metrics have on the underlying dispersion of active returns in those funds that employ such adjustments and those that do not. That said, overall. PCA indices are valid contenders to existing commodity futures based benchmarks.

Table 5.3: Standard deviation and sector loadings of rolling factors

	Factor 1		Factor 2		Factor 3		Factor 4		Factor 5	
	Mean	σ	**Mean**	σ	**Mean**	σ	**Mean**	σ	**Mean**	σ
Brent	0.93	1.31%	0.19	5.15%	0.17	2.92%	0.10	2.41%	0.11	3.26%
Gasoil	0.59	3.11%	0.26	7.92%	0.09	1.95%	0.07	3.31%	0.12	3.26%
Gasoline	0.84	1.25%	0.18	7.92%	0.13	2.82%	0.07	2.46%	0.06	2.04%
Heating oil	0.89	0.54%	0.18	4.50%	0.15	1.81%	0.06	2.33%	0.09	2.73%
Natural gas	0.27	7.68%	0.03	1.33%	0.07	3.55%	0.05	1.62%	0.04	1.79%
WTI	0.88	3.48%	0.21	5.83%	0.17	3.32%	0.10	2.10%	0.10	2.58%
Energy	0.73	23.91%	0.18	8.82%	0.13	4.85%	0.07	3.06%	0.09	3.86%
Corn	0.19	5.12%	0.13	4.25%	0.61	4.28%	0.09	4.38%	0.09	9.48%
HR wheat	0.10	4.69%	0.09	2.84%	0.95	0.74%	0.03	1.75%	0.05	1.79%
Soybeans	0.26	4.82%	0.20	6.23%	0.46	2.91%	0.09	3.95%	0.12	6.81%
SR wheat	0.10	4.29%	0.09	2.66%	0.98	0.60%	0.02	2.10%	0.05	1.80%
Grains	0.16	8.13%	0.13	6.06%	0.75	22.14%	0.06	4.47%	0.08	6.71%
Aluminum	0.19	3.34%	0.72	1.17%	0.11	2.43%	0.09	5.58%	0.14	3.91%
Copper	0.21	4.50%	0.82	0.54%	0.14	3.95%	0.11	4.30%	0.16	3.77%
Lead	0.14	4.57%	0.73	4.19%	0.10	3.21%	0.04	2.46%	0.08	3.25%
Nickel	0.17	5.17%	0.66	2.63%	0.11	4.08%	0.08	2.16%	0.08	1.88%
Zinc	0.12	4.42%	0.83	1.36%	0.10	2.56%	0.06	4.23%	0.11	2.94%
Metals*	0.16	5.48%	0.75	6.99%	0.11	3.62%	0.08	4.66%	0.12	4.41%
Feeder cattle	0.08	5.15%	0.04	3.00%	-0.09	5.42%	0.71	24.62%	0.10	29.90%
Lean hogs	0.07	0.89%	0.08	2.64%	0.09	3.47%	0.31	10.48%	0.04	8.55%
Live cattle	0.08	3.41%	0.05	2.11%	0.12	5.56%	0.89	28.18%	0.11	22.75%
Livestock	0.08	3.60%	0.06	3.12%	0.04	10.35%	0.63	32.89%	0.08	22.12%
Gold	0.16	3.03%	0.23	4.44%	0.13	1.57%	0.07	25.58%	0.70	23.40%
Silver	0.18	1.97%	0.29	2.34%	0.15	2.38%	0.12	24.41%	0.81	25.96%
Metals**	0.17	2.76%	0.26	4.86%	0.14	2.42%	0.09	24.83%	0.76	24.98%
Cocoa	0.14	5.66%	0.21	3.65%	0.13	3.23%	0.07	4.78%	0.14	6.22%
Coffee	0.15	4.34%	0.20	4.62%	0.20	4.78%	0.10	3.04%	0.13	3.41%
Cotton	0.19	3.38%	0.21	5.43%	0.25	3.94%	0.09	3.91%	0.10	3.58%
Sugar	0.18	3.79%	0.20	2.75%	0.22	3.99%	0.09	3.52%	0.08	5.04%
Softs	0.17	4.64%	0.20	4.18%	0.20	6.12%	0.09	3.93%	0.12	5.25%

This figure displays means and standard deviations for individual commodity and sector loadings on the first five factors corresponding to the rolling factor analysis, 18/1/2008-12/1/2016, with daily observations.

Table 5.4: Descriptive Statistics for Factor Model Index

	Panel 1. Contender index				Panel 2. indices			
	Min	**Max**	**Mean**	**Alpha**	**BCOM**	**CRY**	**GSCI**	**E.W.**
Brent	6.32%	6.95%	6.51%	0.16%	6.59%	0.00%	20.43%	4.17%
Gasoil	3.61%	4.54%	4.11%	0.20%	0.00%	0.00%	5.82%	4.17%
Heating oil	5.87%	6.52%	5.85%	0.21%	3.28%	5.00%	5.21%	4.17%
Gasoline	5.38%	6.23%	6.22%	0.15%	3.39%	5.00%	5.31%	4.17%
Natural gas	2.68%	2.76%	0.27%	0.84%	7.49%	6.00%	3.24%	4.17%
WTI	5.86%	6.96%	6.14%	0.30%	5.95%	23.00%	23.04%	4.17%
Energy	28.30%	30.50%	29.10%	0.77%	26.70%	44.00%	63.05%	25.00%
Corn	3.76%	4.55%	4.28%	0.22%	8.13%	6.00%	4.23%	4.17%
HR wheat	5.97%	7.73%	6.63%	0.29%	1.40%	0.00%	0.88%	4.17%
Soybeans	2.76%	3.48%	3.08%	0.75%	5.68%	6.00%	2.95%	4.17%
SR wheat	6.36%	7.98%	6.81%	0.29%	3.95%	1.00%	3.53%	4.17%
Grains	19.52%	21.48%	20.80%	0.48%	24.82%	13.00%	11.59%	16.67%
Aluminum	4.44%	5.03%	4.83%	0.18%	5.29%	6.00%	2.88%	4.17%
Copper	5.23%	5.61%	5.45%	0.13%	8.24%	6.00%	3.85%	4.17%
Lead	4.06%	5.20%	4.86%	0.36%	0.00%	0.00%	0.60%	4.17%
Nickel	3.73%	4.60%	4.38%	0.25%	2.37%	1.00%	0.70%	4.17%
Zinc	5.19%	5.75%	5.57%	0.17%	2.94%	0.00%	0.88%	4.17%
Industrial metals	22.90%	26.11%	25.09%	1.02%	18.84%	13.00%	8.91%	20.83%
Feeder cattle	5.57%	7.42%	5.93%	0.52%	0.00%	0.00%	1.55%	4.17%
Lean hogs	0.00%	0.00%	0.00%	0.00%	2.77%	1.00%	2.30%	4.17%
Live cattle	5.745%	7.84%	7.03%	0.44%	4.19%	6.00%	4.79%	4.17%
Live stock	12.60%	13.82%	12.96%	0.29%	6.96%	7.00%	8.64%	12.5%
Gold	5.36%	6.93%	5.66%	0.39%	12.6%	6.00%	3.24%	4.17%
Silver	5.85%	6.90%	6.40%	0.27%	4.37%	1.00%	0.41%	4.17%
Precious metals	11.50%	13.83%	12.05%	0.59%	16.97%	7.00%	3.65%	8.33%
Cocoa	0.00%	0.00%	0.00%	0.00%	0.00%	5.00%	0.45%	4.17%
Coffee	0.00%	0.00%	0.00%	0.00%	2.05%	5.00%	0.94%	4.17%
Cotton	0.00%	0.00%	0.00%	0.00%	1.38%	5.00%	1.19%	4.17%
Sugar	0.00%	0.00%	0.00%	0.00%	2.27%	5.00%	1.59%	4.17%
Softs	0.00%	0.00%	0.00%	0.00%	5.70%	16.00%	4.17%	16.67%

This table displays descriptive statistics for individual commodity and sector weights in the commodity proposed contender index 2007 - 2016. It also shows the weights for the twenty commodities considered.

Table 5.5: Time horizons of sector exposure in Factor Model Index

Panel 1. Non-zero weight frequency in the proposed contender index index

Brent	100.00%	Corn	100.00%	Aluminum	100.00%	Feeder cattle	100.00%	Gold	100.00%	Cocoa	0.00%
Gasoil	100.00%	HR wheat	100.00%	Copper	100.00%	Lean hogs	0.00%	Silver	100.00%	Coffee	0.00%
Heating oil	100.00%	Soybeans	95.00%	Lead	100.00%	Live cattle	100.00%			Cotton	0.00%
Gasoline	100.00%	SR wheat	100.00%	Nickel	100.00%					Sugar	0.00%
Natural gas	10.00%			Zinc	100.00%						
WTI	100.00%										
Energy	100.00%	**Grains**	100.00%	**Industrial metals**	100.00%	**Live stock**	100.00%	**Precious**	100.00%	**Softs**	0.00%

Panel 2. Number of commodities in the proposed contender index index

	Min	**Max**	**Mean**	σ
proposed contender index	18	19	18.05	22.36 %

This table shows the proportion of time, over the whole period considered in the study 2007 - 2016, that individual commodities and corresponding sector have non-zero weight in the commodity proposed contender index.

Table 5.6: Calendar returns of Factor Model Index

Panel 1. proposed contender index monthly returns

	Jan.	Feb.	Mar.	Apr.	May.	Jun.	Jul.	Aug.	Sep.	Oct.	Nov.	Dec.	Total
2007	-1.35%	3.65%	0.11%	3.04%	-1.70%	1.07%	4.03%	-1.42%	*5.67%	2.25%	-3.05%	2.10%	14.94%
2008	2.79%	**12.73%	-5.70%	1.17%	0.27%	5.37%	-4.72%	-5.79%	-10.11%	*-23.44%	-10.43%	-4.20%	-38.09%
2009	-0.38%	-3.99%	5.22%	1.36%	***16.00%	-2.01%	3.72%	-0.69%	1.38%	4.79%	3.85%	0.06%	31.83%
2010	-7.57%	7.43%	2.48%	3.58%	-8.21%	-0.61%	*10.05%	-0.77%	**8.13%	4.11%	1.10%	**9.87%	31.31%
2011	1.10%	4.16%	2.29%	4.27%	-4.88%	-4.40%	5.00%	0.54%	**-14.78%	6.77%	-1.32%	-0.77%	-3.95%
2012	6.80%	3.96%	-2.88%	-0.45%	**-7.56%	3.35%	2.91%	4.54%	4.16%	-4.55%	2.86%	-1.96%	10.62%
2013	3.28%	***-6.53%	-2.06%	-3.84%	0.48%	-3.94%	2.21%	3.92%	-1.84%	-0.02%	-2.32%	0.25%	-10.41%
2014	-1.23%	***5.05%	0.97%	1.49%	-1.10%	2.28%	-2.49%	-0.74%	**-6.06%	-0.79%	-4.70%	-7.41%	-14.38%
2015	-4.70%	3.79%	-2.34%	6.14%	-3.05%	0.73%	***-11.80%	-0.58%	-4.12%	1.36%	**-7.88%	-3.94%	-24.50%
2016	-1.87%	0.80%	3.21%	7.20%	0.24%	1.10%	-5.31%	-1.36%	3.33%	-0.27%	4.45%	1.39%	13.06%

Panel 2. indices absolute performance

	Contender index	BCOM	CRY	GSCI	E.W.
Ann. return	0.25%	-6.23%	-4.57%	-0.81%	-0.52%
Ann. vol.	18.70%	17.77%	18.76%	24.34%	16.88%
Sharpe ratio	0.01	-0.35	-0.24	-0.03	-0.03
Max drawdown	51.85%	69.37%	67.26%	69.47%	48.96%

This table shows calendar returns for the commodity proposed contender index from 2007 to 2016. Significant negative and positive performance is highlighted. Summary statistics show the absolute returns.

Table 5.7: Performance metrics of the Commodity PCA Factor Model Index

GRS F value	1.949	1.787	0.453	0.659
P value	0.012	0.008	0.095	0.034
Alpha	0.03%	0.02%	0.00%	0.00%
Beta	0.98	0.92	0.69	1.08
R-squared	87.60%	85.32%	80.94%	95.40%
Annualized Alpha	7.01%	4.80%	0.52%	1.03%
Correlation	0.94	0.94	0.92	0.98
Correlation p-value	0.00%	0.00%	0.00%	0.00%
Tracking Error	6.59%	7.32%	11.10%	4.25%
Active Premium	6.48%	4.82%	1.07%	0.77%
Information Ratio	0.98	0.66	0.10	0.18

This table displays relative performance metrics for the commodity proposed contender index vs. the Bloomberg commodity index (BCOM), the Thomson Reuters/Core, Commodity CRB Commodity Index (CRY), the Standard and Poor's-Goldman Sachs' commodity index (GSCI) as well an equally weighted portfolio compared to the proposed contender index, to BCOM Contender index, to CRY Rel, to GSCI Rel and to Equal Weighted. The table reports the GRS test statistic for the null hypothesis and the p-value of the GRS statistic.

5.8 Conclusion

This chapter addressed the issues and challenges in making a PCA commodities index. It presented a factor model based index built using eigenvector weights as a novel way to address the current bias towards production based methodology. The results were used to develop a factor model reflecting the groupings of the proposed contender index.

The approach to constructing a contender index, was presented. it was argued this better fits a broad range of desired commodity dynamics. The resultant index incorporates five common factors for twenty commodity futures, resulting in a time-varying weighted index with desirable risk-adjusted return and replication characteristics compared to existing production-based commodity indices.

The chapter presented the results of this PCA contender index and its characteristics against the leading commodity futures indices. The wide difference in tracking error identified between the proposed contender index and the GCSI, BCOM and CYI shows that the resulting approach has important asset allocation implications. The results showed that the proposed competitor index has a risk-adjusted active return compared to the entire reporting period and the majority of sub-periods. The tracking error is significant, suggesting that the methodology provided a very different interpretation of the weighting criteria.

The results showed the PCA index is capable of risk decomposition. As such, it can be used to evaluate how much of a commodity portfolio's return can be attributable to common factor exposures. This fulfills one of the criteria of appropriateness, as presented in chapter one.

It was shown that the current weighting approach uses production-weighted or trade volume approaches. It was concluded that these do not fully account for the return characteristics of a broad range of commodities. In this sense, existing commodity indices are sub-optimal benchmarks for measuring the performance of passive long-only commodity portfolios.

It was observed that the financialization of commodities requires that the asset allocation decision must be measured and evaluated. The method proposed in this chapter has the advantage that it is not dependent on the production process. By using statistical groupings on futures contracts it is possible to identify the common factors and the specific sensitivities of the individual contracts. Existing commodity indices use weightings for production and trading volume.

In summary, commodities require a suitable reference index and existing indices are shown to be sub-optimal. Understanding the way in which commodities are grouped for investment purposes has important implications for practical fund management. The calculation of a PCA derived commodity index avoids the production bias of existing indices. Such an index is therefore a useful addition to understanding the nature of commodity futures and adds to the literature on index construction. The next chapter will investigate if the findings can be equally useful for hedge fund indices.

Chapter 6

Hedge fund benchmarks, challenges and refinements

This essay investigates the specific issues related to benchmarks designed to measure hedge fund performance. Hedge funds are a "de-jure" asset class, albeit not a de-facto one. Commonly used peer group indices are sub-optimal, as this method is net of fees and subject to survivorship bias. It employs principal component analysis (PCA) to elicit insights into how strategies relate to one another. It applies the GRS test to a PCA derived factor model. The PCA results showed an improvement in sub-sector classification but the complexity of hedge funds and their frequent strategy regime change show that bespoke benchmarks are superior. The result is important because hedge funds use benchmarks to justify their marketing and also to determine their fees. That said, individual performance hurdle rates are not good comparison benchmarks and there still exits a need to compare cross-sectional returns between the competing funds. The PCA approach suggests that synthetic indices may be part of the solution, as these can be constructed based on strategy risk factor replication as a way of measuring gross performance. This essay's contribution is in showing how hedge fund styles can be better grouped and how an index can be constructed that is not nett of fees, as is the case in peer group indices.

6.1 Introduction

This chapter investigates benchmarks from the perspective of investment in hedge funds. The main aim of this chapter is to test the appropriateness of the principal component analysis (PCA) index approach. Firstly, it provides background information on the asset class. Then it tests the mean variance efficiency of the PCA derived index against existing commercial indices. The main empirical is the GRS test. Hedge fund benchmark involve many dimensions and these are not all captured in existing indices. A researcher has to deal with funds as proxies and the underlying use of derivatives. It builds on the insights extracted from the earlier chapters and investigates whether PCA derived indices can address the identified shortcomings in existing peer group indices. These are that peer group hedge fund indicies are (1) not replicable, (2) net of fees, and (3) poorly sub-grouped.

Definition 6.1.1. Hedge funds are largely unregulated investment funds formed as a private limited partnership - Stevenson (2011). Their unregulated nature tends to mean that they employ varying strategies and financial instruments.

Benchmarks are central to the reward system of hedge funds. Such funds are incentivised to generate returns in excess of a chosen benchmark. As a result, benchmark specification is practically as well as academically important. Unfortunately, the private nature of the hedge fund investment vehicle results in poor transparency on benchmark selection and the underlying instruments within them. Research into the appropriateness of benchmarks is therefore of scholarly importance.

According to the Hedge Fund Research database the Assets Under Management as at the 4th Quarter 2019 in hedge funds was USD 3.194 trillion and in hedge fund of funds at that point was USD 290.4 billion, all subject to a benchmark. The challenges faced in measuring the performance of these assets is to (1) reflect divergent strategies, (2) represent funds that trade infrequently and (3) overcome investment restrictions (such as large minimum size lots and investment restrictions). A PCA derived method, as explained in chapter four, ensures clear statistical choices are made and avoids the fee and fund size problems.

The literature, as introduced in chapter two, shows that existing hedge fund indices face a number of issues in their reliability. These are summed up by Hedges (2003). The biggest concern is the survivability of the underlying funds. As shown in chapter three, a peer group hedge fund index has to determine the weights of its constituent funds. This involves a trade off between a bias in favor of the largest such funds and what investors wish to gain exposure to. It also involves a trade-off between absolute return and high relative return. This is sub-optimal.

Hedge funds are not an alternative asset class in a traditional sense but, as shown in chapter one, they can be considered one by the uncorrelated nature of their returns to traditional asset classes. A hedge fund can also be defined by its legal structure. In this respect, they are unconstrained collective investment funds that are characterized by a performance fee based on a benchmark. The pre-determined investment benchmark is typically self-selected. Although it is used to determine the fee, it is not useful for cross-sectional comparison. In this respect, the choice of benchmark is central to what a hedge fund is and what it is trying to achieve. That said, as shown by Eling (2009), there is no consensus on what an appropriate benchmark is for the collective evaluation of such a fund, hence the need for further academic research.

There has been much debate on what the most appropriate benchmark is for hedge funds. The literature suggests that it is not possible to use traditional equity or bond indices to benchmark hedge funds for anything other than measurement of their directional strategies. Many hedge funds have both long and short positions and can employ leverage. The issue is that, in aggregate, hedge funds do not act as an homogeneous group. It is clear from these deficiencies and those identified in the literature that further research is needed to establish a more acceptable and appropriate reference value for the valuation of hedge funds at the strategy level.

In summary, the hedge fund industry markets itself on the back of being able to deliver alpha against a benchmark. However, an all encompassing benchmark does not exist. Hedge funds meanwhile leverage beta, borrowing and use derivative instruments. These obscure the peer comparisons and makes it difficult to identify if funds do indeed deliver alpha. Kat (2001) demonstrated that there are substantial difference in returns against benchmarks in respect of funds that claim to follow similar strategies. The prior chapters identified that improvements in benchmarks can be made if index constructors focus on unique risk factors. This chapter investigates whether PCA can help address that through refinements to hedge fund indices.

6.2 Hedge funds as an asset class

Hedge funds are a *de facto* alternative asset class. As shown in the first chapter, they are collective investment vehicles that invest in other assets, and hence are not a *de jure* asset class in their own right. In this respect, pension funds and other institutional investors that have exposure to hedge funds tend to allocate them to the alternative asset section in their portfolios. Institutional motivation is important to benchmark selection as it defines how they attribute the returns from this asset class as explained in chapter three.

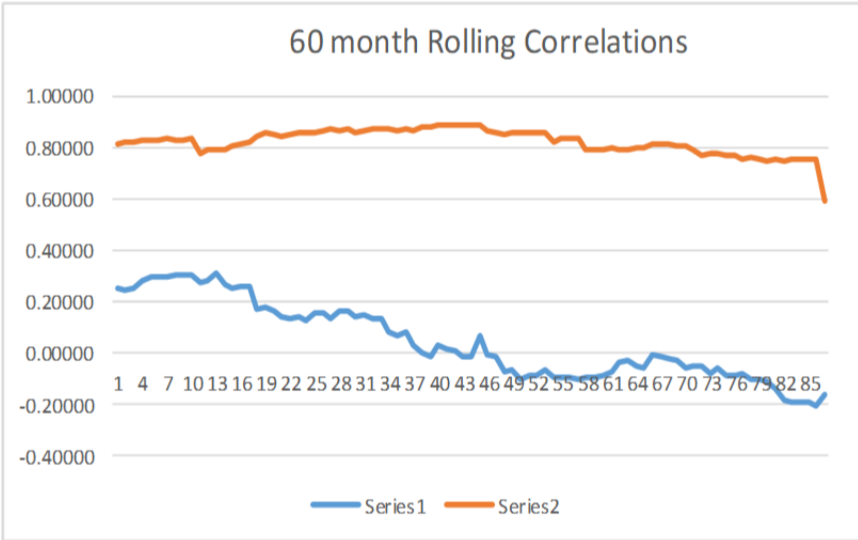
The literature shows little consensus among academics as to which the appropriate benchmark is to use with hedge funds. Hedge Funds of Funds have to be benchmarked to an appropriate and investable index. The indirect nature of the literature reviewed suggests that scientific research would benefit from more research on benchmark selection and construction.

The way that time series vary and co-vary with each other is relevant to the way a hedge fund benchmark is constructed. A PCA based index is a valid contender to peer group indices which do not reflect dynamic strategies. These movements impact the way returns are distributed. In this respect, returns are non-normal and time vary. Kat (2001) shows that the distribution of monthly returns of many hedge fund indices are unusually skewed, have first order correlation and demonstrate kurtosis. Figure 6.1 demonstrates how hedge fund strategies are time varying. It shows the sixty month correlation between the broad hedge fund peer groups and the equity and bond markets. Note that the relationship is not stable. This is particularly the case at a time of market risk at the end of the period, indicating that hedge funds were not hedged.

Research into hedge funds as an asset class is important because the industry makes the claim that individual funds deliver risk adjusted returns relative to their benchmarks. In aggregate, this presents a problem as each hedge fund has a unique benchmark for its performance fee and each investor has a unique but different benchmark for evaluation. Kritzman (1999) shows that almost any manager can demonstrate alpha through a selective choice of benchmark.

Research into the asset class is complicated by the multiplicity of strategies. A breakdown of the common hedge fund strategies is presented in figure 6.2, divided between forward-looking strategies, such as active trading, discretionary thematic and fundamental valuations,

Figure 6.1: Hedge fund 60 month rolling correlation to equities and bonds 31/1/2008-29/2/2020.



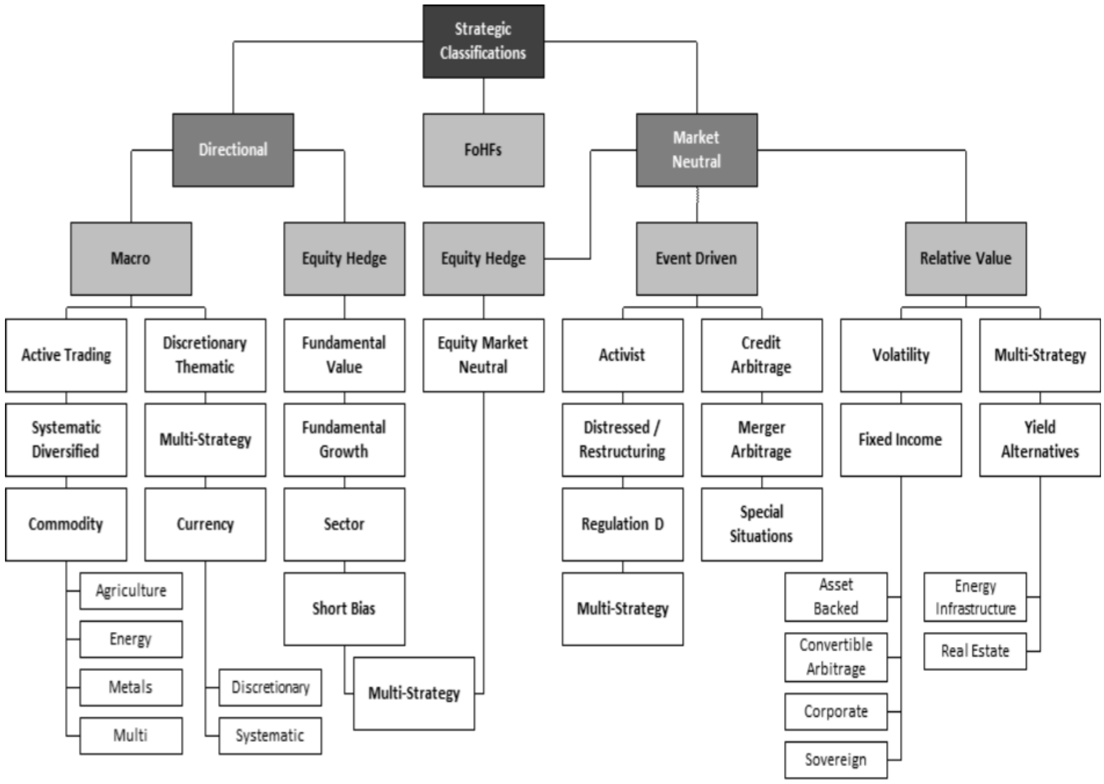
This diagram shows the sixty day rolling correlations between the HFRI Multi Strategy Index from 31/1/2008-29/2/2020 and the FTSE World Bond Index (Series 1) and the MSCI Global Equity Index (Series 2). Although Multi Strategy includes both equity and fixed income strategies, it is clear that there is a high directional correlation to equity markets.

or market-neutral, activist, credit arbitrage, volatility or multi-strategy approaches. Each of these is very distinct in respect of the time series of returns that it generates.

Some insights can be gathered by examining hedge fund of funds. These tend to reflect the asset class. These funds invest in a portfolio of other hedge funds and as such in aggregate behave like a benchmark. That said, they represent a substantial pool of assets that requires appropriate benchmarking in their own right. This is so as to measure and attribute performance. The need for relevant benchmarks as distinct from the fund level is a very real one. At the fund level, appropriate benchmarks are required in order to determine the performance fee. The underlying hedge funds that are included in hedge fund of hedge funds generate substantial management fees and indeed substantially more than that in performance fees. Hedge fund of funds place a further fee on top of this, which further complicates the cost drag that peer group approaches generate.

Hedge funds that are directional are easier to benchmark. Fung and Hsieh (2002b) suggest that the directional component can be risk modeled using long-only conventional indices. In this respect, to measure alpha requires an appropriate market proxy. They point out, however,

Figure 6.2: Hedge fund strategies



This diagram shows the various different hedge fund families and their strategies. Each of these is uncorrelated to the other, complicating benchmark construction. The five broad classifications in Figure 6.2 are used in the factor model proposed in this chapter. There are 35 sub groupings, which complicates the classification of indices. Similarly, there is a mixture of directional, market neutral and leveraged strategies, all of which effect the return distributions.

that this cannot be done for other strategies such as market neutral. The critique of this approach is that it does not allow attribution where funds have mixed assets. PCA indices group underlying instruments, so partially address this issue.

The necessity for appropriate hedge fund benchmarking gives rise to three important supplementary research questions to those raised in chapter one:

- How do you construct a hedge fund benchmark for cross sectional comparison?
- How do you define hedge fund alpha?
- How do you account for fees charged on the underlying strategies?

These questions address the lack of homogeneity in the returns generating process between different hedge funds. The current way of addressing them is through peer group assessment. As observed in chapter one, unlike traditional long only assets, hedge funds do not have a normal distribution of returns. Some strategies, for example, are market neutral and return only positive or negative alpha (minus fees). Others are directional but leveraged, making their returns higher beta than traditional long funds. Some are event driven, which tend to result in spikes in performance as and when the returns are generated.

A different approach to address the non-linearity of returns from a benchmark perspective was proposed by Glosten and Jagannathan (1994). They worked from the starting principal that performance measurement against a benchmark should assign a number to a manager's added value to a portfolio compared to the initial cash flow net of management fees. From there, they approached value added as a contingent claim on an index portfolio. In doing this it is possible to estimate that value using a series of options. Whilst interesting, this can be critiqued as too complex for practical fund management and not able to provide adequate attribution.

The need for a time dependent PCA approach is underlined by the nature of hedge fund returns, Mitchell and Pulvino (2001a) and Agarwal and Naik (2004a) demonstrate that hedge funds generate significant skewness, manifesting itself as tail risk. They observe that the multiplicity of strategies hedge funds deploy to beat their chosen performance benchmarks complicates evaluation. As with leverage, this is captured in a PCA index where the first principal component at time *t+1* may appear as the second or the third principal component at time *t+2*. The interpretation of results of the factor weights, however, requires care when compared over different time horizons as hedge funds can frequently change strategy.

PCA addresses non-linearity by grouping together the principal components. The evaluation of hedge funds compared to appropriate benchmarks is complicated by the challenges presented by non-linear time series. A summary of such features is provided by Kat and Brooks (2001). They found hedge fund indices typically had first order serial correlation with the stock market when based on monthly data. As pointed out earlier, they further found that the returns distribution exhibited both skewness and kurtosis.

PCA also highlights the factor driving hedge fund returns. As benchmarks are used to explain and decompose investment strategies, Chow et al. (2011) (2011) suggest alternative indices be used based on factor risks. They concluded that academic synthesis on benchmarks for investment performance has largely been based on traditional long only linear time series focusing narrowly on market, value, and size factors.

Another focus of the literature on factors is the separation of alpha and beta. It has proved difficult to identify if a hedge fund is exposed to systemic risk other than by comparison to traditional long only indices. As a result, such funds typically select long only indices in order to have a performance hurdle. Benchmarks that capture the non-linear nature of these returns, other than through manager style, have not been the subject of scholarly literature. The variables, such as the risk free rate, the volatility of the underlying instruments and the market return are difficult to incorporate into benchmarks.

One consequence of the non-linearity, and ability of hedge fund managers to go both short and long, is that hedge fund alpha is difficult to calculate against traditional benchmarks. In a number of papers, buy-and-hold strategies have been used as a benchmark to address this issue. Others have used the excess returns above the risk-free interest rate. Titman and Tiu (2011) is typical of the later. They group funds by characteristics in order to see if the funds hedge their factor exposure. Whilst such approaches may well demonstrate correlation to, and performance against a benchmark, they avoid directly addressing the aforementioned non-linearity. As such, alpha as a concept in hedge funds remains elusive despite regularly being claimed against traditional long only benchmarks by the managers themselves.

Another challenge in constructing hedge fund benchmarks is that they have a relatively short history of time series returns, typically less than twenty years of monthly observations. The fact that hedge funds are largely priced monthly also presents a data problem. Existing indices are not calculated on a daily basis. The *HFRI index*, for example, is calculated three times a month, and re balanced only once a year. This suggests a PCA derived synthetic index, which can be priced daily, has a clear advantage over existing peer group indices.

An interesting approach to benchmarking was proposed by by Agarwal and Naik (2004a). They developed a cash flow model that addressed the non-linearity present. They demonstrated that hedge fund payout characteristics tended to reflect a short put option on the market index (in all but momentum and manager arbitrage strategies). An index that incorporates this can be used to market timing ability. This requires creating an options strategy called a straddle. This approach is difficult to replicate and as such is conceptually more interesting than it is in practice from a benchmark perspective.

In a similar vein, Fung and Hsieh (2011) investigated the properties of being both long and short strategies. Once again, their results can deliver a benchmark to measure directional alpha. They argued that a trend following strategy has a similar return profile as an option. A derivative in this way provides a right to buy the underlying instrument at the low point in the life of the option. Similarly, a put option holder has the right to sell at the highest point. Based on this concept, they developed three trends following risk factors, for bonds, commodities and currency that can be used as benchmarks. These can further be combined with equity or bond orientated risk factors, namely the equity market factor, the size spread factor, the bond market factor and the credit spread factor. This results in a seven factor model. In a later extension of their work, they added an emerging market index as a factor, creating an eight factor model. How many factors to choose is a debatable point. The model in this chapter is based on five, as this is the current mode.

6.2.1 Peer group indices

The most widely used technique for benchmarking hedge funds against one another is peer group assessment. Hunter, Kandel, and Wermers (2014) investigated peer groups using mutual funds. The results are equally applicable to hedge funds and fund of hedge funds. They conclude that peer groupings suffer from three key shortcomings

- their time series are stated net of fees,
- they are not investable either at a specific point in time or in a specific unit size,
- the time series are also typically non-linear and illiquid.

Despite the drawbacks, a number of commercial providers produce peer group benchmark indices. There are six index families in common usage. These include *CSFB/Tremont, Altvest, Zurich Capital Markets, Hennessee, Van Hedge Advisors* , *Evaluation Associates Capital Markets*, and *Hedge Fund Research*. The characteristics of these commercial index families are documented in Amenc and Martellini (2002). These families break down categories into sub-indices based on manager style. A hedge fund benchmark, however, can also be based on the individual fund time series as is the case with a PCA derived index. That is, it can be bottom up as well as top down

Peer group indices are not able to benchmark absolute return strategies. As hedge funds purport to deliver alpha, risk adjusted out-performance, there is a requirement to extend a benchmark from a simple hurdle rate to one that takes account of underlying strategy. Whilst peer reviewed benchmarks do have sub-indices constructed by strategy, the weighting method is unrepresentative of either the returns or indeed the full sample of funds available. Commercial hedge fund targets tend to be hurdle rate orientated, and as such do not fit this criteria

One of the critiques of peer group indices is that they are constructed by first identifying the constituent funds. This part of the process is highly subjective due to the large number of funds in the universe, despite it being a rules-based approach. In practice, the selection is tilted towards fund size, which is not what investors necessarily want. As a result, there is a concern in the literature that the larger the fund size gets the less flexibility there is to generate alpha. At the same time, the peer group selection approach favours multi fund firms over single strategy hedge funds, This is because the stables have greater marketing budgets to support the data supply to the index producer. The PCA approach, explained earlier in chapter four, avoids this bias as each fund is evaluated on its return time series alone.

The peer group approach presents a challenge to both academics and practitioners because such benchmarks are difficult to replicate. Investing in the underlying funds is further complicated by regulations designed to protect investors as well as the terms and conditions of the funds themselves. In particular, investment in hedge funds is subject to liquidity considerations. These include the unit size of investment and conditions of subscription and/or redemption. Peer group benchmarks also suffer from distortions due to new fund launches, survivorship bias, changes in the underlying mangers assets under management, and fund re-classifications.

There is also a double charging issue. Peer group benchmarks measure the total return of the underlying net asset value. They therefore incorporate not only alpha but fees and charges that are present in the underlying instruments. Traditional benchmarks are stated gross, not net of costs. The PCA derived index addresses this deficiency. It uses common components in the return time series of existing hedge funds to determine the weights of the different funds and strategies. In this way it reflects the risk factors they exhibit.

In summary, PCA derived indices have a number of advantages over long only strategy based peer groupings. They can be used to take into account derivatives and leverage in the underlying strategies. This is because the first principal component at time *t+1* and that at time *t+2* do not have optimal weighting outcomes. They can also better handle directional funds. They avoid the double counting of fees and are easier to replicate.

6.3 Data

In order to construct a PCA contender index, the Lipper TASS database was used to identify specific fund characteristics based on a research sample period from January 1st 1996 to Jan 1st 2016. The database is the most comprehensive in the hedge fund universe, containing more than 350 fields of data over 3900 hedge funds and 300 plus Commodity Trading Advisors. It also includes data on over 11,000 graveyard funds that have liquidated or stopped reporting. This was chosen in preference to other hedge fund databases, such as that provided by Hedge Fund Research, on the basis of depth of coverage of the Lipper TASS database. As at January 2020, HFM Indices tracked over 18,000 funds, so is larger. That said, it cannot be used to adjust for survivorship bias.

If hedge funds do not perform, investors withdraw funds and they collapse. This biases the sample towards those funds that perform well, as their history remains and the history of those that close gets deleted. As a result, one of the biggest issues in peer group hedge fund benchmarking is a survivorship bias. When a fund closes or merges with another its track record does not survive. Survivorship bias is an issue for all indices. With hedge funds, academics are constrained by the limited data to investigate the entire set of risk factors. The lack of surviving funds is particularly large for such funds. Indeed, according to Grindblatt and Titman (1992), survivorship is one of the biggest issues that peer group indices face. The allied issue of incorrect classification of hedge fund styles is pertinent. In this respect, the literature includes research on de-listing from a database, the impact of database choice and the revision of previously reported returns. In this respect, the survivorship issue is widely researched and is well documented in Brown et al. (1992). As there is no predictability of when a fund will close, it can only be addressed by cleaning the data after the fact. This means that a benchmark made up of peer funds has to be frequently re-balanced, which adds to its tracking error.

A reduction in sample size reduces the robustness of the results. That said, the Lipper TASS database is the best available for academic and commercial researchers. Even so, according to Posthuma and Van der Sluis (2003), more than 50 percent of all returns in the Lipper TASS database are what they call back-filled returns, submitted after the fact by hedge Funds.

One of the advantages of the Lipper TASS database is that it is linked to the most widely used commercial index, the *Credit Suisse Hedge Fund Index*. This index is divided into ten investment strategy sub-categories, namely convertible arbitrage, dedicated short bias, emerging markets, equity market neutral, event-driven, fixed rate arbitrage, global macro, long-short equity, managed futures and multi-strategy. The average total return of hedge funds in the the Lipper TASS database was broken down by Hasanhodzic and Lo (2008). They found six factors and manager-specific alpha for 1,610 hedge funds from February 1986 to September 2005.

In addition to the database, Bloomberg was used to further identify specific fund characteristics and news flow. This enabled not only style and volatility but also the analysis of scenarios. It has been difficult to identify the effect of individual factors because many are correlated. The number of risk factors that explain cross-sectional time series differences is very broad. As a result, the data was prepared on the same basis as Posthuma and Van der Sluis (2003). They investigated the period January 1994 to June 2009. An interesting finding of theirs was that the commonality of hedge fund income doubled during the study period.

As explained in chapter 2 the survivorship of funds in a sample is a major issue for hedge fund researchers. In order to minimize survivorship bias, a sample of the 200 largest hedge funds was chosen based on the completeness of their time series. This resulted in 3,700 hedge funds not being included in the sample from a historical perspective out of a total of 14,900 funds. The sample misses funds that were closed within the sample years and does not included new funds that were started after the sample period starts. Ideally, a larger sample would have been preferable and the sample selection criteria affect the reliability of the results. Future research can address this issue by investigating the principal components of new and failed funds.

The sample period includes the credit crisis of 2008. Kaiser and Haberfelner (2012) observe that hedge fund databases were heavily distorted by the survivorship of particular strategies during this period. This is why the Lipper TASS database is used. As a database for hedge funds, it has a considerable range of live and dead hedge funds. Fung and Hsieh (2001b) claim survivorship bias to be at three percent per annum using the Lipper TASS database. They distinguish between natural and spurious bias, the former being due to the life-cycle of the fund and the later due to data quality from reporting issues. The survivorship bias was very evident in the sample used. That said, according to Posthuma and Van der Sluis (2003) more than 50 per cent of all returns in the Lipper TASS database has back-filled returns, submitted after the fact by hedge funds. Despite the precautions, the survivorship issues suggest the results will exhibit an upward bias in the return time series.

6.4 Method

Using the aforementioned data, a PCA index was constructed using the method as detailed in chapter four. An orthogonal transformation was done on the sample and the co-variance matrix then is used to derive factors from the hedge funds returns. The results are re-grouped into weights based on eigenvalues in order to create the index. The PCA contender index was then constructed using the following model.

$$\zeta_r i, t = \sum_{K=1}^{m} \beta_k, t, Rk, t + \sum_{l=m+1}^{m+n} \beta_l, t, Ri, t + \left(1 - \sum_{k=1}^{m} \beta l, t\right) R_t^f \tag{6.1}$$

Where

- $\zeta_r i, t$ = Return on the contender benchmark
- β_k = Beta of the asset factor

This equation groups the common components together to form a factor model. The outcome summarizes correlations between the various hedge fund strategies, providing insight into those sub-groupings that produce similar factor returns. This approach produces a practical tool that enables (1) identification of common drivers of individual strategy returns, (2) construction of a factor model, and (3) the ability to determine a mean variant efficient portfolio that takes into account liquidity.

An example of this was the Fung and Hsieh (2004a) risk-based approach. While a factor approach does not affect the actual selection of managers performance fee benchmarks, it does provide insight into the manager's assessment of managerial skills.

The model uses PCA to identify the commonality and determine the factors. The technique is different from that traditionally used in factor analysis as it is a data reduction method. In effect it summarizes the hedge funds with strong inter-correlations to various drivers of return. It is an exploratory technique. The original variables are transformed into a linear set that has a strong relationship. The Kaisers rule is applied with a cut-off of one.

After deriving the principal component from the dataset, the next step is to construct the proposed contender index. This is done by deriving the constituent weights for each period using factor analysis implemented using oblique rotations. The proposed contender index return is therefore a weighted average of the returns of the derived portfolio constituents. It is a five-factor vector model variant to the model above. It is described as such:

$$\zeta_r i, t = \beta 1\mathbf{v}\eta_t + \beta 2\mathbf{v}\theta_t + \beta 3\mathbf{v}\chi_t + \beta 4\mathbf{v}\tau_t + \beta 5\mathbf{v}\mu_t \tag{6.2}$$

Where:

- $\zeta_r i, t$ = Proposed contender index, the excess return of sector commodity portfolio i in month t, item **v** = an eigenvector
- η = Broad multi strategy complex
- θ = Equity based complex
- χ = Market Neutral complex
- τ = Event Driven complex
- μ = Relative value complex

Using the vector method, the index is constructed in a stepwise fashion with a three-year observation period along with a six-month re-balancing frequency conducted at each re-balancing date. Table 6.1 shows the resultant standard deviations and factor loadings of the end result.

A GRS test is used to review the validity of the resultant index and its limitations. This is based on the null hypothesis introduced in chapter one and is shown below. The null hypothesis suggests that it is not possible to out-perform the existing indices on a risk adjusted basis.

Null Hypothesis 3.

$$H_0 : GRST_t = \eta_i - \eta_j = 0$$

Where:

- $\eta_i = \mu_i/\sigma_i$ = Mean excess returns i (over a risk-free rate)/standard deviations of excess returns i
- $\eta_j = \mu_j/\sigma_j$ = Mean excess returns j (over a risk-free rate)/standard deviations of excess returns j
- $GRST_t$ = is made up of an estimator of the Sharpe ratio difference, the standard error of the difference estimator, the Sharpe ratios and the correlation between the excess returns of strategy i and j.

Table 6.1: Standard deviation and sector loadings of rolling factors

	Factor 1		Factor 2		Factor 3		Factor 4		Factor 5	
	Mean	σ	**Mean**	σ	**Mean**	σ	**Mean**	σ	**Mean**	σ
Broad multi strategy	0.73	23.91%	0.18	8.82%	0.13	4.85%	0.07	3.06%	0.09	3.86%
Equity based	0.16	8.13%	0.13	6.06%	0.75	22.14%	0.06	4.47%	0.08	6.71%
Market neutral	0.16	5.48%	0.75	6.99%	0.11	3.62%	0.08	4.66%	0.12	4.41%
Event Driven	0.08	3.60%	0.06	3.12%	0.04	10.35%	0.63	32.89%	0.08	22.12%
Relative value	0.17	2.76%	0.26	4.86%	0.14	2.42%	0.09	24.83%	0.76	24.98%

This figure displays means and standard deviations for individual hedge fund style grouping loadings on the first five factors corresponding to the rolling factor analysis, with daily observations. where there is a high factor loading, such as with factor three and equity neutral hedge funds, it is assumed that those factors correspond with the strategy.

6.5 Results

The results of the GRS test are provided in table 6.2. As per the model, five factors were identified corresponding to five investment strategies. The GRS test does not reject the H_0 null hypothesis of zero risk-adjusted returns with a p-value of 0.180 for the HFMI Index 0.201 and the *Credit Suisse Hedge Fund Index*. The proposed index is therefore not necessarily a good substitute. It should be pointed out, however, that the concept of risk adjustment with hedge funds is problematic.

Table 6.2: Performance metrics of the Hedge Fund PCA Factor Model Index

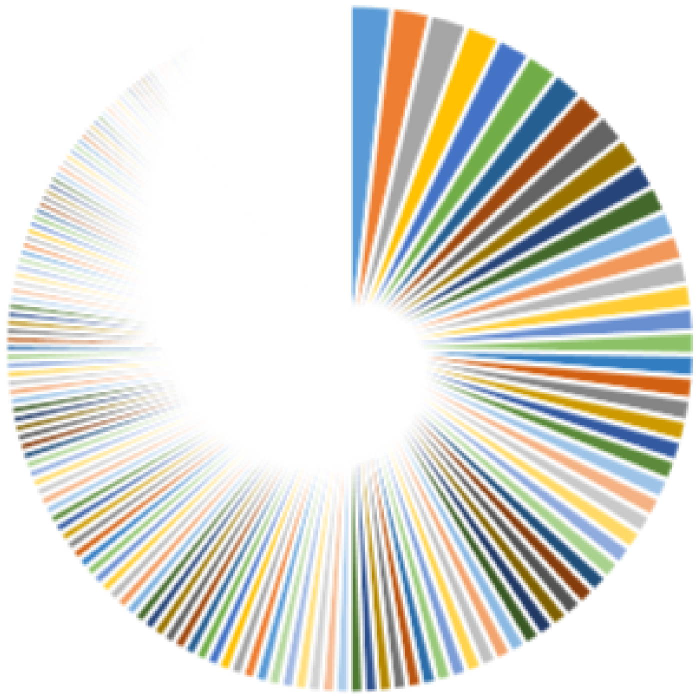
Result	HFMI Index	Credit Suisse hedge fund Index
GRS F value	1.27	1.13
P value	0.180	0.201
Alpha	0.04%	0.03%
Beta	1.03	1.02
R-squared	85.60%	83.32%
Annualized Alpha	3.21%	3.00%
Correlation	0.94	0.94
Correlation p-value	0.00%	0.00%
Tracking Error	3.59%	3.32%
Active Premium	3.45%	3.20%
Information Ratio	0.24	0.13

This table displays the GRS F and P values and the relative performance metrics for the hedge fund proposed contender index vs. the HFMI Index and the Credit Suisse hedge fund Index.

Note the low alpha and near market beta. This suggests that in aggregate the PCA index, and the two indices to which it is compared, deliver average returns. That said, there is a tracking error in excess of three, suggesting that there is clear divergence of index constituents and weights. It is suggested that the lack of significant positive or negative alpha and high or low beta is a reflection of hedge funds in aggregate taking positions against one another.

The initial eigenvalues resulted in a nicely diversified contender index weight output which is a good result from an index perspective. The weights are shown in figure 6.3, ranging from 1.95 per cent to as low as 0.06 per cent. The lack of concentration in the weights demonstrates that a PCA approach can produce adequate diversification. The nature of the factors could

Figure 6.3: Initial Eigenvalues from exploratory data on hedge fund monthly returns

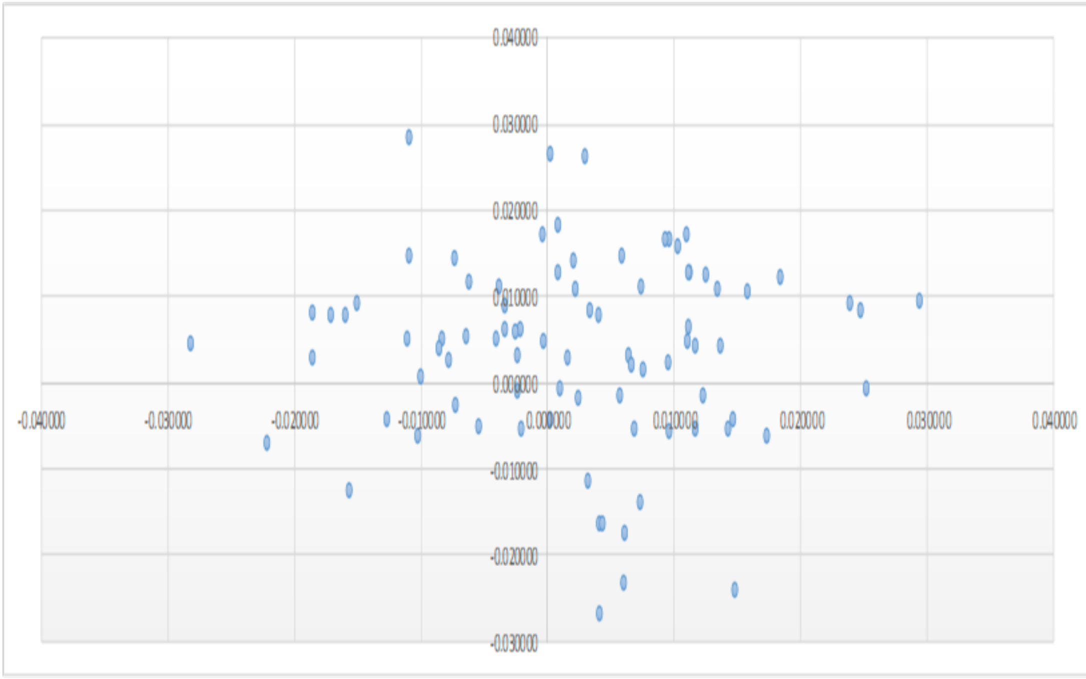


This figure shows both the benchmark, allocation, selection and interaction effects interact in diagrammatic fashion. There are 199 constituent funds with initial eigenvalue weights ranging from 3.9 percent to 0.13 percent and variances of 1.96 to 0.07. It is presented to show that the first principal component derived weights that result in a diversified portfolio, a basic pre-requisite of an appropriate index. These weights are very different from the asset under management or equal weight approaches that a peer group index generates.

not be identified. This was due to the number of different influences and as a result was not possible to derive anything more than a fund level weighting. To illustrate the monthly returns, Figure 6.4 shows the dispersion of the returns of the HFRI Multi Strategy fund versus the PCA index. This confirms there is little cross over between the two index approaches. Referring back to the literature on peer groups, this suggests that the PCA index approach better captures the desired return.

It was hoped that the results would produce a robust contender index. The results for the hedge fund universe produced a disappointing *Kaiser-Meyer-Olkin Measure* of Sampling Adequacy figure of 0.585, well below the optimal 1.00 but above the 0.5 value identified by Cerny and Kaiser (1977) as requiring remedial action.

Figure 6.4: HFRI Multi Strategy Index monthly returns versus PCA Index 1/1/2013-1/2/2020



This figure compares the HFRI Multi Strategy Index monthly returns versus PCA Index. As can be seen, there little correlation suggesting that the HFRI Multi Strategy Index does not capture a common set of returns.

It is postulated that the weak results reflect the diversity of hedge fund styles and that there are some classification issues at the instrument level which is confirmed by the literature. The investigation did, however, provide some insights. The first is that the sub-groupings need to be more refined than simply by strategy. The second is that, as with fixed income grouping by credit, hedge funds should be grouped by leverage. The third is that hedge funds reflect multiple underlying assets and as such a multi factor approach might prove a useful avenue of investigation.

Analysis of the monthly returns identifies a number of issues. These relate to the time varying nature of hedge fund strategies and their underlying liquidity. The linear components of the time series appeared to have curvilinear relationships. The lack of liquidity of the underlying investments during some market events also appeared to impact returns. In order to explain this, it is observed that hedge funds invest in exotic and illiquid securities, which according to Getmansky, Lo, and Makarov (2004) give rise to problems with the time series of returns.

The contender index was tested in a number of scenarios. These were based on Fung and Hsieh (2002a) who identify eight standard asset classes and tested the way hedge funds co-varied with them. The reason given is that the underlying hedge funds are all driven by different risk factors. The results of the scenario analysis are presented in the table 6.3 and the individual tests defined in Appendix B and tabulated in Appendix C. The scenarios show that a PCA index does not prove robust enough to create an index that delivers positive alpha in all scenarios.

It is concluded that PCA derived indices can be used as an alternative to peer group indices but that more research has to be done at the individual strategy level. As highlighted in the literature review in chapter two, the most common approach to benchmarking cross sectional hedge fund performance is to use peer groups. The high dispersion in the performance of individual funds, however, means this approach is best suited to benchmark fund of fund returns. Investors use these to achieve strategy diversification. Due to the large number of strategies and funds, peer group indices are often sub-divided by process.

Figure 6.5 illustrates how the PCA index compares to a *Peer Group Index Spiderweb* using the Likkert scale introduced in chapter one. The outer bounds of this are more dispersed than those of the peer group, suggesting the former is more appropriate.

The Index Spiderweb demonstrates that the PCA index results should be viewed in the context of peer indices. As explained that peer group benchmarks have numerous failings. These include the fact that they are net of fees and are not replicable. This presents a benchmarking problem. Despite the GRS test results, the PCA derived index was shown by the Index Spiderweb to be a usable alternative, albeit not perfect at differentiating between within sample style differences between hedge funds. This is because with hedge funds, a benchmark can be as simple as a hurdle rate. The PCA index is more sophisticated, facilitating attribution and the creation of composite benchmarks. Funds of hedge funds pursuing a diversified investment approach (by creating a portfolio of hedge funds) but must also demonstrate that they provide added value in the selection of managers. A synthetic PCA derived index can help address this issue.

In summary, the results indicate that a PCA index provides a solution to the infrequent pricing of peer group funds but it can't capture hedge funds time varying style shifts. As such,

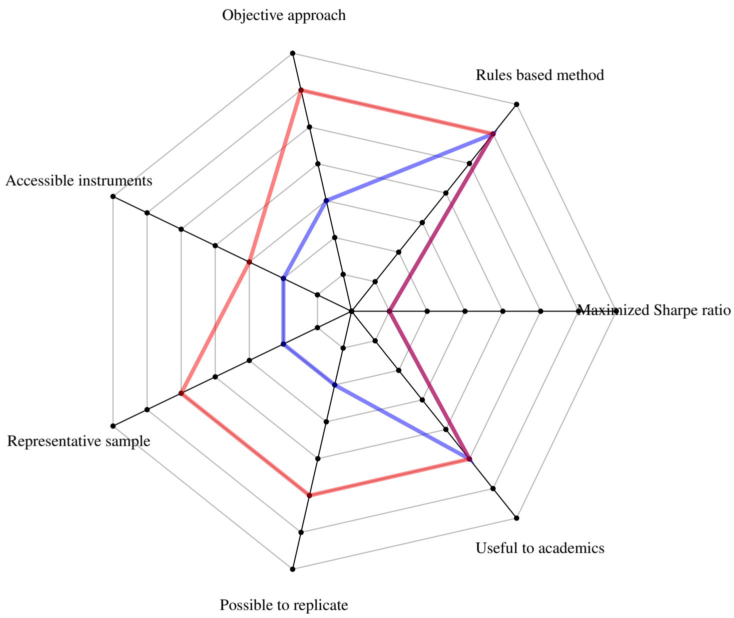
Table 6.3: PCA Index performance during market stress scenarios

	MSCI USA	MSCI NA USD	WORLD EX US	MSCI EM	US GOVT.BOND
Burst of the "Dot com bubble"	0.07974	0.07759	0.00297	-0.02839	0.02497
The end of equity bear market	0.12231	0.12193	0.10643	0.10881	-0.00667
The first Gulf war	0.02319	0.02153	-0.01310	-0.02614	-0.00885
The Russia Financial Crisis	-0.18923	-0.19934	-0.27037	-0.39196	0.04107
The Lehman Default	-0.20718	-0.21494	-0.20433	-0.21446	-0.00816
Equity rebound	0.31733	0.33555	0.39545	0.53549	-0.03158
Oil slump	-0.10033	-0.10172	-0.13159	-0.14161	0.01843
The US DEBT Ceiling crisis	0.06324	0.06266	0.04622	0.03863	0.02774
Libyan civil war	-0.01677	-0.01393	0.00034	0.00472	0.00780
Japan earthquake	-0.02887	-0.02957	-0.07337	-0.03817	0.00751
Greek Crisis	-0.01677	-0.01393	0.00034	0.00472	0.00780
Yuan devaluation	-0.09142	-0.09118	-0.10085	-0.11513	-0.00028

	Global Govt Ex US	EURO/USD 1MTH	USD	Gold spot LDN	PCA Index
Burst of the "Dot com bubble"	0.00472	0.02082	0.03060	0.00385	0.00726
The end of equity bear market	-0.00196	-0.28927	0.02441	-0.02909	0.00652
The first Gulf war	-0.00340	-0.00687	0.05426	-0.00689	-0.00387
The Russia Financial Crisis	1.03820	-0.09993	0.07339	-0.02151	-0.27310
The Lehman Default	1.03208	-0.12365	0.03955	-0.06074	-0.28022
Equity rebound	0.04208	0.06175	-0.12576	0.03229	0.06203
Oil slump	0.31667	-0.06508	0.04733	0.00091	-0.01963
The US DEBT Ceiling crisis	-0.65486	-0.24980	-0.09724	0.01156	0.13337
Libyan civil war	-0.03125	-0.02461	-0.01122	0.00533	0.002034
Japan earthquake	0.00000	0.03043	-0.00030	0.00293	0.00601
Greek Crisis	-0.03125	-0.02461	-0.01122	0.00533	0.00442
Yuan devaluation	-0.00555	0.00000	-0.01002	-0.00759	-0.00796

This table shows the PCA Index performance (10) versus the following asset classes during periods of market stress. (1) US Equities: MSCI North American Equities, (2) Non-US Equities: MSCI non-US Equities, (3) Emerging Market Equities: IFC Emerging Markets, (4) US Government Bonds: JPM organ US Government Bonds, (5) Non-US Government Bonds: JP Morgan non-US Government Bonds, (6) One-month Eurodollar Deposit Return: One-month Eurodollar deposit rate of the previous month, (7) Spot Gold: London morning fixing (8) US Dollar Index: Federal Reserve Traded and (9) Spot Gold.

Figure 6.5: PCA Hedge Fund versus Peer Group Index Spiderweb.



This figure shows the Hedge Fund Peer Group Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. The PCA Hedge Fund Index has higher scores than the Peer Group Indices on all metrics. That said, it is clearly sub-optimal not scoring 7 on any of the criteria of appropriateness.

PCA can be used to generate an index and it is postulated that this may be used to create synthetic hedge fund returns, a topic covered next. Such synthetic benchmarks already exist in the form of the *HFR Risk Parity Indices*. This series is designed to reflect the performance of the universe of managers based on targeting risk levels. the *HFR Risk Parity Indices* does not, however, break funds into sub-groupings based on style, something which PCA indices do.

6.6 Discussion: PCA indices and synthetic hedge fund exposure

Having introduced PCA indices, this section speculates on how they can be used as the basis for synthetic hedge fund replication. The PCA contender index replicates five hedge fund factor risks. It can therefore theoretically be used to construct synthetic portfolios based on the PCA identified eigenvectors. These portfolios could also be a *de facto* proxy benchmark. As a result, such portfolios can be used as an alternative to critiqued peer group indices.

Synthetic time series are generated using financial futures, a low cost alternative to direct investment. Futures allow an index to be created in the same way as the prior chapter using futures to replicate the returns of the underlying time series. Synthetic returns, as they are called, are easier to create and back test, than those actually produced and reported on in a delayed fashion once a month by hedge funds. This property means the proposed contender index improves on the current benchmark construction method (which relies on data that is less transparent).

Synthetic hedge fund factor replication is pioneered in a series of papers. Roncalli (2007) replicates the *HFRI Fund Composite Index*. This series produces the closest indices to the PCA method. They create a time series of intra-day indices based on an optimized composition of liquid positions. These mimic a representative hedge fund. The *HFRI Indices* are sample portfolios for equities and fixed income securities that are based on the most liquid holdings of hedge funds. They are designed to closely follow and to capture the breadth of hedge fund industry trends across all strategies and regions. The indices are equally weighted while the constituent funds of the *HFRI Asset Indices* are weighted by the assets under management (reported by each fund for the previous month).

A synthetic approach is valid because the proposed returns of the PCA contender index are statistically very similar to the returns of hedge funds, but without the usual drawbacks of liquidity, fee, capacity or fund management issues. In terms of synthesising their returns, it is generally accepted that hedge fund returns can be separated into alpha and beta components. The proposed contender index in this chapter could, theoretically, separate beta risk factors. This is because PCA identifies beta factors that are distinct from alpha. This means that the resultant proposed PCA derived index is useful when it comes to skill attribution.

The literature supports such an extension of the use of PCA indices. Papanikolaou, Lo, and Wong (2001) show that securities and complex derivatives can be synthetically replicated by dynamic trading strategies called epsilon-arbitrage. While the method is worth consideration, the replicating strategies are quite complicated and as such cannot be used to build a benchmark that can be easily replicated. Indeed, PCA indices are easier to replicate. The epsilon-arbitrage approach synthesizes risk profiles and replaces standard rolling panel regressions with a Kalman filter. As a result, the risk-return profiles then share several properties with those of the hedge fund indices in much the same way as proposed for the PCA index.

According to Wallerstein, Tuchschmid, and Zaker (2010) there are three main approaches to hedge fund replication. The first is based on linear regression analysis of returns, the second is based on dynamic trading techniques to replicate a payoff, and the third is reverse engineering. The closest to this research approach is linear regression analysis proposed by Jaeger and Christian (2005) and Hasanhodzic and Lo (2008).

The synthetic approach to benchmark construction was also favored by Hasanhodzic and Lo (2008) who constructed linear clones of individual hedge funds using the same Lipper TASS database. They created synthetic hedge funds based on six common risk factors, namely credit, US dollar, bonds, volatility, commodity exposure and alpha. They calculated the weights of these factors by estimation, effectively reflecting the returns of individual hedge funds on the risk factors.

In summary, factor model based benchmarks can also be derived from common components. These can be used to create portfolios that can mimic underlying hedge fund strategies. There are some papers in the literature that have pursued this line of thought with hedge funds.

6.7 Conclusion

This chapter addressed the benchmarking of hedge funds. A PCA derived index was presented as an alternative and tested for optimality using the GRS test. The results indicated that the PCA index did not reject the null hypothesis of risk adjusted performance. The factors identified through the PCA approach, however, were shown to have a useful explanatory role. It built on factor indices, introduced in chapter three. They have the premise that an asset class risk premia exists and that this can be captured in a systematic low cost way by an index. Hedge fund strategies aim to exploit multiple risk factors and as such PCA can be used to decompose the eigenvalue of a hedge fund returns matrix and thereby capture such factors.

It was shown that as a variable reduction strategy, PCA is closely related to hedge fund factors. In this respect, PCA preserves the orthogonal nature of the underlying hedge funds which factor analysis cant do. This gives it an advantage in respect of hedge funds sub-groupings. The returns related to strategy leverage are therefore more likely to be grouped together. As shown in chapter five, with PCA the first eigenvector is the dominant one and is a proxy for the market factor.

The chapter used the Lipper Tass database. This is the one most commonly used to benchmark hedge funds by academics and addresses some of the concerns about survivorship bias. Individual funds typically use their own custom benchmarks to determine their performance fee, so the PCA approach was useful in focusing on how the funds covaried with one another rather than performed against their own benchmark. The PCA analysis showed that due to the non-linearity of the underlying time series of hedge funds, traditional risk-return ratios should be treated with caution.

The discussion on PCA indices led to a recommendation to use PCA to create synthetic benchmarks for hedge funds. This addresses the peer group benchmark issues and allows for daily pricing and gross of fee returns. In this way, indices can be created that can be replicated by investors.

In conclusion, the chapter's empirical results confirmed that it is possible to group the returns of hedge funds through the use of PCA. Superior performance attribution is important because it enables investors to evaluate investment strategies. The time series is independent of the full universe of underlying funds, and/or the Capital Asset Pricing Model. The contender index is both investable and low cost. The GRS test, similarly, proved informative. The results show the PCA approach is a valid index for hedge funds but not a replacement.

Chapter 7

Real estate benchmarks, challenges and refinements,

This essay explores what makes an appropriate real estate benchmark. It analyses real estate returns using principal component analysis (PCA) and investigates the problems associated with measuring real estate price discovery. It provides useful insight into how to construct an index in this illiquid asset class. Real estate only generates a reliable pricing event when a property changes hand in open market. Such transactions are infrequent. Real estate indices are therefore not reflective of passive investment outcomes. The PCA derived index approach is tested on real estate Investment Trusts. As a result of the investigation, weaknesses in the current approach to benchmarking real estate returns is identified. The PCA approach did not, however prove to be a statistically reliable alternative to existing indices. The lack of liquidity and homogeneity are identified as significant constraints to the construction of representative indices. Rental income, a part of total return, and the costs associated with it are found to be impediments to the use of the PCA approach. The results of the PCA analysis show that there remain significant barriers to benchmark measurement and attribution in real estate as an investment asset class.

7.1 Introduction

This chapter addresses investment benchmarks from the perspective of investment in real estate. The main aim of this chapter is to test the appropriateness of the principal component analysis (PCA) index approach as applies to commodities. The literature suggest property indices are important evaluation tools. There is not, however, a representative liquid index that reflect the dynamic nature of the asset. The closest proxies are indices based on Real Estate Investment Trusts (REIT's), which are listed with daily pricing on some stock exchanges.

Definition 7.1.1. Real estate is an interest vested in an item of real property; (more generally) buildings or housing in general - Stevenson (2011).

A real estate benchmark involves many dimensions that are mostly captured in existing indices. The literature, however, identifies that real estate indices face representative issues as a result of the illiquidity and infrequent pricing of property. They only price known transactions, failing to reflect the true nature of the underlying market. This essay therefore investigates whether principal component analysis can address the gap. In doing this, it reviews the appropriateness of this approach for real estate index construction. In order to do this, a REIT PCA index is built, and its superiority to existing indices is tested using a GRS test as proposed by Gibbons, Ross, and Shanken (1989).

Real estate is one of the largest investment asset classes. There are an estimated two billion private dwellings in the world, and an unmeasured amount of commercial properties. According to the *MSCI Annual Update (2018)* on the size of the professionally managed global real estate investment market, the size of the professionally managed global real estate investment market was USD 8.5 trillion in 2017. As such, the need for a clear benchmark is apparent.

There is both a private and public market place for real estate, both of which have poor price discovery. In the former, transactions are done either at an auction or by private treatise. In the later, indirect investments are made by property investment vehicles such as REIT's or Real Estate Operating Companies (REOE's). Overall, price discovery in REIT's is better than REOE's, as the later are listed on a stock exchange.

REIT's are used as proxies for real estate in a number of scholarly studies. As collective investment vehicles, REIT's are deemed alternative assets in their own right. That said, their returns are less correlated to those of traditional property investments, although their underlying investments are in physical assets. This investigation adds to what is known about the shortcomings of existing index method. It builds on the insights extracted from the literature on real estate indices and the issues associated with their construction.

Geltner (2001) identified the characteristics that a real estate benchmark should have. Most important of these is that they should be useful for both practitioners and academics. The desired index characteristics are fairly similar to those of traditional asset classes, namely the requirements identified in chapter three. The other characteristics are that a benchmark should be measurable, investable, pre-defined, appropriate, and should not be susceptible to manipulation. For a benchmark to be useful to academics it should also be transparent, accessible and have incorporate risk. With real estate, risk includes the additional dimension of leverage.

The lack of homogeneity and the liquidity in physical real estate are the main reason that many of the optimal benchmark requirements identified in the literature are problematic. As a result of this, existing real estate benchmarks tend to focus on investment property. This is defined as assets that are owned in order to generate investment returns. Such investment related real estate is distinct from owner-occupied and leased private property, as documented by Maxwell and Saint-Pierre (1998).

The issue for benchmarking is that such indices can measure achieved prices, but they cannot be replicated. PCA indices, however, can be replicated. This characteristic addresses the index challenges documented by Geltner, MacGregor, and Schwann (2003) when benchmarking real estate, namely to:

- find a liquid proxy with which to base measurement upon.
- classify the type of property in a comparable way.
- provide a methodical and less subjective way to value properties.

It is clear that these challenges make identification of a liquid proxy problematic. As explained, this is addressed through the use of REIT's as a proxy. This is because these instruments are listed. The characteristics, style and performance of REIT's are documented by Cheng and Roulac (2007) and Liang and McIntosh (1998), amongst others. They demonstrate that REIT's can be used to measure changes in demand in real estate investment. They find they do not correlate with the demand for physical property.

In summary, as the real estate investment universe is comprised of very different underlying assets, there is a need for differentiation and appropriate indices. To achieve this, the specific investment indices should be able to address different real estate investment strategies. These include commercial, office and residential. REIT's have segmented strategies. The common components of these need to be identified in order to see if the classifications of existing indices are reflective of their variance and co-variance. To do this, PCA is applied to a sample of REIT's and an index is constructed based on the eigenvalues.

7.2 Background

Although real estate is an academic sub-genre in its own right, there is not a great deal of literature on real estate benchmarks. That said, Sirmans and Worzala (2003) produced a good critical review of the literature on international direct investment in real estate, which indirectly relates to the benchmarks used. They start with a traditional chronological approach and end with highlighting the literature that uses a traditional finance framework. Most papers, they point out, use the lens of variance in returns to examine international real estate as an alternative asset class. Like most traditional asset class studies on diversification, they point out that the literature shows the financialization of real estate improves risk reward. The caveat is made that these gains are often reduced when currency risk is taken into account. In other words, greater investment in the asset class makes its pricing more efficient. They highlight the need for better real estate benchmarks.

According to the literature, a major challenge in creating an appropriate benchmark for this asset class is the subjective nature of real estate valuations. They are made without a common set of assumptions and are based on estimates from a number of locations and property functions. The outputs are not comparable, which makes objective measurement difficult. A review of valuation and appraisal methods by Pagourtzi et al. (2003) highlights these issues. They note that valuation surveys are expensive and time consuming. It is not uncommon for there to be only one investment valuation a year. This hinders the pricing of real estate index constituents. Whilst REIT's also face the same issues on the valuation of their underlying assets, the literature suggests that the price discovery process in the stock market, based on the Efficient Market Hypothesis, reflects known information on those assets.

The infrequent pricing of underlying property will always present an issue for index constructors and as such has been subject to academic enquiry. It was investigated by Schwann (1998). The chain linking approach to index construction was introduced in chapter three (section 3.3.10). He suggests addressing the the lack of pricing information by linking current transactions to previous transactions.

The geographic dispersion of real estate also presents benchmarking issues and makes the management of investment portfolios difficult. Diewert, Fox, and Shimizu (2016) review existing international real estate indices from a location perspective and conclude that they are generally biased and don't provide adequate sub-level indices to be reflective. This provides another motivation for the testing of PCA derived indices

There is the lack of uniformity in the formula used by investors to construct the various physical property based benchmarks and real estate indices. The equity index methods introduced in chapter three are all applicable, although total return of listed investments should be replaced by the capitalization rate. This is equal to the annual net operating income divided by the market value. The formula is shown below.

$$\text{Capitalization Rate} = \frac{\text{annual net operating income}}{\text{cost (or value)}} \tag{7.1}$$

The capitalization rate is subject to great variation between properties because its dependent on future income. It is also sensitive to variations in the discount rate. Jud and Winkler (1995) developed a model that explains how capitalization rates are impacted by the weighted average cost of capital and the capital asset pricing model.

Another challenge in benchmarking physical property investment is that real estate types differ. The different categories need to be aligned with an index in order to provide meaningful attribution. This is especially relevant because property can be owned by institutions, investment funds or individuals, all of whom have different objectives. Institutions more commonly invest in office buildings, shopping centers and industrial properties. Individuals prefer residential properties. Existing benchmarks therefore differentiate between commercial and residential properties and investment.

Creating a diversified benchmark that is replicable is further complicated by the unit cost of buildings. It is not easy to sub-divide these. This make real estate investment diversification difficult and therefore equally difficult to capture in a replicable benchmark. Further, as a result of infrequent transactions there tends to be limited historical time series observations and hence no reliable method or model for calculating risk in real estate investments. Recall that identified risk as an important component of an appropriate benchmark. Addressing these and the earlier challenges will need a few industry developments. These include the development of a real estate futures market. That entails a greater securitization of commercial real estate.

7.3 Real estate as an asset class

Real estate has always been considered an asset class in its own right, so it is somewhat surprising that the state of benchmarking is less developed in this alternative asset class than many others. Indeed, according to Andonov, Eichholtz, and Kok (2015), real estate is the most significant alternative investment category. They point out that returns are driven by size, type of investment and geography. In addition to being one of the largest parts of most individuals net worth, real estate investment also has an important role in multi-asset institutional portfolios. As such, it requires proper benchmarking as a separate asset class. Real estate, as Chen, Roll, and Ross (1986) point out, is a long duration asset. As such, it is suitable for pension fund investment and benchmarks need to be long term in nature to reflect this. A long term time series can iron out the inefficiencies in short term real estate price discovery. Although a distinct asset class, the lack of a liquid futures market and/or other hedging instruments is one of the biggest impediments to devising an appropriate index.

The literature suggests a large number of factors impact real estate returns. There are many such variables, a few of these are detailed in figure 7.1. These range from size to location. Attribution of these variables is a requirement of a suitable benchmark. Some of the elements of return are idiosyncratic, for example being awarded a building permit can result in revaluation. Even a change in availability of utilities, such as water connection, a refuge collection, and electricity has an impact on returns. These are difficult to capture in any benchmark.

Figure 7.1: Factors impacting real estate returns.

- Location. Address and district, census or administrative authority.
- Number of rooms. The room counts with the possible listing of certain room types, For example, commercial offices and / or bedrooms or bathrooms for residential buildings.
- Number of units after use. In the case of residential real estate, a unit is normally an apartment with only one owner.
- Physical condition. This refers to the interior and exterior maintenance, and covers evidence of damage and other maintenance issues.
- Purpose. The use of real estate unit, such as apartment, shop or shop, factory or government offices.
- Rental information. The amounts paid to rent the property in total or in the form of the cost per unit of room and / or square meter.
- Size. The square meters of the structure and the plot.
- Term. The status of the occupant, ie owner (property or lease), private leasing or rental in the public sector.
- Type. Concrete, wood and brick.
- Unit, construction. For the living area, the single-family or semi-detached townhouse or the apartment can be used.
- Free units. Whether the property is empty or not.

Although all these variables impact valuation, Pagourtzi et al. (2003) note that the date and value of the last comparable sale of a property is the biggest factor used to determine current market value. This can be captured in physical indices but is not suited to the PCA approach. The problem us that pricing has to be done manually by an expert who know the variables that contribute to the market price. That is a subjective rather than objective input and therefore difficult to incorporate into index design.

A note of caution needs to be expressed on the use of PCA. REIT's are priced by equity investors. As a result, they exhibit an equity market risk factor. Within PCA, this partial determinant of their price will theoretically manifest itself in the size of the first eigenvalue, thereby distorting the weights of a PCA derived factor index.

As an asset class, the return that REIT investments deliver investors is typically believed to be superior to fixed income, which like real estate has a running yield and more similar to equity investments. Despite this, Newell and Marzuki (2016) find UK REIT's delivered poor risk adjusted returns compared to equities between 2007–2014. This suggests that a PCA derived index based on REIT's might also prove sub-optimal compared to an index based on physical transactions.

7.4 Real Estate Investment Trusts (REIT's)

As stated, a REIT is a listed collective investment vehicle. REIT's allow investors to gain real estate exposure without direct ownership. REITS's however, are not geographically diversified. They can trade at a discount of premium to net asset value, which can have implications for index weight and the reflection of the underlying net asset value. Indirect real estate investment can be done in a variety of other ways, including but not limited to unlisted wholesale funds, direct private funds, property syndicates and unlisted retail funds. That said, the listed nature of the REIT's makes them easier to use in the construction of indices.

The classification of REIT's as an asset class is subject to much discussion. For example, Corgel, Mcintosh, and Ott (1995) point out that returns on such instruments should be reflective of the underlying asset in theory, but in practice they are not. They are in effect as securitized claims without the transaction costs. Their time series returns therefore have a different dimension, as already flagged. They referenced a study by Gyourko and Keim (1992) to illustrate this. That study shows that REIT returns are highly correlated with the home price series of the *National Association of Realtors*, a US benchmark. This provides evidence that a theoretical link between home prices and the securitized property market exists. This study also found that the betas of real estate firms, a measure of how an individual real estate prices move when market moves, differ across different market segments.

7.4.1 Impact of REIT corporate structure (Leverage)

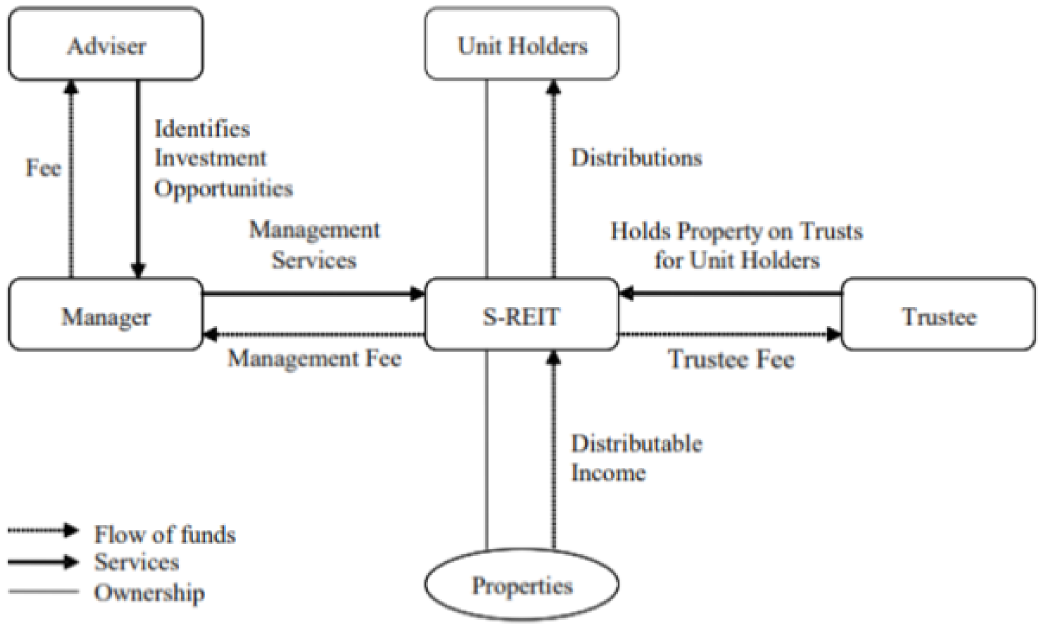
The literature suggests that there is a caveat that applies when using REIT's as a proxy. That is that they employ leverage in their operating structure. Although REIT's are collective investment vehicles the leverage in their corporate structure that makes their returns differ from those of physical property. In addition to the leverage, REIT's also pay dividends which alters the way returns are measured.

The REIT corporate structure is shown in figure 7.2. There has been scholarly investigation of leverage in REIT's. Feng, Ghosh, and Sirmans (2007) evaluate whether the leverage ratio is temporary or persistent. They find that REIT's with high market to book ratios have persistently high leverage ratios. This may be because REIT's that have good growth prospects are able to raise funds through debt issuance.

Scott (1990) and Mei and Liu (1994) further clarify the nature of REIT returns. They suggest returns lag real estate capitalization rates but lead un-smoothed appraisal generated returns. In other words, there is an element of forward prediction discounted into the price of the equity. The returns do, however, share a strong common component to real estate fundamentals which justifies their use in this chapter. That said, these authors claim REIT returns are considered noisier than traditional real estate returns. They find that the leverage in this sector is amongst the highest of that found in standard industries.

Ibbotson and Kaplan (2000) were the first to compare REIT benchmark returns with other investments. They point out that non-risk factors influence its time series, such as tax shelters, imperfect marketability and information costs. They address this element of returns through what they term *New Equilibrium Theory*. Research on REIT's by Burns and Epley (1982) and Ennis and Burik (1991) corroborate this. It indicates that the addition of real estate as an asset

Figure 7.2: REIT Structure



This figure shows the way a REIT is structured. It is a collective investment vehicle that charges a management fee for holding property pn behalf of investors. The REIT is listed on a stock exchange. Its price dynamic is determined by buyers and sellers of the equity and as such is an indirect investment in real estate. Source: CFA Institute 2011

class improves the mean variance efficiency of a diversified portfolio. One has to be aware, however, that real estate is typically purchased using a mortgage and as such most real estate investment involves some form of leverage.

7.4.2 Skewness

The corporate structure and the leverage in REIT's skews their returns. Lizieri and Ward (2001) suggest that the skewed nature of real estate returns also impacts their benchmark returns. This is relevant for determining what makes an appropriate benchmark.

Academic research has shown that returns are non-normal in both the US and international property markets and that there is evidence of skewness and kurtosis. This was demonstrated in physical property by Young and Graff (1995), and in the REIT returns by Mei (2000). The later suggested the distribution could be stated based on their log characteristics as follows:

$$\varphi(t) = \begin{cases} i\delta t - |ct| \left[1 - i\beta sgn(t) tan(\pi\alpha]/2\right], for\alpha \neq 1 \\ i\delta t \, |ct| \left[1 + i\beta(2/\pi) sgn(t) logl[t]\right], for\alpha = 1 \end{cases} \tag{7.2}$$

Where:

c = the scale parameter

δ = the location parameter

β = the skewness parameter

α = the characteristic variable

In order to illustrate this, table 7.1 shows the skew and kurtosis of the *FTSE REIT NAREIT Index* from US from Dec 1971 and July 2018. The results show, as expected, that there is a marked -0.945 negative skew. There is also a 7.11 result for kurtosis. This implies investors will experience occasional unusually high returns on the upside.

Table 7.1: REITFTSE NAREIT US Real Estate Index - Skew and kurtosis

REITFTSE NAREIT U.S. Real Estate Index					
Series December 1971 - July 2018					
	N	Min	Max	Mean	σ
V2	222	-30.22	27.97	1.05	5.72
Valid N (listwise)	222				
Skewness and Kurtosis					
V2					
N		222			
Skewness		-0.945332583			
Std. Error of Skewness		0.163302645			
Kurtosis		7.119642511			
Std. Error of Kurtosis		0.325189711			

This table shows the analysis of the skew and kurtosis of the largest 15 US REIT's. This is a measure of the asymmetry of the probability distribution of the REIT's price returns around their mean. The data exhibits a clear skew in returns which is possibly explained by the leverage inherent in the REIT instrument.

In summary, the distribution of property returns appears to behave in a different way from traditional assets. A PCA index can be constructed despite the presence of skewed returns but its time series would likely be non-normal.

7.5 Benchmarking approaches

The many approaches to index construction were detailed in chapter four. Ferrero and Morri (2004) investigated how one would go about benchmarking real estate investment. They identified two broad categories, those used for (1) performance and research and (2) those used for direct and indirect indexes according to their construction method. Figure 7.3 depicts these differences from the perspective of the investor. The key differentiation being between equity and debt on the one axis and public and private on the other.

Figure 7.3: Public and private real estate investment exposure.

	Equity	Debt
Private	Direct Investment in Real Estate	Direct Investment in Real Estate Mortgages
Public	Real Estate Investment Trusts (REITs) Real Estate Operating Companies (REOCs)	Commercial Mortgage-Backed Securities (CMBS)

This figure shows the different ways an investor can gain exposure to movements in real estate prices. Individuals are represented in the private rows, either owning outright or with a mortgage. The more public forms of ownership include through REIT's, REOC's an CMBS. Source: Ferrero and Morri (2004)

Real estate benchmarks fall into two categories, (1) a reflection of actual transactions and (2) an index based on the performance of collective investment vehicles that invest in real estate. Typically, these can be broken down into two broad investment groupings.

1. Individual investors who focus on house prices. They are focused on *residential property prices*, and *dwelling prices*. These terms are used interchangeably to describe such residential properties. They are benchmarked using the repeat sales, multiplicative chain and/or multiplicative chain method.

2. Institutional investors who focus on collective investment vehicles. These investors often view direct real estate investment as time consuming and costly and use proxies. The larger ones may invest directly. Those that use proxies invest in REIT's and/or Mortgage Backed Securities. The pricing of these collective investments is related to the cash flows from real estate investments and is not related to their capital value. These collective vehicles require proper benchmarking in the form of a liquid index.

The literature review in chapter two showed that the market for physical real estate is not efficient. Risk was shown by Young and Graff (1995) to be heteroscedastic. They observe investment in real estate is slow, expensive and cumbersome. Also, that many of the basic assumptions behind the Capital Asset Pricing Model do not hold for the asset class. Real estate, for example, is not traded in a central marketplace; transactions are privately negotiated, there are few participants and only occasional trades; public availability of information is limited, and it is difficult to observe the conditions that achieved transaction prices. At the same time, almost all direct real estate involves transaction costs, stamp duty and often capital gains taxes. Real estate is, as a result, unique and heterogeneous. It is not possible to gather adequate data on the influence of market participants or predict pricing with any certainty.

The literature highlights that the aforementioned inefficiencies present issues in aligning real estate benchmarks with the traditional market risk factors. A suggested approach to address this, through the cross section of returns, was presented by Elton and Gruber (1997). They estimate real estate risk components. This was achieved by firstly calculating the monthly equal weighted holding period variance in return for each individual property, and then secondly, the variance of a diversified portfolio. They then compared the non-diversified return variance to the diversified variance. They conclude that if this ratio is low, traditional systematic influences are low, and visa versa. This adjustment is not necessary in a daily priced PCA index built using REIT's.

The estimated rate of return (the capitalization rate) is the most basic benchmark for real estate. It can be depicted as the annual net operating income stated as a function of its cost or value. A benchmark can therefore use these to estimate a property's value at the end of its holding period. As a benchmark, however, this approach is critiqued as it does not take account of leverage, the time horizon or future cash flows from property improvements or degradation.

Whether real estate is invested directly or collectively, the literature suggests that the key considerations that determine the price are the type of property, any historic transaction price, the market value of similar properties, and the appraised value. As mentioned, there are benchmark construction issues in respect of each of these and there is a trade-off between random measurement error and temporal lag bias. There are also issues with valuation reporting and property revaluation frequencies.

The main commercial real estate index families that cover the performance of listed commercial entities are *Alpha Shares, Dow Jones, FTSE, Northern Trust, Standard and Poor's* and *Wisdom Tree*. Each of these providers cater for different types of real estate investors. As such, these providers have a host of sub-indices based on property type. The *FTSE NAREIT US Real Estate Index* Series, which are used in the analysis in this chapter, are a range of REIT focused indexes that measure price changes for listed commercial real estate funds. The sub-indices constituents are market cap-weighted and reviewed quarterly, calculated in US dollars, Euros, British pounds and Japanese Yen.

The volatility of price movements was addressed by Bailey, Muth, and Nourse (1963b). They proposed an alternative regression method for real estate price index construction. They argued their method reduces quality differences by introducing qualitative characteristics in the period effects. This could be said for PCA indices as well. Their model is essentially a regression whereby:

$$R^{itt'} = \frac{B_{t'}}{B_t} X^{U_{iit'}}, or \tag{7.3}$$

$$^{itt'} = b_t + bt' + U_{iit'} \tag{7.4}$$

Where: $R^{itt'}$= the ratio of the final sales price in the period

t' = initial sales price in the period t for the ith pair of transactions with initial and final sales in the two periods.

$B_{t'}$ and B_t are the true but unknown indices for period t and t', respectively, where $t = 0, 1, ..., T-1$

The geographic and functional difference between properties makes real estate a more idiosyncratic asset class than any other. Benchmarks have to be geographic in nature. Also, each property is unique, they have different unit sizes. This is why real estate transactions have poor transparency and asymmetric information. The time series is therefore heterogeneous. When combined with the illiquidity and high costs of either sale or valuation, the characteristics results in a skewed time series of return for the asset class and the resulting benchmark index. That said, the returns can still be broken down into rental income and capital.

Rental is an important part of the return on a real estate investment benchmark but one that is hard to capture. There is no homogeneity in rental payments, They can either be paid in advance or in arrears, monthly or quarterly. Direct costs need to be subtracted from returns. These include insurance, maintenance, ground rent and repairs. The many indirect costs are lumpy and unpredictable. A PCA index would capture rental as it uses gross returns, however there would be no ability to attribute this element of the return.

Capital growth and or decline is also an important benchmark return contributor. This capital element is difficult to estimate. Holding periods vary and contractual negotiations take a long time. Real estate sales are poor reference points for valuation due to the long lag between transactions. As a result, benchmarks typically rely on subjective valuation. Ball et al. (1995) identify a number of issues related to this process, Lizieri and Ward (2001) suggest, therefore, that the complexity is such that one should use un-adjusted returns over a time window surrounding the valuation date, using a Bayesian adjustment as new information comes to light.

It is possible to create synthetic returns for real estate. That said, this cannot be done through the futures market. It can be achieved by estimating a yield from public rental data and then aggregating the results into a geographic index. Another approach to create synthetic returns is based on a portfolio of representative properties and then using extrapolation techniques. Neither approach is perfect from an index construction point of view.

Despite the challenges, there are a number of commercial indices which can be used for benchmarking real estate investment. These are (1) direct indices such as the *Case-Shiller Index* and the *Halifax Index* series and (2) collective indices that are based on REIT's such as the *Bloomberg REIT Indices*. These are addressed next.

7.5.1 Repeat sales method - The Case-Shiller Index

The *Case-Shiller Index*, developed by Case and Shiller (1988), uses the repeat sales method. It provides an index of sales prices of the same unit of property over time. Its approach is used to estimate changes in real estate prices at an aggregate level over a long time period.

The *Case-Shiller Index* is one of the better known real estate indices. The base price index for the index an observed price followed by subsequent changes of the prices on individual properties. The index is calculated by the price change between two sales of the same family house. A potential drawback of this widely used index is that it is reported only monthly rather than daily. The PCA index method addresses this, as it is based on quoted REIT's.

The multiplicative chain method is used where properties are sold infrequently. This method is adapted for real estate by Wyngarden (1927). The resulting index is a geometric mean of price relatives based on houses sold. This is calculated on an original base period comparing those properties sold during the same period. Then a price index is created for all pairs of transactions and then adjusted by the index of the base period. The price index for the second period then becomes the geometric mean of all price relatives.

The multiplicative chain method is often critiqued as it neglects information about earlier periods where there had been no final sale. At the same time, the calculation of the standard error or the time series is difficult. The critique can be overcome by a process called *mix adjustment*. Using this, a group of properties is divided into sub-groups or cells according to characteristics, such as location or number of bedrooms. An average price is then calculated for each cell. This allows for the calculation off sub-indices based on property type (for example, a three-bedroom property, with a garden and garage, in Glasgow). The resulting index can then be calculated as a weighted combination of the various sub-indices.

7.5.2 Halifax House Price Index

In the United Kingdom, the *Halifax House Price Index* is the most cited measure of home prices. It is based on a commercial mortgage provider's transaction data and has been calculated since January 1983. The index is also cited in the economics and finance literature. Wood (2005) provides a comparison of this index with other UK residential house price indices.

The historical data makes this a particularly useful index, as does the depth of observations. It is based on using datasets comprising five monthly indices, sixty-five quarterly indices and fifty two annual indices. These cover twelve regions which all have their distinct price dynamics.

7.5.3 Artificial Intelligence based indices

The advent of online valuation services, like *Zoopla* for retail property, could well change the benchmarking dynamics of real estate in the future. These services utilize Artificial Intelligence and Machine Learning to give an updated estimate of value based on recent sales of similar properties and local price trends.

Abidoye (2017) suggests that it is possible to accurately price property by using artificial intelligence. This approach could then be used to link current transactions in similar geographies using transactional histories. This would allow for better pricing of real estate in the future and thereby facilitate real time indices. Indeed, there are commercial tools that are proving very accurate current pricing of properties based on statistical input. That said, artificial intelligence is not an immediate solution for the improvement of real estate indices.

An index could be constructed that captures such price dynamics once the results become more reliable. The outputs will, however, always be determined by the referencing test data-set from which it bases its assumptions.

7.5.4 Real Estate Investment Trust Indices

In the literature, REIT indices are the most widely used. There have been a few cross sectional studies on the existing real estate indices. Quan and Titman (1997), for example, use a number of such benchmarks to compare the performance characteristics of real estate markets with their respective stock markets. These included the *JLW all-cities index* and the *Frank Russell Company indices*. Stevenson (2011), meanwhile, proxy-ed US direct real estate with *NCREIF*, US indirect real estate with the *NAREIT equity index*, UK direct real estate by the *JLW Property Index* and UK indirect real estate with UK property companies. Titman and Warga (1986) took the approach of building benchmarks from a sample of REITs. They made these these into equal and value weighted indices

The broad categories of institutional real estate are *Commercial* and *Retail* and these are reflected in the REIT index families. REIT index families also have three major sub-categories based on their approach to investment. These are direct investment in properties (equity REIT's), the purchase of mortgage debt on income producing properties (mortgage REIT's), and a combination of both (hybrid REIT's). Much of the literature has segregated the market in this way. That said, as investment in the sector has grown, REIT's are now further defined based on functional type of property, such as hotel, office, industrial, residential, healthcare, retail, and a broad classification of diversified REIT's.

In addition to the REIT classifications, real estate can be further sub-categorized based on its unique differentials, but as yet no benchmark has been able to capture this granularity. If this trend for segmentation continues, PCA may have a greater role in differentiation of sub-groupings going forward.

The traditional market capitalization methodology presents problems for REIT's. Using such an approach, a weighted REIT index would overweight or underweight based on the legal entity, thereby not taking into account the financial structuring that includes debt, minority interest, and preferred stock. As a result, a market capitalization weighted REIT index would re-balance by selling off and buying into positions that reflect only a portion of the value of the underlying assets. An enterprise value weighted index would better reflect the economic business cycle. The *Nasdaq ETRE Indices* address this issue by being free floating enterprise value weighted.

According to Clarke et al. (2015) there is some evidence that within the time series of REIT's, broad equity factors are becoming less important than specific real estate factors. This is also positive for future usage of PCA as it can be used to identify such factors. That said, as the analysis shall show, they are yet to manifest themselves as sufficiently differentiating at present.

As an alternative asset class, real estate is a tangible asset, offering a reasonably forecastable rental income stream. The time series tends to deliver long-term positive real returns by virtue of the rental contract linkage to inflation. This inflation-linked income stream, combined with the lack of liquidity and high transaction costs, typically make real estate investments long-term in nature. Real estate benchmarks are not just useful for investors but also macroeconomic analysis. In this respect the asset class is a major component of both public and private wealth. As a result, understanding real estate price changes is important for institutions that not only invest in but have large lending exposure (both direct and indirect) to real estate. This is because the volatility of price movements impacts their margins.

7.5.5 REIT index methodology

The index method for REIT's is the same as for listed equities, as explained in chapter three. In this way, it is possible to construct REIT indices which are priced daily. The listed real estate vehicles offer investors transparent, frequent and regulated pricing. The most common index method used in their construction is the Laspeyres index, also introduced in chapter three. This is calculated using a weighted arithmetic average of the real estate price relatives and the values of the earlier periods to determine the weights.

$$P_L = \frac{\sum (p_t \cdot q_0)}{\sum (p_0 \cdot q_0)} \tag{7.5}$$

Where:

P_L = is the price of real estate asset L (L = 1,. . . ,m) in period t

Pio = Price of real estate asset i (i = 1,. . . ,m) in the base period

All the index construction issues are compounded by real estate's idiosyncratic return characteristics, something benchmarks need reflect. A paper by Gorton, Hayashi, and Rouwenhorst (2013) reports on the mean returns, standard deviations and correlation coefficients. They concede that getting a reliable long term understanding of real estate prices and returns is problematic. The benchmarks they used as a result was a combination of two data-sets. In the US, they used the *EAI survey* and in the UK they used the *NCREIF Index*, the *JLW Property Index* and the *IPD index.* This chain-linking is often the only way to get an idea of real estate returns over time.

7.6 Data

Data was obtained on 89 REIT's whose shares were traded in the United States on the New York Stock Exchange, the American Stock Exchange or the National Association of Securities Dealers Exchange between 01/01/2001 - 01/01/2018. All data was obtained using datastream. REIT's which did not have a continuous history were excluded. This reduced the number from 225 listings.

The monthly return of each REIT was calculated, measured as a percentage change in the stock price and the return from its dividend yield. Descriptive statistics of the data for

1/01/2001 - 01/06/2018 are shown in tables 7.2. The key REIT's include a number of specialist areas such as diversified, healthcare and hotel REIT's. A correlation matrix is shown at the end in appendix C.

Table 7.2: Descriptive statistics of REIT daily rate of return 01/06/2008 - 01/06/2018 (* denoting incomplete time series)

Name	Symbol	Mean	Stan Dev	Minimum	Maximum
SP500 HEALTHCARE REITS *	SP5SXTL	0.00355	0.05448	-0.09259	0.11297
SP500 HOTEL AND RESORT REITS *	SP5SXT4	0.00786	0.06558	-0.10629	0.18878
SP500 DIVERSIFIED REIT'S *	SP5IXD4	0.01275	0.11430	-0.40536	0.41068
SP500 INDUSTRIAL REIT'S	SP5IXI4	0.03527	0.42979	-0.80076	4.48780
SP500 OFFICE REIT'S	SP5IXO4	0.00733	0.08376	-0.33089	0.33938
SP500 RESIDENTIAL REIT'S	SP5IXH4	0.01105	0.08371	-0.29493	0.27914
SP500 RETAIL REIT'S	SP5IXR4	0.00820	0.10579	-0.46577	0.46207
SP500 SPECIALIZED REIT'S	SP5IXS4	0.01141	0.07651	-0.28087	0.34865
SP 500	SPCOMP	0.00863	0.05108	-0.16638	0.15987

A Kaiser-Meyer-Olkin test was applied to the REIT sample 01/01/2001 - 01/01/2018 in order to see how suited the data was for factor analysis. The results are shown in table 7.3 This suggests that REIT's are a clear homogeneous group. As such, basing an index on them is justified. The result does not, however, say anything about how representative REIT's are to the underlying asset. A Bartlett's test was also done to test whether the assumption of equal variances was true. The variables was selected and the open descriptions were checked for the KMO and Bartlet test. The KMO came in at 0.96, which was extremely close to one and hence a good result. The Barlett test also proved significant. The results confirm the data is suitable for use in a PCA index.

Table 7.3: Kaiser-Meyer-Olkin and Bartlett's Test REIT Sample 01/01/2001 - 01/01/2018

KMO and Bartlett's Test		
Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		0.966
Bartlett's Test of Sphericity	Approx. Chi-Square	24360.85
	df	3916
	Sig.	0

This table shows the results of the Kaiser-Meyer-Olkin test on the REIT sample01/01/2001 - 01/01/2018 which shows how suited the data is for Factor Analysis. and Bartlett's Test which checks that the assumption of equal variances is true before running a one way ANOVA regression.

7.7 Method

PCA was proposed as a benchmarking approach in chapter three and the method is adopted from there. A PCA index was constructed. The data was subject to dimension reduction. This was done on the top 89 REIT's by market capitalization in the Standard and Poor's 500. They were selected from the period 1/1/2001-1/1/2018 and only included those that had continuous histories. The aim was to investigate if PCA could provide explanatory sub-groupings for the asset class. The number is chosen based on their surviorship and uninterrupted time series. A PCA contender index was constructed and the constituent weights for each period were derived using factor analysis. This was implemented using oblique rotations.

The resulting PCA contender index is a weighted average of the returns of the derived portfolio constituents. It can be described by a five-factor vector model. This includes the key existing classifications of residential, retail and office property and the two key strategies employed by REIT's, namely equity and mortgage exposure. The model is detailed below:

$$\zeta_r i, t = \beta 1 \mathbf{v} \eta_t + \beta 2 \mathbf{v} \theta_t + \beta 3 \mathbf{v} \chi_t + \beta 4 \mathbf{v} \tau_t + \beta 5 \mathbf{v} \mu_t \tag{7.6}$$

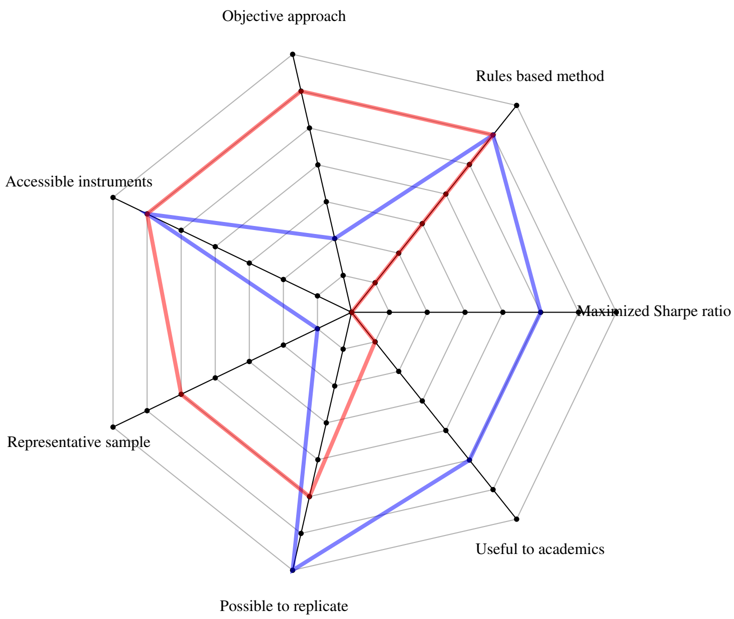
Where:

- $\zeta_r i, t$ = Proposed contender index, the excess return of sector commodity portfolio i in month t, item **v** = an eigenvector
- η = Industrial REIT's complex
- θ = Office REIT's complex
- χ = Residential REIT's complex
- τ = Equity REIT's complex
- μ = Specialized REIT's complex

The index is constructed using this vector factor model based in a step-wise fashion. A three-year observation period is used, along with a six-month re-balancing frequency conducted at each re-balancing date. Table 7.4 shows the factor loadings and the standard deviations.

The resultant index is mapped on the subjective *Real Estate Index Spiderweb* in figure 7.4. This compares the Likert score of the outcome of this approach, as explained in chapter one, with the traditional REIT indices mentioned earlier. Note that it both index approaches have strengths that the other lacks.

Figure 7.4: PCA Real Estate versus Real Estate Investment Trust Index Spiderweb.



This figure shows the PCA Real Estate Index versus Real Estate Investment Trust Index Spiderweb. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. The diagram dimensions are sub-optimal in both instances.

The Likert score for representative sample is only one, reflecting the earlier discussion about the exposure to the equity market rather than the real estate market. The PCA index scores better on being able to replicate and its usefulness due to it being based on liquid instruments.

Table 7.4: Standard deviation and sector loadings of rolling factors

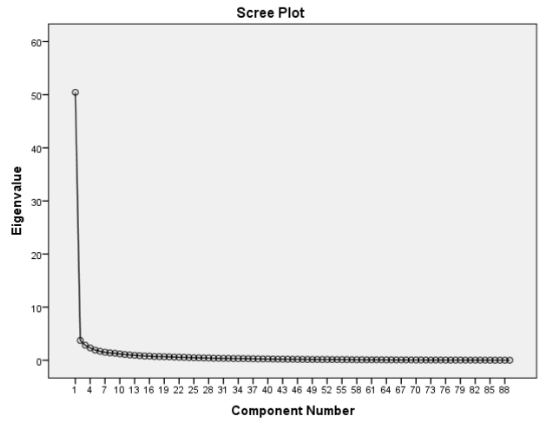
	Factor 1		**Factor 2**		**Factor 3**		**Factor 4**		**Factor 5**	
	Mean	σ	**Mean**	σ	**Mean**	σ	**Mean**	σ	**Mean**	σ
Industrial REIT's complex	0.73*	23.91%	0.18	8.82%	0.13	4.85%	0.07	3.06%	0.09	3.86%
Office REIT's complex	0.16	8.13%	0.13	6.06%	0.75*	22.14%	0.06	4.47%	0.08	6.71%
Residential REIT's complex	0.16	5.48%	0.75*	6.99%	0.11	3.62%	0.08	4.66%	0.12	4.41%
Equity REIT's complex	0.08	3.60%	0.06	3.12%	0.04	10.35%	0.63*	32.89%	0.08	22.12%
Specialized REIT complex	0.17	2.76%	0.26	4.86%	0.14	2.42%	0.09	24.83%	0.76*	24.98%

*This figure displays means and standard deviations for individual hedge fund style grouping loadings on the first five factors corresponding to the rolling factor analysis, with daily observations. * depicts which factor loading a complex is most exposed to.*

7.8 Results

The data file was analyzed using the IBM SPSS dimension reduction software package and the result analysed. One extraction was conducted and the resultant scree plot is shown in figure 7.5. This reports strong evidence of a market factor, as predicted by the literature. All the REIT's have commonalities to this factor at 0.5 or above and many are at 0.8 indicating a confirmed positive result.

Figure 7.5: Eigenvalue scree plot of REIT Sample 01/01/2001 - 01/01/2018



This figure shows the eigenvalues. Note that there are a large number of these but that only the first is significant, representing the market factor common to all equities.

The dominance of the stock-market factor suggests that REIT's are not a good proxy for physical real estate assets. Table 7.5 details the two components of real estate returns identified by this process of extraction and rotation. This also highlights a large stock-market factor.

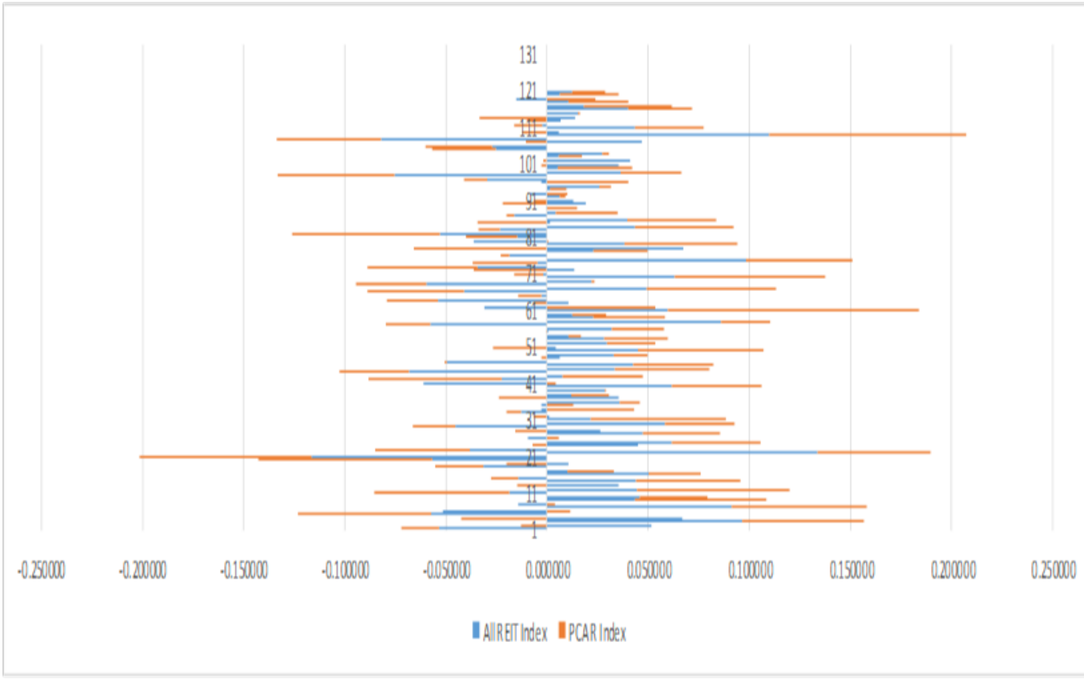
Table 7.5: The two components of Real Estate returns identified by extraction and rotation

Component	1	2
1	0.950653	0.310255
2	-0.31025	0.950653
Extraction Method: Principal Component Analysis.		
Rotation Method: Varimax with Kaiser Normalization.		

It is concluded a PCA index is not a good proxy for physical real estate. It was, however, further investigated if it has merits as a REIT index. To do this, the total variance explained was calculated. The results are shown in table 7.6. The table reports that 95 percent of the total variance was explained by the equity market. This confirms that the PCA analysis is not a good indicator of the broader physical real estate market, but that it is reflective of equity risk.

A time series was then generated and is shown graphically in figure 7.6. This shows that the returns are typically more directionally pronounced than the *All REIT Index*. This is also not surprising as it is done by reducing the principal components to five factors after accounting for the dominant market factor. As such, the model is just a sub-set residual of the total universe.

Figure 7.6: All REIT Index monthly returns versus PCA Index 01/01/2001 - 01/01/2018



This figure compares the REIT Index monthly returns versus PCA Index. As can be seen, there is a high degree of directional correlation which is derived from the equity market factor. The base scale is in basis points, ranging from -0.25 to 0.25.

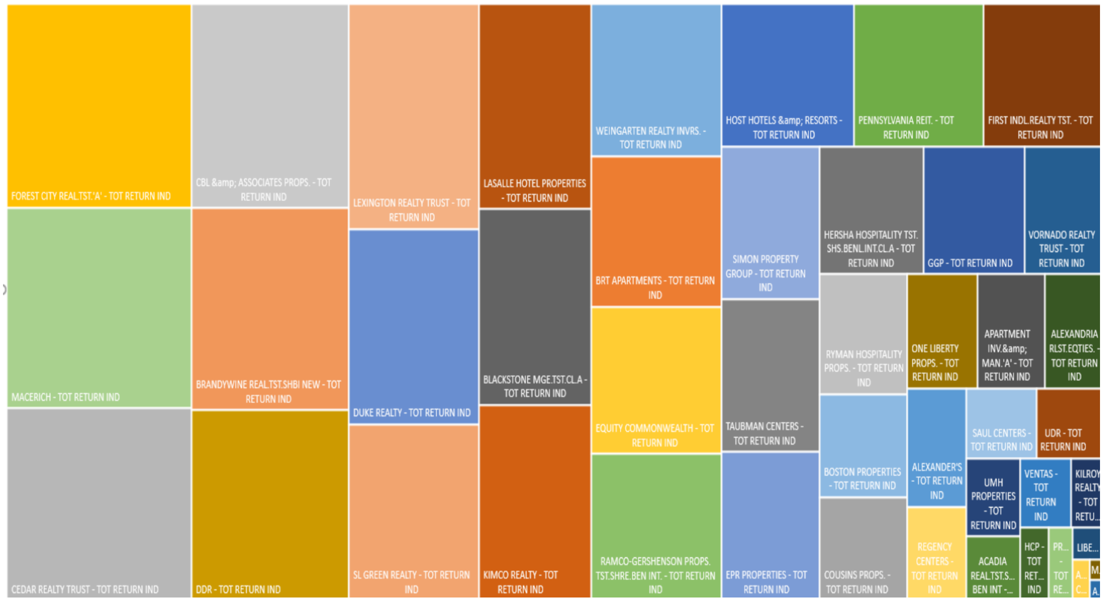
Tables 7.7 through 7.10 report descriptive statistics on the monthly returns of each REIT sector and the REIT benchmark, as well as the monthly returns of the NYSE index and the 30-day treasury bill.[1]. The total variance explained by the five factors and the initial eigenvalues are reported in table 7.6.

The open score is saved as a variable in the default regression and used to check the display factor score. The final results are then used to select the first factors from the data view for

[1]Healthcare and Hotels start from 1/4/2014; Diversified only goes up to 1/1/2015

use as the factor index weights. This results in twelve factor weightings, five of which are significant. This is more than the seven existing REIT sub-sectors of hotels, offices, industrial, residential, healthcare, retail, and diversified. Only five are applied to the factor model. This result may be attributed to the difference in underlying leverage in the capital structure of REIT's. Table 7.6 reports the total variance explained.

Figure 7.7: Weighting tree for PCS derived REIT Index



This figure shows the individual REITS with a box volume depicting their weight. The weighting results file is too big to present in tabular form.

As shown in Figure 7.7, the weights of the contender index are well diversified. Twelve of the REIT's, however, represented 50 percent of the benchmark. The results are cross-checked for evidence of the common market factors and the two components of real estate returns. The 0.95 score indicates that REIT's are a distinct sub-category of equities. The first principal components explain 50.47 of the cross-sectional variation.

In order to demonstrate the effect of tracking error on REIT indices in the same alternative asset class, table 7.8 presents the tracking error of the PCA contender index to the *All REIT Index*, *REIT 50 Index*, *EQUITY REIT Index* and the *Mortgage REIT's Index*. As can be seem the tracking error between these similar assets is between 2.66 percent and 2.44 percent per annum.

Table 7.9. reports the 10 year cross sectional regression on the time series regression on

Performance of PCA REIT Index relative to the daily REIT sector 1/6/08 - 1/06/18. Table 7.10 reports the 5 year cross sectional time series regression on Performance of PCA REIT Index relative to the daily REIT sector 500 using Jensen alpha 1/6/13 - 1/06/18. Note the PCA index performs poorly against the office and retail REIT's.

Overall, the results, are mixed with both positive and negative alphas, so the contender index is deemed not statistically strong enough. It is concluded that the suggested PCA method does not improve on existing benchmarks in real estate. That said, the sample size and lack of geographical diversification could all be said to be factors in the under-performance. The contribution in this analysis is therefore that a larger and more diversified sample is required but also in confirming the presence of a market factor through the first eigenvector, representing a 55 percent common component. Table 7.10 reports the information ratio. This is 0.10, which is considered a poor risk adjusted result. The GRS test is used to examine the mean-variance efficiency of the PCA derived index against the REIT index. The sample is tested based on the null hypothesis introduced in chapter one:

$$H_0 : GRST_t = \eta_i - \eta_j = 0$$

Where:

- $\eta_i = \mu_i / \sigma_i$ = Mean excess returns i (over a risk-free rate)/standard deviations of excess returns i
- $\eta_j = \mu_j / \sigma_j$ = Mean excess returns j (over a risk-free rate)/standard deviations of excess returns j
- $GRST_t$ = is made up of an estimator of the Sharpe ratio difference, the standard error of the difference estimator, the Sharpe ratios and the correlation between the excess returns of strategy i and j.

The resultant GRS F statistic is 0.28, as illustrated in table 7.11. The existing index also has a high p-value of 0.747. It is therefore not possible to reject the H0 for the PCA index. It is found that the test fails to reject the null hypothesis. That is, there is no improvement to the tangent index portfolio on a risk adjusted basis.

The tracking error of 2.85 and active premium of 2.48 indicate the index is sufficiently different from the contender index. That said, the difference is in the weights rather than the instruments. REIT indices are based on market capitalization weights.

Table 7.6 reports the total variance explained by each of the eigenvalues. As can be seen, although there are 15 factors identified, the top five are significant enough to use in the factor model. Table 7.7 illustrates the PCA individual components.

Overall, the results demonstrated that PCA does not add much value in the context of benchmarking REIT's. This is because of the equity market factor exposure, the leverage in their structure and the lack of representation of the overall physical property universe. The equity nature of REIT's partly explains this and as such the results do not invalidate the PCA approach for other alternative asset classes.

Table 7.6: Principal Component Analysis - Total variance explained

Component	Eigenvalues			Extraction			Rotation		
	Total	% of Var	Cum %	Total	% of Var	Cum %	Total	% of Var	Cum %
1	8.392606	55.95071	55.95071	8.392606	55.95071	55.95071	7.782061	51.88041	51.88041
2	2.049817	13.66545	69.61616	2.049817	13.66545	69.61616	2.660362	17.73575	69.61616
3	0.859486	5.729908	75.34606						
4	0.751611	5.010737	80.3568						
5	0.612788	4.085255	84.44206						
6	0.529091	3.52727	87.96933						
7	0.358143	2.387622	90.35695						
8	0.290344	1.935627	92.29257						
9	0.240813	1.60542	93.89799						
10	0.230767	1.538445	95.43644						
11	0.206179	1.374528	96.81097						
12	0.186796	1.245308	98.05628						
13	0.131843	0.878953	98.93523						
14	0.081468	0.54312	99.47835						
15	0.078248	0.521652	100						

This table presents the findings of the Principal Component Analysis based on the total variance explained of each of the eigenvalues.

Table 7.7: Total Variance Explained - Components and initial Eigenvalues

A	B			C			D		
	E	F	G	H	I	J	K	L	M
1	50.437	56.671	56.671	50.437	56.671	56.671	25.908	29.11	29.11
2	3.732	4.194	60.864	3.732	4.194	60.864	20.932	23.519	52.629
3	2.841	3.192	64.056	2.841	3.192	64.056	4.124	4.633	57.262
4	2.325	2.612	66.668	2.325	2.612	66.668	3.163	3.554	60.816
5	1.92	2.157	68.825	1.92	2.157	68.825	3.126	3.512	64.328
6	1.69	1.899	70.724	1.69	1.899	70.724	2.909	3.268	67.597
7	1.509	1.696	72.419	1.509	1.696	72.419	2.172	2.44	70.037
8	1.394	1.567	73.986	1.394	1.567	73.986	2.127	2.39	72.426
9	1.322	1.486	75.472	1.322	1.486	75.472	1.9	2.134	74.56
10	1.197	1.345	76.817	1.197	1.345	76.817	1.417	1.592	76.152
11	1.103	1.239	78.056	1.103	1.239	78.056	1.37	1.539	77.691
12	1.002	1.125	79.181	1.002	1.125	79.181	1.326	1.49	79.181

Where: A= Component B=Initial Eigenvalues C=Extraction Sums of Squared Loadings D=Rotation Sums of Squared Loadings E=Total F=Percent of Variance G=Cumulative Percent H=Total I=Percent of Variance J=Cumulative Percent K=Total L=Percent of Variance M=Cumulative Percent

Table 7.8: Illustrative Tracking Error of the PCA Index to various REIT indices June 2008 - June 2018

PCA Index	Return to All REIT	Return to REIT 50	Return to EQ.REIT	Return to MORT.REIT	equity	mort
Ave (%):	0.34	0.21	0.32	0.24	**-0.03**	**0.08**
Periodic TE (%):					2.85	2.67
Period/Yr:					0.83	0.83
Yr TE (%):					**2.60**	**2.44**

Tracking error is a measure of the absolute return difference between the new index and the incumbent indices, where the measure of returns differ from the most commonly used benchmark. This method is considered to be useful in such evaluation. Tracking error is often used to express the efficiency of Exchange Traded Funds that track an index. It is also possible to uses Standard Deviation to measure tracking error.

Table 7.9: 10yr Cross sectional time series regression on Performance of PCA REIT Index relative to the daily REIT sector 1/6/08 - 1/06/18.

Dependent variable	Alpha	t Stat	Beta	t-stat	R2	F-value	Dubin Watson
SP500 INDUSTRIAL REIT'S	0.0047	0.13	3.539	5.06	17.69	25.57	2.15037
SP500 OFFICE REIT'S	-0.0034	-0.67	1.2434	12.69	57.49	160.93	2.31078
SP500 RESIDENTIAL REIT'S	0.00106	0.19	1.158	10.89	49.94	118.7	2.21879
SP500 RETAIL REIT'S	-0.00491	-0.74	1.52	11.78	53.83	138.74	2.27536
SP500 SPECIALIZED REIT'S	0.00131	0.3	1.1698	13.64	60.98	185.99	2.36277

Table 7.10: 5yr Cross sectional time series regression on Performance of PCA REIT Index relative to the daily REIT sector 500 using Jensen alpha 1/6/13 - 1/06/18.

Dependent variable	Alpha	t Stat	Beta	t-stat	R2	F-value	Dubin Watson
SP500 INDUSTRIAL REIT'S	0.00262	0.41	0.768	4.1	22.21	16.84	2.13256
SP500 OFFICE REIT'S	-0.00286	-0.49	0.751	4.37	24.49	19.13	2.0362
SP500 RESIDENTIAL REIT'S	0.0035	0.6	0.419	2.45	9.22	6	2.40366
SP500 RETAIL REIT'S	-0.00335	-0.52	0.582	3.11	14.1	9.68	2.04169
SP500 SPECIALIZED REIT'S	0.00118	0.25	0.562	4.15	22.57	17.2	1.96239

Table 7.11: Performance metrics of the REIT PCA Factor Model Index

GRS F value	0.28
P value	0.747
Tracking Error	2.85%
Active Premium	2.48%
Information Ratio	0.10

This table displays relative performance metrics for the REIT proposed contender index vs. All REIT Index.

I summary, the empirical evidence shows that PCA indices are limited in application in REIT's. The contribution of this chapter, therefore, is to document the challenges faced by real estate index constructors and the limitations in applying the PCA technique.

7.9 Limitations

There is a limitation to the use of the PCA in the grouping of real estate types. This is due to the lack of liquidity. In this respect, REIT'S are poor substitutes for physical property transactions. This does not invalidate the earlier findings in the literature that REIT PCA derived indices are useful, as they address the lack of liquidity by a focus on listed instruments. That said, REIT's cannot measure underlying real estate prices, lack uniformity and have leveraged capital structure. They do, however, provide a daily record of net asset value. This provides a linkage to physical property. This essay tests the PCA indices against existing indices using the Gibbons, Ross, and Shanken (1989) GRS test introduced in chapter one.

7.10 Conclusion

This chapter investigated the appropriateness of existing real estate indices. It was shown that although these benchmarks have practical use in the investment profession, there are some shortcomings in respect to their usage in measuring investment returns of physical property. This is the result of their lack of liquidity and the extent of the universe they represent. The PCA index approach was explored as a solution, as explained in chapter four, *on the use of principal components in index construction.* It was concluded, however, that its application was limited due to the nature of REIT's used as a proxy for physical ownership. The equity nature of these instruments resulted in the first principal component dominating. This reflects the systematic risk of the equity market rather than the real estate market.

This chapter uses REIT's as a liquid proxy for the real estate market. The many facets of real estate investment, both public and private, are shown to make this a poor choice. It is suggested that this is the biggest area for potential future index improvement. As a result of the drawbacks of using REIT's as a proxy, it is shown that there is a need for bespoke benchmarks for real estate. It was concluded that the different securities and sub-groupings within these asset classes need to be properly classified in order to facilitate attribution.

An index is generated according to PCA loadings. The individual principal components used were a linear combination of REIT input variables and were the weighted sum of these variables. The results proved inconclusive because the market factor was more dominant than the real estate factor.

The PCA showed that there was some differentiation between the REIT sub-classifications. That said, the group proved to be highly correlated and exhibiting a strong equity market factor. This negated their use in a PCA index as it violated the requirement of an index to be reflective of the asset class that it is designed to measure, namely the underlying physical real estate. The ability to divide real estate between investment categories, ownership types and yield buckets proved elusive. This highlights a limitation in the principal component approach but also, more importantly, highlighted the lack of a liquid proxy for real estate.

As a result of the difficulties in measuring private real estate investment, it is not expected that there is much progress that can be made on indices in this sub-section of the market until the price discovery process is improved. Publicly listed REIT's are secondary by nature and suffer from the drawback that they are structured in a way that introduces leverage and additional costs, both issues from a benchmark perspective. It is anticipated that developments in artificial intelligence and machine learning will be able to address the lack of price discovery in the future.

In summary, a number of areas where benchmarking techniques for real estate can be improved upon are, however, identified. The essay thereby provides a contribution to the body of knowledge on the subject. PCA proves a useful tool to interpret REIT returns but does not provide a viable contender index with which to measure the true nature of physical real estate returns.

Chapter 8

Conclusion: Context, theory and practice.

This final chapter collates the conclusions and findings on benchmarks and shows how they relate to alternative assets. Overall, the collective chapters identify a rich field of theory supporting benchmarking practice and index construction methodology. The chapter reviews the aims and objectives presented by the research problem through the lens of Principal Component Analysis (PCA). It shows how the thesis addresses the appropriateness of existing benchmarks and identifies what characteristics are required for an alternative asset benchmark. It shows how the literature places PCA in context to both benchmark theory and practice. In this respect, how weights can be built using PCA eigenvalues. This chapter presents the overall contribution of such PCA derived indices. The collective findings provide an extension to the existing knowledge on how benchmarking should be done in practice. The thesis in its entirety provides a taxonomy on benchmarks and how to overcome the index construction challenges identified in the literature. The unique challenges within the commodity, hedge fund and real estate asset classes are reviewed within the empirical chapters. These conclude that there is a the need for bespoke solutions to benchmarking in alternative assets. This chapter, therefore, summarizes the practical insights and suggestions that support these considerations and suggests potential areas of future academic research.

8.1 Introduction

This thesis has provided a codification, a classification and an extension of benchmarking techniques for alternative assets. It fills a number of gaps identified in the literature and provides a theoretical framework for the use of principal component analysis (PCA) indices to help address these. The collective essays show how PCA can be used to create an appropriate index and refine existing construction techniques. The theoretical contribution extends the literature on benchmarks and provides a novel tool to overcome the identified problems of (1) the use of proxies to gain exposure to alternative assets, (2) the use of peer groups to measure fund returns and (3) the lack of liquidity in specific alternative assets.

The thesis details alternative asset sub-groupings. They are determined using PCA and the way the various instruments correlate and co-vary with each other. It is demonstrated that the j^{th} principal component of the alternative asset instruments in an investment universe, which is expressed as $\mathbf{a}_j'\mathbf{X}$: $\mathrm{Cov}(\mathbf{a}_j'\mathbf{X}, X_i) = \mathrm{Cov}(\mathbf{a}_j'\mathbf{X}, \mathbf{b}'\mathbf{X})$, where $\mathbf{b}' = [0 \cdots 0\, 1\, 0 \cdots 0]$, with the 1 in the i^{th} position, $= \mathbf{a}_j'\mathbf{\Sigma}\mathbf{b} = \mathbf{b}'\mathbf{\Sigma}\mathbf{a}_j = \mathbf{b}'\lambda_j\mathbf{a}_j = \lambda_j$ times the i^{th} component of $\mathbf{a}_j = \lambda_j a_{ij}$, can be made into an investment index suitable for alternative asset classes. The resultant index is based on the *loading* of X_i on the j^{th} component and this correlation corresponds to the weight assigned $\zeta_r i_t$.

The presented PCA index is derived from the constituent weights for each discrete time period. This is done using factor analysis implemented by using oblique rotations. The index weights are the weighted average of the returns of the derived portfolio constituents. A factor vector model can then be made and presented as follows:

$$\zeta_r i_t = \beta 1 \mathbf{v} F_t + \beta 2 \mathbf{v} F_t + ...\beta X \mathbf{v} F_t \tag{8.1}$$

Where:

- $\zeta_r i_t$ = the roposed contender index, the excess return of portfolio i in month t.
- $\mathbf{v}$ = an eigenvector.
- F = the factors as identified by the eigenvalues.

This factor model, and the method to construct the proposed PCA index, is detailed in chapter four. In effect, it is constructed in a stepwise fashion with a defined observation period. This is then combined with a re-balancing frequency which is conducted at a pre-specified date. The rules are in accordance with the standard approach to index maintenance, as detailed in chapter three.

The PCA index method was used to gain insights into three major alternative asset classes, namely commodities, hedge funds and real estate. Its use as a statistical method is well documented in other sciences. A major advantage of the approach is the fact that prior knowledge of the asset class is not required in order to break down its instruments into sub-components. The three empirical chapters tested the appropriateness of the principal component analysis (PCA) index approach and tested the mean variance efficiency of the PCA derived index against existing commercial indices.

As an introduction to the PCA Index, the first chapters investigated how alternative assets are identified and defined and what insights or improvements the approach can bring to benchmarking. They explored the universe of available instruments and what gaps the literature has in respect of benchmark selection and appropriateness. They set the scene and the theoretical framework and methodology. The PCA indices were tested on specific alternative asset classes, namely commodities, hedge funds and real estate. The application of PCA in these sectors demonstrated that there are unique issues but also commonalities. The research question, whether a PCA contender index can deliver superior risk adjusted returns, was tested on each of the alternative asset classes under investigation. In order to do this, the indices were subject to the GRS test.

The evidence presented in this thesis suggests a PCA approach is particularly helpful for commodities. The analysis from hedge funds and real estate investment trusts was less conclusive. Nevertheless, the PCA gives insights and the conclusions, as summarized in this chapter, can be used to suggest refinements to the construction of benchmarks.

8.2 Reviewing the aims and objectives

The aims and objectives of the thesis were clearly stated in the first chapter. That is, to determine an appropriate benchmark for alternative assets. It is highlighted that the literature does not answer this question. A key aim of the thesis was, therefore, to clarify this in the context of appropriateness as defined by Bailey, Richards, and Tierney (1990). In this respect, the fundamental way benchmarks and indices are formulated are different. Indices can be used as valid benchmarks in certain circumstances. The research questions were framed around this and whether PCA indices can be appropriate and useful for alternative assets. The chapters re-contextualized existing benchmark techniques and suggested how PCA can be used in index construction. The method was used to construct indices for commodities, hedge funds and real estate. The null hypothesis, introduced in section 1.4.2, was then tested using the GRS test of Gibson and Schwartz (1990). This was done to see if it was more appropriate for the evaluation, attribution and appraisal of commodity, hedge funds and real estate relative to existing indices.

The theoretical framework was introduced in chapters three and four. It was shown how the PCA approach can determine the sub-groupings of the instruments and illustrated how weights can be derived from the first principal component of alternative asset returns and therefore used to construct an index. This represents a major contribution of the thesis extending the literature on index construction and showing how the PCA theory can be extended to benchmarking alternative assets.

The definition of alternative assets was shown to be a broad one. It includes all those investment instruments not classified as traditional. That said, the alternative asset classes reviewed in chapters five, six and seven, are the major ones, namely commodities, hedge funds and real estate. Some insights from other alternative asset classes, such as high yield debt and private equity, were also explored in chapter two.

The objectives stemmed from gaps in the literature. In this respect, the literature review began with Bailey (1992) who made clear what the fundamental properties of a benchmark are. These are that they are (1) unambiguous as to weights and instruments, (2) investable, to the extent that they can help passively, (3) measurable on a regular and frequent basis, (4) appropriate to a fund manager;s style basis, (5) reflective of the investment objectives, (6) and specified in advance. From the literature, eight questions provided insights:

- Question one asked whether benchmarks are fit for purpose? It was concluded that the have a good grounding in theory but can be improved.
- Question two asked how benchmarks are used for identifying skill. It was concluded that the CAPM provides a strong support for mean variance indices.
- Question three asked if it is possible to persistently out-perform a benchmark. The answer is that it is extremely difficult. This suggests that if a contender index is found to have superior risk adjusted characteristics to an incumbent, that it is likely a better index.
- Question four asked how liquidity affects benchmark selection? The literature shows that liquidity commands a premium and that illiquid instruments make for poor index constituents.
- Question five asked how leverage distorts benchmark returns. The literature suggests that it skews the returns. As such, indices that incorporate leverage will not have the normal distribution that the CAPM assumes.
- Question six asked what the consequences of benchmark misspecification are. The literature suggests that this results in unreliable research and attribution outcomes. This provides a justification for benchmark enhancement.
- Question seven asked how benchmark superiority is determined. The answer ties back to the identification of a superior index in chapter one and suggests a good way to do this is to use the GRS test. The various tests to measure benchmarks were presented in order to demonstrate how remedy of these deficiencies can be achieved using both quantitative and subjective approaches.
- Question eight focused on what the priority areas were for alternative asset constructors. The literature suggests this is liquidity and addressing the proxies.

The link between equity indices and the mean variance market proxy was explained. This was used with a theoretical framework to investigate the appropriateness of commercial alternative asset indices and the presented PCA index. The various chapters addressed the question of appropriateness through index properties and the concept of mean variance. This links the asset class indices with optimal outcomes as suggested in the literature. This facilitates the measurement of risk adjusted performance, an under-researched concept in alternative assets, and links it to the PCA derived factor index.

The PCA index is a novel addition to the index literature on alternative assets. The construction technique was applied to each of three separate alternative assets and the empirical findings add to the literature on each of these sub-groups. It was shown that there is no clearly defined subset of assets that can be called alternative. As such, in documenting best practice and recommending improvements to the body of knowledge this thesis contributes to the theory behind benchmark construction and choice by detailing and expanding on what the most appropriate benchmark approach is for alternative assets. The concept behind benchmarking was introduced in the first chapter and the subsequent two chapters reviewed the literature from different perspectives. In these chapters, it was concluded that indices were preferable as benchmarks for alternative assets due to their ability to attribute returns.

Throughout the thesis it was shown that research into benchmarks is important. Inappropriate benchmarks give a misleading picture of manager performance. Good benchmarks help better assess fund managers. Benchmarks are also instrumental in shaping investment contracts that typically have an incentive compensation component based on their out-performance. In terms of appropriateness, peer group indices fell short of what is required as defined by a full score on the Likert scale as introduced in chapter one. The chapters on commodities, hedge funds and real estate delved further into this line of inquiry.

8.3 Chapter structure

The structure of the thesis allowed the chapters to collectively build the case for improvement and refinement in the construction of alternative asset indices. There proved to be no single recommendation suitable for all three asset classes investigated in depth. The need, however, to individually improve the construction method for commodities, hedge funds and real estate was clear. The implications for the testing of asset pricing and the measurement of risk and return are important.

As a result of financialization, and the trend for greater institutional investment in alternative assets, any suggested benchmark must take investibility into account. In this way passive investors can buy exposure to the index constituents through collective funds. That said, alternative assets are by definition difficult to trade and it was shown that such replication is not as easy as it is in traditional assets. As the resulting index must be appropriate for both investors and their managers, this can result in challenges. In this respect, it is perceived that the more sophisticated the benchmark, the better one can understand the cross-section of such returns. In the context of alternative asset classes this is relevant because the risk factors are different. These were detailed in each of the empirical chapters.

An observation from the explanatory background to benchmarks in the first chapter is that their selection is complicated because the definition of alternative assets is by exclusion. In other words, practitioners define them as anything that is not a traditional asset class is an alternative asset class. For this reason, there has been little academic focus on the definition of the investment set. The term is all encompassing and vague, incorporating a set of potential investments that are not traditional. The term can include, for example, physical assets, such as real estate, complex methods of investing, hybrid debt, private equity and even off-benchmark geographic regions. In this respect, it is opined that an alternative asset class should have common features and have returns that exhibit a reasonable degree of correlation. In the context of alternative assets, this is important because investors select such assets based on distinct diversification and return expectations.

Chapter one introduced the research problem and the first three chapters built on this to set the scene. It is clear from these that a focus on risk factors is an important aspect of the way alternative assets should be benchmarked. As a result of the widespread usage of the Capital Asset Pricing Model, the market proxy has become the *de-facto* benchmarking tool of traditional assets. This default was shown to be not appropriate for alternative assets and as such practical fund management is incorrectly using the wrong benchmarks and passive funds and indexed to inappropriate weightings. It is concluded that a more bespoke approach to risk factors is warranted.

In chapter two the literature on investment benchmarks was explored from the perspective of what contribution and insights the wider finance literature can bring to the topic. In this respect, indices have been used extensively for the measurement of returns of various time series and the testing of capital asset pricing models. chapter two also critiqued the use of peer group indices as inappropriate. The suggested refinements and replacement for such peer group indices in the various chapters are a key contribution of this thesis. Addressing investability through synthetic replication is another.

In chapter three, it was shown that indices have evolved to give the user a return time series that is discrete, but not always appropriate. In academic terms they deliver a continuous state process. This needs to be captured by indices, something that PCA can address. Using PCA, various types of benchmark were addressed to identify a better fit. chapter three further demonstrated that a benchmark requires an appropriate calculation method. Its trade-offs need to maximize the utility of investor options. It should deliver adequate diversification. Correctly conceived, benchmarks play a key role in fund managers selection and evaluation. They can be used to analyze past performance in the context of the investment process. In this way, benchmarks are used to asses a fund manager's skill. Benchmarks are also used in marketing investment products as well as in demonstrating compliance with regulations.

The fourth chapter introduced PCA derived indices and explained their theoretical construct. The indices are generated based on PCA loadings. Each principal component was a linear combination of the alternative asset instrument under investigation. The weighted sum of the index universe was turned into a price time series and the return measured. It was argued that this approach was worthy of investigation because existing alternative asset indices were shown to have a methodology that evolved from equity indices.

The thesis links what is known about equity indices with what is known about alternative assets. The broad equity indices are justified by capital asset and factor models. That said, these models cannot be used in the same way for alternative assets as they are used in equity indices. In equities, the Capital Asset Pricing Model is used to justify market capitalization indices. This model is consumption based. In contrast, (1) many existing commodity indices are production based, (2) hedge fund indices would not be diversified if based on the size of assets under management, and (3) real estate indices are not reflective of the housing stock available for investment. A PCA approach avoids having to use capitalization, focusing instead on co-variances to determine weights.

The thesis addressed practical as well as academic issues through an empirical approach in chapters five, six and seven. The rapid growth in alternative investments has meant that benchmarks have to address practical as well as academic issues. Improvements and enhancements should be positioned within the context of financial market theory. In this respect, one of the key observations is that any contender index should be a measure of risk as well as return. In the context of alternative asset classes, this is important because they are brought into a portfolio as a diversifying asset.

The thesis viewed indices through the lens of Modern Portfolio Theory. The Sharpe ratio developed by Sharpe (1994), similar in nature to a t-statistic, was used to convey the risk and reward inherent in an index. It measures the mean return minus the risk free rate and thus illustrates the statistical significance of contender indices. As an extension of its application, the GRS test was introduced in chapter one and two. This is used in the empirical chapters to test the PCA derived indices using the null hypothesis $\mathrm{H}_0 : GRST_t = \eta_i - \eta_j = 0$.

In the context of existing commodity indices, the null hypothesis that the contender PCA cannot outperform on a risk adjusted basis is rejected. With hedge funds and REIT's it is not. This suggests a PCA index is potentially more appropriate for the measurement of investment strategies that utilize commodity futures.

In summary, it was demonstrated that there are a number of deficiencies in existing benchmarks as pertaining to alternative asset classes. These include shortfalls in reflecting the asset class dynamics, proxy instruments, and investibility. Some of the properties of the PCA approach address these deficiencies but it does not prove a panacea. The presented chapters collectively conclude that alternative assets should have a bespoke solution. The idiosyncratic nature of alternative asset proxies, specifically commodity futures, hedge funds and real estate investment trusts, require tailored measurement tools. It was shown how PCA can help achieve this goal and how it can be used to identify shortcomings in order for performance to be benchmarked in such assets so as to determine the success or failure of an investment strategy.

8.4 Contribution

This thesis fills a gap in the literature by providing a taxonomy of benchmarking and index construction knowledge. Its contribution is as a collection of tailored recommendations that can make a benchmark more appropriate for the measurement of specific alternative asset investment returns. In this respect, it is suggested that it is possible to design more optimal benchmarks than those indices that currently exist. PCA is used to demonstrate this. It is argued that the technique helps in the understanding of alternative assets through their coefficients.

This thesis extends the literature in three ways. It is found that PCA derived indices can be used to benchmark commodity futures in an optimal mean variance way. It is found that hedge funds strategies can be synthesised using PCA and be used to form strategy level indices. It is further found that the use of PCA is limited in benchmarking real estate due to the illiquid nature of the asset.

The non-linearity of the time series of returns of alternative assets is addressed by the PCA technique through a focus on the common components in the alternative asset time series. It allows for the creation of a weighting protocol that can be used as an index that takes alternative asset return idiosyncrasies into account. This approach addresses the skewness and kurtosis present in the cross-section of returns in alternative assets.

Specifically, a number of failings were identified in existing commercial indices for each of the asset classes investigated. These include (1) the widely used peer group approach to hedge fund benchmarking; (2) the use of capitalization proxies rather than consumption proxies in commodity futures benchmarks; and (3) the lack of homogeneity in the underlying real estate exposure of Real Estate Investment Trusts.

PCA derived indices are presented as a way to part address all three issues. Using the spiderwebs diagrams, the method is evaluated using the eight tests for index numbers proposed by Fisher (1911) Eichhorn (1976). In this respect, it is shown that:

1. PCA derived indices have proportionality as to investment objectives.

2. PCA derived indices have proportionality as to cross-section of returns.

3. PCA derived indices can be used to determine instruments.

4. PCA derived indices can be used to determine constituents.

5. PCA derived indices can be be adapted for entry and withdrawal of index constituents.

6. PCA derived indices can have a changing base.

7. PCA derived indices can accommodate a changing unit of measurement (termed the *Durability Test)*

8.4.1 Context

The introductory chapter was a codification of the state of the art in benchmarking and index construction. In this respect, it puts benchmarking in context. The empirical chapters were in essence a characterization of the time series of returns of each of the major alternative asset classes. The wider contribution is in drawing specific conclusions about the uniqueness of each alternative asset class investigated, namely commodities, hedge funds and real estate.

Research into benchmarks is important because regardless of whether they are active or passive, professional investors should use some sort of benchmark. In this respect, an investor is defined as a person who makes an investment decision either for himself or on behalf of someone else. As a result, it was observed that in order to identify the appropriateness of a benchmark for alternative assets, it is also necessary to define these users and their objectives. This is because the information the users seek from them defines whether the benchmark should be an index, with its granular detail, or simply a hurdle rate, based on a quantitatively stated objective. In this respect, it was broadly observed that investors in commodities see them as a hedge for inflation, investors in hedge funds see them in the light of capital preservation and investors in real estate see them in the light of the stability of the returns over time.

It was noted that there are a broad number of uses and users of benchmarks. With third party mandates, for example, it is the responsibility of plan sponsors to determine the type of benchmark used. The fund manager, in this instance, requires the benchmark in order to reference his investment strategy and manage it according to the mandate that he has been given. Mandates are awarded but plan sponsors have other goals and awards are made with benchmarks assigned to multiple investors (and typically structured in separate legal entities).

Such a mandate may be either a passive or an active one. Increasingly, as has been observed, they include alternative assets such as commodities, hedge funds and real estate.

The basic premise behind this thesis is that a benchmark provides the user with a good starting point from which the time series can be observed and measured. That said, there are a multitude of different return measures that can be employed versus the benchmark and these are explained in this chapter. Although investment in different asset classes have different risk and returns, the mathematical method should follow pre-defined rules.

If various securities in a time series are related, they could be said to be an asset class. One definition of an asset class, then, is that it is a broad collection of securities that vary and co-vary in similar ways. This is why chapter six treats hedge funds as an asset class. It is acknowledged that a case can be made that they are collective investment vehicles that invest in underlying asset classes but the literature is shown to treat them as a *de facto* asset class. This discussion is picked up in the chapter in chapter five on hedge fund benchmarks, challenges and refinements.

It was noted that correlation is impacted by the skew and kurtosis. This is due to the unique nature of the financial instruments that underlay them. The PCA derived indices are based on sub-groupings and as such unaffected by this statistical property. It was further noted that illiquid trading is a feature of alternative investments. As a result, characteristic matching in the form of PCA eigenvectors prove superior in explaining the risk components in these asset classes. The literature on liquidity broadly concluded that it can have a major effect on pricing and on investments required holding period return. The literature also suggests that the effect of liquidity on re-balancing is important. To address these issues (1) futures are used as a proxy in chapter five, (2) synthetic indices are proposed in chapter six and, (3) REIT's are used as a proxy in chapter seven. These proxies are reviewed and critiqued through the lens of PCA.

In summary, principal components are quantitative, inductive, and observational. They can be used to establish a benchmark for alternative investments. It identifies common components which allow for a different weighting regime. The size of the assets being professionally managed in these alternative asset classes justifies the adoption of more appropriate benchmarks. Such professional investment requires optimal indices. The GRS test is therefore used in this thesis to test PCA derived indices. It is concluded that there is potential for such indices to be used in practice.

8.4.2 Theory

Most of the conclusions of this thesis are corroboration, extension and elaboration of existing index construction theory and method. The chapters include practical as well as theoretical contributions. The importance of refining the methodology and specifically tailoring it for alternative investments was demonstrated. The theoretical contribution from this is the observation that an assessment of an index for such investments has to be seen in the context of the skew and non-normal nature of the returns.

The PCA method uses a two-step process of sorting results for expected returns in one dimension. Subsequently, the co-movements are isolated, which are evaluated in the cross-section of the returns of the alternative asset class to determine index weights. The basic idea of PCA is that the underlying observed random variables can be expressed as linear functions of common factors. It was concluded that the principal component approach is an alternative way of addressing the underlying structure in the time series from which an index can be constructed.

The application of PCA using these steps results in a contribution that was identified in chapter four, on the identification and classification of alternative assets. This illustrates how the use of PCA departs from current index construction approaches and can be used to build index weights from those instruments that are closely associated with each other. This method is not currently used in practice but is fairly common in other disciplines. The PCA is undertaken without any prior assumptions about the existing universe. Using the method identified in the literature and explained at length in chapter four. In identifying the common principal components, this approach goes back to basics in the way that alternative assets are grouped.

The literature review in chapter two contributes to the collation of academic knowledge on the theory of benchmarking . chapter three shows how it can be applied in practice. In this way, the various chapters critique existing theory which draw on the use of benchmarks and indices from a number of studies. In doing this, the thesis illustrates the issues and the omissions that needed to be addressed from the perspective of alternative assets.

It is shown that the theory of index construction is well defined for equities. In this respect, an index can be constructed using a simple, weighted, geometric, arithmetic, harmonic or exponential average, and the weights assigned accordingly. Indeed, it is shown that they can also be constructed using a combination of them all. The PCA method built on this and was investigated for appropriateness using Jensens Alpha and the GRS test to determine if these were more optimal in construction than existing indices.

Theory was examined from the broad finance literature. As most academic investigations into portfolio performance, market efficiency and pricing make use of benchmarks, the focus was on those papers that made a contribution to the understanding of what makes a suitable benchmark for alternative assets. The prior research was introduced in the context of what it means from the perspective of benchmarking for alternative assets. In this respect there are similarities between the literature on benchmarks, the PCA method and conventional asset pricing.

A factor model was developed. This was derived from the eigenvalues is expressed as follows. It has applicability for commodities, hedge funds and real estate.

$$\zeta_r i, t = \beta 1 \mathbf{v} \eta_t + \beta 2 \mathbf{v} \theta_t + \beta 3 \mathbf{v} \chi_t + \beta 4 \mathbf{v} \tau_t + \beta 5 \mathbf{v} \mu_t \tag{8.2}$$

Where for commodities:

- $\zeta_r i, t$ = Proposed contender index, the excess return of sector commodity portfolio i in month t, item **v** = an eigenvector
- η = Grain complex
- θ = Industrial metals complex
- χ = Livestock complex
- τ = Energy complex
- μ = Precious metals complex

Where for hedge funds:

- $\zeta_r i, t$ = Proposed contender index, the excess return of sector commodity portfolio i in month t, item **v** = an eigenvector

- η = Broad multi strategy complex
- θ = Equity based complex
- χ = Market Neutral complex
- τ = Event Driven complex
- μ = Relative value complex

Where for REITs':

- $\zeta_r i, t$ = Proposed contender index, the excess return of sector commodity portfolio i in month t, item **v** = an eigenvector
- η = Residential housing complex
- θ = Retail complex
- χ = Office complex
- τ = Equity REIT complex
- μ = Mortgage REIT complex

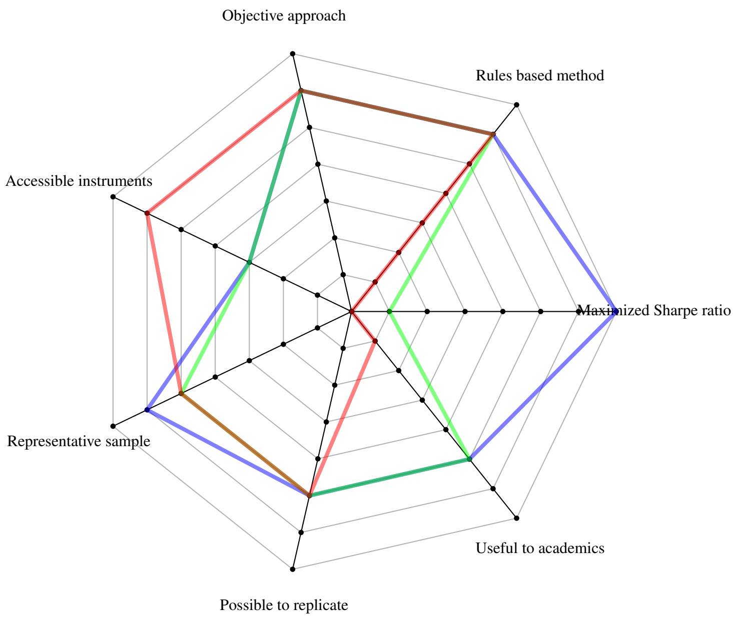
In summary, these PCA derived factor models can be used to determine weights for indices constructed for commodities, hedge funds and real estate. This finding does not contradict existing models. It draws together existing PCA index concepts from other disciplines and shows how the combination is useful and practically relevant for alternative assets.

8.4.3 Practice

In a practical context, existing benchmarks for alternative assets were critiqued. The widespread use of peer group indices of funds exposed to the underlying asset was discussed. The returns for such benchmarks are net of fees, which does not give an objective measurement. Part of the contribution of this thesis is therefore to recommend against the usage of peer group benchmark when evaluating alternative assets.

Figure 8.1 shows the outcome on a spiderweb diagram of superimposing the PCA derived indices for each of the alternative asset classes. This was presented in the respective chapters. In aggregate, the approach has clear strengths, although it is not optimal in all circumstances.The discussion throughout the thesis demonstrates that the various sub groupings in alternative assets could be better defined. This would in turn provide for better attribution of returns. On a similar note, the discussion on trade-offs can be extended to practice with alternative assets, investability is a more important trade-off than breadth. Value judgments play a greater role in instrument selection in the trade off against rules as a result of surivorship bias.

Figure 8.1: Combined PCA Commodities, PCA Hedge Fund and PCA Real Estate Index Spiderweb.



This figure overlays the PCA Commodities (blue), Hedge Fund (green) and Real Estate Index (red) Spiderwebs. Representation of 7 Dimensions, 7-Notch Scale, and identified subjective index inclusion criteria. The diagram shows that although useful, none of the applications of PCA Indices are optimal. The most significant Likert score is dimension seven, the high Sharpe ratio of the PCA derived commodity index.

The empirical investigation in chapters five, six and seven demonstrates the concept. PCA indices are explained in the chapter on method. The subsequent chapters show that it can be used more widely in the determination of common components between instruments within asset classes. In that context, Exchange Traded Funds could be developed that reflect the investment weights of such contender indices and synthetic funds can be created to mimic factor returns.

The later represents a practical contribution. The use of passive synthetic hedge fund indices, as discussed in chapter six, would avoid the double charging of fees by fund of funds. This is because the factor vector model that forms the basis of the PCA index weights can be used to gain exposure to the underlying traits in hedge funds. This alleviates the need to buy the underlying funds.

8.5 Limitations

The many different approaches to benchmarking illustrate that there are limitations in every approach. By documenting the issues and the many solutions this thesis sheds light on the trade-offs that are inherent in index construction. As such, the biggest limitation is that PCA is that it is a subset of the many index approaches that can be used. It is not a substitute for each and every alternative asset, as was illustrated by the chapter on real estate.

A documented limitation is the number of instruments in a PCA index. The instruments can be interpreted as uncorrelated risk sources inherent in the asset class data set. An index of 25 instruments contains 25 uncorrelated risk sources. Unfortunately, the eigenvalues of principal components typically decrease quickly and the higher numbered principal components have relatively small eigenvalues. As such, an index of 100 instruments would see the relevance of principal components decline rapidly. The method is therefore suited to smaller number of instruments such as was demonstrated in the chapter on commodity futures.

The proposed PCA approach has other drawbacks. One of theses limitations is that the PCA index approach does not incorporate short positions. With alternative assets, there is the ability to go both long and short. This changes the risk return dynamics. Investment strategies that pursue such approaches would require very different benchmarks. There is a nascent but growing demand for such benchmarks which should be the subject of further scholarly inquiry.

Another limitation is that the PCA index approach does not give exposure to the underlying asset, merely a derivative thereof. One of the uses of a benchmark is to obtain a passive representation for the index. The ability to get (1) a passive representation of physical cash commodities, (2) invest in all hedge funds (due to gates and other restrictions), and (3) gain direct real estate exposure (rather than collective investment funds), remains elusive.

The empirical chapters further showed that PCA has its limitations based on the sort of proxy used. It is a contender index for commodity futures, has useful properties in identifying hedge fund sub-groupings, but it does not appear to be useful in REIT evaluation.

8.6 Suggestions for further research

Whilst this thesis has proposed a way to regroup and weight the components of risk and return in alternative assets there are still identified areas that would be worthy of further research. Primary amongst these is the leverage inherent in alternative assets. This needs to have a method to be properly measured and the impact on benchmark risk and return needs to be assessed. At present, no indices do this. Leverage skews the returns and exaggerates the beta of a portfolio relative to an index. It is recommended more research be conducted in this area of benchmarking

It was suggested in the limitations section that a further avenue of potential future research would be into short positions and indices that reflect these. The extent of any short position will always remain problematic, making such indices harder to replicate. Where futures are used, the ability to short is enhanced. In a related fashion, the way funds treat stock lending costs should also be subject to further research.

It was suggested in the literature review that further research should be conducted on the attribution of alternative asset returns. One insight that comes from the use of factors to determine weights is that this approach could be used to measure the skill of fund managers. For example, the difference between weights between hard and soft commodities. With hedge funds, the responsiveness of funds to regime change. With US REIT's, the geographic exposure is relevant. One could regress regional growth rates with real estate returns to give factor exposure to the fastest growing areas. Similar investigation could be done on risk exposures using dummy variables to economic or physical shocks.

It was suggested in chapter five on commodities that further research be done on indices to be based on the physical market rather than the futures market. The biggest impediment to this is price discovery, physical delivery and recording of multiple venues of trading. Hopefully, as technology advances can address these issues in the future.

It was suggested in chapter six that further research be done on synthetic replication of hedge fund returns. This could potentially lead the way for products with similar characteristics but lower costs. The chapter on real estate identified a need for further research to be done to differentiate between rental and capital returns in real estate.

It was suggested in chapter seven that further research should be conducted on the other ways that investors can gain exposure to real estate, such as commercial mortgage backed securities or physical property. The need to professionally manage such exposure with proper benchmarks is self evident.

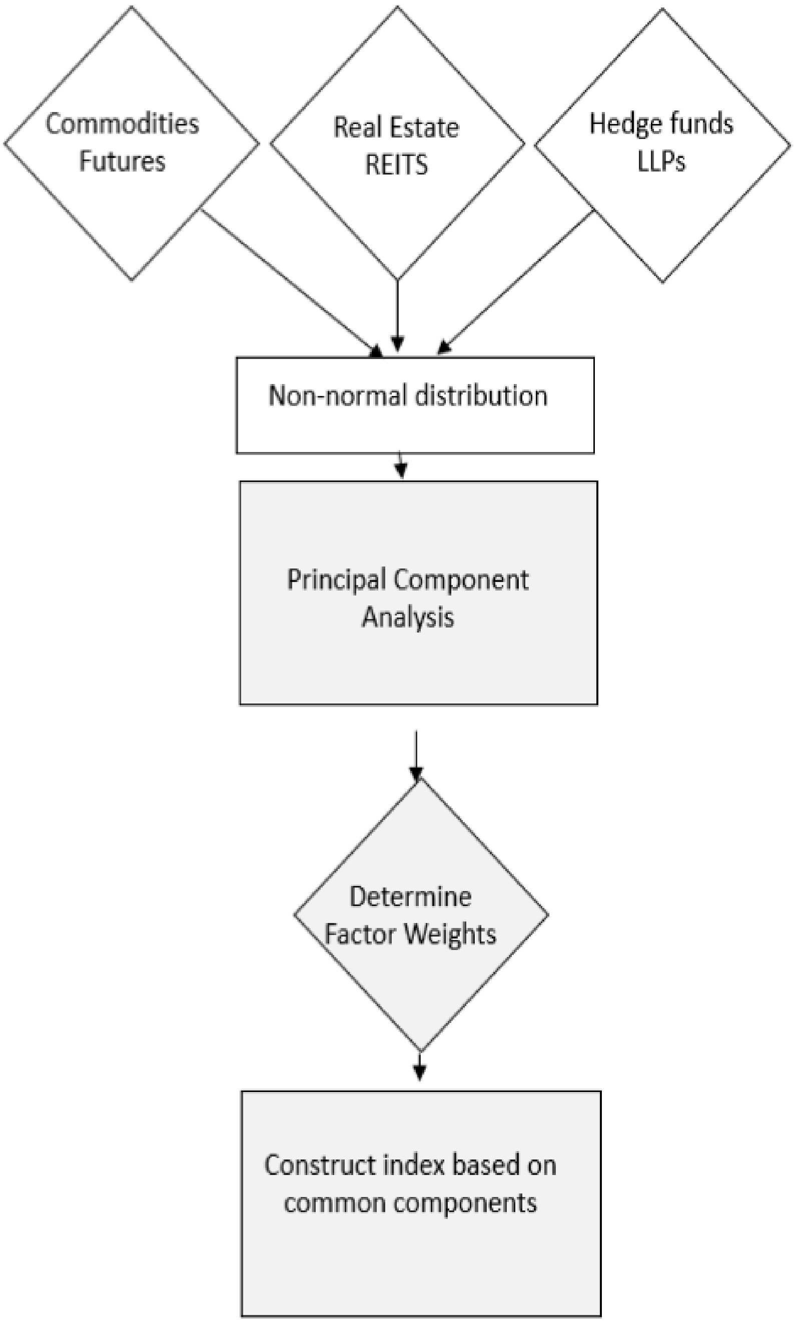
Whilst this thesis ruled out absolute and price series benchmarks, there is little research into these areas. As the motive for investment in many alternative assets is to preserve capital or beat inflation, this is clearly an area that would also benefit from further research.

8.7 Conclusion

The thesis and the chapters herein provide a taxonomy of observed benchmark phenomena to give a rich understanding of how alternative asset time series of returns can be measured and attributed. It fills gaps in the literature on benchmarks and provides insights into how existing method might be improved for alternative assets using PCA derived indices. Its review of knowledge gives the user a better practical understanding of benchmarking and what can be achieved for the alternative asset class. The flow diagram in figure 8.2 summarises its contribution and what the PCA index method brings to the field of index construction.

The chapters show that alternative asset instruments are best measured by the use of an index built from the top down. PCA does this by assigning weights based on the factor vector eigenvalues. The PCA approach produces a separate genre of indices. Using the method, it is possible to identify some of the unique characteristics of the various alternative asset classes. The key sub-groupings, commodities, hedge funds and real estate, proved to have idiosyncratic features that demand specific treatment from the perspective of index construction. It was shown from both the literature and the empirical chapters that a bespoke approach is required.

Figure 8.2: Flow Diagram of Principal Component approach to index construction



This figure shows the way the same approach to identifying common components can be applied to the three alternative assets investigated in this thesis.

This thesis used a GRS test to empirically test PCA contender indices. The results suggest that the PCA approach is superior to some capitalization based commodity indices. They also show PCA indices may provide an alternative to peer group based hedge fund indices. It is concluded that PCA may prove a useful way to synthesise factor exposure, particularly for hedge funds.

The findings in this thesis are important because there is no commonly agreed view of what is an appropriate benchmark for alternative assets. It highlights that alternative assets have a number of unique features that need to be captured in their benchmarking. These include instrument, leverage, and liquidity. These have to be catered for in the construction process in order to make the benchmark appropriate. It was shown that a PCA index does this and helps fill a number of gaps in the literature. As a result, academics can gain valuable insights through the use of PCA derived indices in the analysis and attribution of alternative asset class returns.

Bibliography

Abhyankar, A., Copeland, L., and Wong, W. (1997). "Uncovering nonlinear structure in real-time stock-market indexes: The S & P 500, the DAX, the Nikkei 225, and the FTSE-100". In: *Journal of Business and Economic Statistics* 15.1, pp. 1–14. ISSN: 15372707. DOI: 10.1080/07350015.1997.10524681.

Acharya, V. and Pedersen, L. (2005). "Asset pricing with liquidity risk". In: *Journal of Financial Economics* 77.2, pp. 375–410. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2004.06.007. URL: http://www.scopus.com/inward/record.url?eid=2-s2.0-22144449932&partnerID=tZOtx3y1.

Admati, A. and Pfleiderer, P. (July 1997). "Does It All Add Up? Benchmarks and the Compensation of Active Portfolio Managers". In: *The Journal of Business* 70.3, pp. 323–350. ISSN: 0021-9398. DOI: 10.1086/209721. URL: https://www.jstor.org/stable/10.1086/209721.

Agarwal, V. and Naik, N. (2000). "Multi-Period Performance Persistence Analysis of Hedge Funds". In: *The Journal of Financial and Quantitative Analysis* 35.3, p. 327. ISSN: 00221090. DOI: 10.2307/2676207.

— (2004a). "Risks and Portfolio Decisions Involving Hedge Funds". In: *Review of Financial Studies* 17.1, pp. 63–98. ISSN: 0893-9454. DOI: 10.1093/rfs/hhg044. URL: https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/hhg044.

— (2004b). "Risks and Portfolio Decisions involving Hedge Funds." In: *Review of Financial Studies*, 17, p. 63.

— (2015). "Hedge Funds: A Survey of the Academic Literature". In: *Unpublished paper: ssrn.com 2650919*. URL: http://ssrn.com/abstract=2650919.

Agnew, J. and Szykman, L. (2005). "Asset Allocation and Information Overload: The Influence of Information Display, Asset Choice, and Investor Experience". In: *Journal of Behavioral Finance*. ISSN: 1542-7560. DOI: 10.1207/s15427579jpfm0602{_}2.

Amenc, N., Goltz, F., and Le Sourd, V. (2009). "The performance of characteristics-based indices". In: *European Financial Management* 15.2, pp. 241–278. ISSN: 13547798. DOI: 10.1111/j.1468-036X.2008.00468.x.

Amenc, N., Goltz, F., and Lodh, A. (2012). "Choose Your Betas: Benchmarking Alternative Equity Index Strategies". In: *The Journal of Portfolio Management* 39.1, pp. 89–111. ISSN: 00954918. DOI: 10.3905/jpm.2012.39.1.088. URL: http://www.iijournals.com/doi/abs/10.3905/jpm.2012.39.1.088#sthash.kBEGVZ2X.dpbs.

Amenc, N. and Martellini, L. (2002). "Portfolio Optimization and Hedge Fund Style Allocation Decisions". In: *The Journal of Alternative Investments* 5.02, pp. 7–20. ISSN: 1520-3255. DOI: 10.3905/jai.2002.319051.

Amihud, Y. (2002). "Illiquidity and stock returns: Cross-section and time-series effects". In: *Journal of Financial Markets* 5.1, pp. 31–56. ISSN: 13864181. DOI: 10.1016/S1386-4181(01)00024-6.

Amihud, Y. and Goyenko, R. (2013). "Mutual Fund's R2 as predictor of performance". In: *Review of Financial Studies* 26.3, pp. 667–694. ISSN: 08939454. DOI: 10.1093/rfs/hhs182.

Amihud, Y. and Mendelson, H. (2015). "The Pricing of Illiquidity as a Characteristic and as Risk". In: *Multinational Finance Journal,* 19.

Amihud, Y., Mendelson, H., and Pedersen, L. (2005). "Liquidity and Asset Prices". In: *Foundations and Trends® in Finance* 1.4, pp. 269–364. ISSN: 1567-2395. DOI: 10.1561/0500000003. URL: http://www.nowpublishers.com/article/Details/FIN-003.

Andonov, A., Eichholtz, P., and Kok, N. (Jan. 2015). "Intermediated investment management in private markets: Evidence from pension fund investments in real estate". In: *Journal of Financial Markets* 22, pp. 73–103. ISSN: 13864181. DOI: 10.1016/j.finmar.

2014.11.002. URL: http://linkinghub.elsevier.com/retrieve/pii/S1386418114000780.

Ang, A., Papanikolaou, D., and Westerfield, M. (2014). "Portfolio Choice with Illiquid Assets". In: *Management Science* 60.11, pp. 2737–2761. ISSN: 0025-1909. DOI: 10.1287/mnsc.2014.1986. URL: http://pubsonline.informs.org/doi/abs/10.1287/mnsc.2014.1986.

Ang, A., Rhodes-Kropf, M., and Zhao, R. (2008). "Do Funds-of-Funds Deserve Their Fees-on-Fees?" In: *Journal of Investment Management* 6.4, pp. 34–58. ISSN: 15459144. DOI: 10.3386/w13944. URL: http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=36151325&site=ehost-live.

Anson, M. (2004). "Strategic versus Tactical Asset Allocation". In: *The Journal of Portfolio Management* 30.2, pp. 8–22. ISSN: 0095-4918. DOI: 10.3905/jpm.2004.319926. URL: http://www.iijournals.com/doi/abs/10.3905/jpm.2004.319926.

Aragon, G. and Ferson, W. (2006). *Portfolio performance evaluation*. DOI: 10.1561/0500000015.

Arango, L., Arias, F., and Flórez, A. (2012). "Determinants of commodity prices". In: *Applied Economics* 44.2, pp. 135–145. ISSN: 00036846. DOI: 10.1080/00036846.2010.500273.

Arnott, R. (May 2011). "Better Beta Explained: Demystifying Alternative Equity Index Strategies". In: *The Journal of Index Investing* 2.1, pp. 51–58. ISSN: 2154-7238. DOI: 10.3905/jii.2011.2.1.051. URL: http://jii.iijournals.com/lookup/doi/10.3905/jii.2011.2.1.051.

Arnott, R., Hsu, J., and Moore, P. (2005). "Fundamental Indexation". In: *Financial Analysts Journal* 61.2, pp. 83–99. ISSN: 0015-198X. URL: http://www.jstor.org/stable/4480658.

Arnott, R. and Lovell, R. (Feb. 1993). "Rebalancing". In: *The Journal of Investing* 2.1, pp. 5–10. ISSN: 1068-0896. DOI: 10.3905/joi.2.1.5. URL: http://joi.iijournals.com/lookup/doi/10.3905/joi.2.1.5.

Arrow, K. (1964). "The role of securities in the optimal allocation of risk-bearing." In: *The Review of Economic Studies* 31(2), pp.

Atilgan, Y., Bali, T., and Demirtas, O. (2013). "The performance of hedge fund indices". In: *Borsa Istanbul Review* 13.3, pp. 30–52. ISSN: 22148450. DOI: `10.1016/j.bir.2013.10.007`.

Bacon, C. (2011). *Practical portfolio performance measurement and attribution*. Vol. 568. John Wiley & Sons. URL: `https://books.google.co.uk/books?hl=en&lr=&id=ng1nMo_cIW0C&oi=fnd&pg=PT10&dq=Bacon,+Carl+R,+2011,+Practical+portfolio+performance+measurement+and+attribution,+volume+568+(John+Wiley+%26+Sons).&ots=9YUy8XFUWD&sig=18b1OSUTQrrac79nDQe8SZpJ4p8`.

— (2015). "Does Your Benchmark Measure Up?" In: *CFA Institute, CFA Pubs*.

Bailey, J. (1992a). "Are Manager Universes Acceptable Performance Benchmarks?" In: *The Journal of Portfolio Management*, pp. 9–13.

— (1992b). "Are Manager Unvierses Acceptable Performance Benchmarks?" In: *The Journal of Portfolio Management, 18(3), pp.9-13.*

— (1992c). "Evaluating Benchmark Quality". In: *Financial Analysts Journal* 48.3, pp. 33–39. DOI: `10.2469/faj.v48.n3.33`.

Bailey, J. and Arnott, R. (Nov. 1986). "Cluster Analysis and Manager Selection". In: *Financial Analysts Journal* 42.6, pp. 20–28. ISSN: 0015-198X. DOI: `10.2469/faj.v42.n6.20`.

Bailey, J., Richards, T., and Tierney, D. (1990). *Benchmark Portfolios: Concept and Design."*. Edited by Frank J. Fabozz. New York: Harper & Row.

— (1998). *Benchmark Portfolios and the Manager/Plan Sponsor Relationship.*

Bailey, M., Muth, R., and Nourse, H. (1963a). "A Regression Method for Real Estate Price Index Construction". In: *Journal of the American Statistical Association* 58.304, pp. 933–942. ISSN: 1537274X. DOI: `10.1080/01621459.1963.10480679`.

— (Dec. 1963b). "A Regression Method for Real Estate Price Index Construction". In: *Journal of the American Statistical Association* 58.304, pp. 933–942. ISSN: 0162-1459. DOI: `10.1080/01621459.1963.10480679`. URL: `http://www.tandfonline.com/doi/abs/10.1080/01621459.1963.10480679`.

Bailey, W. and Chan, K. (1993). "Macroeconomic influences and the variability of the commodity futures basis." In: *The Journal of Finance*, pp. 555–573.

Bali, T., Brown, S., and Caglayan, M. (2011). "Do hedge funds' exposures to risk factors predict their future returns?" In: *Journal of Financial Economics* 101.1, pp. 36–68. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2011.02.008.

— (2014). "Macroeconomic risk and hedge fund returns". In: *Journal of Financial Economics* 114.1, pp. 1–19. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2014.06.008.

Ball, R. et al. (1995). *Problems in measuring portfolio performance An application to contrarian investment strategies*. Tech. rep. URL: http://faculty.chicagobooth.edu/ray.ball/research/Papers/1995%20Problems%20Measuring%20Contrarian%20Strategy%20Returns.pdf.

Bange, M., Khang, K., and Miller, T. (June 2008). "Benchmarking the performance of recommended allocations to equities, bonds, and cash by international investment houses". In: *Journal of Empirical Finance* 15.3, pp. 363–386. ISSN: 0927-5398. DOI: 10.1016/J.JEMPFIN.2007.07.005. URL: https://www.sciencedirect.com/science/article/pii/S0927539807000990#fig3.

Barlett, M. (1948). "A note on the Statistical estimation of supply and demand relations from time series." In: *Econometrica (pre-1986); Evanston*. URL: https://search.proquest.com/docview/214686196/abstract/7DAF774910B74ABDPQ/1?accountid=14116.

Barnhill, T. and Maxwell, W. (Mar. 2002). "Modeling correlated market and credit risk in fixed income portfolios". In: *Journal of Banking & Finance* 26.2-3, pp. 347–374. ISSN: 0378-4266. DOI: 10.1016/S0378-4266(01)00226-6. URL: https://www.sciencedirect.com/science/article/pii/S0378426601002266.

Basak, S. and Pavlova, A. (2016). "A Model of Financialization of Commodities". In: *Journal of Finance* 71.4. ISSN: 15406261. DOI: 10.1111/jofi.12408.

Ben-Zion, U. and Shalit, S. (Sept. 1975). "Title: Size, leverage, and dividend record as determinants of equity risk". In: *The Journal of Finance* 30.4, pp. 1015–1026. DOI: 10.1111/j.1540-6261.1975.tb01018.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1975.tb01018.x.

Bessembinder, H. (Oct. 1992). "Systematic Risk, Hedging Pressure, and Risk Premiums in Futures Markets". In: *Review of Financial Studies* 5.4, pp. 637–667. ISSN: 0893-9454.

DOI: 10.1093/rfs/5.4.637. URL: https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/5.4.637.

Bhardwaj, G., Gorton, G., and Rouwenhorst, G. (2014). "Fooling some of the people all of the time: The inefficient performance and persistence of commodity trading advisers." In: *Review of Financial Studies* 27, 3099-3132.

Black, F., Jensen, M., and Scholes, M. (1972). *The Capital Asset Pricing Model: Some Empirical Tests*. Vol. 81, pp. 79–121. ISBN: 0002269600. DOI: 10.2139/ssrn.908569. URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=908569&rec=1&srcabs=350100.

Blake, D., Lehmann, B., and Timmermann, A. (Sept. 2002). "Performance clustering and incentives in the UK pension fund industry". In: *Journal of Asset Management* 3.2, pp. 173–194. ISSN: 1470-8272. DOI: 10.1057/palgrave.jam.2240073. URL: http://link.springer.com/10.1057/palgrave.jam.2240073.

Blake, D. and Timmermann, A. (1998). "Mutual Fund Performance: Evidence from the UK". In: *European Finance Review* 2.1, pp. 57–77. ISSN: 1572-3097, 1573-692X. DOI: 10.1023/A:1009729630606. arXiv: 0005074v1 [astro-ph]. URL: http://rof.oxfordjournals.org/content/2/1/57%5Cnhttp://rof.oxfordjournals.org/content/2/1/57.short.

Bloomberg Indexes (2018). *The Bloomberg Commodity Index: A Primer on Index Calculation and Performance*. Tech. rep. URL: https://data.bloomberglp.com/indices/sites/2/2016/01/BCOM-Calculation-Primer.pdf.

Blume, M. (Sept. 1974). "Unbiased Estimators of Long-Run Expected Rates of Return". In: *Journal of the American Statistical Association* 69.347, pp. 634–638. ISSN: 0162-1459. DOI: 10.1080/01621459.1974.10480180. URL: http://www.tandfonline.com/doi/abs/10.1080/01621459.1974.10480180.

Bodie, Z., Kane, A., and Marcus, A. (2014). *Investments*. 10e.: McGr. McGraw-Hill Education.

Bollen, N. and Pool, V. (2009). "Do hedge fund managers misreport returns? evidence from the pooled distribution". In: *Journal of Finance* 64.5, pp. 2257–2288. ISSN: 00221082. DOI: 10.1111/j.1540-6261.2009.01500.x.

Borensztein, E. and Mauro, P. (2004). *Book Manuscript*.

Brennan, M. and Schwartz, E. (1985). "On the Geometric Mean Index: A Note." In: *Journal of Financial and Quantitative Analysis* 20.1, pp. 119–122. ISSN: 00221090. DOI: 10.2307/2330681. URL: http://search.ebscohost.com/login.aspx?direct=true%7B&%7Ddb=bth%7B&%7DAN=5720861%7B&%7Dsite=ehost-live.

Brinson, G., Hood, R., and Beebower, G. (1995). "Determinants of Portfolio Performance". In: *Financial Analysts Journal* 51.1, pp. 133–138. ISSN: 0015-198X. DOI: 10.2469/faj.v51.n1.1869. URL: http://www.cfapubs.org/doi/full/10.2469/faj.v51.n1.1869.

Broby, D. (2007). *A guide to Equity Index Construction.* London: Risk Books.

— (2010). *A Guide to Fund Management*. Risk Books, p. 289. ISBN: 1906348189. URL: https://www.amazon.co.uk/Guide-Fund-Management-Daniel-Broby-ebook/dp/B00HHDJSR4/ref=sr_1_cc_1?s=aps&ie=UTF8&qid=1535099207&sr=1-1-catcorr&keywords=daniel+broby.

— (Aug. 2011). "Equity Index Construction". In: *The Journal of Index Investing* 2.2, pp. 36–39. ISSN: 2154-7238. DOI: 10.3905/jii.2011.2.2.036. URL: http://jii.iijournals.com/lookup/doi/10.3905/jii.2011.2.2.036.

— (2017). "Investment Benchmarks: Their Ontological and Epistemological Roots". In: *International Journal of Humanities and Social Science* 7.2.

Broby, D., Bautleac, O., and McKenzie, A. (2017). "A factor model index (FMI) benchmark for commodity investment." In: *Working draft*. URL: https://pure.strath.ac.uk/portal/files/63673909/Broby_IJHSS_2017_Investment_benchmarks_their_ontological.pdf.

Broby, D. and Lochhead, M. (2002). "What is the Appropriate Index Construction Methodology for African Equity Investment?" In: *Journal of Financial Transformation*.

Broby, D., McKenzie, A., and Bautleac, O. (2017). "A factor model index (FMI) benchmark for commodity investment." In: *Unpublished paper (Submission pending)*.

Broby, D. et al. (2016). "The impact of the (2011) devaluation of the swiss franc on eurozone equity benchmark diversification". In: *International Journal of Economics and Financial Issues* 6.3. ISSN: 21464138.

Brooks, C. and Kat, H. (2002). “The statistical properties of hedge fund index returns and their implications for investors.” In: *Journal of Alternative Investments* 5 (Fall),

Brown, S. and Goetzmann, W. (June 1995a). “Performance Persistence”. In: *The Journal of Finance* 50.2, pp. 679–698. DOI: 10.1111/j.1540-6261.1995.tb04800.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1995.tb04800.x.

— (June 1995b). “Performance Persistence”. In: *The Journal of Finance* 50.2, pp. 679–698. ISSN: 00221082. DOI: 10.1111/j.1540-6261.1995.tb04800.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1995.tb04800.x.

— (Jan. 2003a). “Hedge Funds with Style”. In: *The Journal of Portfolio Management* 29.2, pp. 101–112. ISSN: 0095-4918. DOI: 10.3905/jpm.2003.319877. URL: http://jpm.iijournals.com/lookup/doi/10.3905/jpm.2003.319877.

— (2003b). “Hedge funds with style. ." The Journal of Portfolio Management 29(2), pp.101-112.” In: *The Journal of Portfolio Management* 29(2), pp.

Brown, S. and Warner, J. (1980). “Measuring security price performance”. In: *Journal of Financial Economics* 8(3).

Brown, S. et al. (Oct. 1992). “Survivorship Bias in Performance Studies”. In: *Review of Financial Studies* 5.4, pp. 553–580. ISSN: 0893-9454. DOI: 10.1093/rfs/5.4.553. URL: https://academic.oup.com/rfs/article-lookup/doi/10.1093/rfs/5.4.553.

Burns, W. and Epley, D. (Apr. 1982). “The performance of portfolios of REITS + stocks”. In: *The Journal of Portfolio Management* 8.3, pp. 37–42. ISSN: 0095-4918. DOI: 10.3905/jpm.1982.408866. URL: http://jpm.iijournals.com/lookup/doi/10.3905/jpm.1982.408866.

Busse, S., Brümmer, B., and Ihle, R. (2012). “Busse, S., Brümmer, B. and Ihle, R., 2011. Emerging linkages between price volatilities in energy and agricultural markets.” In: *Safeguarding food security in volatile global markets.*, pp. 107–121.

Bussiere, M., Hoerova, M., and Klaus, B. (2014). “Commonality in Hedge Fund Returns.” In: *ECB Working Paper Series March.*

Campbell, J. and Cochrane, J. (2000). "Explaining the poor performance of consumption-based asset pricing models". In: *Journal of Finance* 55.6, pp. 2863–2878. ISSN: 00221082. DOI: 10.1111/0022-1082.00310.

Campbell, R. (2010). "Art as an Alternative Asset Class". In: *Journal of Wine Economics* 5.February, pp. 87–118. ISSN: 1931-4361. DOI: 10.1017/S1931436100001395.

Cao, D. and Teiletche, J. (2007). "Reconsidering asset allocation involving illiquid assets." In: *Journal of Asset Management* 267-282.

Cargill, T. and Rausser, G. (Sept. 1975). "Temporal price behavior in commodity futures markets." In: *The Journal of Finance* 30.4, pp. 1043–1053. DOI: 10.1111/j.1540-6261.1975.tb01020.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1975.tb01020.x.

Carhart, M. (1997). "On persistence in mutual fund performance." In: *The Journal of finance* 52(1), pp.

Carpenter, J. and Lynch, A. (Dec. 1999). "Survivorship bias and attrition effects in measures of performance persistence". In: *Journal of Financial Economics* 54.3, pp. 337–374. ISSN: 0304-405X. DOI: 10.1016/S0304-405X(99)00040-9. URL: https://www.sciencedirect.com/science/article/pii/S0304405X99000409.

Case, B. and Quigley, J. (1991). "The dynamics of real estate prices." In: *The Review of Economics and Statistics,* pp.50-58.

Case, K. and Shiller, R. (Feb. 1988). *The Efficiency of the Market for Single-Family Homes*. Tech. rep. Cambridge, MA: National Bureau of Economic Research. DOI: 10.3386/w2506. URL: http://www.nber.org/papers/w2506.pdf.

Cerny, B. and Kaiser, H. (1977). "A Study Of A Measure Of Sampling Adequacy For Factor-Analytic Correlation Matrices". In: *Multivariate Behavioral Research* 12.1, pp. 43–47. ISSN: 0027-3171. DOI: 10.1207/s15327906mbr1201{_}3. URL: https://www.tandfonline.com/action/journalInformation?journalCode=hmbr20.

Chance, W. (Feb. 1966). "A Note on the Origins of Index Numbers". In: *The Review of Economics and Statistics* 48.1, p. 108. ISSN: 00346535. DOI: 10.2307/1924869. URL: https://www.jstor.org/stable/1924869?origin=crossref.

Chang, K. (May 2011). "The nonlinear effects of expected and unexpected components of monetary policy on the dynamics of REIT returns". In: *Economic Modelling* 28.3, pp. 911–920. ISSN: 0264-9993. DOI: 10.1016/J.ECONMOD.2010.11.003. URL: https://www.sciencedirect.com/science/article/pii/S0264999310002257#f0010.

Chao, Y. and Wu, C. (1996). "Principal component-based weighted indices and a framework to evaluate indices: Results from the Medical Expenditure Panel Survey 1996 to 2011". In: DOI: 10.1371/journal.pone.0183997. URL: https://meps.ahrq.gov/data_stats/download_data_files.jsp..

Chaves, D., Kalesnik, V., and Little, B. (2011). "The Information in the Term Structure of Commodity Futures". In: *Available at SSRN 1982200* July 2016. ISSN: 1556-5068. DOI: 10.2139/ssrn.1982200. URL: http://papers.ssrn.com/sol3/papers.cfm?abstract_id=1982200.

Chen, N., Roll, R., and Ross, S. (1986). "Economic Forces and the Stock Market: Testing the APT and Alternative Asset Pricing Theories". In: *Journal of Business* 59.3, pp. 383–403.

Cheng, I. and Xiong, W. (2014). "Financialization of Commodity Markets". In: *Annual Review of Financial Economics* 6.1, pp. 419–441. ISSN: 1941-1367. DOI: 10.1146/annurev-financial-110613-034432. URL: http://www.annualreviews.org/doi/10.1146/annurev-financial-110613-034432.

Cheng, P. and Roulac, S. (2007). "REIT Characteristics and Predictability". In: *International Real Estate Review*. URL: https://core.ac.uk/download/pdf/7074726.pdf.

Chow, T. et al. (2011). "A survey of alternative equity index strategies." In: *Financial Analysts Journal,* 67(5), pp.

Christoffersen, A., Olesen, K., and Christoffersen, P. (2014). "Factor Structure in Commodity Futures Return and Volatility". In: *Rotman School of Management Working Paper, (2495779).* URL: http://ssrn.com/abstract=2495779.

Clarke, C. et al. (2015). "The Level, Slope and Curve Factor Model for Stocks". In:

Collin-Dufresn, P. and Goldstein, R. (2001). "The determinants of credit spread changes." In: *The Journal of Finance* 56, pp. 2177–2207.

Connor, J. and Rossiter, R. (Jan. 2005). "Wavelet Transforms and Commodity Prices". In: *Studies in Nonlinear Dynamics & Econometrics* 9.1. ISSN: 1558-3708. DOI: 10.2202/1558-3708.1170. URL: https://www.degruyter.com/view/j/snde.2005.9.1/snde.2005.9.1.1170/snde.2005.9.1.1170.xml.

Conover, M., Broby, D., and Cariño, D. (2013). "Introduction to Benchmarks". In: *CIPM Readings, CFA Institute*.

Corgel, J., Mcintosh, W., and Ott, S. (1995). "Real Estate Investment Trusts: A Review of the Financial Economics Literature Part of the Finance and Financial Management Commons, and the Real Estate Commons". In: *The Scholarly Commons*. URL: http://scholarship.sha.cornell.edu/articleshttp://scholarship.sha.cornell.edu/articles/1093.

Cremers, K. and Petajisto, A. (2009). "How active is your fund manager? A new measure that predicts performance." In: *The Review of Financial Studies* 22(9), pp.

Cumming, D., Hassand, H., and Schweizer, D. (2014). "Strategic asset allocation and the role of alternative investments." In: *European Financial Management* 20(3), pp.

Cumming, D., Helge Hass, L., and Schweizer, D. (2013a). "Private equity benchmarks and portfolio optimization". In: *Journal of Banking and Finance* 37.9, pp. 3515–3528. ISSN: 03784266. DOI: 10.1016/j.jbankfin.2013.04.010.

— (Sept. 2013b). "Private equity benchmarks and portfolio optimization". In: *Journal of Banking & Finance* 37.9, pp. 3515–3528. ISSN: 0378-4266. DOI: 10.1016/J.JBANKFIN.2013.04.010. URL: https://www.sciencedirect.com/science/article/pii/S0378426613001933#f0015.

Daniel, K. et al. (1997). "Measuring Mutual Fund Performance with Benchmarks". In: *the Journal of Finance* 52.3, pp. 1035–1058. ISSN: 00221082. DOI: 10.1111/j.1540-6261.1997.tb02724.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1997.tb02724.x.

De Souza, C. and Gokcan, S. (Jan. 2004). "Hedge Fund Investing". In: *The Journal of Wealth Management* 6.4, pp. 52–73. ISSN: 1534-7524. DOI: 10.3905/jwm.2004.391058. URL: http://jwm.iijournals.com/lookup/doi/10.3905/jwm.2004.391058.

DeMiguel, V., Garlappi, L., and Uppal, R. (2009). "Optimal versus naive diversification: How inefficient is the 1/N portfolio strategy?" In: *Review of Financial Studies* (22(5), 19.

Denvir, E. and Hutson, E. (Feb. 2006). "The performance and diversification benefits of funds of hedge funds". In: *Journal of International Financial Markets, Institutions and Money* 16.1, pp. 4–22. ISSN: 1042-4431. DOI: 10.1016/J.INTFIN.2004.12.002. URL: https://www.sciencedirect.com/science/article/pii/S1042443105000193#fig1.

Dhillon, U. and Johnson, H. (1991). "Changes in the Standard and Poor's 500 List". In: *The Journal of Business*, pp. 75–85. URL: https://www.jstor.org/stable/pdf/2353073.pdf.

Diesinger, P., Kraft, H., and Seifried, F. (Sept. 2010). "Asset allocation and liquidity breakdowns: what if your broker does not answer the phone?" In: *Finance and Stochastics* 14.3, pp. 343–374. ISSN: 0949-2984. DOI: 10.1007/s00780-008-0085-5. URL: http://link.springer.com/10.1007/s00780-008-0085-5.

Diewert, E. (1998). *Index Number Issues in the Consumer Price Index*. Tech. rep., pp. 47–58. URL: https://pubs.aeaweb.org/doi/pdfplus/10.1257/jep.12.1.47.

Diewert, E., Fox, K., and Shimizu, C. (Dec. 2016). "Commercial Property Price Indexes and the System of National Accounts." In: *Journal of Economic Surveys* 30.5, pp. 913–943. ISSN: 14676419. DOI: 10.1111/joes.12117.

Dimson, E. (1979). *Risk Measurement when shares are subject to infrequent trading*. Tech. rep. London Business School. URL: https://s3.amazonaws.com/academia.edu.documents/46386849/0304-405x_2879_2990013-820160610-22134-1d2i8po.pdf?AWSAccessKeyId=AKIAIWOWYYGZ2Y53UL3A&Expires=1536232101&Signature=gaON%2BQ89L1yaKrxvPPxDUGEoO4Q%3D&response-content-disposition=inline%3B%20filename.

Domanski, D. and Heath, A. (2007). "Financial investors and commodity markets". In: *BIS Quarterly Review* March, pp. 53–67.

Dor, A. and Jagannathan, R. (2002). "Understanding mutual fund and hedge fund styles using return based style analysis". In: *National Bureau of Economic Research.* (No. w9111.

Duarte, D., Jefferson, F., and Yu, F. (2007). “Risk and return in fixed incomearbitrage: nickels in front of a steamroller?” In: *Review of Financial Studies* 20 770-811.

Dunne, P., Moore, M., and Portes, R. (2007). “Benchmark Status in Fixed-Income Asset Markets”. In: *Journal of Business Finance & Accounting* 34.9, pp. 306–686. DOI: 10.1111/j.1468-5957.2007.02039.x. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1468-5957.2007.02039.x.

Dunsby, A. and Nelson, K. (2010). “An evolution from passive to active indexes A Brief History Of Commodities Indexes”. In: *Journal of Indexes*. URL: www.journalofindexes.com.

Dutot, C. (1738). *Reflexions Politiques sur les Finances et le Commerce*. URL: https://www.jstor.org/stable/pdf/1924869.pdf.

Dybvig, P. and Ross, S. (June 1985). “The Analytics of Performance Measurement Using a Security Market Line”. In: *The Journal of Finance* 40.2, pp. 401–416. ISSN: 00221082. DOI: 10.1111/j.1540-6261.1985.tb04964.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1985.tb04964.x.

Edmister, R., Graham, S., and Pirie, W. (Sept. 1994). “Excess returns of index replacement stocks: Evidence of liquidity and substitutability”. In: *Journal of Financial Research* 17.3, pp. 333–346. DOI: 10.1111/j.1475-6803.1994.tb00196.x. URL: http://doi.wiley.com/10.1111/j.1475-6803.1994.tb00196.x.

Edwards, F. and Caglayan, M. (2001). “Hedge fund performance and manager skill.” In: *Journal of Futures Markets* 21(11),pp.

Eichhorn, W. (1976). “Fisher Tests Revisited”. In: *Econometrica* 44, No 2.March. URL: https://search.proquest.com/docview/214667366?pq-origsite=gscholar.

Eling, M. (Mar. 2009). “Does Hedge Fund Performance Persist? Overview and New Empirical Evidence”. In: *European Financial Management* 15.2, pp. 362–401. ISSN: 13547798. DOI: 10.1111/j.1468-036X.2008.00471.x. URL: http://doi.wiley.com/10.1111/j.1468-036X.2008.00471.x.

Eling, M. and Schuhmacher, F. (2006). “Wahl des Performancemasses einen Einfluss auf die Beurteilung von Hedgefonds-Indizes? - Choosing the performance mass influences the assessment of hedge fund indices?” In: *Kredit und Kapital* 39 (3), 41.

Eling, M. and Schuhmacher, F. (2007). “Does the choice of performance measure influence the evaluation of hedge funds?” In: *Journal of Banking and Finance*. DOI: 10.1016/j.jbankfin.2006.09.015.

Eling, M. et al. (2010). “Skewness in hedge funds returns: Classical skewness coefficients vs Azzalini’s skewness parameter”. In: *International Journal of Managerial Finance* 6.4, pp. 290–304. ISSN: 17439132. DOI: 10.1108/17439131011074459.

Elton, E. and Gruber, M. (1992). “Portfolio analysis with a nonnormal multi-index return-generating process”. In: *Review of Quantitative Finance and Accounting* 2.1, pp. 5–16. ISSN: 0924865X. DOI: 10.1007/BF00243981.

— (1997). “Modern portfolio theory, 1950 to date.” In: *Journal of Banking & Finance* 21 (1743):

Elton, E., Gruber, M., and Blake, C. (1999). “Common factors in active and passive portfolios.” In: *European Finance Review* 3(1), pp.5.

Elton, E., Gruber, M., and Das, S. (1993). “Efficiency with Costly Information: A Reinterpretation of Evidence from Managed Portfolios,.” In: *Review of Financial Studies* 6, 1-22.

Elton, E., Gruber, M., and Rentzler, J. (1987). *Professionally Managed, Publicly Traded Commodity Funds*. DOI: 10.2307/2352810. URL: https://www.jstor.org/stable/2352810.

Enderle, F., Pope, B., and Siegel, L. (Feb. 2003a). “Broad-Capitalization Indexes of the U.S. Equity Market”. In: *The Journal of Investing* 12.1, pp. 11–22. ISSN: 1068-0896. DOI: 10.3905/joi.2003.319529. URL: http://joi.iijournals.com/lookup/doi/10.3905/joi.2003.319529.

— (2003b). “Broad-capitalization indexes of the US equity market.” In: *The Journal of Investing*, pp. 11–22.

Ennis, R. and Burik, P. (May 1991). “Pension Fund Real Estate Investment Under a Simple Equilibrium Pricing Model”. In: *Financial Analysts Journal* 47.3, pp. 20–30. ISSN: 0015-198X. DOI: 10.2469/faj.v47.n3.20. URL: http://www.cfapubs.org/doi/full/10.2469/faj.v47.n3.20.

Erb, C. and Harvey, C. (2006). “The Strategic and Tactical Value o f Commodity Futures”. In: *Financial Analysts Journal* 62.2.

Escobari, D. and Jafarinejad, M. (May 2016). "Date stamping bubbles in Real Estate Investment Trusts". In: *The Quarterly Review of Economics and Finance* 60, pp. 224–230. ISSN: 1062-9769. DOI: 10.1016/J.QREF.2015.10.003. URL: https://www.sciencedirect.com/science/article/pii/S1062976915001027#fig0020.

Fabozzi, F. and Kaiser, D. (2011). *The Handbook of Commodity Investing*. ISBN: 9780470117644. DOI: 10.1002/9781118267004.

Fama, E. (1972). "Components of investment performance." In: *The Journal of Finance* 27(3), pp.

Fama, E. and French, K. (1992). "The cross-section of expected stock returns." In: *Journal of Finance*.

— (2010). "Luck versus Skill in the Cross-Section of Mutual Fund Returns". In: *The Journal of Finance* LXV.5. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.2010.01598.x.

Feng, Z., Ghosh, C., and Sirmans, C. (2007). "On the capital structure of Real Estate Investment Trusts (REITs)". In: *Journal of Real Estate Finance and Economics*. ISBN: 1114600790052. DOI: 10.1007/s11146-007-9005-2.

Fernandez-Perez, A. et al. (2015). "The Pricing of Skewness in Commodity Futures Markets: Risk or Lottery?" In: *FMA Conference paper*. URL: http://www.fmaconferences.org/Vegas/Papers/Skewness_EDHEC_January_2016.pdf.

Ferrero, C. and Morri, G. (June 2004). "Real estate indexes and performance benchmarks". In: *ERES*. URL: https://ideas.repec.org/p/arz/wpaper/eres2004_137.html.

Fisher, J., Geltner, D., and Webb, R. (1994). "Value indices of commercial real estate: a comparison of index construction methods." In: *The Journal of Real Estate Finance and Economics* 9(2), pp.1.

Fletcher, J. (2017). "An examination of the benefits of factor investing in the UK stock returns". In: *Working Paper (Strathclyde Business School)*.

Fletcher, J. and Marshall, A. (Dec. 2005). "An empirical examination of the benefits of international diversification". In: *Journal of International Financial Markets, Institutions* *and Money* 15.5, pp. 455–468. ISSN: 10424431. DOI: 10.1016/j.intfin.2004.11.002. URL: http://linkinghub.elsevier.com/retrieve/pii/S1042443105000065.

Flury, B. (1984). "Common principal components in k groups". In: *Journal of the American Statistical Association* 79.388, pp. 892–898. URL: http://amstat.tandfonline.com/doi/abs/10.1080/01621459.1984.10477108.

Fong, K. and Lee, D. (2008). "Benchmarking benchmarks: measuring characteristic selectivity using portfolio holdings data." In: *Accounting & Finance* 48.5: 761-.

Fraley, C. and Raftery, A. (n.d.). *How Many Clusters? Which Clustering Method? Answers Via Model-Based Cluster Analysis*. Tech. rep. URL: https://academic.oup.com/comjnl/article-abstract/41/8/578/360856.

Frazzini, A., Friedman, J., and Pomorski, L. (Mar. 2016). "Deactivating Active Share". In: *Financial Analysts Journal* 72.2, pp. 14–21. ISSN: 0015-198X. DOI: 10.2469/faj.v72.n2.2. URL: http://www.cfapubs.org/doi/full/10.2469/faj.v72.n2.2.

Frino, A., Gallagher, D., and Oetomo, T. (2005). "The Index Tracking Strategies of Passive and Enhanced Index Equity Funds". In: *Australian Journal of Management*. ISSN: 03128962. DOI: 10.1177/031289620503000103.

FTSE Russell. (2017). "Corporate Actions and Events Guide for Non-Market Capitalisation Weighted Indices". In: July, pp. 1–29.

Fung, W. and Hsieh, D. (1999). "A primer on hedge funds". In: *Journal of Empirical Finance* 6.3, pp. 309–331. ISSN: 09275398. DOI: 10.1016/S0927-5398(99)00006-7. URL: http://linkinghub.elsevier.com/retrieve/pii/S0927539899000067.

— (2001a). "The risk in hedge fund strategies: Theory and evidence from trend followers". In: *Review of Financial Studies* 14.2, pp. 313–341. ISSN: 08939454. DOI: 10.1093/rfs/14.2.313.

— (2001b). "The Society for Financial Studies The Risk in Hedge Fund Strategies: Theory and Evidence from Trend Followers". In: *Source: The Review of Financial Studies* 14.2, pp. 313–341. URL: http://www.jstor.org/stable/2696743%20http://about.jstor.org/terms.

Fung, W. and Hsieh, D. (2002a). “Hedge-Fund Benchmarks: Information Content and Biases”. In: *Financial Analysts Journal* 58159210.1, pp. 22–34. URL: http://www.jstor.org%20http://www.jstor.org/stable/4480366%20http://www.jstor.org/page/.

— (2002b). “Risk in Fixed-Income Hedge Fund Styles”. In: *The Journal of Fixed Income* 12.2, pp. 6–27. ISSN: 1059-8596. DOI: 10.3905/jfi.2002.319321.

— (2004a). “CFA Institute Hedge Fund Benchmarks: A Risk-Based Approach”. In: *Financial Analysts Journal* 60159210.5, pp. 65–80. URL: http://www.jstor.org%20http://www.jstor.org/stable/4480604%20http://www.jstor.org/page/.

— (2004b). *Hedge fund benchmarks: A risk-based approach*. DOI: http://dx.doi.org/10.2469/faj.v60.n5.2657.

— (2011). “The risk in hedge fund strategies: Theory and evidence from long/short equity hedge funds”. In: *Journal of Empirical Finance* 18.4, pp. 547–569. ISSN: 09275398. DOI: 10.1016/j.jempfin.2011.04.001.

Gao, G. and Rossi, A. (2013). “Understanding the Sources of Risk Underlying the Cross-Section of Commodity Returns”. In: *Working paper*.

Gargano, A. and Timmermann, A. (2014). “Forecasting commodity price indexes using macroeconomic and financial predictors”. In: *International Journal of Forecasting* 30.3, pp. 825–843. ISSN: 01692070. DOI: 10.1016/j.ijforecast.2013.09.003.

Geltner, D. (1993). “Temporal Aggregation in Real Estate Return Indices”. In: *Real Estate Economics* 21.2, pp. 141–166. ISSN: 15406229. DOI: 10.1111/1540-6229.00605.

— (2001). “Ideal Research and Benchmark Indexes in Private Real Estate: Some Conclusions from the RERI/PREA Technical Report.” In: *Briefings in real estate finance.* Henry Stewart Publications. Chap. Vol.17(4), URL: http://suprimo.lib.strath.ac.uk/primo_library/libweb/action/openurl?date=2001&aulast=Geltner&issue=4&isSerivcesPage=true&spage=17&dscnt=2&auinit=D&title=Real+estate+finance&atitle=Ideal+Research+and+Benchmark+Indexes+in+Private+Real+Estate%3A+Some+Conclus.

Geltner, D. and Ling, D. (2009). “Considerations in the design and construction of investment real estate research indices.” In: *Journal of Real Estate Research.*

Geltner, D., MacGregor, A., and Schwann, G. (2003). "Appraisal smoothing and price discovery in real estate markets." In: *Urban Studies* 40(5-6), p.

Getmansky, M., Lo, C., and Makarov, I. (2004). "An econometric model of serial correlation and illiquidity in hedge fund returns." In: *Journal of Financial Economics, 74(3), pp.529-609.*

Giacomini, E., Ling, D., and Naranjo, A. (2017). "REIT Leverage and Return Performance: Keep Your Eye on the Target". In: 4, pp. 930–978. DOI: 10.1111/1540-6229.12179. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/1540-6229.12179.

Gibbons, M., Ross, S., and Shanken, J. (1989). "A test of the efficiency of a given portfolio." In: *Econometrica: Journal of the Econometric Society* pp.1121-11.

Gibson, R. and Schwartz, E. (1990). "Stochastic convenience yield and the pricing of oil contingent claims". In: *The Journal of Finance* 45(3), 959.

Glosten, L. and Jagannathan, R. (1994). "A contingent claim approach to performance evaluation." In: *Journal of Empirical Finance 1(2), pp.133-160.* Journal of.

Goetzmann, W., Ingersoll, J., and Ross, S. (2003). "High-Water Marks and Hedge Fund Management Contracts". In: *Journal of Finance* LVIII no 4.August.

Gorton, G. and Geert, G. (2006). "Facts and fantasies about commodity futures". In: *Financial Analysts Journal* 62.2, pp. 47–68. ISSN: 0015198X. DOI: 10.2469/faj.v62.n2.4083.

Gorton, G., Hayashi, F., and Rouwenhorst, G. (2013). "The fundamentals of commodity futures returns". In: *Review of Finance* 17.1, pp. 35–105. ISSN: 15723097. DOI: 10.1093/rof/rfs019.

Goyenko, R., Holden, C., and Trzcinka, C. (2009). "Do liquidity measures measure liquidity?" In: *Journal of Financial Economics* 92.2, pp. 153–181. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2008.06.002.

Graham, S. and Pirie, W. (June 1994). "Index fund rebalancing and market efficiency". In: *Journal of Economics and Finance* 18.2, pp. 219–229. ISSN: 1055-0925. DOI: 10.1007/BF02920524. URL: http://link.springer.com/10.1007/BF02920524.

Granger, C. and Ding, Z. (1995). "Some properties of absolute return: An alternative measure of risk." In: *Annales d'Economie et de Statistique* pp.67-91.

Granger, C. and Joyeux, R. (Jan. 1980). "An introduction to long memory time series models and fractional differencing". In: *Journal of Time Series Analysis* 1.1, pp. 15–29. DOI: 10.1111/j.1467-9892.1980.tb00297.x. URL: http://doi.wiley.com/10.1111/j.1467-9892.1980.tb00297.x.

Granito, M. (1987). "The problem with bond index funds." In: *The Journal of Portfolio Management,* 13(4), pp.

Greer, R. (2000). "The Nature of Commodity Index Returns". In: *The Journal of Alternative Investments* Summer. URL: http://www.priceasset.com/documents/Naure%20of%20Commodity%20Index%20Returns.pdf.

Gregoriou, G. and Kooli, M. (2011). *Hedge fund replication*. Ed. by Springer, pp. 1–193. ISBN: 9780230358317. DOI: 10.1057/9780230358317.

Grinblatt, M. and Titman, S. (1993). "Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns Performance Measurement without Benchmarks: An Examination of Mutual Fund Returns*". In: *The Journal of Business* 66159210.1, pp. 47–68. URL: http://www.jstor.org%20http://www.jstor.org/stable/2353341%20http://www.jstor.org/page/.

— (1994). "A study of monthly mutual fund returns and performance evaluation techniques." In: *Journal of ﬁGrinblatt, M. and Titman, S. (1994) 'A study of monthly mutual fund returns and performance evaluation techniques.', Journal of financial and quantitative analysis, (29(03), pp.419-444.).nancial and quantitative analysis* 29(03), pp.419-444.

Grindblatt, M. and Titman, S. (Dec. 1992). "The Persistence of Mutual Fund Performance". In: *The Journal of Finance* 47.5, pp. 1977–1984. DOI: 10.1111/j.1540-6261.1992.tb04692.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1992.tb04692.x.

Groot, W. and Swinkels, L. (June 2008). "Incorporating uncertainty about alternative assets in strategic pension fund asset allocation". In: *Pensions: An International Journal* 13.1-2, pp. 71–77. DOI: 10.1057/palgrave.pm.5950063. URL: http://link.springer.com/10.1057/palgrave.pm.5950063.

Gruber, M., Agrawal, D., and Mann, C. (2004). "Factors affecting the valuation of corporate bonds." In: *Journal of Banking & Finance* 28(11), pp.

Gyourko, J. and Keim, D. (1992). "What Does the Stock Market Tell Us About Real Estate Returns?" In: *Real Estate Economics* 20.3, pp. 457–485. ISSN: 15406229. DOI: `10.1111/1540-6229.00591`.

Hafner, R. and Wallmeier, M. (2007). "Volatility as an asset class: European evidence". In: *European Journal of Finance* 13.7, pp. 621–644. ISSN: 1351847X. DOI: `10.1080/13518470701380142`.

Hamdan, R. et al. (2016). "A Primer on Alternative Risk Premia". In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: `10.2139/ssrn.2766850`.

Hamilton, J. and Wu, J. (2015). "Effects of index-fund investing on commodity futures prices". In: *International Economic Review* 56.1. ISSN: 14682354. DOI: `10.1111/iere.12099`.

Hammond, P. and Subramanian, R. (2013). *Building Best Practices Benchmarks for Global Equities."*. Tech. rep.

Harvey, C. and Siddique, A. (2000). *Conditional Skewness in Asset Pricing Tests; Conditional Skewness in Asset Pricing Tests*. Tech. rep. 3. DOI: `10.1111/0022-1082.00247`. URL: `https://onlinelibrary.wiley.com/doi/pdf/10.1111/0022-1082.00247`.

Harvey, C., Zhu, Y., and Liu, H. (2016). ". . . . and the cross-section expected returns." In: *Review of Financial Studies* 29, 5-68.

Hasanhodzic, J. and Lo, A. (2008). "Can Hedge-Fund Returns Be Replicated?: The Linear Case." In: *CFA Digest.*

Haugen, R. and Baker, N. (1991). "The efficient market inefficiency of capitalization–weighted stock portfolios." In: *The Journal of Portfolio Management* 17(3), pp.

Hedges, J. (2003). "Hedge fund indices". In: *Journal of Financial Transformation*, pp. 52–56. ISSN: 1520-3255. DOI: `10.3905/jai.2001.318995`.

Heinkel, R. and Stoughton, N. (1994). "The Dynamics of Portfolio Management Contracts." In: *Review of Financial Studies 351–387.*

Hendrciks, D., Jaydendu, P., and Richard, Z. (Mar. 1993). "Hot Hands in Mutual Funds: Short-Run Persistence of Relative Performance, 1974-1988". In: *The Journal of Finance* 48.1, pp. 93–130. DOI: 10.1111/j.1540-6261.1993.tb04703.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1993.tb04703.x.

Hendrickson, A. and White, P. (1964). "Promax: a quick method for rotation to oblique simple structure". In: *British Journal of Mathematical and Statistical Psychology* 17.1, pp. 65–70. URL: http://onlinelibrary.wiley.com/doi/10.1111/j.2044-8317.1964.tb00244.x/full.

Hoag, J. (May 1980). "Towards Indices of Real Estate Value and Return". In: *The Journal of Finance* 35.2, pp. 569–580. ISSN: 00221082. DOI: 10.1111/j.1540-6261.1980.tb02189.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1980.tb02189.x.

Hoernemann, J., Junkans, D., and Zarate, C. (2005). "Strategic asset allocation and other determinants of portfolio returns." In: *The Journal of Wealth Management, 8(3), pp.26-38.*

Houweling, P. and Zundert, J. (2017). "Factor Investing in the Corporate Bond Market ." In: *Financial Analysts Journal* Vol. 73.

Hubert, M., Rousseeuw, P., and Verdonck, T. (2009). "Robust PCA for skewed data and its outlier map". In: *Computational Statistics and Data Analysis*. ISSN: 01679473. DOI: 10.1016/j.csda.2008.05.027.

Hunter, D., Kandel, S., and Wermers, R. (2014). "Mutual fund performance evaluation with active peer benchmarks." In: *Journal of Financial Economics* 112(1), pp.

Ibbotson, R. (2012). *Ibbotson SBBI Classic Yearbook*. URL: http://corporate.morningstar.com/ib/asp/subject.aspx?xmlfile=1414.xml.

Ibbotson, R. and Kaplan, P. (2000). "Does asset allocation policy explain 40, 90, or 100 percent of perforance?" In: *Financial Analysts Journal* pp.26-33.

Idzorek, T. and Kowara, M. (2013). "Factor-Based Asset Allocation vs. Asset-Class-Based Asset Allocation". In: *Financial Analyst Journal* 69.3, pp. 19–30.

Inderst, G. (June 2011). *Infrastructure as an Asset Class*. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=1860947.

Indexes, Russell (2008). "Russell indexes". In: *Futures* December.

Indices, Dow Jones and Methodology, Index (2015). "Methodology". In: October.

Ippolito, R. (1989). “Efficiency with costly information: A study of mutual fund performance, 1965–1984.” In: *The Quarterly Journal of Economics* 104(1), pp.

Irwin, S. and Sanders, D. (2011). “Index funds, financialization, and commodity futures markets”. In: *Applied Economic Perspectives and Policy*. ISSN: 20405790. DOI: 10.1093/aepp/ppq032.

Irwin, S., Sanders, D., and Merrin, R. (Aug. 2009). “Devil or Angel? The Role of Speculation in the Recent Commodity Price Boom (and Bust)”. In: *Journal of Agricultural and Applied Economics* 41.02, pp. 377–391. ISSN: 1074-0708. DOI: 10.1017/S1074070800002856. URL: https://www.cambridge.org/core/product/identifier/S1074070800002856/type/journal_article.

Jacobs, H., Muller, S., and Weber, M. (2014). “How should individual investors diversify? An empirical evaluation of alternative asset allocation policies."”. In: *Journal of Financial Markets* 19, pp.62-.

Jacobs, H., Müller, S., and Weber, M. (June 2014). “How should individual investors diversify? An empirical evaluation of alternative asset allocation policies”. In: *Journal of Financial Markets* 19, pp. 62–85. ISSN: 1386-4181. DOI: 10.1016/J.FINMAR.2013.07.004. URL: https://www.sciencedirect.com/science/article/pii/S1386418113000451.

Jaeger, L. and Christian, W. (2005). “Factor Modelling and Benchmarking of Hedge Funds: Can passive investments in hedge fund strategies deliver?” In: *Partners Group, Baar-Zug*.

Jagannathan, R., Malakhov, A., and Novikov, D. (2010). “Do hot hands exist among hedge fund managers? An empirical evaluation.” In: *The Journal of Finance* 65(1), pp.

Jeffers, J. (1967). “Two Case Studies in the Application of Principal Component Analysis”. In: *Applied Statistics*. ISSN: 00359254. DOI: 10.2307/2985919.

Jensen, M. (1968). “The performance of mutual funds in the period 1945–1964.” In: The Journal of Finance.

Jensen, M. and Black Scholes, M. (1972). “The capital asset pricing model: Some empirical tests”. In:

Jevons, S. (June 1865). “On the Variation of Prices and the Value of the Currency since 1782”. In: *Journal of the Statistical Society of London* 28.2, p. 294. ISSN: 09595341. DOI: 10.

2307/2338419. URL: https://www.jstor.org/stable/10.2307/2338419?origin=crossref.

Jianping, M. and Moses, M. (2002). “Art as an Investment and the Underperformance of Masterpieces”. In: *The American Economic Review* Vol. 92, N. URL: https://www.jstor.org/stable/pdf/3083271.pdf.

Jobson, J. and Korkie, B. (1980). “Estimation for Markowitz Efficient Portfolios”. In: *Journal of the American Statistical Association,* Vol. 75, N, pp. 544–554. URL: https://www.jstor.org/stable/pdf/2287643.pdf.

Jolliffe, A. (2002). *Principal Component Analysis*. Springer Series in Statistics. New York: Springer-Verlag. ISBN: 0-387-95442-2. DOI: 10.1007/b98835. URL: http://link.springer.com/10.1007/b98835.

Jorion, P. (2007). “Value at Risk: The New Benchmark for Managing Financial Risk”. In: *New York* 0071464956, p. 602. ISSN: 0932-5026. DOI: 10.1007/s00362-009-0296-7. URL: http://www.lavoisier.fr/livre/notice.asp?id=OOLWR2AAAO3OWX.

Kaiser, D. and Haberfelner, F. (2012). “Hedge fund biases after the financial crisis”. In: *Managerial Finance* 38.1, pp. 27–43. ISSN: 0307-4358. DOI: 10.1108/03074351211188349.

Kaiser, H. (1958). “The varimax criterion for analytic rotation in factor analysis”. In: *Psychometrika* 23.3, pp. 187–200. URL: http://www.springerlink.com/index/977786TH55M62671.pdf.

Kaplan, P. and Knowles, J. (2004). “Kappa: A generalized downside risk-adjusted performance measure.” In: *Journal of Performance Measurement* 8, pp.42-5.

Kaplan, S. and Ruback, R. (1995). *The Valuation of Cash Flow Forecasts: An Empirical Analysis*. Tech. rep. 4, pp. 1059–1093. URL: https://www.jstor.org/stable/pdf/2329344.pdf.

Kaplan, S. and Schoar, A. (2005). “Private equity performance: Returns, persistence, and capital flows.” In: *The Journal of Finance.* 60(4), pp.

Kat, H. (2001). *The Statistical Properties of Hedge Fund Index returns*. Tech. rep. URL: www.ismacentre.rdg.ac.uk.

Kat, H. and Brooks, C. (2001). “The Statistical Properties of Hedge Fund Index Returns and their Implications for Investors”. In: *ISMA Centre Discussion Papers In Finance*.

Keating, C. and Shadwick, W. (2002). “A universal performance measure”. In: *Journal of performance measurement* 6(3), pp.5.

Kraus, A. and Litzenberger, R. (Sept. 1976). “Skewness Preference and the valuation of Risk Assets”. In: *The Journal of Finance* 31.4, pp. 1085–1100. DOI: 10.1111/j.1540-6261.1976.tb01961.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1976.tb01961.x.

Kritzman, M. (June 1999). “Toward Defining an Asset Class”. In: *The Journal of Alternative Investments* 2.1, pp. 79–82. ISSN: 1520-3255. DOI: 10.3905/jai.1999.318919. URL: http://jai.iijournals.com/lookup/doi/10.3905/jai.1999.318919.

Kryzanowski, L. and Rahman, A. (2008). “Portfolio performance ambiguity and benchmark inefficiency revisited”. In: *Journal of Asset Management* 9.5, pp. 321–332. ISSN: 1470-8272. DOI: 10.1057/jam.2008.33.

Kuenzi, D. (2003). “Strategy benchmarks.” In: *The Journal of Portfolio Management,* 29(2), pp.

Kuhle, J. (1987). “Portfolio diversification and return benefits—common stock vs. real estate investment trusts (REIT’s)”. In: *The Journal of Real Estate Research.* Vol.2(2), URL: http://suprimo.lib.strath.ac.uk/primo_library/libweb/action/openurl?date=1987&aulast=Kuhle&issue=2&isSerivcesPage=true&spage=1&title=The+journal+of+real+estate+research.&dscnt=2&auinit=J&atitle=Portfolio+diversification+and+return+benefits%E2%80%94common+.

Kuhn, D. and Luenberger, D. (Feb. 2010). “Analysis of the rebalancing frequency in log-optimal portfolio selection”. In: *Quantitative Finance* 10.2, pp. 221–234. ISSN: 1469-7688. DOI: 10.1080/14697680802629400. URL: https://www.tandfonline.com/doi/full/10.1080/14697680802629400.

Lam, M. (2014). “Is portfolio re-balancing good for investors?” PhD thesis. University of Northern British Columbia, Thesis. URL: https://core.ac.uk/download/pdf/84873313.pdf.

Lanne, M. and Pentti, S. (Dec. 2007). “Modeling Conditional Skewness in Stock Returns”. In: *The European Journal of Finance* 13.8, pp. 691–704. ISSN: 1351-847X. DOI: 10.1080/13518470701538608. URL: http://www.tandfonline.com/doi/abs/10.1080/13518470701538608.

Laspeyres, K. (Jan. 1871). “IX. - The calculation of a medium commodity price increase.” In: *Jahrbücher für Nationalökonomie und Statistik* 16.1, pp. 296–318. ISSN: 2366-049X. DOI: 10.1515/jbnst-1871-0124. URL: https://www.degruyter.com/view/j/jbnst.1871.16.issue-1/jbnst-1871-0124/jbnst-1871-0124.xml.

Lawley, D. and Maxwell, A. (1971). *Factor Analysis as a Statistical Method.* Butterworths.

Lehmann, B. and Modest, D. (1987). “Mutual fund performance evaluation: a comparison of benchmarks and benchmark comparisons”. In: *The Journal of Finance* 42.2, pp. 233–265. URL: http://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1987.tb02566.x/full.

Lennep, D. van et al. (Sept. 2004). “Weather Derivatives”. In: *The Journal of Alternative Investments* 7.2, pp. 65–74. ISSN: 1520-3255. DOI: 10.3905/jai.2004.439650. URL: http://jai.iijournals.com/lookup/doi/10.3905/jai.2004.439650.

Levine, A. et al. (2016). “Commodities for the Long Run”. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.2856435. URL: https://www.ssrn.com/abstract=2856435.

Levy, H. and Levy, M. (2014). “The benefits of differential variance-based constraints in portfolio optimization.” In: *European Journal of Operational Research*, pp. 372–381.

Liang, B. (2003). “The Accuracy of Hedge Fund Returns”. In: *The Journal of Portfolio Management* 29.3, pp. 111–122. ISSN: 0095-4918. DOI: 10.3905/jpm.2003.319889.

Liang, Y. and McIntosh, W. (1998). “REIT Style and Performance.” In: *Journal of Real Estate Portfolio Management: 1998, Vol. 4, No. 1, pp. 69-78.*, pp. 69–78.

Lin, T. and Phalippou, L. (2007). *Estimating the performance and risk exposure of private equity funds: A new methodology*. Tech. rep. URL: https://www.researchgate.net/publication/228903045.

Lintnert, J. (1965). "Risk and Maximal Gains from Diversification". In: *The Journal of Financial Security Prices.* DOI: 10.1111/j.1540-6261.1965.tb02930.x. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1540-6261.1965.tb02930.x.

Lizieri, C. and Ward, C. (Jan. 2001). "The distribution of commercial real estate returns". In: *Return Distributions in Finance*, pp. 47–74. DOI: 10.1016/B978-075064751-9.50004-7. URL: https://www.sciencedirect.com/science/article/pii/B9780750647519500047.

Lobosco, A. (1999). "Style/risk-adjusted performance." In: *The Journal of Portfolio Management.* 25(3), pp.

Lu, K. and Mei, J. (1999). "The return distributions of property shares in emerging markets." In: *Journal of Real Estate Portfolio Management* 5(2), pp.1.

Maillard, S., Roncalli, T., and Teïletche, J. (2010). "The properties of equally weighted risk contribution portfolios." In: *The Journal of Portfolio Management* 36(4), pp.60-70.

Malkiel, B. (June 1995). "Returns from Investing in Equity Mutual Funds 1971 to 1991". In: *The Journal of Finance* 50.2, pp. 549–572. ISSN: 00221082. DOI: 10.1111/j.1540-6261.1995.tb04795.x. URL: http://doi.wiley.com/10.1111/j.1540-6261.1995.tb04795.x.

Markham, J. (1991). "Manipulation of commodity futures prices-the unprosecutable crime". In: *Yale journal on regulation.* Vol.8, p.2. URL: http://suprimo.lib.strath.ac.uk/primo_library/libweb/action/openurl?date=1991&aulast=Markham&isSerivcesPage=true&spage=281&dscnt=2&auinit=JW&title=Yale+journal+on+regulation.&atitle=Manipulation+of+commodity+futures+prices-the+unprosecutable+crime&url_ctx.

Markowitz, H. (1952). "Portfolio selection". In: *The Journal of Finance* 7.60, pp. 77–91. ISSN: 00221082. DOI: 10.1111/j.1540-6261.1952.tb01525.x.

Marshall, M (1996). “Sampling for qualitative research”. In: *Family Practice* 13.6, pp. 522–525. ISSN: 02632136. DOI: 10.1093/fampra/13.6.522.

Martellini, L. (2012). “Toward the design of better equity benchmarks: Rehabilitating the tangency portfolio from modern portfolio theory.” In: *The Journal of Portfolio Management* 34(4), pp.

Maxwell, K. and Saint-Pierre, P. (1998). “Benchmarking real estate investment performance: the application of real estate indices.” In: *Journal of Property Management* 63(3), pp.

McCarthy, D. and Spurgin, R. (1998). “A review of hedge fund performance benchmarks.” In: *The Journal of Alternative Investments.* 1(1), pp.1.

Meade, N. and Salkin, G. (1989). “Index funds-construction and performance measurement.” In: *Journal of the Operational Research Society* pp.871-879.

Mei, J. (2000). *Conditional Risk Premiums of Asian Real Estate Stocks*. Tech. rep. 3. URL: https://link.springer.com/content/pdf/10.1023/A:1012008004402.pdf.

Mei, J. and Lee, Y. (1994). “Is There a Real-Estate Factor Premium”. In: *Journal of Real Estate Finance and Economics* 9.2, pp. 113–126. ISSN: 0895-5638. DOI: 10.1007/bf01099970.

Mei, J. and Liu, C. (1994). “The Predictability of Real Estate Returns and Market Timing Part of the Finance and Financial Management Commons, and the Real Estate Commons The Predictability of Real Estate Returns and Market Timing”. In: *The Scholarly Commons*. URL: http://scholarship.sha.cornell.edu/articleshttp://scholarship.sha.cornell.edu/articles/299.

Meric, I., Ratner, M., and Meric, G. (Jan. 2008). “Co-movements of sector index returns in the world’s major stock markets in bull and bear markets: Portfolio diversification implications”. In: *International Review of Financial Analysis* 17.1, pp. 156–177. ISSN: 1057-5219. DOI: 10.1016/J.IRFA.2005.12.001. URL: https://www.sciencedirect.com/science/article/pii/S1057521906000147.

Miffre, J. (2012). “Comparing First, Second and Third Generation Commodity Indices”. In: *SSRN Electronic Journal*. ISSN: 1556-5068. DOI: 10.2139/ssrn.2132949. URL: http://www.ssrn.com/abstract=2132949.

Miffre, J. (2016). *Long-short commodity investing: A review of the literature*. DOI: 10.1016/j.jcomm.2016.01.001.

Mitchell, M. and Pulvino, T. (2001a). “Characteristics of risk and return in risk arbitrage.” In: *Journal of Finance* 56, 2135-2.

— (2001b). *No Title*.

Mongras, P. and Marchal-Domrat, C. (2006). “Commodities: an asset class in their own right?” In: *Banque de France, Financial Stability Review*. URL: www.banque-france.fr.

Morrison, F. (2016). “Setting the New Benchmark: EU Regulation on Financial Benchmarks Background”. In:

Mossin, J. (Oct. 1966). “Equilibrium in a Capital Asset Market”. In: *Econometrica* 34.4, p. 768. ISSN: 00129682. DOI: 10.2307/1910098. URL: https://www.jstor.org/stable/1910098?origin=crossref.

MSCI (2014). “MSCI Corporate Events Methodology”. In: May.

MSCI Inc. (2018). “MSCI Global Investable Market Indexes Methodology”. In: *Indexes* March. URL: http://www.msci.com/eqb/methodology/meth_docs/MSCI_Jan2015_GIMIMethodology_vf.pdf.

Mundlak, Y. (Jan. 1978). “On the Pooling of Time Series and Cross Section Data”. In: *Econometrica* 46.1, p. 69. ISSN: 00129682. DOI: 10.2307/1913646. URL: https://www.jstor.org/stable/1913646?origin=crossref.

Newell, G. and MacFarlane, J. (1996). “Risk estimation and appraisal-smoothing in UK property returns.” In: *Journal of property Research* 13 (1), 1–.

Newell, G. and Marzuki, M. (2016). “The significance and performance of UK-REITs in a mixed-asset portfolio”. In: *Journal of European Real Estate Research*. ISSN: 17539277. DOI: 10.1108/JERER-08-2015-0032.

Ng, S. and Ruge-Murcia, F. (2000). *Explaining the Persistence of Commodity Prices*. Tech. rep., pp. 149–171. URL: https://link.springer.com/content/pdf/10.1023/A:1008713823410.pdf.

Olmo, J. and Sanso-Navarro, M. (2012). “Forecasting the performance of hedge fund styles”. In: *Journal of Banking and Finance* 36.8, pp. 2351–2365. ISSN: 03784266. DOI: 10.1016/j.jbankfin.2012.04.016.

Pagourtzi, E. et al. (2003). “Real estate appraisal: a review of valuation methods.” In: *Journal of Property Investment & Finance* 21(4), pp.383-401.

Papanikolaou, D., Lo, K., and Wong, A. (2001). “No Title”. In: *Operations Research* Vol. 49, N.

Partovi, H and Caputo, M (2004). *Principal Portfolios: Recasting the Efficient Frontier*. Tech. rep.

Pasini, G. (2017). “A Principal Component Analysis for stock portfolio management”. In: *International Journal of Pure and Applied Mathematics* 115.1, pp. 153–167. ISSN: 1314-3395. DOI: 10.12732/ijpam.v115i1.12.

Pasquini, M. and Serva, M. (July 1999). “Multiscaling and clustering of volatility”. In: *Physica A: Statistical Mechanics and its Applications* 269.1, pp. 140–147. ISSN: 0378-4371. DOI: 10.1016/S0378-4371(99)00088-6. URL: https://www.sciencedirect.com/science/article/pii/S0378437199000886.

Pearson, K. (1901). “On lines and planes of closest fit to systems of points in space”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*. ISSN: 1941-5982. DOI: 10.1080/14786440109462720.

Pedersen, N., Page, S., and He, F. (2013). *Asset Allocation: Risk Models for Alternative Investments*. Tech. rep. 3. URL: www.cfapubs.org.

Pesaran, H. and Yamagata, T. (2017). “Testing for Alpha in Linear Factor Pricing Models with a Large Number of Securities”. In: URL: http://www.econ.cam.ac.uk/people-files/emeritus/mhp1/wp17/PY_LFPM_11_March_2017_Paper.pdf.

Petajisto, A. (Mar. 2011). “The index premium and its hidden cost for index funds”. In: *Journal of Empirical Finance* 18.2, pp. 271–288. ISSN: 0927-5398. DOI: 10.1016/J.JEMPFIN.2010.10.002. URL: https://www.sciencedirect.com/science/article/pii/S0927539810000745#f0005.

Polson, N. and Tew, B. (2000). “Bayesian portfolio selection: An empirical analysis of the S&P 500 index 1970–1996.” In: *Journal of Business & Economic Statistics* 18(2), pp.

Pope, P. and Yadav, P. (1994). “Discovering errors in tracking error.” In: *The Journal of Portfolio Management,* 20(2), pp.

Posthuma, N. and Van der Sluis, A. (2003). "A reality check on hedge funds returns." In: *Vrije Universiteit. SSRN 438840.*

Quan, D. and Titman, S. (May 1997). "Commercial Real Estate Prices and Stock Market Returns: An International Analysis". In: *Financial Analysts Journal* 53.3, pp. 21–34. ISSN: 0015-198X. DOI: 10.2469/faj.v53.n3.2082. URL: http://www.cfapubs.org/doi/full/10.2469/faj.v53.n3.2082.

Rallis, G., Miffre, J., and Fuertes, A. (2013). "Strategic and tactical roles of enhanced commodity indices". In: *Journal of Futures Markets*. ISSN: 02707314. DOI: 10.1002/fut.21571.

Ramsey, J. (1969). "Tests for specification errors in classical linear least-squares regression analysis." In: *Journal of the Royal Statistical Society. Series B (Methodological)* pp.350-371.

Rao, R. (1979). "Separation theorems for singular values of matrices and their applications in multivariate analysis". In: *Journal of Multivariate Analysis* 9.3, pp. 362–377. ISSN: 10957243. DOI: 10.1016/0047-259X(79)90094-0.

Reinganum, M. (1981). "Misspecification of capital asset pricing". In: *Journal of Financial Economics* 9.1, pp. 19–46. ISSN: 0304405X. DOI: 10.1016/0304-405X(81)90019-2.

Rennie, E. and Cowhey, T. (1990). "The Successful Use of Benchmark Portfolios: A Case Study". In: *Financial Analyst Journal*.

Ribeiro, R. and Hodges, D. (2004). "A Two-Factor Model for Commodity Prices and Futures Valuation *". In: *Unpublished paper*.

Roll, R. (1977). "A critique of the asset pricing theory's tests Part I: On past and potential testability of the theory". In: *Journal of Financial Economics* 4.2, pp. 129–176. ISSN: 0304405X. DOI: 10.1016/0304-405X(77)90009-5.

Roncalli, T. (2007). "An Alternative Approach to Alternative Beta". In: *Unpublished paper*.

— (2017). "Alternative Risk Premia: What Do We Know?" In: *Factor Investing*. Elsevier, pp. 227–264. DOI: 10.1016/B978-1-78548-201-4.50010-6. URL: http://linkinghub.elsevier.com/retrieve/pii/B9781785482014500106.

Rosenberg, B. and Marathe, V. (1975). *Tests of capital asset pricing hypotheses*. 32).

Ross, S. (1976). “The Arbitrage Theory of Capital Asset Pricing”. In: *Journal of Economic Theory* 13, pp. 341–360. ISSN: 00220531. DOI: 10.1016/0022-0531(76)90046-6. URL: http://economy.njau.edu.cn:8011/files/7/3/The%20Arbitrage%20Theory%20of%20Capital%20Asset%20Pricing.pdf.

Ross, S. and Zisler, R. (1991). “Risk and Return in Real Estate.” In: *Journal of Real Estate Finance and Economics* 4.2, pp. 175–190. ISSN: 08955638. URL: http://search.ebscohost.com/login.aspx?direct=true%7B&%7Ddb=bth%7B&%7DAN=17295379%7B&%7Dsite=ehost-live.

Sadka, R. (2010). “Liquidity risk and the cross-section of hedge-fund returns”. In: *Journal of Financial Economics* 98.1, pp. 54–71. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2010.05.001.

Sanders, D. and Irwin, S. (2012). “A reappraisal of investing in commodity futures markets”. In: *Applied Economic Perspectives and Policy*. ISSN: 20405790. DOI: 10.1093/aepp/pps026.

“schneeweis-spurgin - comparisons of commodity and managed futures benchmark indexes -1997” (1997). In:

Schneeweis, T. and Spurgin, R. (1997). “Survivor bias in commodity trading advisor performance”. In:

Schwann, G. (1998). “A Real Estate Price Index for Thin Markets”. In: *Journal of Real Estate Finance and Economics* 163, pp. 269–287. ISSN: 08955638. DOI: 10.1023/A:1007719513787.

Scott, L. (1990). “Do Prices Reflect Market Fundamentals in Real Estate Markets?” In: *Tile Journal of Real Estate Finance and Exonomics* 3, pp. 5–23. URL: https://link.springer.com/content/pdf/10.1007/BF00153703.pdf.

Sensoy, B. (2009). “Performance evaluation and self-designated benchmark indexes in the mutual fund industry”. In: *Journal of Financial Economics* 92.1, pp. 25–39. ISSN: 0304405X. DOI: 10.1016/j.jfineco.2008.02.011.

Sharpe, W. (1964). “Capital Asset Prices: A Theory of Market Equilibrium under conditions of risk.” In: *The journal of finance* 19.3, pp. 425–442. ISSN: 00221082. DOI: 10.2307/2329297. URL: http://www.jstor.org/stable/2329297?origin=

crossref%5Cnhttp://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1982.tb03586.x/abstract%5Cnhttp://onlinelibrary.wiley.com/doi/10.1111/j.1540-6261.1964.tb02865.x/full.

Sharpe, W. (1992). "Asset Allocatiom: Management Style and Performance Measurement". In: *Journal of Portfolio Management*, pp. 7–19. URL: http://www.stanford.edu/~wfsharpe/art/sa/sa.htm.

— (1994). "The Sharpe ratio". In: *The Journal of Portfolio Managementortfolio management* 21.1, pp. 49–58. URL: http://www.iijournals.com/doi/pdfplus/10.3905/jpm.1994.409501@jpm40.1994.21.issue-1.

Shlens, J. (2005). "A Tutorial on Principal Component Analysis". In: *Preprint arXiv:1404.1100* 51, p. 52. ISSN: 00219991. DOI: 10.1.1.115.3503. URL: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.115.3503&rep=rep1&type=pdf.

Siegel, L. (2003). "Benchmarks and investment management". In: *Research Foundation Publications* 2003.1, pp. 1–136. URL: http://www.cfapubs.org/doi/pdf/10.2470/rf.v2003.n1.3922.

Sirmans, C. and Worzala, E. (May 2003). "International Direct Real Estate Investment: A Review of the Literature". In: *Urban Studies* 40.5-6, pp. 1081–1114. ISSN: 0042-0980. DOI: 10.1080/0042098032000074335. URL: http://journals.sagepub.com/doi/10.1080/0042098032000074335.

Snoussi, I. and Hellara, S. (2004). "Performance Evaluation of Hedge Funds: A Literature Review of Theoretical and Empirical Issues and Contributions"". In: *Finance India* 18, p.587.

Sortino, F. and Van der Meer, R. (1991). "Downside risk." In: *The Journal of Portfolio Management* 17(4), pp.

Standard and Poors Dow Jones Indices. (2018). *Standard and Pooos GSCI Methodology*. Tech. rep. URL: https://www.spindices.com/documents/methodologies/methodology-sp-gsci.pdf.

Statman, M. (2006). "Socially Responsible Indexes". In: *The Journal of Portfolio Management* 32.3, pp. 100–109. ISSN: 0095-4918. DOI: 10.3905/jpm.2006.628411.

Stevenson, A. (2011). *Oxford Dictionary of English.* 3rd ed. Oxford University Press. ISBN: 0199571120.

Stigler, G. (Aug. 1950). "The Development of Utility Theory." In: *Journal of Political Economy* 58.4, pp. 307–327. ISSN: 0022-3808. DOI: 10.1086/256962. URL: https://www.journals.uchicago.edu/doi/10.1086/256962.

STOXX (2013). "The benchmark industry An introduction and outlook". In: *STOXX-Deutshe Borse White Paper*.

Su, Y., Lau, M., and Chau, F. (Mar. 2013). *Alternative Investments: Commodity Futures and Strategic Assset Allocation*. Ed. by H. Kent Baker and Greg Filbeck. Hoboken, NJ, USA: John Wiley & Sons, Inc., Chapter 20. ISBN: 9781118656501. DOI: 10.1002/9781118656501. URL: http://doi.wiley.com/10.1002/9781118656501.

Szymanowska, M. et al. (2014). "An Anatomy of Commodity Futures Risk Premia; An Anatomy of Commodity Futures Risk Premia". In: *The Journal of Finance* LXIX.1. DOI: 10.1111/jofi.12096. URL: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jofi.12096.

Tang, K. and Xiong, W. (2012). "Index investment and the financialization of commodities". In: *Financial Analysts Journal*. ISSN: 0015198X. DOI: 10.2469/faj.v68.n6.5.

Terhaar, K., Staub, R., and Singer, B. (2003). "Appropriate policy allocation for alternative investments." In: *The Journal of Portfolio Management,* 29(3), pp.

Till, H. and Gunzberg, J. (2005). "Absolute returns in commodity (natural resource) futures investments." In: *Hedge Fund Investment Management*, p. 25.

Titman, S. and Tiu, C. (2011). "Do the best hedge funds hedge?" In: *Review of Financial Studies* 24(1), pp.

Titman, S. and Warga, A. (Sept. 1986). "Risk and the Performance of Real Estate Investment Trusts: A Multiple Index Approach". In: *Real Estate Economics* 14.3, pp. 414–431. ISSN: 1080-8620. DOI: 10.1111/1540-6229.00395. URL: http://doi.wiley.com/10.1111/1540-6229.00395.

Tobin, M. (1958). "Liquidity preference as behavior towards risk". In: *The Review of Economic Studies*, p. 25.

Treynor, J. (1965). “How to rate management of investment funds.” In: *Harvard business review* 43(1), pp.

Treynor, J. and Black, F. (1973). “How to use security analysis to improve portfolio selection.” In: *The Journal of Business* 46(1), pp.

Tudor, D. and Cao, B. (Feb. 2012). “The absolute returns of hedge funds”. In: *Managerial Finance* 38.3, pp. 280–302. ISSN: 17587743. DOI: 10.1108/03074351211201424.

Viole, F. and Nawrocki, D. (June 2013). “An analysis of heterogeneous utility benchmarks in a zero return environment”. In: *International Review of Financial Analysis* 28, pp. 190–198. ISSN: 1057-5219. DOI: 10.1016/J.IRFA.2013.02.014. URL: https://www.sciencedirect.com/science/article/pii/S1057521913000318.

Wallerstein, E., Tuchschmid, N., and Zaker, S. (2010). “How do hedge fund clones manage the real world?” In: *Journal of Alternative Investments*. ISSN: 15203255. DOI: 10.3905/JAI.2010.12.3.037.

Walsh, C. (1901). “The measurement of general exchange-value”. In: URL: https://scholar.google.co.uk/scholar?hl=en&as_sdt=0%2C5&q=walsh+1901&btnG=#d=gs_cit&p=&u=%2Fscholar%3Fq%3Dinfo%3AQog0Aw_i0RUJ%3Ascholar.google.com%2F%26output%3Dcite%26scirp%3D0%26hl%3Den.

Wang, Z. (1998). “Efficiency loss and constraints on portfolio holdings1.” In: *Journal of Financial Economics*, pp. 359–375.

Waring, B. and Siegel, L. (Mar. 2006). “The Myth of the Absolute-Return Investor”. In: *Financial Analysts Journal* 62.2, pp. 14–21. ISSN: 0015-198X. DOI: 10.2469/faj.v62.n2.4080. URL: http://www.cfapubs.org/doi/full/10.2469/faj.v62.n2.4080.

Wilson, J. and Jones, C. (2002). “An analysis of the S&P 500 index and Cowles’s extensions: Price indexes and stock returns, 1870–1999.” In: *The Journal of Business* 75(3), pp.

Wold, S. (1978). *Cross-validatory estimation of the number of components in factor and principal components models*. - [s. l.]: Technometrics, pp. 397–405.

Wyngarden, H. (1927). “An Index of Local Real Estate Prices.” In: *University of Michigan, School of business administration, Bureau of business research.* URL: https://www.

philadelphiafed.org/-/media/research-and-data/publications/working-papers/1998/wp98-21r.pdf.

Yadav, P., Pope, P., and Paudyal, K. (Apr. 1994). "Threshold Autoregressive Modelling in Finance: The Price differences of equivalent assets". In: *Mathematical Finance* 4.2, pp. 205–221. DOI: 10.1111/j.1467-9965.1994.tb00058.x. URL: http://doi.wiley.com/10.1111/j.1467-9965.1994.tb00058.x.

Yang, L., Rea, W., and Rea, A. (2015). *Identifying Highly Correlated Stocks Using the Last Few Principal Components*. Tech. rep. URL: http://www.sirca.org.au/.

Yasin, M. (2002). "The theory and practice of benchmarking: then and now."". In: *Benchmarking International* 9(3), pp.2.

Yin, L. and Han, L. (2015). "Co-movements in commodity prices: Global, sectoral and commodity-specific factors". In: *Economics Letters* 126, pp. 96–100. DOI: 10.1016/j.econlet.2014.11.027. URL: http://dx.doi.org/10.1016/j.econlet.2014.11.027.

Young, M. and Graff, R. (1995). *Real Estate Is Not Normal: A Fresh Look at Real Estate Return Distributions*. Tech. rep., pp. 225–259. URL: https://link.springer.com/content/pdf/10.1007/BF01096940.pdf.

Zaremba, A. (2015). "Is financialization killing commodity investments?" In: *Journal of Alternative Investments* 18.1, pp. 61–89. ISSN: 15203255. DOI: 10.3905/jai.2015.18.1.066.

Appendices

Appendix A

Alternative Asset Indices

Table A.1: The key alternative asset indices

Index name	Asset	Float adj	Weighting	Rebalanced
Dow Jones AIG CI	Commodity	No	Liq + Pro	Annual
Morningstar CI	Commodity	No	Mag + Mom	Annual
S&P GSCI	Commodity	No	Pro	Monthly
GSCI	Commodity	No	Pro	Monthly
Reuters-CBR	Commodity	No	Equal	Monthly
DJ Hedge Fund	Hedge Fund	No	Cap	Monthly
FTSE Hedge Fund	Hedge Fund	No	Inv	Annual
CASAM Hedge	Hedge Fund	No	Equal	Monthly
CSFB Tremont	Hedge Fund	No	Cap	Quarterly
Hennesse Group	Hedge Fund	No	Equal	Monthly
HFR	Hedge Fund	No	Equal	Monthly
MSCI Hedge	Hedge Fund	No	Med	Quarterly
Van Hedge	Hedge Fund	No	Equal	Monthly
DJ Whilshire	REIT	Yes	Cap	Quarterly
FTSE NAREIT	REIT	Yes	Cap	Quarterly
MSCI US REIT	REIT	Yes	Cap	Semi Annual
S&P Citi Global	REIT	Yes	Cap	Quarterly

Table showing indices from the major commercial providers: Dow Jones, (DJ) Credit Suisse (CSFB), Morningstar, Goldman Sachs (GSCI) Standard and Poors (S&p), Finnacila Times (FTSE). Column breakdown shows index abbreviation, Type of alternative asset, whether it is float adjusted, its weighting method [Liq = Liquid, Pro = Production, Mag = Magnitude, Mom = Momentum, Inv = Investibility, Equ = Equal, Cap = Capitalization]and frequency of rebalancing.

Appendix B

Matlab code for GRS test

B.1 Function [grs,pgrs];

Runs the mean-variance efficiency tests of Gibbons et al(1989)

Input: r is a (T,N) matrix of index excess returns; F is a (T,K) matrix of K PCA factor index excess returns

Output: grs is the Gibbons et al(1989) F test and pgrs is the p value;

Aar is $A|alpha|, AarisA|a|/A|r|$, and Aase is

$Ase(alpha)^2/A(alpha^2)$.

B.2 Code

r ='A2:A122'; *cells based on file imported* Rf='B2:B122'; F='C2:c122';

```
[T,N]=size(r);
[K]=size(F,2);
lhv=r;
rhv=[ones(T,1),F];
[bv,sebv,R2,R2adj]=lists(lhv,rhv);
alpha=bv(1,:)';
sealpha=sebv(1,:)';
talpha=alpha./sealpha;
```

```
resids=lhv-rhv*bv;
Ve=cov(resids,1);
uf=mean(F)';
Vf=cov(F,1);
grs=[(T-N-K)/N]*[(alpha'*inv(Ve)*alpha)/(1+uf'*inv(Vf)*uf)];
pgrs=1-cdf('f',grs,N,T-N-K);
```
To get results
```
disp[grs,pgrs]
```

Appendix C

Hedge fund stress scenarios

Burst of the "Dot com bubble" 10/03/2000 – 10/04/2000: This was an equity event that impacted technology, telecommunications and media stocks. These were heavily invested by hedge funds that employed momentum strategies. These include event driven and directional funds. Developed markets had more of these types of stocks than emerging markets.

The end of equity bear market 09/10/2002 – 09/11/2002: This was a severe bear market that saw broad equity prices decline by 30 percent. As such, long bias hedge funds would have been disadvantaged and short bias hedge funds advantaged.

The first Gulf war The first 01/03/2003 – 01/04/2003: The first Gulf war which started on 20/3/2003 was widely discounted by markets in the two weeks prior to the outbreak. The impact was mostly felt on the crude oil price. This impacted those hedge funds with commodity focused trading strategies the most.

The Russia Financial Crisis between 07/08/2008 – 06/10/2008: The Russia Financial Crisis between 07/08/2008 – 06/10/2008 saw credit spreads widen in unprecedented magnitude. This hurt fixed income funds and in particular credit, arbitrage or event focused hedge funds.

The Lehman Default 15/9/2008 – 14/10/2008: The Lehman Default signalled the credit crisis of 2008. It unfolded from 15/9/2008 – 14/10/2008. This impacted all hedge fund strategies, in particular highlighting illiquid instruments and counterparty risk.

Equity rebound 4/3/2009 – 1/6/2009: The implementation of monetary policy initiatives to ease the credit crisis resulted in an Equity market rebound of 20 per cent in 2009 between 4/3/2009 – 1/6/2009. Hedge fund managers were most pessimistic and had begun to reduce risk. As such, all strategies were not as exposed to the upside as they could have been. Long bias strategies benefited and short bias strategies were harmed.

Oil slump 3/5/2010 – 20/5/2010: The price of oil dropped 20per cent due to concerns that European nations would reduce their budget deficits and thereby slow their economies 3/5/2010 – 20/5/2010. This impacted both macro hedge fund strategies and commodity strategies.

The US DEBT Ceiling crisis - 22/7/2010 – 8/8/2011: The US debt ceiling is a statutory maximum the federal government is allowed to borrow. A political impasse occurred where the Republicans refused to raise the ceiling. This led to the Debt Ceiling crisis 22/7/2010 – 8/8/2011 where potentially the US AAA rating was at risk. Sovereign government bonds therefore performed extremely poorly during this period. This impacted fixed income strategies.

Libyan civil war 14/2/2011 – 23/2/2011: The 17th February Revolution in 2011 started a civil war in Libya between forces loyal to Colonel Gadaffi and the opposition to his rule. The period 14/2/2011 – 23/2/2011 saw oil price volatility reminiscent of the Gulf war. This impacted commodity based hedge fund strategies.

Japan earthquake 10/3/2011 – 15/3/2011: An 9.1 earthquake hit Japan on 11 March, 2011. The earthquake impacted Far Eastern equities, particularly Japan due to the damage done to nuclear facilities 10/3/2011 – 15/3/2011 which stopped 6.8per cent of GDP production in the country. Hedge funds with exposure to Far Eastern Markets were impacted negatively.

Greek Crisis - 14/2/2011 – 23/2/2011: In the aftermath of the credit crisis, Greece faced a sovereign crisis because due to structural issues it could no longer service its debt. The Greek Crisis 22/6/2015 – 8/7/2015 caused a widening of bind spreads, particularly in peripheral European countries. This impacted financial equities, and high yield fixed income hedge fund strategies.

Yuan devaluation 01/08/2015 – 01/09/2015: The Chinese economy was the fastest growing in the world when the government decided to devalue the Yuan. The absolute devaluation was modest but markets took this to be a signal that more devaluations were to come. This impacted equity markets and commodity markets negatively during the period 01/08/2015 – 01/09/2015. Emerging market, commodity related and directional hedge funds suffered as a result.

Appendix D

Correlation tables - Hedge fund

Table D.1: PCA Index correlation table - The Russia Financial Crisis between 07/08/2008 – 06/10/2008

USA	NA USD	WORLD EX	MSCI EM	US GOVT	Global Gov	EURO/USD	USD	Gold	PCA	
1										
0.997743892	1									
0.513831492	0.547624573	1								
0.593008625	0.623999237	0.910645797	1							
-0.753826053	-0.775468972	-0.673032915	-0.726755951	1						
-0.556778223	-0.58131103	-0.776440847	-0.791295352	0.727442619	1					
0.377272545	0.36406931	-0.07755167	-0.018333196	-0.153826736	-0.185172582	1				
0.444045155	0.428551143	0.20196278	0.196409831	-0.256109316	-0.143909632	0.025544736	1			
0.057959897	0.015742972	-0.405898185	-0.253884116	0.156783496	0.372391236	-0.007509563	0.104788245	1		
-0.665711907	-0.672703707	-0.419636806	-0.499541682	0.792000408	0.604350725	-0.484772203	-0.047935781	0.093551532	1	
-0.214485917	-0.209526118	0.085517009	0.06433098	-0.058217656	-0.137344872	-0.237678187	0.090268235	-0.090555108	0.089842742	1

This table shows the correlation between the MSCI USA MSCI NORTH AMERICA USD MSCI WORLD EX US USD MSCI EM USD JPM UNITED STATES GOVT.BOND USD ICE BofAML Global Government Excluding the US Index US EUROUSD DEP 1 MONTH (BID,LDN) (D) Gold spot LDN 22.00 hthis researchs Euro/USD US NOMINAL DOLLAR MAJOR CURR INDEX. (C) PCA Index. All rows on X axis correspond to rows on Y axis above.

Table D.2: PCA Index correlation table - The Lehman Default 15/9/2008 – 14/10/2008

USA	NA USD	WORLD EX	MSCI EM	US GOVT	Global Gov	EURO/USD	USD	Gold	EuroUSD	
1										
0.997661923 1										
0.532790442	0.571048991	1								
0.53556683	0.565033784	0.905385588	1							
-0.426540416	-0.462140436	-0.488369027	-0.59315733	1						
-0.358770436	-0.375529734	-0.512976707	-0.710997101	0.707668395	1					
0.358874982	0.353685175	-0.082139493	-0.104246772	-0.085688776	-0.067937637	1				
0.541888723	0.517875452	0.105524422	0.264706125	-0.240924233	-0.343889934	0.044062716	1			
-0.089163581	-0.135311496	-0.618830199	-0.411557789	0.208175738	0.268023056	-0.03331751	0.203826812	1		
-0.342310732	-0.352771595	-0.14584767	-0.30557541	0.778730894	0.623186967	-0.377102486	-0.106602947	-0.060189703	1	
-0.437163447	-0.397960647	0.024474855	0.061750995	-0.238720031	-0.128333301	-0.142191513	-0.156213858	-0.202238188	-0.028684528	1

This table shows the correlation between the MSCI USA MSCI NORTH AMERICA USD MSCI WORLD EX US USD MSCI EM USD JPM UNITED STATES GOVT.BOND USD ICE BofAML Global Government Excluding the US Index US EUROUSD DEP 1 MONTH (BID,LDN) (D) Gold spot LDN 22.00 hthis researchs Euro/USD US NOMINAL DOLLAR MAJOR CURR INDEX PCA Index. All rows on X axis correspond to rows on Y axis above.

Table D.3: PCA Index correlation table - Equity rebound 4/3/2009 – 1/6/2009

USA	NA USD	WORLD EX	MSCI EM	US GOVT	Global Gov	EURO/USD	USD	Gold	EuroUSD	
1										
0.998823403	1									
0.581817478	0.608651328	1								
0.632494289	0.652307409	0.850863716	1							
-0.208971409	-0.209511903	-0.22678451	-0.241447232	1						
-0.291402269	-0.300038174	-0.409070071	-0.394744386	0.349213028	1					
-0.039336979	-0.036210052	0.044863473	0.066287266	-0.003447697	0.084866464	1				
0.495229327	0.48916513	0.38216342	0.358512652	-0.378566268	-0.451916914	0.024350866	1			
-0.218828206	-0.247101433	-0.699031703	-0.485201055	0.020726126	-0.01358687	0.031884941	-0.138716768	1		
-0.190380083	-0.18583835	-0.106981413	-0.132645906	0.913448048	0.248227278	0.042487796	-0.329221778	-0.027965044	1	
0.149528761	0.143838456	0.047180267	0.022879821	-0.128559498	-0.011263769	0.017834185	0.113552726	0.056656856	-0.095620464	1

This table shows the correlation between the MSCI USA MSCI NORTH AMERICA USD MSCI WORLD EX US USD MSCI EM USD JPM UNITED STATES GOVT.BOND USD ICE BofAML Global Government Excluding the US Index US EUROUSD DEP 1 MONTH (BID,LDN) (D) Gold spot LDN 22.00 hthis researchs Euro/USD US NOMINAL DOLLAR MAJOR CURR INDEX PCA Index.. All rows on X axis correspond to rows on Y axis above.

Table D.4: Oil slump 3/5/2010 – 20/5/2010

USA	NA USD	WORLD EX	MSCI EM	US GOVT	Global Gov	EURO/USD	USD	Gold	EuroUSD	
1										
0.99855485	1									
0.784673845	0.804086496	1								
0.752440745	0.777903081	0.934964528	1							
-0.794459234	-0.777224383	-0.515453178	-0.47056186	1						
0.523274258	0.531376772	0.753178943	0.656268549	-0.233144588	1					
-0.175742851	-0.187032559	-0.320365913	-0.206516842	-0.258416724	-0.529360339	1				
0.17553731	0.137854704	0.021518088	-0.115850851	-0.45822571	0.265973569	0.178926734	1			
-0.449770955	-0.462831224	-0.743352924	-0.653056373	0.537900671	-0.474746223	-0.192731771	-0.163093355	1		
-0.396540557	-0.372374046	-0.052738349	0.014763869	0.794978378	0.222788698	-0.618070127	-0.44285916	0.359114741	1	
-0.228526785	-0.244822553	-0.327804344	-0.304940571	0.126788718	-0.118788928	0.143442829	-0.005296842	0.241836132	-0.141220166	1

This table shows the correlation between the MSCI USA MSCI NORTH AMERICA USD MSCI WORLD EX US USD MSCI EM USD JPM UNITED STATES GOVT.BOND USD ICE BofAML Global Government Excluding the US Index US EUROUSD DEP 1 MONTH (BID,LDN) (D) Gold spot LDN 22.00 hthis researchs Euro/USD US NOMINAL DOLLAR MAJOR CURR INDEX PCA Index.. All rows on X axis correspond to rows on Y axis above.

Table D.5: PCA Index correlation table - The US DEBT Ceiling crisis - 22/7/2010 – 8/8/2011

USA	NA USD	WORLD EX	MSCI EM	US GOVT	Global Gov	EURO/USD	USD	Gold	EuroUSD	
1										
0.997504981	1									
0.692542347	0.718952991	1								
0.638039849	0.660939876	0.81388893	1							
-0.492064402	-0.487973073	-0.276370105	-0.242411125	1						
-0.382783983	-0.381695783	-0.319250092	-0.286675562	0.439411604	1					
-0.014309712	-0.01099072	-0.006722843	-0.016441457	-0.032464117	0.072622304	1				
0.332755099	0.311645262	0.258403673	0.180547031	-0.274720845	-0.224467326	-0.009328538	1			
-0.436699312	-0.463276732	-0.755097414	-0.519924581	0.103810743	0.003733442	0.003898306	-0.230151304	1		
-0.362649378	-0.355591788	-0.111624715	-0.061815948	0.961908518	0.375275869	-0.029449127	-0.225009993	0.01297095	1	
-0.139819104	-0.133214957	-0.080783386	-0.084650455	0.17290848	0.041052219	0.025339988	-0.071161209	0.010544009	0.157227015	1

This table shows the correlation between the MSCI USA MSCI NORTH AMERICA USD MSCI WORLD EX US USD MSCI EM USD JPM UNITED STATES GOVT.BOND USD ICE BofAML Global Government Excluding the US Index US EUROUSD DEP 1 MONTH (BID,LDN) (D) Gold spot LDN 22.00 hthis researchs Euro/USD US NOMINAL DOLLAR MAJOR CURR INDEX PCA Index. All rows on X axis correspond to rows on Y axis above.

Table D.6: Libyan civil war 14/2/2011 – 23/2/2011

USA	NA USD	WORLD EX	MSCI EM	US GOVT	Global Gov	EURO/USD	USD	Gold	EuroUSD	
1										
0.998641615	1									
0.765048104	0.785176197	1								
0.878127138	0.881073479	0.733126967	1							
-0.691805005	-0.662026639	-0.270182802	-0.560054434	1						
-0.462467561	-0.45969207	-0.37474355	-0.750761245	0.479163888	1					
0.206795271	0.230841517	0.381958028	0.251792113	0.360359385	-0.130174127	1				
-0.234836444	-0.232207392	0.095862414	-0.203573211	0.222119159	-0.098432064	-0.059967651	1			
-0.597525773	-0.587511543	-0.615550638	-0.410107165	0.44316067	0.003462324	0.272735269	-0.272947743	1		
-0.593785564	-0.560008057	-0.143710979	-0.484311995	0.984614314	0.491695216	0.443183426	0.137729497	0.398721555	1	
0.19939777	0.216169093	0.519741951	0.058155803	0.283554992	0.497307119	0.168185716	-0.02611003	-0.581670773	0.394238831	1

This table shows the correlation between the MSCI USA MSCI NORTH AMERICA USD MSCI WORLD EX US USD MSCI EM USD JPM UNITED STATES GOVT.BOND USD ICE BofAML Global Government Excluding the US Index US EUROUSD DEP 1 MONTH (BID,LDN) (D) Gold spot LDN 22.00 hthis researchs Euro/USD US NOMINAL DOLLAR MAJOR CURR INDEX PCA Index.. All rows on X axis correspond to rows on Y axis above

Table D.7: PCA Index correlation table - Japan earthquake 10/3/2011 – 15/3/2011

USA	NA USD	WORLD EX	MSCI EM	US GOVT	Global Gov	EURO/USD	USD	Gold	EuroUSD
1									
0.999279755	1								
0.69470222	0.686938707	1							
0.487994463	0.511878176	0.629292069	1						
-0.818775052	-0.814571933	-0.194473878	-0.001899931	1					
0.962537734	0.956261096	0.854756054	0.511247269	-0.673166613	1				
-0.131573397	-0.137255968	-0.741445219	-0.744230022	-0.461350332	-0.327338812	1			
-0.910058471	-0.924663002	-0.624850953	-0.760658108	0.650007163	-0.839658883	0.279294365	1		
-0.565286237	-0.558020046	0.130750025	0.318434066	0.935728666	-0.387397289	-0.741527379	0.360962026	1	
0.714236828	0.733355349	0.723981631	0.959415032	-0.268664011	0.718710054	-0.634467674	-0.905373261	0.069146774	1

This table shows the correlation between the MSCI USA MSCI NORTH AMERICA USD MSCI WORLD EX US USD MSCI EM USD JPM UNITED STATES GOVT.BOND USD ICE BofAML Global Government Excluding the US Index US EUROUSD DEP 1 MONTH (BID,LDN) (D) Gold spot LDN 22.00 hthis researchs Euro/USD US NOMINAL DOLLAR MAJOR CURR INDEX PCA Index. All rows on X axis correspond to rows on Y axis above

Table D.8: Greek Crisis - 14/2/2011 – 23/2/2011

USA	NA USD	WORLD EX	MSCI EM	US GOVT	Global Gov	EURO/USD	USD	Gold	EuroUSD	
1										
0.998641615	1									
0.765048104	0.785176197	1								
0.878127138	0.881073479	0.733126967	1							
-0.691805005	-0.662026639	-0.270182802	-0.560054434	1						
-0.462467561	-0.45969207	-0.37474355	-0.750761245	0.479163888	1					
0.206795271	0.230841517	0.381958028	0.251792113	0.360359385	-0.130174127	1				
-0.234836444	-0.232207392	0.095862414	-0.203573211	0.222119159	-0.098432064	-0.059967651	1			
-0.597525773	-0.587511543	-0.615550638	-0.410107165	0.44316067	0.003462324	0.272735269	-0.272947743	1		
-0.593785564	-0.560008057	-0.143710979	-0.484311995	0.984614314	0.491695216	0.443183426	0.137729497	0.398721555	1	
0.19939777	0.216169093	0.519741951	0.058155803	0.283554992	0.497307119	0.168185716	-0.02611003	-0.581670773	0.394238831	1

This table shows the correlation between the MSCI USA MSCI NORTH AMERICA USD MSCI WORLD EX US USD MSCI EM USD JPM UNITED STATES GOVT.BOND USD ICE BofAML Global Government Excluding the US Index US EUROUSD DEP 1 MONTH (BID,LDN) (D) Gold spot LDN 22.00 hthis researchs Euro/USD US NOMINAL DOLLAR MAJOR CURR INDEX PCA Index. All rows on X axis correspond to rows on Y axis above.

Appendix E

Correlation - REIT’s

Table E.1: PCA Index correlation table - REIT Sample 01/06/2008 - 01/06/2018

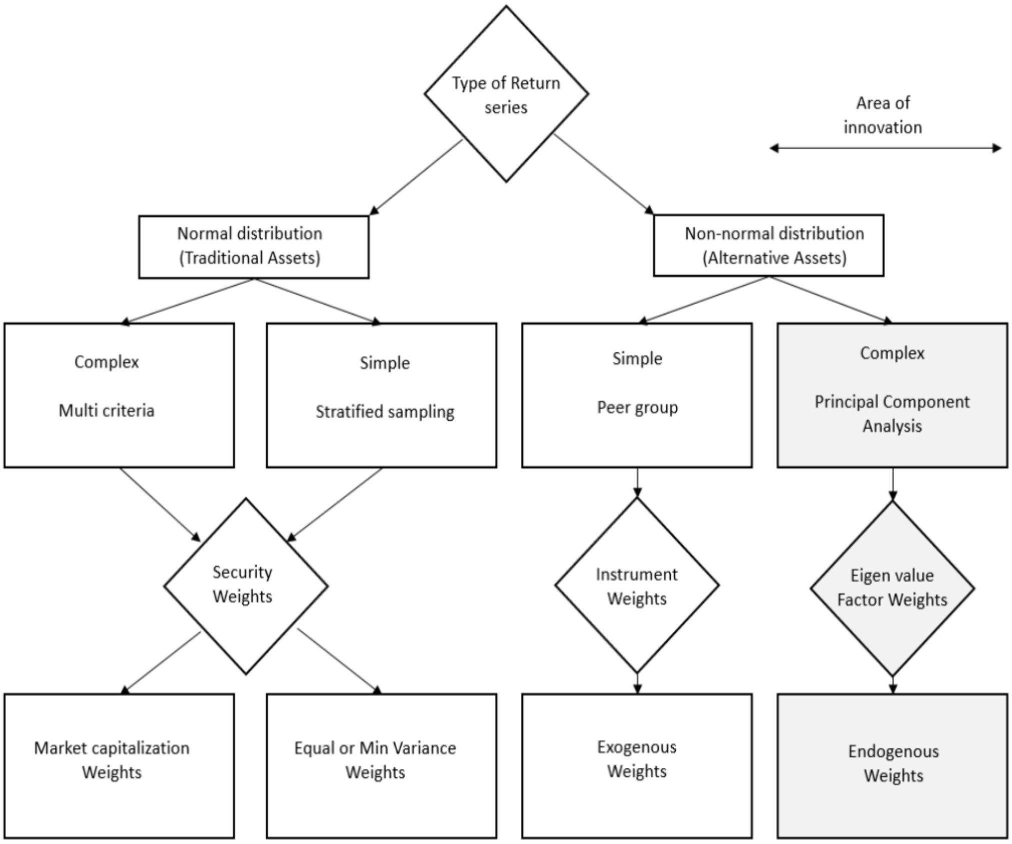
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Correlation	1.00	0.18	0.59	0.16	0.29	0.22	0.23	0.31	0.26	0.14	0.18	0.21	0.24	0.67	0.21	
	REIT2	0.18	1.00	0.34	0.81	0.26	0.84	0.50	0.80	0.83	0.73	0.72	0.82	0.88	0.21	0.80
	REIT3	0.59	0.34	1.00	0.24	0.28	0.34	0.22	0.30	0.30	0.27	0.32	0.30	0.33	0.43	0.30
	REIT4	0.16	0.81	0.24	1.00	0.22	0.76	0.43	0.74	0.73	0.72	0.51	0.79	0.77	0.17	0.76
	REIT5	0.29	0.26	0.28	0.22	1.00	0.25	0.26	0.24	0.31	0.12	0.20	0.21	0.28	0.53	0.21
	REIT6	0.22	0.84	0.34	0.76	0.25	1.00	0.55	0.76	0.77	0.68	0.67	0.80	0.80	0.22	0.76
	REIT7	0.23	0.50	0.22	0.43	0.26	0.55	1.00	0.54	0.55	0.47	0.43	0.41	0.55	0.17	0.39
	REIT8	0.31	0.80	0.30	0.74	0.24	0.76	0.54	1.00	0.91	0.68	0.63	0.71	0.81	0.27	0.71
	REIT9	0.26	0.83	0.30	0.73	0.31	0.77	0.55	0.91	1.00	0.68	0.68	0.71	0.83	0.28	0.69
	REIT10	0.14	0.73	0.27	0.72	0.12	0.68	0.47	0.68	0.68	1.00	0.48	0.80	0.72	0.14	0.75
	REIT11	0.18	0.72	0.32	0.51	0.20	0.67	0.43	0.63	0.68	0.48	1.00	0.56	0.64	0.18	0.51
	REIT12	0.21	0.82	0.30	0.79	0.21	0.80	0.41	0.71	0.71	0.80	0.56	1.00	0.74	0.22	0.77
	REIT10	0.24	0.88	0.33	0.77	0.28	0.80	0.55	0.81	0.83	0.72	0.64	0.74	1.00	0.24	0.75
	REIT14	0.67	0.21	0.43	0.17	0.53	0.22	0.17	0.27	0.28	0.14	0.18	0.22	0.24	1.00	0.20
	REIT15	0.21	0.80	0.30	0.76	0.21	0.76	0.39	0.71	0.69	0.75	0.51	0.77	0.75	0.20	1.00
Sig. (1-tailed)		0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	
	REIT2	0.01		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	REIT3	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	REIT4	0.01	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	REIT5	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00	0.00
	REIT6	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	REIT7	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00
	REIT8	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00	0.00
	REIT9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00	0.00	0.00
	REIT10	0.02	0.00	0.00	0.00	0.05	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.02	0.00
	REIT11	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00	0.00
	REIT12	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00	0.00
	REIT13	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00		0.00	0.00
	REIT14	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.02	0.00	0.00	0.00		0.00
	REIT15	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	

This table presents a correlation matrix of listed REIT's from the United States between 01/01/2001 - 01/01/2018.

Appendix F

Thesis contribution in context

Figure F.1: Illustration of the innovation proposed by the PCA index method



This figure shows the contribution that this method makes to the field of index construction. The grey shaded boxes illustrate the PCA Index method. The unshaded boxes show traditional methods of index construction technique.