

# CHOOSING AN INDEX FOR THE US RESIDENTIAL REAL ESTATE CAPITAL GUARANTEED INDEX-LINKED NOTE – AN EMPIRICAL STUDY

Kansantaloustiede  
Maisterin tutkinnon tutkielma  
Kaisa Kivipelto  
2009

**Kansantaloustieteen laitos**  
HELSINGIN KAUPPAKORKEAKOULU  
HELSINKI SCHOOL OF ECONOMICS



HELSINKI SCHOOL OF ECONOMICS  
Department of Economics



**CHOOSING AN INDEX FOR THE US RESIDENTIAL REAL ESTATE CAPITAL  
GUARANTEED INDEX-LINKED NOTE – AN EMPIRICAL STUDY**

Master's Thesis in Economics  
Kaisa Hannele Kivipelto  
Spring term 2009

Approved by the Head of the Economics Department \_\_\_/\_\_\_ 200\_\_\_ and awarded  
the grade \_\_\_\_\_

## ABSTRACT

The main goal of this thesis was to determine what is a good index. Through a theoretical discussion, empirical analysis and consideration of practical aspects, the best index for a case study structured product, the US Residential Real Estate Capital Guaranteed Index-Linked Note, was chosen. The main motivation behind this study was that it is essential to understand how the construction and characteristics of an index affect the performance of the index, especially if one is creating or purchasing an instrument whose return depends on the index. A case study was chosen to give this paper a practical application, and the chosen case study was a capital guaranteed index-linked note because of their unique return possibilities. The research concentrates mainly on choosing the index, but also describes how the case study product is constructed.

The thesis included two research questions. The first research question was how is an index constructed. First, indices were examined from a theoretical point of view via studying index theories and index formulae. The material used was academic literature on index theory, as well as material published by financial institutions. The conclusion of this discussion was that there is no unified opinion upon which is the most accurate or suitable index formula for this case study. The research then took an empirical approach. The two candidates for the index-linked product were the Dow Jones U.S. Real Estate Index and the Philadelphia Housing Sector Index. The Dow Jones Index was more suitable from a practical point of view, as the existence of an exchange traded fund linked to this index makes the options market more liquid. On the other hand, the contents of the Philadelphia Housing Sector index suit the product better – the Philadelphia Index measures the performance of residential real estate sector while the Dow Jones index contains only commercial real estate.

The empirical study utilised time series data acquired from the Bloomberg database. Because of the better availability of options for the Dow Jones Index, it was chosen for closer statistical testing. A regression was run using EViews comparing the monthly returns of the Dow Jones Index and the S&P/Case-Shiller Home Price Composite-20 Index, which was chosen to represent the return development of the housing market. The Dow Jones Index did in fact have a strong positive correlation with the S&P/Case-Shiller Composite index, and thus it was selected as the best index for the case study product. The empirical study, nonetheless, requires further statistical testing to solidify the results.

The second research question was how is a capital guaranteed index-linked note constructed. To answer this question, pricing aspects, different types of structured products and investor and issuer goals were discussed. The study utilised some academic references, marketing materials from financial institutions and an interview from a structured product professional. The construction of the case study was kept to a basic form, consisting of a zero coupon bond and a long position in American call options on the underlying, the iShares Dow Jones U.S. Real Estate Index Exchange Traded Fund. The final conclusions were that even though the projected returns seem promising, the current low interest rate, high volatility, low confidence in financial institutions environment makes the timing poor for issuing the case study product.

**Key words:** Index theory, Laspeyres index, Irving Fisher, structured product, index-linked note, capital guarantee, stock market index, S&P/Case-Shiller home price index, real estate, the United States, derivative, zero coupon bond

**TABLE OF CONTENTS**

1	INTRODUCTION.....	1
1.1	Goals and motivation for the study .....	1
1.1.1	Background.....	1
1.1.2	Motivation.....	1
1.1.3	The goals of the research.....	3
1.2	Structure .....	4
1.3	Research questions and methods .....	5
1.3.1	How is an index constructed?.....	5
1.3.2	How is a capital guaranteed index-linked note constructed? .....	6
1.4	Focus and limitations.....	6
1.5	Results of the research .....	7
2	INDEX THEORY .....	9
2.1	Index background.....	9
2.2	Origins and history of index theory .....	9
2.3	Main index theories and formulae.....	11
2.3.1	Main index theories .....	11
2.3.2	Fisher's seven tests .....	14
2.3.3	The most common index formulae .....	15
2.4	Evaluation of the index theories and formulae.....	24
3	CASE STUDY INDICES.....	27
3.1	The United States residential real estate sector .....	27
3.2	US real estate indices .....	28
3.2.1	Dow Jones U.S. Real Estate Index.....	30
3.2.2	Philadelphia Housing Sector Index.....	33
3.2.3	S&P/Case-Shiller Composite Home Price Indices .....	35
3.3	Concluding remarks on the theoretical and practical aspects of indices .....	38
4	EMPIRICAL ANALYSIS .....	41
4.1	Description of the data set.....	41
4.1.1	Data set.....	41
4.1.2	Descriptive statistics.....	42
4.1.3	Returns of the indices.....	43
4.1.4	Correlations.....	44
4.2	Empirical analysis .....	46

4.2.1	Background.....	46
4.2.2	Model specification.....	47
4.2.3	Data analysis .....	47
4.2.4	Testing for the stationarity in the time series .....	53
4.2.5	Empirical analysis after adjusting for non-stationarity .....	57
4.3	Results of the empirical analysis .....	59
5	CONSTRUCTING THE US RESIDENTIAL REAL ESTATE CAPITAL GUARANTEED INDEX-LINKED NOTE .....	61
5.1	Background.....	61
5.1.1	Structured products in general .....	61
5.1.2	Different types of structured products.....	63
5.1.3	Factors affecting the construction of a structured product .....	65
5.1.4	Typical investor's characteristics .....	69
5.1.5	Overview of the pros and cons of capital guaranteed index-linked notes.....	73
5.2	The case study product.....	75
5.2.1	Properties of the case study product .....	75
5.2.2	The purpose of the product .....	78
5.3	Putting together the product .....	79
5.3.1	The maturity .....	79
5.3.2	The structure and pricing.....	80
5.3.3	Other characteristics and possible limitations.....	82
5.3.4	The final structure and projections of future performance.....	83
6	CONCLUSION .....	86
6.1	Overview of the thesis and answers to the research questions .....	86
6.2	Issues for further consideration .....	88
	BIBLIOGRAPHY.....	89
	APPENDIX 1 .....	93
	APPENDIX 2.....	96
	APPENDIX 3.....	97

**LIST OF TABLES AND FIGURES**

Table 1: Philadelphia Housing Sector Index contents on 05.12.2008 .....	34
Table 2: Bloomberg tickers for the indices .....	41
Table 3: Statistical characteristics of the data set .....	42
Table 4: Annual returns .....	43
Table 5: Correlations between the monthly returns of the indices .....	44
Table 6: Ranked correlations between the monthly returns of the indices .....	44
Table 7: Correlations between the quarterly returns of the indices .....	45
Table 8: Ranked correlations between the quarterly returns of the indices .....	45
Table 9: Correlations between the annual returns of the indices.....	45
Table 10: Ranked correlations between the annual returns of the indices .....	45
Table 11: Example calculation from the regression equation.....	52
Table 12: Low and high interest rate scenarios .....	66
Table 13: The pros and cons of capital guaranteed index-linked notes.....	74
Table 14: The iShares Dow Jones U.S. Real Estate Index ETF call options.....	81
Table 15: Projected performance for five price scenarios .....	84
Figure 1: Historical prices of all indices, 2002-2008.....	30
Figure 2: The annual returns of the indices under study, years 2003-2007.....	44
Figure 3: Regression output .....	48
Figure 4: The White Heteroscedasticity test .....	50
Figure 5: Serial correlation - residual graph .....	51
Figure 6: The augmented Dickey-Fuller test for the Dow Jones U.S. Real Estate Index.....	54
Figure 7: The augmented Dickey-Fuller test for S&P/Case-Shiller Composite-20.....	55
Figure 8: The ADF test for S&P/Case-Shiller Composite-20, 1 <sup>st</sup> difference.....	56
Figure 9: Regression output, adjusted for non-stationarity.....	57
Figure 10: Typical choices made by investors in a low volatility market .....	70
Figure 11: Typical choices made by investors in a high volatility market.....	71
Figure 12: Effect of market volatility and market direction on investor's choice of product .....	72

## 1 INTRODUCTION

### 1.1 Goals and motivation for the study

#### 1.1.1 Background

The financial world is surrounded by, or more accurately, embedded by indices. They are more often than not taken for granted – it is usually assumed they are the correct and the best presentation of the underlying phenomenon. It is also often taken that an index is merely the aggregate of the values of a group of instruments. However, in economics we find indices which measure a variety of phenomena, in order to show the true nature of things. For example, consumer price indices aim to encapsulate the changes in price levels of consumer goods over time. Often also the aim is to compare the development of these prices between different countries. With financial indices, also, there can be many more factors at play than the market price of the instrument. The mathematical aspect of indices is more complex than one would imagine – rarely is a pure non-weighted average used to calculate an index. The inputs of the index as well as the method of calculations are the factors, which determine the performance of the index. Therefore, when either selling or buying, or purely examining a product whose performance depends on the performance of an index, it is essential to understand how the index in question is constructed and, most importantly, does the index reflect reality.

#### 1.1.2 Motivation

This thesis has structured products as the underlying theme because they are an extremely interesting and continuously growing sector in the financial product field. A structured product is a packaged financial instrument, which is constructed of two or more financial instruments such as bonds, equity, derivatives, indices or currencies. Derivatives, in short, are financial contracts such as options and forwards – financial instruments whose value depends on the price development of some underlying instrument. There is a vast variety of different kinds of structured products, this thesis concentrate on a basic type of a structured product, a capital guaranteed index-linked note. Capital guaranteed structured products provide

excellent investment opportunities for clients with low risk appetite, but with a strong market view in some specific market. Structured products are attractive for issuers as well, as they provide them with financing for the maturity of the product.

Structured products suit investors for a wide variety of reasons, for example for institutional clients who have restrictions on the risk of the investments or private clients for tax reasons. Using an index as the basis of the structured product, as opposed to using one or few companies, commodities or currencies, leads to a higher degree of diversification, and thus lowers the risk of the investment. Also, the performance of the investment is easier and simpler to follow for the client. Thus, capital guaranteed index-linked notes are common in banks' product range and popular among investors, and therefore also used as the case study for this thesis.

In constructing a financial product, the return of which depends on the performance of an index, it is essential to know how that index is constructed, what it truly measures and what factors may affect it. With indices, which use the market price of the instrument as the factor which is aggregated, should the different instruments be weighted somehow? If they should be weighted, should it be done based on market capitalisation, past or estimated future growth rate, exposure to some business segment or market, or based on some other factor? It can also be questioned whether the market price of the stock is the best measure of the instruments performance. For example, if one wants to know a group of companies' growth potential, and wants to construct an index to evaluate this, should the only factor included in the index be past growth of the market price? According to financial theory, also factors such as debt to equity ratio, market share, the number of employees as well as efficiency and profitability figures reflect growth potential of a company.

Since the performance of an index-linked note is mostly dependent to the index, it is also important to know whether the index correctly and in an undistorted way represents the value development of the underlying basket on instruments. From an investor's point of view, it is essential to know if the construction of the index hinders somehow the return for the investor. The construction of the index is the second most important factor affecting the development of the index. There are a variety of ways to calculate an index from the underlying basket of instruments, and naturally, the way the index is calculated is a crucial determinant of the development of the

index. Therefore, when either constructing or purchasing index-linked structured products, it is of utmost importance to know and understand how the underlying index functions, on what principles it is built and is it really suited for the product in question.

### 1.1.3 The goals of the research

The goal of this research is, thus, to make a thorough examination of indices: what kinds of indices there are, what their characteristics are and how their construction affects their performance. Most importantly, this research aims to discover how indices should be constructed so that they truly and accurately reflect the development of the underlying basket of instruments, in what ever terms wanted. With the help of theoretical and empirical analysis, an index that fulfils specifically the needs of the index-linked instrument can be chosen. Such an instrument will ideally give an investor a return which accurately reflects the development of the underlying basket of instruments. This research will ideally give the adequate tools to consider and construct different kinds of indices tailored for different kinds of purposes, also beyond this case study.

Although this research can be utilised for many different kinds of products and for indices in general, a case study was chosen for this research. This is to provide the research a practical application, so that possibly the research could be utilised in issuing an actual US residential real estate capital guaranteed index-linked note. One of the main goals is, thus, to illustrate how such a structured product is constructed, and what aspects need to be considered in the process. The US residential real estate sector was chosen as the underlying basket of stocks because of the great upside potential that lies in the sector after the recent dramatic bursting of the real estate price bubble along with the subprime crisis. The values of especially residential real estate have declined from their peak levels in the US. However, the starting point of this product suggestion is that the real estate prices may be beginning to reach their bottom. Especially the residential real estate sector, which has taken the largest hit in the latest bear market and recessionary pressures, might soon become a buyer's market and may present a good investment opportunity. There is of course a relatively high risk in this market, but as this product is capital guaranteed, it is a very tempting profit opportunity for any client.

Moreover, the residential real estate sector provides an interesting point of view to constructing this kind of a product. The index-linked notes are usually constructed based on options on ETF's, or exchange traded funds. There is only one major index, the development of which is based dominantly on the US residential real estate sector, and doesn't include other real estate sub sectors. This index is the Philadelphia Housing Sector Index. There is, nonetheless, no exchange traded fund based on this index. The most commonly used ETF for this kind of a product would be the iShares Dow Jones U.S. Real Estate Index ETF. However, the Dow Jones U.S. real estate Index includes mainly companies and real estate Investment Trusts investing in commercial real estate. Thus, basing a US residential real estate index-linked note on the Dow Jones index would in theory not capture perfectly the actual development of the US residential real estate sector. This thesis will look into what would be the solution to this problem. Can the Philadelphia Housing Sector Index be used for the product despite the lack of an ETF, or does the Dow Jones Real Estate Index reflect the returns of the residential real estate sector well enough to justify its use for the product.

## 1.2 Structure

To begin with, the two research questions and the results of the thesis in short are covered. The actual study is divided into two main sections. The first section of this study, chapters 2, 3 and 4, will concentrate on index theory, go over the indices considered for the case study and perform an empirical analysis of the indices. The theoretical background to indices is discussed; however, the main emphasis is on the empirical analysis. After the analysis of the first part of the thesis, an index specifically suited for the case study, US residential real estate capital guaranteed index-linked note will be chosen. The second section, chapter 5, will more briefly go over the mechanics of constructing the case study product. The conclusion will reiterate the results of the study, as well as make some projections about the future of indices, structured products and especially the future performance of the case study product.

### 1.3 Research questions and methods

#### 1.3.1 How is an index constructed?

A starting question for this study is: What does theory suggest as the most suited indices for different purposes? After answering this question, the study can go on to determine which index is best suited for the case study product, the US residential real estate capital guaranteed index-linked note. The criteria for choosing the best index is twofold: analytical and practical criteria. Analytically the best index is the one that represents reality most accurately, as defined by index theory or empirical testing. Practically the best index is the one that could in real life be used for the case study product, taking into account the availability and price of index options, exchange traded funds or other vehicles. This study will look into a variety of index theories. Their background and mechanics will be clarified. Discussion of index theory will provide a basis upon which the empirical analysis is built. The real world is then taken into account in the analysis by looking at the existing United States Real estate indices. Their index formulae will be introduced, and they will be compared and contrasted to the theoretical index formulae.

In order to link this research question to the next one, it needs to be established which index should be chosen for the case study. Index theory has a multitude of views on what is the best index formula, and this discussion will be included. However, the final choice will be made based on an empirical analysis, with the assistance of Standard & Poor's/Case-Shiller composite home price indices. They are a collection of indices, not stock market indices, but indices measuring directly changes in house prices in various regions in the United States. They will be described more carefully later on in the paper. It is assumed in this thesis that the S&P/Case-Shiller indices represent the housing market in the United States accurately. The data set used is closing price time series acquired from the Bloomberg database, and it has been converted into monthly return figures. Data for the dates 3.7.2002 – 31.10.2008 is used in the analysis. The monthly return figures for the S&P/Case-Shiller indices will be compared to the monthly returns of the two United States real estate indices, the Philadelphia Housing Price Index and the Dow Jones U.S. Real Estate Index. From a theoretical point of view, out of these two stock market indices, the index which correlates the best with the S&P/Case-Shiller

indices represents reality best, and thus would be the most correct index for our index-linked note. However, in the final decision of choosing the index for the case study, one needs to also take into account how the choice of the index affects the pricing of the product.

### 1.3.2 How is a capital guaranteed index-linked note constructed?

After discovering the best suited index for this case study product, it will be discussed how a capital guaranteed index-linked product in general and specifically this product can and should be constructed. In the process of illustrating how the case study product is constructed, it will be evaluated how the choice of index affects the pricing, mechanics and the sales potential of the product. To conclude this section, it will be examined what future references should be made from this research when building other similar products.

The method used for this case study section will be analysis of existing product structures and their qualities, and discussing with the help of literature on the topic the most suitable structure for this product. The purpose of this section is to build a product, which could be realistically created. References will be made to theoretical recommendations, but practicalities of the market situation, available instruments and marketability aspects will be the major elements defining the construction of the case study. This section will rely heavily on a professional's interview.

### 1.4 Focus and limitations

The research will concentrate mostly on index theory and the empirical analysis of data. There will be a relatively concise description of structured products, and especially index-linked notes. The focus will be on describing the general mechanism and pricing issues of structured products. The actual construction of the case study will also be brief and descriptive, as most of the analysis and research will be made in the discovering and constructing the index. Only a short overview of US real estate sector and its history and present status will be presented. The empirical study will include a regression analysis and some of the most common statistical tests. The results of the statistical tests will imply that further testing is needed, which is however out of the scope for this thesis.

Even though the focus is on the theoretical aspects of indices and on the empirical suitability of the indices, in the construction of the case study, realities of availability and pricing of instruments will be taken into discussion. In other words, if the theoretically and empirically preferable index cannot be used for the product for some practical reasons, the next best, available and usable solution will be chosen for the product. Naturally, the reasons behind and consequences of this choice will be thoroughly discussed. The construction of the product will follow general market information, and is not entirely realistic, especially as many costs which are not identified by issuers publicly are not included.

## 1.5 Results of the research

The first task undertaken by this paper is to discuss index theory, and whether there is a theoretically supreme index formula or index calculation method, that should be recommended for the case study product. The paper goes over the main theoretical approaches as well as the most common index formulae and the types of stock market index calculation methods. As a result of an overview of different index theorists' research, it is concluded that there is no theoretically supreme index that can be recommended for this case study. It is deduced then, that since this paper takes a descriptive approach to index theory, the best route to follow is to undertake an empirical analysis of historical index return time series.

The empirical analysis is thus utilised to choose between the Dow Jones and Philadelphia indices. Time series used is price data for the Dow Jones U.S. Real Estate Index, Philadelphia Housing Sector Index, and three S&P/Case-Shiller composites, the 20 largest metropolitan areas, the 10 largest metropolitan areas and the US national composite. The price data is available for the time period 3.7.2002 – 31.10.2008, and for the analysis, the prices are converted into daily returns. In the analysis it is taken into account that from the point of view of constructing and pricing the case study product, the Dow Jones index is more suited because there is an exchange traded fund linked to it. Even though the focus of the Dow Jones index is on commercial real estate, if the Dow Jones index correlates better with 'reality', the S&P/Case-Shiller home price index, it should definitely be used. However, if the Philadelphia index correlates better with 'reality', then it needs to be evaluated which

weighs more – the accuracy of the index or the pricing of the product in the final choice between the two indices.

The simple correlations show that the Philadelphia Housing Sector Index has better correlation with the S&P/Case-Shiller composites. However, it is concluded that the Philadelphia Housing Sector Index would make the product prohibitively expensive because of the lack of longer maturity options in the market. A regression analysis is then used to determine whether the correlation between the Dow Jones U.S. Real Estate Index and the S&P/Case-Shiller Composite-20 is strong enough that the use of the Dow Jones index could be justified.

Based on the simple regression, the relationship is strongly positive with high correlation and sufficient statistical significance. The time series data, however, is also tested to reveal the level of stationarity of the time series. It turns out that the monthly return time series of the S&P/Case-Shiller Composite-20 is not stationary. The correction for this problem goes beyond the scope of this thesis, and thus, keeping this error in mind, the initial regression is chosen to illustrate the true relationship.

The final section deals with the actual construction of the product. First, background to structured products is introduced, after which the thesis turns to building the product with the help of an interview with a structured product professional. The final product is a 2-year-maturity index-linked note, which takes a long position in an American call option on the iShares Dow Jones U.S. Real Estate Index with a strike of 30 USD. The product also includes a zero coupon bond, priced at 94,95 EUR and thus 5,05 EUR is left over for the derivative. The lock-in date for calculating the payoff is at maturity. Even though from an investor's point of view, a capital guaranteed product could very be lucrative after all the market turmoil, it is the consequences of the market turmoil that may have led to making the product too expensive. The low interest rates decrease the amount of funds available for the derivative, and the high volatility environment makes the derivatives expensive. All in all, it is concluded that now may not be the right time for issuing the case study product.

## 2 INDEX THEORY

### 2.1 Index background

The word index comes from Latin, meaning an indicator. Indices are in fact a useful way of representing a variety of economic and financial phenomena, in a way indicating aggregated changes in for example prices, costs or quantities. For example, one may look at the stock prices at market close for the five largest Finnish companies, as measured by market capitalisation at the moment. After some days of observations, one may still have a quite good picture about the collective price development of this basket of companies. However, as more observations are added, either with more days or more companies, or both, it becomes more difficult to get a good picture of the collective price development of the basket of companies. This is where an index becomes extremely useful.

An index is essentially a tool that enables us to “compress the information included in a time series into a more illustrative form” (Säynevirta 1991, 5). A good index combines and condenses any number of time series into a single number, but while doing so loses as little information as possible. Index itself can also be a time series, if it describes the procession of some phenomenon over time. There are a variety of different types of indices: for example, price indices, quantity indices and cost indices. Throughout this study, unless otherwise mentioned, indices refer to price indices. Using indices makes observing and analysing different kinds of economic, financial, demographic, technical and other kinds of phenomena over time much simpler. Index theory has in fact arisen from the search for ways to study and analyse larger and larger entities. (Säynevirta 1991)

### 2.2 Origins and history of index theory

Classical index theory is divided into the tabular standard (fixed basket approach), statistical approach, test approach, Divisia index approach and economic approach (Diewert 1993). Säynevirta divides these classical approaches into two main categories, descriptive and analytical research (Säynevirta 1991, 9). This paper will follow Säynevirta's classification, but where appropriate, use also Diewert's analysis

as complementary material. The details of all these classical approaches will be analysed more thoroughly later, but in short terms, Säynevirta's classification is as follows: the analytical approach uses some kind of an economic theory as the basis for the assumptions that are used when creating the index. If the index behaves as can be expected by the theory, then the calculated index reflects reality accurately. However, usually the theoretical assumptions do not hold fully in the real world, and as a consequence, indices calculated using this theoretical background do not match the ideal indices as defined by theory. On the opposite, the descriptive research approach concentrates on the characteristics of the data, such as mean, median or other statistical feature. The goal is not to create or validate a theory, but rather to observe the actual nature of the data. Even though the theoretical background can be generated, it is not necessary. (Säynevirta 1991) The descriptive and analytical approaches will be discussed in more detail in the coming sections.

The basic division for indices is the division between fixed-base and non-fixed base indices. This division arises from the way the indices are constructed. Fixed-base indices are calculated so that the first observation in the index time series is the value to which all the consequent values in the index time series are compared. Non-fixed-base indices, on the other hand, are calculated using the chained method: each observation in the time series is compared to the previous observation, and not to the first observation every time. To pursue the issue further, there are a variety of ways the observation points of the index time series are generated. The value of the index at any point of time depends on the mathematical formula used to create the single index figure from a group of values, as well as on whether the different values of the group are weighted or not. Also, the types of mathematical comparisons that can be made between the two different observations from the index time series are multiple.

As mentioned before, this paper will concentrate on price indices. There are, however, also quantity indices. Moreover, indices are most commonly used for comparisons over time, but they can also be used for comparison between different groups. (Säynevirta 1991) An example could be a living standard index between countries, or regions within countries. However, this paper will concentrate on price indices that measure the development of the prices of a basket of goods, or in this case, stocks, over time.

## 2.3 Main index theories and formulae

### 2.3.1 Main index theories

The first approach to index number theory was the fixed basket approach, or the tabular standard as referred to by Diewert (1993). The earliest evidence of this approach has been the book *Chronicon Preciosum* by William Fleetwood, the Bishop of Ely, in 1707. He calculated the effect of prices changing from year 1460 to 1707 by keeping the basket of goods constant (the basket included 5 quarters of wheat, 4 hogsheads of beer and 6 yards of cloth). The simple way to calculate such an index is to use the following formula:

$$P(p_2, p_1) = \frac{p_2 \times q}{p_1 \times q} \quad (1)$$

This approach has had a number of followers over the years, including index number theorists and economists such as Edgeworth, Marshall and Pigou. Also Laspeyres and Paasche indices result from this school of thought, with Laspeyres using the quantities of the base period, and Paasche using the quantities of the current period. These two index formulae will be illustrated in more detail later. Also the Edgeworth index will be explained in more detail. Edgeworth was one of the theorists, who began to argue that if the quantities of goods in the basket are different between the base period and the current period, some kind of an average of these two quantities should be used in the calculation of the index. In 1925, Edgeworth chose to use an arithmetic average of the quantities and for example Sidgwick and Bowley at the end of the 19<sup>th</sup> century, decided to use an average of the Paasche and Laspeyres indices. The Edgeworth index, as well as the Walsh index which uses the geometric average of quantities, are illustrated in more detail later. (Diewert 1993) This progress within the fixed basket approach finally led to the emergence of two different approaches: test and economic approaches (Diewert 1993), or descriptive and analytical approaches, as defined by Säynevirta (1991).

Analytical or theoretical approach begins with a theory, such as a theory of demand for a good. Based on the assumptions of this theory, the index is constructed. It is assumed that the data in use, when turned into an index, behaves as implied by the

theory. If this holds, the calculated index reflects the true index. There are, however, a variety of issues that lead to the actual index not reflecting the theoretically correct index fully. These issues include, among others, quality change in the data and changing of the theoretical assumptions over time. (Säynevirta 1991) The economic approach is one of the analytical approaches to index theory. Diewert outlines that the economic approach to indices “relies on the assumption of optimizing behaviour on the part of economic agents: utility maximizing or expenditure minimizing behaviour on the part of consumers and profit maximizing or cost minimizing behaviour on the part of producers” (Diewert 1993, 44). Economic theorists examine prices and quantities with the help of economic theories – the main purpose is not to find the ideal index formula, but rather it is the point to establish the index on some economic theories, such as a demand function. Economic price indices, for example, aim at reflecting the ratio of required minimum costs in different price situations so that the level of utility is kept constant. (Säynevirta 1991)

One of the main issues when dealing with economic price indices is the homothetic preferences case. Price indices are dependent on price, quantity and utility, that is often assumed at a set, maximum level. However, in reality, different indices differ because they are dependent on the utility level. In order to eliminate the effect of changes in the utility level on the price index, consumer's preferences need to be homothetic. Preferences are homothetic when the elasticity of demand for all the goods is one, or in other words, consumer values all goods similarly. Moreover, economic quantity indices are also dependent in the level of utility or the preferences of the consumer. Therefore, preferences have to be homothetic for quantity indices also, which holds if Fisher's proportionality test holds. In short, if the quantities in the index change by some scalar, the resulting index number has to change by the same scalar as well. This requirement of homothetic preferences is not rational in reality – elasticities of demand vary for different goods and for different consumers. (Säynevirta 1991, 20) Moreover, as is noted in the field of microeconomics, many of economic theories are not applicable as such to real phenomena. Using these theories then as the basis for an index formula may not be very useful, as the index may very well not reflect reality.

The descriptive approach to indices takes the realities of the data set as a starting point. There is necessarily no need to create a theory, but a theory may be created

on the characteristics of the data, depending on the nature of the data. Index theorists have different names for this type of an approach: Ragnar Frisch has named it atomistic research, and Paul Samuelson called it statistical approach. Quite often, the descriptive approach settles on merely describing the reality, and not so much theorising it. (Säynevirta 1991) This paper represents the descriptive research approach.

There are historically two types of approaches within descriptive research. The older, stochastic thinking is represented by early index theorists such as Jevons and Edgeworth. This approach takes individual price changes as random numbers that do not reflect the average change in prices. The price index, on the other hand, is a more reliable measure of the average change in the prices over time. (Säynevirta 1991). According to Diewert (1993), Jevons used his index to approximate how increases in the money supply affect prices. He recommended the use of a geometric mean of the price ratios. Edgeworth moved into the descriptive approach camp from the fixed basket approach, and recommended the use of the median of price ratios as the best estimator of price change. (Diewert 1993).

The other subgroup within descriptive research is the test theoretical approach, represented by economists and index theorists such as Irving Fisher and C.M. Walsh. The test theoretical approach attempts to find an index formula, which fulfils certain conditions. According to Diewert (1993, 38), "the origins of the test approach are rooted in the more or less casual observations of the early workers in the index number field on their favourite index number formulae or those of their competitors." The most prominent figure within this approach is Irving Fisher, who published one of his founding works in 1922, "The Making of Index Numbers". Fisher criticised the statistical or stochastic approach for the assumption that all prices move proportionally (Diewert 1993). Fisher's goal was to establish one index formula, which would suit all purposes and measurements. (Säynevirta 1991, 11). Fisher's classic tests were in fact partly developed by earlier or contemporary index theorists, such as Walsh and Pierson (Diewert 1993, 38). Fisher's tests, which should be fulfilled by the perfect index, are explained now and also used in the evaluation of the real estate indices.

### 2.3.2 Fisher's seven tests

In his famous 1922 work, "The Making of Index Numbers", Irving Fisher aimed to create the perfect index which could be used for any data set, and which would always correctly represent the data. The starting point for his tests is the price index,  $P(p^1, p^2)$ , comparing prices between the base period  $t^0$  and comparison period  $t^1$ . For simplicity, quantities are ignored for now. The Fisher's seven tests are: the unit of measurement, proportionality, identity, circular, time reversal, factor reversal and determinateness tests. (Säynevirta 1991, 11-12)

In addition to his formal tests, Fisher also has argued that a good index formula should be consistent when aggregated. The starting point for his argument is that a specific index formula is used to calculate a set of index numbers from portions of a larger data set. (Säynevirta 1991) An example could be the price indices for all the different industries in the Helsinki Stock Exchange. Then, the same index formula is used to calculate an index for the whole data set, or all the companies in the Helsinki Stock Exchange, using the already calculated index numbers for the different industries. In other words, the total economy price index is calculated using the industry index numbers, and not the actual underlying data set. The resulting index is then compared to a total economy index calculated with the same index formula, but now using the underlying data set, and not the industry index numbers. If the index formula 'aggregates consistently', the price index number for all the companies in the Helsinki Stock Exchange is the same whether it is calculated directly from all the prices of the companies, or via the industry specific indices.

The Fisher's tests are historically an important part of index theory, and may be very useful when examining indices. However, reportedly no index has fulfilled all of these tests. Moreover, the most commonly used index formulas, Paasche and Laspeyres, do not fulfil the two tests which Fisher outlined as the most important tests to fulfil: time reversal and factor reversal tests. Also, it has been proven by later index theorists that Fisher's tests contradict each other. Moreover, even Fisher himself later came to the conclusion that circular test is not theoretically correct. (Säynevirta 1991, 14). Each of these tests is explained in more detail in appendix 1.

### 2.3.3 The most common index formulae

Before going over the most common index formulae, some conditions and notations used in this section and beyond are explained. The most basic kind of index is the weighted average of the price and quantity ratios. The weights used for the calculation can be the values or value weights of either the base period or the observation period. These concepts as well as other basic index working tools are represented in appendix 2, Index Basics. (Säynevirta 1991)

Here are a few brief examples of index calculations. For example, the price of the new basket of goods using the prices of the base year is:

$$p^0 \times q^1 = \sum p_i^0 q_i^1 \quad (2)$$

Another example is a price index formula, where the weights are based on the base year values. This kind of an index is called a fixed basket price index as discussed before:

$$\frac{p^1 \times q^0}{p^0 \times q^0} \quad (3)$$

These notations and conditions will be used in describing the classical and most commonly used indices, as well as throughout this paper. Now we can turn to describing briefly the major indices and their formulae. The most common fixed-base indices are examined here, and the non-fixed-base indices and their formulae can be found in appendix 3.

#### 1) Laspeyres index

The Laspeyres index, one of the most used and well known indices, was created by E. Laspeyres in the 19<sup>th</sup> century. It is a fixed basket price index, meaning that only the price changes over time, and weights are held constant at the base year level. The Laspeyres price index is represented mathematically as follows:

$$P_0^1(La) = \frac{p^1 \times q^0}{p^0 \times q^0} \quad (4)$$

The formula thus gives us the effect of a price change when a fixed basket of goods is purchased with today's prices as compared to yesterday's prices. (Säynevirta 1991, 21) In practice, the formula used for calculation is formed as follows (using the index basic conditions in Appendix 2):

$$\begin{aligned} P_0^1(La) &= \frac{p^1 \times q^0}{p^0 \times q^0} = \frac{\sum p_i^1 q_i^0}{\sum p_i^0 q_i^0} = \frac{\sum v_i^0 (p_i^1 / p_i^0)}{\sum v_i^0} \\ &= \frac{\sum v_i^0 \Pi_0^1(i)}{\sum v_i^0} = \frac{\sum v_i^0 \Pi_0^1(i)}{V^0} = \sum w_i^0 \Pi_0^1(i) \end{aligned} \quad (5)$$

In words, the Laspeyres price index is the sum of the multiples of the base year weight of good  $i$  in the basket and the price ratio of prices from period 0 and 1. To reiterate, the weight of a good in the basket is the ratio of its value, price multiplied by quantity, compared to the value of the whole basket. Laspeyres is a simple and rather widely used basic index, but it does not fulfil the Fisher time and factor reversal tests. It is, however, consistent when aggregated. (Säynevirta 1991, 20-22) The main issue with the Laspeyres index formula is that it does not allow for substitution of goods in the basket. Keeping the basket of goods fixed is not a reasonable requirement for any price index, and the requirement becomes more unreasonable as time from the base year increases. Especially, when considering a stock price index, companies that should be included in the index may emerge, or other companies in the index may go bankrupt. These changes need to be taken into account in the calculation of an index, and the Laspeyres index, at least as such, does not allow for such changes.

## 2) Paasche index

As the Laspeyres index, the Paasche index was also developed in the 19<sup>th</sup> century by Herman Paasche, a German economist. The Paasche index is similar to the Laspeyres index, with the main exception that each price observation is compared to the previous period's observation, and the weights are those of the current period.

Therefore, when using the Paasche index, the weights of each good in the basket need to be known at the calculation moment. The Paasche index is represented mathematically as follows:

$$P_0^1(\text{Pa}) = \frac{p^1 \times q^1}{p^0 \times q^1} \quad (6)$$

The formula shows how much more or less the new basket costs with the prices of this period as compared to the prices of previous period. The practical formula for calculation is derived again using the index basic conditions from Appendix 2:

$$\begin{aligned} P_0^1(\text{Pa}) &= \frac{p^1 \times q^1}{p^0 \times q^1} = \frac{\sum p_i^1 q_i^1}{\sum p_i^0 q_i^1} = \frac{1}{\sum v_i^1 (p_i^0 / p_i^1) / \sum v_i^1} \\ &= \frac{1}{\sum [v_i^1 / \Pi_0^1(i)] / V^1} = \frac{1}{\sum [w_i^1 / \Pi_0^1(i)]} \end{aligned} \quad (7)$$

Since the Laspeyres index was a base-period, weighted, arithmetic average of price ratios, the Paasche index is a current period, weighted, harmonic average of price ratios. As the Laspeyres index, the Paasche index does not fulfil Fisher's time reversal and factor reversal tests, but is consistent when aggregated. (Säynevirta 1991)

### 3) Fisher ideal index

The Fisher ideal index is the geometric average of the Laspeyres and Paasche indices, represented mathematically as follows:

$$P_0^1(\text{F}) = \sqrt{P_0^1(\text{La})P_0^1(\text{Pa})} \quad (8)$$

However, as opposed to the Laspeyres and Paasche indices, the Fisher ideal index takes into account quantity change, and also fulfils the time and factor reversal tests, as outlined by Fisher himself. It also satisfies the determinateness test and partly also the proportionality test. In relevance to Fisher's requirements, it does not

however meet the requirement of being consistent when aggregated. (Säynevirta 1991)

#### 4) Edgeworth index

The Edgeworth index is similar to the Fisher ideal index in that it is also a type of an average of the Laspeyres and Paasche indices, but not only of the price indices but also quantity indices. The proof for this is not shown here however, as quantity indices are not relevant for this thesis. The difference between the Edgeworth index and the Laspeyres and Paasche indices is that for the Edgeworth index, the weights used are the averages of the new and old basket of goods:  $1/2(q^0 + q^1)$ . The formula used for calculation is:

$$P_0^1(\text{Ed}) = \frac{p^1 \times 1/2(q^0 + q^1)}{p^0 \times 1/2(q^0 + q^1)} = \frac{p^1 \times (q^0 + q^1)}{p^0 \times (q^0 + q^1)} = \frac{\sum p_i^1(q_i^0 + q_i^1)}{\sum p_i^0(q_i^0 + q_i^1)} \quad (9)$$

The Edgeworth index satisfies the time reversal test, but not the factor reversal test. (Säynevirta 1991)

#### 5) Walsh index

The Walsh price index, established in 1901, is one of the indices that have their base in the fixed basket index approach. It has been set up to remove the problem of the changing quantities over time, and it uses the geometric mean of the quantities in two periods to attempt to solve it. (Diewert 1993, 36)

$$P_w(p^0, p^1, q^0, q^1) = \frac{\sum (q^0 q^1)^{1/2} \times p^2}{\sum (q^0 q^1)^{1/2} \times p^1} \quad (10)$$

The major non-fixed-base indices are introduced in appendix 3.

## 6) Stock market indices

There are so many different kinds of stock market indices across the globe today, that instead of examining a few index calculation formulae, this thesis will examine stock market indices from a more descriptive perspective. Many stock market indices use common index formulae such as Laspeyres or Walsh price index formulae as the basis of calculation, but they have also other characteristics not discussed in conjunction with the index formulae. These characteristics, such as the divisor, different weighting strategies and the calculation of price and return indices are explained next.

An equally weighted index is exactly that – all stocks in the index have equal weight in the calculation. Thus, a 10 % price change in the price of a small company, priced at \$1,50 per share has the same effect on the index number as a 10 % price change in a share priced at \$40 per share. The smaller the price of the share, the more probable large percentage movements are (10 % of \$1,50 is 15 cents, while 10 % of \$40 is 4 dollars). A price weighted index accounts for this problem, as it weights the stocks in the index based on the price. The price, however, is a per share measure, and therefore is affected by the number of outstanding shares. An index using price weighting does not reflect accurately the development of company values if the total value of the company is not accounted for. For this reason, majority of stock market indices are market capitalisation weighted. This means that the weight of one stock in the index is determined by its market capitalisation, in other words, the number of company's outstanding shares multiplied by the share price. If there are 5 companies in an index, and the market capitalisations are \$0,5 million, \$1 million, \$1,5 million, \$2 million and \$5 million (total index market capitalisation is thus \$10 million), then the weights of these companies in the index calculation are 5%, 10%, 15%, 20% and 50%, respectively.

The aforementioned market capitalisation is full market capitalisation. Some indices, such as Dow Jones US indices, use free-float market capitalisation weighted calculation for their indices. This means that "a company's outstanding shares are adjusted by block ownership to reflect only truly tradable and investable shares" (Guide to the Dow Jones Global Indexes 2008, 17) The types of block ownership that are excluded from the float are cross ownership (shares owned by other

companies), government ownership (shares owned by central or municipal governments), private ownership (shares owned by individuals, families, trusts or foundations) and restricted shares (shares not allowed to be traded during certain period). (Guide to the Dow Jones Global Indexes 2008) One of the aims of using float-adjusted market capitalisation is to make it possible to create exchange traded funds linked to the index and to enable investors to mimic the index. There are no problems if the free-float factor, the availability of free-floating shares as a portion of total outstanding shares is high. Problems arise when the free-floating factor is very low. Assume a very large company, which because of its market capitalisation has a 20 % weight in the index. However, only a small portion of that company's shares are available in the market. As a result, only a small group of investors would be able to purchase a sufficient amount of shares if they were linking the investment portfolio to the index. This demand might even lead to a spiral where the price is pushed up, increasing the market capitalisation and thus increasing the company's weight in the index even more. Therefore, especially with indices that are used actively in portfolio management or have an exchange traded fund linked to them, may benefit from using a float-adjusted market capitalisation weighting.

A modified market capitalisation weighted index is very common in the market. It works just as a market capitalisation weighted index, with the exception that the weight of one share in the index has an upper limit. For example, even though according to market capitalisation the weight of a company would be 40 % in the index, the maximum weight may be set at 10 %. This is to ensure that the index is a diversified measure of the whole market. There is also yet another way of weighting the shares in the index called attribute weighting or fundamental weighting. An attribute weighted index uses other criteria than price or market capitalisation in determining the weight of one share in the index. The criteria based on which it is determined how to weigh the shares needs to be set in advance. Often the attributes include factors, such as company's price-to-earnings or price-to-book ratio, dividends or dividend ratio or other key figures. One could also imagine a socially responsible investing index, which determines the weight of a share in the index based on how large a percentage of the company's income comes from environmental, ethical or social projects.

There are two major types of stock market indices, the return index and the price index. In the return index, it is assumed that the dividend is reinvested into the index basket of stocks, and the index is therefore calculated accordingly. A price index, on the other hand, ignores the effect of regular dividends. At least the major indices are calculated in both price and total return method. All other kinds of corporate actions, as well as any other changes that affect the stocks in the index but are not regular price changes need to be taken into account even when calculating a pure price index. Meanwhile, the index should however only represent the changes in prices, and any other phenomena should not show in the index number. This effect can be achieved by using an index divisor.

According to the Guide to the Dow Jones Global Indexes (2008), there are two main reasons why stock market indices use divisors. The guide explains that the market capitalisation of an index is affected by company level and market level phenomena. Changes at company level refer to changes in the market capitalisation of a company due to “share changes caused by corporate actions such as takeovers, secondary offerings, repurchase programs, rights offerings and spinoffs” (Guide to the Dow Jones Global Indexes 2008, 11) Changes at the market level refer mainly to companies entering and exiting the market, and thus the index.

According to the Dow Jones publication, the divisor is an adjustment factor which “links each successive weighted basket of securities in the index with the preceding basket” (Guide to the Dow Jones Global Indexes 2008, 11). This creates a chaining effect. If there are no changes in company or market level, the divisor remains unchanged. The character of the index divisor is best illustrated by its formula as presented by Dow Jones. The divisor formula may differ between different agencies, but Dow Jones, being one of the largest agencies, provides a credible example:

$$D_{t+1} = D_t \times \frac{\sum (p_i \times q_i) \pm \Delta MC_{t+1}}{\sum (p_i \times q_i)}$$

Where

$D_t$  = Divisor at time t

$D_{t+1}$  = Divisor at time t + 1

$p_i$  = Stock price of company i at time t (11)

$q_i$  = Number of shares of company i at time t

+  $\Delta MC_{t+1}$  = Add new components' market capitalisation or use adjusted market capitalisation if corporate action

-  $\Delta MC_{t+1}$  = Subtract market capitalisation of companies to be deleted

The adjusted market capitalisation refers to corporate actions, with which one needs to calculate an adjusted price or new number of shares, depending on the action.

For example, in the case of a split (for example, the holder of a company's shares receives 3 new shares for each share that she holds), the adjusted price is calculated with the following formula:

$$\text{Adjusted price} = \text{Close price at time } t \times \frac{\text{Number of current shares}}{\text{Number of new shares}} \quad (12)$$

A split does not create more value; rather, it merely divides the cake (company's market value) into smaller bits. Therefore, the adjusted price is diluted from the previous price. In the case of a new company entering or exiting the index, the market capitalisation is simply the price of the stock times the number of shares. (Guide to the Dow Jones Global Indexes 2008, 18-19)

A major criticism of the stock market indices today is that they may act as a source of volatility in the market, and even contribute to the emergence of bubbles. Behind this phenomenon lies the popularity of index investing, either directly by investors through building their portfolio based on an index, or through exchange traded funds. As is explained by Landis (2006) in his article "Reinventing the index", current market capitalisation weighted indices are flawed because they are dominated by a handful of companies. Market capitalisation weighted indices rely on the assumption of efficient markets – the price of the stock is a pure and undistorted representation of the company's value and expected future earnings. At least in short run, however,

conditions such as market sentiment, market squeezes and other factors not dependent on the underlying value of the company may affect the price of the stock. Adding index investing to the picture, the result is a phenomenon where in the case of a stock increasing its market capitalisation due to an increase its price, investors and funds following an index will increase the share of the stock in their portfolio. Since index investing is very popular, in some cases the overall increase in the demand on the stock may be significant enough to push the price of the stock further up. This process thus may become a vicious circle, especially if the index is not a modified market capitalisation weighted index. This process works exactly in the same manner if the price of a stock decreases. The stocks in the index may become under- or overpriced, and if the mass of investors using the index investing method is large enough, there may not exist enough investors who look for arbitrage opportunities in the undervalued stocks to balance off the effect of the index investors. Also, the index investors are forced to buy stocks when they are overpriced, and sell them when they are underpriced. (Landis 2006)

It has also been criticised that market capitalisation based indices are backward looking, since they give weight to stocks that have done well in the past. The attribute or fundamental weighted indices, however, are a new group of indices that attempt to look into the future as well. The supporters of these new stock market indices argue that if an investor ties his investments into an index, they can be sure that most of their money is invested in overvalued companies. This is called "the noisy market hypothesis" (Perold 2007, 1). An example of the fundamental indices is group of indices called the Research Affiliates Fundamental Indices, RAFIs. These indices weight holdings based on sales, dividends, cash flow and book value, and they attempt to do what the stock price theoretically should do: show the fundamental value of the company. The idea of fundamental indices has also been taken further with the creation of intelligent indices, which do not only aim at matching the market but also at beating it. The companies in the index are weighted based on a combination of measures of value, price, momentum risk and timeliness. (Landis 2006)

The fundamental indices have, nevertheless, been criticised for a variety of reasons. Perold (2007, 1) begins his criticism by stating that investing in a market capitalisation weighted index is the only strategy that all investors can follow

because “the collective holdings of investors (by definition) aggregate to the market portfolio, for every investor who is underweight a stock, another is overweight that stock, and between them, it is at best a zero-sum game.” Perold (2007) argues that the noisy market hypothesis, beginning with the assumption that the fair value of the company is not known, falsely uses a fixed value of fair value to deduce the probability distribution of market prices using the probability distribution of the pricing error. Perold (2007, 4), arguing for market capitalised indices, states that “the correct analysis is to hold that [current] price fixed and use the probability distribution of the pricing error to deduce the probability distribution of the unknown fair value.” He continues to prove that because market capitalisation does not reveal the undervaluing or overvaluing of stocks, “the random mispricing of stocks does not systematically shift the portfolio weights toward overvalued stocks” (Perold 2007, 4). It has been noted that fundamental indices tilt the portfolio to value stocks due to the nature of weighting. It is concluded by Perold (2007, 5), however, that “if value stocks are systematically mispriced, fundamental indexing may perform well – along with other value-oriented strategies – because it is exploiting this particular inefficiency, not because capitalisation weighting in and of itself, creates a performance bias.”

#### 2.4 Evaluation of the index theories and formulae

This paper illustrated earlier Fisher’s ideal index, and especially the tests he had set up for choosing the best index. It has already been shortly discussed how there is no index that fulfils all of Fisher’s test – not even Fisher’s own ideal index. This matter has been discussed and proven by W. Eichhorn (1976) in his article “Fisher’s test revisited”. Eichhorn mathematically tests all of Fisher’s tests, and shows in his paper that the system of tests set up by Fisher is inconsistent, even though individual tests may work and may provide valuable information about the properties of indices. Indices may fulfil some of Fisher’s test, but no index fulfils all. Since there is no clear ranking between the tests suggested by Fisher or other theorists, Fisher’s test approach cannot be used to decide between different index formulae. Therefore, we must use another theoretical approach to try to find the best index formula.

Wilson (1982) in his article "Appropriate Index Number Formulae for Productivity Measurement at the Plant/Organisational Level" explains how choosing the best index is by no means an easy or straightforward task. First of all, the list of different kinds of index formulae is endless. Irving Fisher alone listed 126 different index formulae, and this paper only illustrates a few of the most common and relevant index formulae. Secondly, there are almost as many criteria for choosing a good index as there are index formulae, and not one criterion is accepted as universal or prevailing. The criteria that is chosen depends dominantly on the index theorist's view point. Fisher had the test theoretical viewpoint, and thus for him the ideal index fulfils the tests he has set up for it. Meanwhile, some index theorist use economic theories as a basis of the index formula and therefore evaluate the performance of the index through the theory.

Wilson (1982) begins his analysis with evaluating the Laspeyres and Paasche index formulae. From an economic theory point of view, the Laspeyres index formula suffers from a substitution bias. As it assumes a constant basket of goods, the Laspeyres index fails to account for the increase in utility as consumers are able to substitute from cheaper goods to more expensive goods if relative prices change. Some economic theorists have argued the Laspeyres and Paasche indices may serve as bounds between which the 'true' index should lay. However, other index theorists such as Samuelson and Swamy have stated that "these formulae only act as limits to the 'true' index as defined from economic theory if the underlying utility... or production... functions are homothetic" (Wilson 1982, 245). However, as was discussed above, preferences are very rarely homothetic in reality. According to Wilson: "Homotheticity is not as unrealistic an assumption in production function work (satisfied by, for example, the Cobb-Douglas function) as for utility functions for consumer price indices" (Wilson 1982, 245). A consumer price index, based on a basket of food and non-durables can be compared to a stock market index, based on a basket of stocks, purchased by investors.

Since the Fisher 'ideal' index is a geometric average of the Paasche and Laspeyres indices, it would be rational to assume that it might give the true index value that lies in between the Paasche and Laspeyres indices. Research shows, however, that this assumption does not hold. Even though Samuelson and Swamy (1974) have shown the Fisher 'ideal' index to be a second-order approximation to a true homothetic

index, if the underlying function is not homothetic, then the index will yield inaccurate results. Overall, it can be questioned whether any index can accurately represent reality, especially if the approach is an economic theory framework: "Economic theory warns that no single index-number formula could correctly characterise the alternative preference and indifference-contour patterns" (Samuelson et al 1974).

Examining what index number theorists have to say about the issue of choosing the best index leads to a situation where one needs to choose a side, so to speak. One needs to determine which approach is the most suitable for the specific index that is being examined. As was already mentioned before, this paper looks at the case study index from a descriptive point of view. What is desired from an index that is used in an index-linked product is that the chosen index characterises reality in the most accurate manner possible. Since we have no economic theory that we can compare the index formula performance with, and since we have concluded that the test theoretical approach is not sufficiently universal to be used for this analysis, the next step is to move into an empirical evaluation of the existing indices. Using index time series, we can compare the historical performance of two stock market index candidates for the product with a home price index measuring directly changes in home prices in the United States. This empirical method gives a point of comparison – the index which from a statistical point of view represents reality better, is the index that should be chosen for the case study product.

### 3 CASE STUDY INDICES

#### 3.1 The United States residential real estate sector

In the United States, house ownership is common. In 2004, 68 % of households owned their own homes, and for majority, housing equity makes up nearly all of the house owner's non-pension assets at retirement (Himmelberg et al 2005, 2). Until recently, the house owners have been witnessing the value of their assets increasing at record breaking pace. In the period from 1975 to 1995, real single-family house prices in the US increased by an average of 0,5 % per year, the same figure for the period 1995-2004 was 3,6 % per year. From the 70's to the 90's, house prices increased overall by 10 %, while from the 90's to the year 2004, real house prices saw a total growth of 40 %. (Himmelberg et al 2005) In 2007, the estimate for the value of residential real estate held by households and non-profit organisations was 22,5 trillion US dollars. Comparing this figure to the market capitalisation of US equities, 19,9 trillion USD and bonds, 29,7 trillion USD, shows that the residential real estate sector should really be recognised as a significant asset class. (S&P/Case-Shiller Home Price Indices Factsheet 2008, 1)

The turn to the home price development, however, came in the wake of the subprime crisis. The Office of Federal Housing Enterprise Oversight (2008), a US governmental institute, states that the year 2007 will go down in history due to "declining house prices, a weak housing sector, and continued deterioration of the performance of subprime mortgages." The subprime crisis led to a process of repricing of risk in mortgage and broader financial markets, resulting in the widening of credit spreads and flight to more secure asset classes and forms of investment (The Office of Federal Housing Enterprise Oversight 2008).

The reasons behind the subprime crisis are multiple, but at the core of the issue lays the house price boom especially rampant during the years 2001-2005. The price boom, according to The Office of Federal Housing Enterprise Oversight (2008), was "fuelled by rapid growth in subprime, Alt-A, and other non-traditional mortgage lending." During this credit boom, the underwriting standards began to deteriorate while it was believed that house prices would keep on growing. However, in 2005

and 2006, house price growth began to slow down, increasing default risk on the securitised mortgages dramatically as suddenly, home owners who had mortgaged their homes to the peak and over saw their home values melt away. The crisis in the subprime loans, the highest default risk group, affected all mortgages as soon after the initial signs of loose lending in the subprime mortgages group led to a “virtual collapse of the primary and secondary markets for subprime, Alt-A, and non-traditional mortgages.” (The Office of Federal Housing Enterprise Oversight 2008) The process of pooling and securitising regular home mortgages that had begun in the 1970's, culminated in 2007 in an unforeseen collapse in prices, returns and confidence that has since caused a global financial crisis, followed by a global economic crisis. It can be questioned, however, whether the house prices have overreacted to the credit crisis. It can also be speculated that when the US and global economy begin to recover from the economic slowdown, house prices will also again begin to increase. An important thing to consider is that it may be that we are now seeing the bottom to United States house prices, which provides an alluring investment opportunity.

### 3.2 US real estate indices

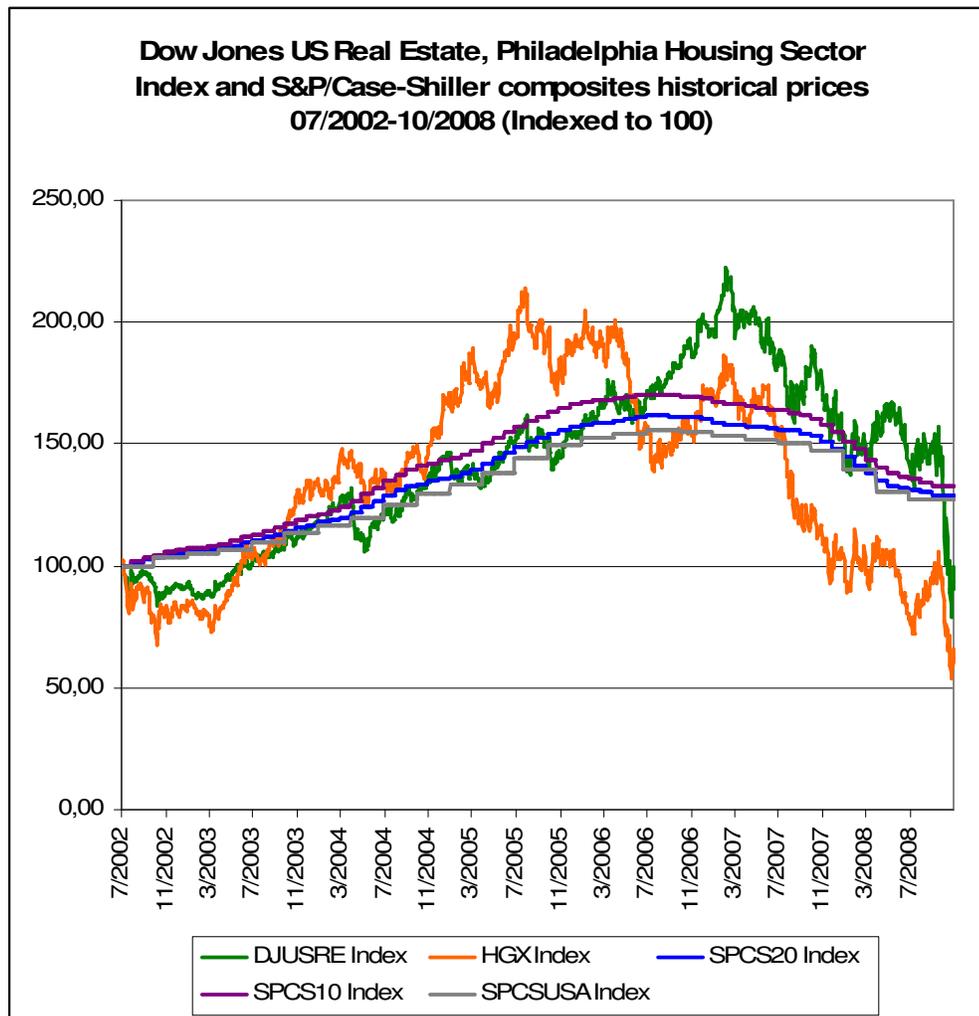
The world is full of indices, varying in focus, calculation method and other aspects. Also, a variety of financial institutions calculate real estate indices from global and regional perspectives, as well as concentrating on purely REIT's, or taking a wider section of companies in the real estate business into the index. There are also indices, which measure changes in the values of property not through the stock market, but directly by using actual property sale prices.

The indices chosen for this analysis are: the Dow Jones U.S. Real Estate Index, the Philadelphia Housing Sector Index and the S&P/Case-Shiller Composite Home Price Indices. The Dow Jones U.S. Real Estate Index was chosen as it is often used as the basis for real estate linked structured products because there is an exchange traded fund linked to it. For example, Nordea recently used the Dow Jones index-linked ETF, iShares Dow Jones U.S. Real Estate Index Fund, as the underlying of the derivative in a structured product (Nordea Markets 2008). The Philadelphia Housing Index has been chosen as it seems to more accurately represent the residential real estate market. As opposed to the Dow Jones U.S. Real Estate Index,

the Philadelphia Housing Sector Index mainly invests in construction companies and other firms which are more or less directly involved in the residential housing market. The Dow Jones index, on the other hand, mainly invests in companies involved in the commercial real estate market, as will be illustrated later.

The S&P/Case-Shiller composite indices have been chosen as a representation of reality. This choice mainly relies on the fact that it is an index which directly measures the housing market price development. As opposed to the stock market indices, the S&P/Case-Shiller indices derive their values of actual house sale prices, and therefore the data has less risk of being distorted by non-housing market aspects. The performance of the stock market indices is dependent on the performance of the companies within the index. In efficient markets, the value of a company in the market should be a pure representation of the expected value of future earnings for the company. For example, a construction company's future earnings are expected to grow along with house prices. In reality, nonetheless, the market value of a company is affected by a multitude of factors – by company related issues such as management or financing position, and by market related issues such as market sentiment, liquidity and market squeezes. Thus, an index which directly measures changes in house prices can be expected to provide a good estimate of the reality of the housing market. Shiller and Case have done groundbreaking research on real estate prices, and on establishing an index that accurately measures the development of home prices. This thesis will shortly discuss the theoretical work done by Shiller and Case in conjunction with introducing the S&P/Case-Shiller Composite Indices. The historical prices of these indices are represented in figure 1 below:

Figure 1: Historical prices of all indices, 2002-2008



### 3.2.1 Dow Jones U.S. Real Estate Index

The Dow Jones U.S. Real Estate Index is “a float-adjusted capitalization-weighted, real-time index that provides a broad measure of the U.S. real estate securities market” (Dow Jones Indexes, Dow Jones U.S. Real Estate Index Factsheet 2008). This index provides a broad measure, because it entails a broad variety of companies. Many of them are real estate Investment Trusts, but the contents also include “companies that invest directly or indirectly in real estate through development, management or ownership, including property agencies” (Dow Jones U.S. Real Estate Index Factsheet 2008). According to the index factsheet, the stock

price of REIT's represents well the true value movements of commercial property, domestic real estate on the other hand is not mentioned. The factsheet states that since "the index is comprised primarily of REITs, the prices of the component stocks reflect changes in lease rates, vacancies, property development and transactions" (Dow Jones U.S. Real Estate Index Factsheet) The index is real-time, as it is calculated every 15 seconds when the market is open, and at the end of the trading day high, low and close prices are published. It is quoted in US dollars. The U.S. Real Estate index is a subindex of the Dow Jones U.S. index, which belongs to the family of indices called the Dow Jones Global Indexes. For this family, more specific information on index calculation is available, which will be presented next.

For each share in the index, the number of shares and the market float are determined every day with the help of stock exchanges, data vendors and companies themselves. Also possible corporate action information is included in the calculation of the divisor. Both a price index and a return index are calculated, and they are both calculated using a Laspeyres formula, although modified to suit the calculation of a stock index. The only difference between the price and return indices is that the divisor is different. In the price index and its divisor, any dividend payments are not taken into account. For the return index, dividend payments are reinvested in the index sample. If the dividend is larger than 10 % of the equity price, it is considered a special dividend, type of a corporate action. As a result the divisor is adjusted accordingly. (Dow Jones U.S. Real Estate Index Factsheet 2008) The index formula for the Dow Jones U.S. Real Estate Index is as follows:

$$\begin{aligned} \text{DJUSRE Index} &= \frac{\sum_{i=1}^n (p_i^t \times q_i^t)}{C^t \times \sum_{i=1}^n (p_i^0 \times q_i^0)} \times \text{Base index value} \\ &= \frac{M^t}{B^t} \times \text{Base index value} = \frac{M^t}{D^t} \end{aligned} \tag{13}$$

Where :

$$D^t = \frac{B^t}{\text{Base index value}} = \text{Divisor at time } t$$

$n$  = Number of stocks in the index

$p_i^t$  = The closing price of stock  $i$  at the base date (December 31, 1991)

$q_i^t$  = The number of shares of company  $i$  at the base date (December 31, 1991)

$p_i^0$  = The price of stock  $i$  at time  $t$

$q_i^0$  = The number of shares of company  $i$  at time  $t$

$C^t$  = The adjustment factor for the base date market capitalisation

$t$  = The time the index is computed

$M^t$  = Market capitalisation of the index at time  $t$

$B^t$  = Adjusted base date market capitalisation of the index at time  $t$

From the index formula we can see that it is a modified version of the Laspeyres formula presented before. The use of divisor and an adjustment factor for the base date market capitalisation remove the problem of substitution of goods, discussed in the theory section. The industry indices, such as the Dow Jones Real Estate Index are constructed through categorising the component stocks of the Dow Jones US Index into 10 industries, 19 supersectors, 41 sectors and 114 subsectors, according to the Industry Classification Benchmark. The Dow Jones US Index is calculated based on these subindices, and since the same formula is used for each index calculation, this group of indices proves what was stated earlier – the Laspeyres index formula is consistent when aggregated.

The Dow Jones U.S. Real Estate Index is a commonly used index, as either a benchmark or a basis for an investment product. Only recently, Nordea (2008) issued a structured note linked to the return of this index. There is also an exchange traded fund linked to the performance of this index, the iShares Dow Jones U.S. Real Estate Index ETF (Bloomberg ticker IYR US). It seeks “investment results that correspond to the performance of the Dow Jones U.S. Real Estate Index” through approximately matching the holdings of the index (Bloomberg 2008 – IYR US Description). The fund’s market capitalisation was on 1.12.2008 1,5 billion US dollars. According to Bloomberg (2008), the fund’s correlation with the index is 1. Thus, even though it is stated that the fund follows approximately the index, in reality it would seem that the return profiles are identical. The longest option contracts linked to the iShares Dow Jones US Real Estate Index Fund (ETF) quoted on the market expire in January 2011 (Bloomberg 2008, IYR US Option Monitor).

In essence, even though the Dow Jones U.S. Real Estate Index aims to reflect the growth of the commercial real estate sector, it has in some cases been used for overall real estate sector or residential real estate sector purposes either in the form of a benchmark or in liaison with a structured product. One reason behind the popularity may be the fact that the Dow Jones company is a trusted and well-known agent in the market. Another, more significant reason may however be the availability of options for such a long maturity. This availability of options is perhaps aided by the fact that they are linked to a liquid underlying, which can also be traded – the ETF. Holders of shares in the ETF may wish to use options for a multitude of purposes, for example to hedge their position, to leverage the position or to take arbitrage opportunities. This market creates more demand for options linked to the ETF. Thus, there may be several different kinds of reasons behind purchasing options on an ETF, but purchasing options on an index results almost purely from speculative purposes. As a result, the markets for options on indices are most probably usually smaller than the markets for options on ETFs, even though the return profiles may be the same. In conclusion, it can be stated that the Dow Jones US Real Estate Index is a very strong candidate for the product due to the good availability of long maturity options. It is in the benefit of the investor that the product is constructed upon instruments which are liquid and efficiently priced in the market, eliminating the need for any kind of risk premiums. It is, however, also in the benefit of the investor that the chosen index reflects the underlying accurately – this issue will be examined in more detail in the coming empirical section.

### 3.2.2 Philadelphia Housing Sector Index

The Philadelphia Housing Sector Index is calculated by the Philadelphia Stock Exchange, which has recently been acquired by the Nasdaq OMX group. It is a modified capitalisation weighted index, and it includes companies whose primary lines of business are construction, development, support and sales relating to the residential housing industry. (Bloomberg Database – HGX Index Description 2008) There is no exchange traded fund linked to this index, but there are options and futures available. The index is quoted in USD. For this index, the index formula used for calculation is not explicitly stated. However, an example of the calculation method is given, from which the formula can be extracted. This example is presented below in table 1:

Table 1: Philadelphia Housing Sector Index contents on 05.12.2008

Company Name	Symbol	Closing			Market Value, USD	Market Percentage
		Price, USD	Index Shares			
Centex Corporation	CTX	11,72	87 138 772	1 021 266 407	2,65 %	
D.R. Horton, Inc.	DHI	8,17	240 854 626	1 967 782 294	5,11 %	
Hovnanian Enterprises Inc	HOV	2,34	206 626 972	483 507 114	1,26 %	
KB Home	KBH	14,20	83 399 787	1 184 276 975	3,08 %	
Lennar Corporation	LEN	9,06	131 496 189	1 191 355 472	3,10 %	
Lennox International, Inc.	LII	27,54	207 879 230	5 724 993 994	14,88 %	
Masco Corporation	MAS	10,18	256 994 223	2 616 201 190	6,80 %	
M.D.C. Holdings, Inc.	MDC	33,57	87 162 069	2 926 030 656	7,60 %	
M/I Homes, Inc.	MHO	11,56	17 781 124	205 549 793	0,53 %	
Meritage Corporation	MTH	13,36	18 903 787	252 554 594	0,66 %	
Palm Harbor Homes, Inc.	PHHM	8,20	18 721 538	153 516 611	0,40 %	
Pulte Homes, Inc.	PHM	11,86	195 801 114	2 322 201 212	6,04 %	
PMI Group, Inc. (The)	PMI	1,82	110 546 444	201 194 528	0,52 %	
Radian Group Inc.	RDN	3,46	66 133 092	228 820 498	0,59 %	
Ryland Group, Inc. (The)	RYL	18,89	63 812 850	1 205 424 736	3,13 %	
Standard Pacific Lp	SPF	2,03	223 452 400	453 608 372	1,18 %	
Temple-Inland Inc.	TIN	2,94	128 876 258	378 896 198	0,98 %	
Toll Brothers Inc.	TOL	22,03	231 155 356	5 092 352 492	13,23 %	
Vulcan Materials Company	VMC	66,75	70 954 869	4 736 237 505	12,31 %	
Weyerhaeuser Company	WY	35,39	173 242 703	6 131 059 259	15,93 %	
<b>SUM</b>			<b>2 620 933 403</b>	<b>38 476 829 900</b>	<b>100,00 %</b>	

In the above table, the contents and the basis for the calculation of the index are presented. The formulas used are presented next:

Market value = Closing price × Index shares

$$\text{Market percentage} = \frac{\text{Stock market value}}{\text{Total market value}} \times 100\%$$

(14)

Divisor on 5.12.2008 = 444 269 565

$$\text{Index value on 5.12.2008} = \frac{\text{Total market value}}{\text{Divisor}} = \frac{38\,476\,829\,900}{444\,269\,565} = 86,61$$

When comparing this calculation method to that of the Dow Jones index, it can be concluded that it is nearly identical. Even though it is not explicitly stated, the index formula used to calculate the Philadelphia Housing Sector Index is a Laspeyres index formula, modified to suit the purposes of stock market index calculation similarly as for the Dow Jones index.

According to Bloomberg (2008) and the Nasdaq OMX group website (2008), there is no exchange-traded fund available on the index. There are, nonetheless, options available directly on the index. According to Bloomberg Options Monitor (2008), in December 2008 the longest available maturity for these options was June 2009. The same information source shows also that even though there are prices quoted until June 2009, there are no volumes for such long dates. For the purposes of an index-linked note, which seeks to gain from longer-term trends, maturities of 6 months or less are not sufficient. Thus, the issuer of an index-linked note who decides to use this index as the underlying would have to rely on over-the-counter options which are tailored to the maturity needs of the issuer. The important point is that there is no secondary market for these kinds of OTC options, as there is for options traded in the options markets. In the case the issuer needs to liquidate the options before maturity, for example if a client decides to redeem the note in advance, the issuer most probably would have a hard time finding any other buyer for the option than the counterparty of the OTC option contract. In this kind of a case, the counterparty has significant bargaining power, which can result in making the unravelling of the note before maturity very expensive.

### 3.2.3 S&P/Case-Shiller Composite Home Price Indices

The S&P/Case-Shiller Home Price Indices “are designed to measure the growth in value of residential real estate in various regions across the United States”, as defined in the S&P/Case-Shiller factsheet (2008, 1). The 20 metropolitan regional indices are calculated monthly, from which 3 composite indices are calculated. One of the composites comprises of 10 of the metropolitan areas, including the largest US cities such as New York, Los Angeles, Chicago, San Francisco and Washington D.C. The composite of 20 includes, in addition to the 10 largest cities, metropolitan areas such as Dallas, Detroit, Minneapolis and Seattle. The third composite is the US national composite, calculated only quarterly, the other composites and metropolitan indices are calculated monthly. It is “a broader composite of single-family home price indices for the nine US Census divisions” (S&P/Case-Shiller Home Price Indices Factsheet 2008, 1)

The calculation of these indices uses a technique called repeat sales pricing method, a methodology constructed by Karl E. Case and Robert J. Shiller to reflect

and measure housing price movement. The method includes collecting data on sale prices of single-family homes in each region, where each sale is one data point. When a specific home is resold, the new sale price and the previous sale price create a sales pair. Sales pairs are designed to “yield the price change for the same house, while holding the quality and size of each house constant” (S&P/Case-Shiller Home Price Indices Methodology 2008, 7). When the differences in the prices of each sales pair are recorded, they can be aggregated into an index. To adjust for quality changes such as remodelling or adding a home addition, in other words, for changes in price caused by other factors than the housing market, the sales pairs are weighted. In other words, the index calculation method compares each sale pair to the average sales pairs in the community. The sale pair’s weight is decreased if it deviates greatly from the average value. Also, the longer the time between the two sales, the less weight will that sales pair have in the index, accounting thus for the probability of physical changes in the property. (S&P/Case-Shiller Home Price Indices Factsheet 2008) Sale prices associated with new construction, condominiums, co-ops/apartments, multi-family dwellings, or other properties which cannot be identified as single-family are not included in the index calculations. Moreover, pairs of sales with very short time intervals are excluded, as well as are unusually low sale prices due to non-arms-length transactions (sold between relatives) (S&P/Case-Shiller Home Price Indices Methodology 2008).

The actual calculation is done monthly, but by using a three-month moving average algorithm. In other words, “home sales pairs are accumulated in rolling three-month periods, on which the repeat sales methodology is applied” (S&P/Case-Shiller Home Price Indices Methodology 2008, 7). For example, “December 2005 index point is based on repeat sales data for October, November and December of 2005” (S&P/Case-Shiller Home Price Indices Methodology 2008, 7). This technique has been chosen to account for the delays in obtaining the sales data and also to keep the sample sizes large enough.

The calculation of the individual metropolitan area home price indices is built on a rather complicated method called the Robust Interval and Value-Weighted Arithmetic Repeat Sales algorithm (Robust IVWARS). This method, due to its complexity, is only very briefly described along with a simple formula illustrating the functioning of the index calculation. Firstly, the reason behind using interval weights

is that it corrects for errors that arise in repeat sales pairs due to the length of time between transactions. The base period for all the indices is January 2000, where index point equals 100. All index points prior to the base period have been estimated with a weighted regression model not illustrated here. After the base period, the index points are estimated using a chain-weighting procedure. This chain-weighting makes any index point “conditional on all previous index points, but independent of all subsequent index points” (S&P/Case-Shiller Home Price Indices Methodology 2008, 22). According to the S&P/Case-Shiller home price indices methodology guide (2008, 25), the purpose of a post-base, chain-weighting procedure is to “limit revisions to recently estimated index points while maintaining accurate estimates of market trends” (S&P/Case-Shiller Home Price Indices Methodology 2008, 25). The formula for the simultaneous, post-base index estimation is presented for illustration here:

$$\text{Index}_t = \frac{\sum_{n \in t} w_n P_{n\tau(2,n)}}{\sum_{n \in t} w_n P_{n\tau(1,n)} / \text{Index}_{\tau(1,n)}}$$

Where :

P = Price

$\tau(2,n)$  = Period of the second sale

$\tau(1,n)$  = Period of the first sale

$n \in t$  = Set of pairs with second sales in period t

(15)

The simple formula for calculating the composite indices is as follows:

$$\text{Index}_{Ct} = \frac{\sum (\text{Index}_{it} / \text{Index}_{i0}) \times V_{i0}}{\text{Divisor}}$$

Where :

$\text{Index}_{Ct}$  = Level of the composite index in period t

$\text{Index}_{it}$  = Level of the home price index for  
a specific metro area i in period t (16)

$\text{Index}_{i0}$  = Level of the home price index in the base period 0

$V_{i0}$  = Aggregate value of housing stock in  
the specific metro area i in the base period 0

The divisor in the equation is chosen so that the measure of aggregate housing value (the numerator in the fraction above) is converted into an index number with the same base value as the metro area indices. The numerator is in fact an estimate of the aggregate value of housing stock for all metro areas in a composite index. (S&P/Case-Shiller Home Price Indices Methodology 2008, 11) According to the S&P/Case-Shiller Home Price Indices Methodology guide (2008, 11), the composite home price indices are “analogous to a cap-weighted equity index, where the aggregate value of housing represents the total capitalisation of all of the metro areas included in the composite.” Even though these index formulae are rather complicated to derive, their final appearance is very similar to that of a simple Laspeyres index and the most common stock market indices.

### 3.3 Concluding remarks on the theoretical and practical aspects of indices

The theoretical discussion came to the conclusion that current theories are very much divided upon the issue of best index formula. It was decided that the choice of the best index formula and the best index should depend on which stock market index most accurately reflects the reality. As can be observed from this chapter, the two stock market indices are very similar in their method of calculation – there is no need for comparing the index formulae as they both use the Laspeyres formula.

However, there are major differences between the two stock market indices. The Dow Jones U.S. Real Estate Index has an exchange traded fund linked to it, upon which there are options available. The option liquidity is larger and maturities are much longer than those of Philadelphia Housing Sector Index-linked options. Since the structure of this index-linked note is very simple, consisting only of a zero coupon bond and a call option on the underlying index, the pricing of the product is highly sensitive to the pricing of the option.

Even though index-linked notes are supposed to be held for the whole maturity, the clients are often given the opportunity to redeem the investment before maturity. This provision of liquidity for the client leads to the issuer being forced to guarantee liquidity for the different parts of the product. The funds from zero coupon bonds need to be invested in relatively short-term maturity products, and the options used need to have a liquid secondary market. The issuer could remove the redemption possibility, but this may severely limit investors' interest in the product. As discussed above, the existence of an exchange traded fund creates a large advantage as compared to options linked directly to the index through increasing the demand for the options.

It may be that whatever the result of the empirical analysis next, the only feasible choice for this index-linked note is the Dow Jones U.S. Real Estate Index due to the option pricing issues. The fact that there are no long maturity indices available in the market for the Philadelphia Housing Sector index would force the issuer of a structured product to use tailored options, which are less liquid and may not be efficiently priced. On the other hand, there are options available in the market for the Dow Jones US Real Estate Index ETF for maturities up to almost two years. The following chapter will evaluate what truly is the relationship between the stock market indices and the S&P/Case-Shiller Home Price composites. Moreover, despite the seemingly poor suitability of the Dow Jones index for this product, the empirical analysis will look into whether the Dow Jones U.S. Real Estate Index has a significant enough correlation with the real movements in the housing prices so that it could be used for this product. After all, it is in the advantage of the client as well as the issuer that the product and its parts are efficiently and this correctly priced in the market.

In order to complement the analysis done in this thesis and in order to bring this mostly theoretical study closer to reality, a structured product professional, who will remain anonymous, has been interviewed. He is an experienced structured product specialist, who has worked for several Finnish investment banks and financial institutions. His expertise includes constructing and analysing a variety of structured products, especially index-linked notes. He concurs with the above statement about the necessity of liquidity in the options market (Interview, 10.2.2009). The existence of an ETF on itself is not a necessity, but he states that a liquid market provides for better options prices. Furthermore, according to the professional, “many counterparties need a certain level of liquidity before they are able/willing to quote on some market” (Interview, 10.2.2009). He also confirms that liquidity guarantees reasonable secondary market prices. (Interview, 10.2.2009) Thus, the best solution in theory may prove to be an unfeasible solution in reality due to factors such as lack of liquidity or other market issues.

## 4 EMPIRICAL ANALYSIS

### 4.1 Description of the data set

#### 4.1.1 Data set

The data set consists of daily, monthly and quarterly closing prices for the Dow Jones U.S. Real Estate Index, the Philadelphia Housing Sector Index and the three S&P/Case-Shiller Composite Home Price Indices. The Case-Shiller Indices are the composites for the 10 largest cities in the United States, the composite for the 20 largest cities and the national composite index. The data is from the time period 3.7.2002 – 31.10.2008 because it is the longest time that there is data available for all the indices (the Philadelphia Housing Sector Index is the youngest). It would have been ideal to find data from early 1990's onwards, since the previous housing price bottom. Nonetheless, since the data was not available for the one of the key indices for that long of a time period, this data set will have to suffice. The source of the data is the Bloomberg Online Database, accessed during the fall 2008. In the remained of this empirical section, the Bloomberg ticker codes will be used for the indices, they are as follows:

Table 2: Bloomberg tickers for the indices

<b>Index name</b>	<b>Bloomberg ticker</b>	<b>Data</b>
Philadelphia Housing Sector Index	HGX Index	Daily
Dow Jones US Real Estate Index	DJUSRE Index	Daily
S&P/Case-Shiller Composite-10 Home Price Index	SPCS10 Index	Monthly
S&P/Case-Shiller Composite-20 Home Price Index	SPCS20 Index	Monthly
S&P/Case-Shiller Composite USA Home Price Index	SPCSUSA Index	Quarterly

All the calculations have been made from rate of return figures, calculated based on the daily, monthly and quarterly close prices. Also annual rate of return figures have been calculated.

## 4.1.2 Descriptive statistics

The most important statistical information on the individual time series is outlined in table below.

Table 3: Statistical characteristics of the data set

	Mean	Median	Standard Deviation (Volatility)	Min	Max
<b>HGX Index</b>					
Daily return	0,00 %	0,00 %	2,0 %	-11,8 %	9,6 %
Monthly return	0,26 %	1,22 %	7,1 %	-18,4 %	15,0 %
Quarterly return	0,94 %	1,19 %	13,3 %	-24,8 %	30,7 %
Annual return	8,33 %	10,80 %	33,5 %	-48,8 %	86,4 %
<b>DJUSRE Index</b>					
Daily return	0,01 %	0,02 %	1,7 %	-13,2 %	18,3 %
Monthly return	0,76 %	1,71 %	4,9 %	-15,5 %	10,9 %
Quarterly return	2,19 %	0,78 %	8,2 %	-14,0 %	14,1 %
Annual return	12,13 %	17,66 %	17,9 %	-29,0 %	44,3 %
<b>SPCS10 Index</b>					
Monthly return	0,38 %	0,70 %	1,2 %	-2,8 %	2,3 %
Quarterly return	1,21 %	2,37 %	3,5 %	-7,3 %	6,9 %
Annual return	6,62 %	12,31 %	12,2 %	-17,7 %	20,5 %
<b>SPCS20 Index</b>					
Monthly return	0,33 %	0,64 %	1,1 %	-2,6 %	2,0 %
Quarterly return	1,08 %	2,05 %	3,2 %	-6,9 %	6,0 %
Annual return	5,93 %	10,57 %	11,0 %	-16,6 %	17,1 %
<b>SPCSUSA Index</b>					
Quarterly return	0,95 %	1,94 %	3,0 %	-6,7 %	4,6 %
Annual return	5,55 %	10,20 %	10,4 %	-15,4 %	15,7 %

One of the most important aspects to look at in the statistical data table is the historical volatility. Even though it is not a guarantee of future volatility, it can be used as an indicator. The volatility affects strongly the pricing of the derivative. The mechanism will be explained later in the construction of the case study product section, but in short: *ceteris paribus*, the higher the volatility, the higher the cost of the option. An option on the Philadelphia Housing Sector Index would be more costly than an option on the Dow Jones Index, assuming all other things are equal. When compared to the S&P/Case-Shiller indices, the volatilities of the stock market indices are significantly higher than those of the direct home price indices. This may imply that the stock market indices are affected by other factors, such as market sentiment. The reason may also lie in the fact that it is much easier to sell or buy a share than a house – a single stock can exchange hands several times a day or even an hour, but the timeline for a house exchanging hands is measured in months

or years. Moreover, we can see that the mean and median returns are almost across the board higher for the Dow Jones Index than the HGX Index, even though it was the HGX Index which had higher volatility. From this set of descriptive statistics it would seem in fact that the HGX returns are closer to the S&P/Case-Shiller composites' returns.

#### 4.1.3 Returns of the indices

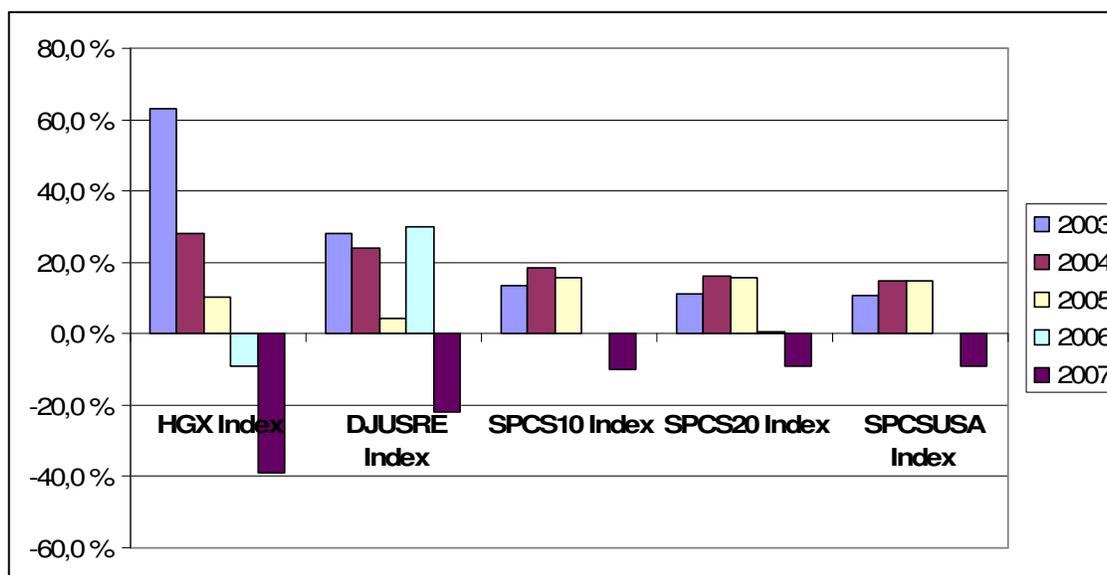
The full year annual returns are illustrated in the table and figure below. From these we can see that the HGX Index has yielded the most in 2003 and the least in 2007 out of all the indices. Examining these figures we can see that the HGX Index in fact has more extreme returns, positive and negative, even though in the descriptive statistics it seemed that the HGX Index had the most moderate returns. Since the HGX Index has been more volatile, these returns seem reasonable, and it may be that when calculating the mean the extreme outcomes cancel each other, leading to a distorted figure.

Table 4: Annual returns

Year	ANNUAL RETURNS				
	HGX Index	DJUSRE Index	SPCS10 Index	SPCS20 Index	SPCSUSA Index
<b>2003</b>	63,0 %	28,4 %	13,4 %	11,4 %	10,7 %
<b>2004</b>	28,1 %	24,2 %	18,7 %	16,2 %	14,6 %
<b>2005</b>	10,4 %	4,1 %	15,9 %	15,5 %	14,7 %
<b>2006</b>	-9,1 %	29,7 %	0,2 %	0,7 %	0,2 %
<b>2007</b>	-38,9 %	-21,7 %	-9,8 %	-9,0 %	-8,9 %

From the figure depicting the annual returns we can also see the peak in the real estate prices in the United States. In 2006, the HGX Index was the first to turn negative, but all the other indices quickly followed in 2007.

Figure 2: The annual returns of the indices under study, years 2003-2007



#### 4.1.4 Correlations

The first step in the empirical analysis is to examine the correlations between the index return time series, generated from the price data from Bloomberg database.

Table 5: Correlations between the monthly returns of the indices

	MONTHLY RETURNS				
	HGX Index	DJUSRE Index	SPCS10 Index	SPCS20 Index	SPCSUSA Index
<b>HGX Index</b>	1,000	0,601	0,192	0,183	0,112
<b>DJUSRE Index</b>	0,601	1,000	0,189	0,195	0,220
<b>SPCS10 Index</b>	0,192	0,189	1,000	0,997	0,533
<b>SPCS20 Index</b>	0,183	0,195	0,997	1,000	0,534
<b>SPCSUSA Index</b>	0,112	0,220	0,533	0,534	1,000

Table 6: Ranked correlations between the monthly returns of the indices

In order of correlation:					
	SPCS10 Index		SPCS20 Index		SPCSUSA Index
1.	<b>HGX Index</b>	0,192	<b>DJUSRE Index</b>	0,196	<b>DJUSRE Index</b>
2.	<b>DJUSRE Index</b>	0,190	<b>HGX Index</b>	0,183	<b>HGX Index</b>
					0,112

From the tables above we can see that the correlations of monthly returns between the HGX Index and the S&P/Case-Shiller Indices are quite low, and except for the SPCS10 Index, they are lower than the same for the DJUSRE Index.

Table 7: Correlations between the quarterly returns of the indices

	QUARTERLY RETURNS				
	HGX Index	DJUSRE Index	SPCS10 Index	SPCS20 Index	SPCSUSA Index
<b>HGX Index</b>	1,000	0,651	0,470	0,448	0,484
<b>DJUSRE Index</b>	0,651	1,000	0,367	0,370	0,416
<b>SPCS10 Index</b>	0,470	0,367	1,000	0,997	0,990
<b>SPCS20 Index</b>	0,448	0,370	0,997	1,000	0,994
<b>SPCSUSA Index</b>	0,484	0,416	0,990	0,994	1,000

Table 8: Ranked correlations between the quarterly returns of the indices

In order of correlation						
	SPCS10 Index		SPCS20 Index		SPCSUSA Index	
1.	<b>HGX Index</b>	0,470	<b>HGX Index</b>	0,448	<b>HGX Index</b>	0,484
2.	<b>DJUSRE Index</b>	0,367	<b>DJUSRE Index</b>	0,370	<b>DJUSRE Index</b>	0,416

For quarterly data the correlations are higher, but still not significant. For this data, however, the HGX Index correlates more with all the composites than the DJUSRE Index. The case study product could use quarterly or even semi-annual evaluation points, which would make the quarterly return data more significant for the purposes of the case study.

Table 9: Correlations between the annual returns of the indices

	ANNUAL RETURNS				
	HGX Index	DJUSRE Index	SPCS10 Index	SPCS20 Index	SPCSUSA Index
<b>HGX Index</b>	1,000	0,728	0,818	0,800	0,449
<b>DJUSRE Index</b>	0,728	1,000	0,712	0,721	0,382
<b>SPCS10 Index</b>	0,818	0,712	1,000	0,997	0,511
<b>SPCS20 Index</b>	0,800	0,721	0,997	1,000	0,513
<b>SPCSUSA Index</b>	0,449	0,382	0,511	0,513	1,000

Table 10: Ranked correlations between the annual returns of the indices

In order of correlation						
	SPCS10 Index		SPCS20 Index		SPCSUSA Index	
1.	<b>HGX Index</b>	0,818	<b>HGX Index</b>	0,800	<b>HGX Index</b>	0,449
2.	<b>DJUSRE Index</b>	0,716	<b>DJUSRE Index</b>	0,721	<b>DJUSRE Index</b>	0,382

The annual returns of the indices correlate much more, as could be expected, than the monthly or quarterly returns. Again, it would seem that the HGX Index follows the composites more closely than the DJUSRE Index. However, the correlation analysis is not sufficient. Further statistical testing needs to be done to estimate the properties of the data set and the relationships between the different time series. Moreover, the stationarity of the time series data needs to be tested also to find out

whether the correlations are reliable estimates and whether there is any kind of a relationship between the returns of these indices.

## 4.2 Empirical analysis

### 4.2.1 Background

Even though in previous discussion on correlations suggested that the Philadelphia Housing Sector Index followed more closely the S&P/Case-Shiller composites, the starting point and initial assumption for this data analysis is that it would be significantly cheaper to construct the index-linked note upon the Dow Jones U.S. Real Estate Index. This is due to the availability of options, as explained in conjunction with the descriptions of the Dow Jones U.S. Real Estate and Philadelphia Housing Sector indices. Even though the Dow Jones U.S. Real Estate Index does not include residential real estate, and thus in theory, its value development and return profile should not match the development of the residential real estate prices. However, the Dow Jones index is used in several residential real estate index-linked notes, and it might even be that using the Philadelphia Housing Sector index with the index-linked note would make the note too expensive to issue. As a reflector of the reality of house prices in the US, the S&P/Case-Shiller Composite-20 Home Price Index is used. The national composite, the S&P/Case-Shiller Composite USA Home Price Index would otherwise be more suitable, but has the problem that the number of observations is limited to 23 because it is calculated only quarterly. The S&P/Case-Shiller Composite-20 Home Price Index is calculated monthly, taking the number of observations to 74. This is still a rather small sample, but is significantly larger than a sample of 23.

Therefore, the focus is to examine whether the relationship between the Dow Jones Real Estate Index time series and the S&P/Case-Shiller 20 composite index is strong enough to justify the use to the Dow Jones index in the note. It is therefore hypothesised here that the commercial real estate and residential real estate, even though separate sectors, are correlated so that their price developments follow each other closely. It should be noted that even if there is a historical relationship between the two, it is not a guarantee of future correlation. This is especially important in this analysis, since not only are we hypothesising on the relationship between two time

series, we are also making assumptions about the relationship between two asset classes.

#### 4.2.2 Model specification

In the regression analysis to follow, the dependent variable will be the Dow Jones U.S. Real Estate Index and the independent variable is the S&P/Case-Shiller Composite-20. Thus, it is hypothesised that the Dow Jones Real Estate Index follows the return profile of the S&P/Case-Shiller Composite-20 with no time lag. No time lag is assumed, because the stock market is taken as efficient. An error term is added to the equation. The relationship and hypothesised signs can be represented as follows:

$$\begin{matrix} (+) \\ DJRE = f(SP20) + e \end{matrix} \quad (17)$$

Where:

DJRE = Monthly return of the Dow Jones U.S. Real Estate Index

SP20 = Monthly return of the S&P/Case-Shiller Composite-20 Home Price Index

The functional form of the regression will thus take the form:

$$DJRE^t = \beta_0 + \beta_1(SP20^t) + e \quad (18)$$

The regression model is taken to be linear, but the error term is included to account for the possibility that the underlying equation is not linear, and also to take into account random variation. (Studenmund 2006, 11)

#### 4.2.3 Data analysis

The data was estimated by using EViews with the Ordinary Least Squares method, figure 3 illustrates the results.

Figure 3: Regression output

Dependent Variable: DJRE  
 Method: Least Squares  
 Date: 01/22/09 Time: 12:40  
 Sample: 1 74  
 Included observations: 74  
 DJRE=C(1)+C(2)\*SP20

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.004487	0.005821	0.770922	0.4433
C(2)	0.886386	0.526185	1.684552	0.0964
R-squared	0.037918	Mean dependent var		0.007398
Adjusted R-squared	0.024556	S.D. dependent var		0.048415
S.E. of regression	0.047817	Akaike info criterion		-3.216203
Sum squared resid	0.164628	Schwarz criterion		-3.153931
Log likelihood	120.9995	Hannan-Quinn criter.		-3.191362
F-statistic	2.837714	Durbin-Watson stat		1.995834
Prob(F-statistic)	0.096406			

The estimated equation is (rounded to three significant figures):

$$\text{DJRE} = 0,004 + 0,886(\text{SP20}) \quad (19)$$

Where:

Standard error = (0,526)

t = 1,685

n = 74

$\bar{R}^2 = 0,025$

The initial reaction is that the overall fit of the equation is rather poor. Both  $R^2$ , the coefficient of determination, and adjusted  $R^2$  very low. This refers to the fact that the equation used does not reflect the underlying relationship accurately. However, this is not the only measure of the goodness of fit of the regression. The F-statistics needs to also be examined.

The null hypothesis of the F-test of overall significance is that all the slope coefficients in the equation (in this case one) equal zero simultaneously. According to Studenmund (2006, 154), "the F-test is really testing the null hypothesis that the

fit of the equation isn't significantly better than that provided by using the mean alone". To estimate the critical F-statistic, degrees of freedom need to be calculated. Here,  $k$  is the number of independent variables and  $n$  is the number of overall observations.

$$\begin{aligned}\text{Degrees of freedom:} \quad v_1 &= k = 1 \\ v_2 &= n - k - 1 = 74 - 1 - 1 = 72\end{aligned}$$

The critical F-statistic at 10 % level of significance is 2,78. The F-statistic of the regression is 2,84, which is higher than the critical value. Thus, the null hypothesis that the coefficients of the equation are zero can barely be rejected, and the overall fit of the equation would seem to be significant at the 90 % level of confidence.

In order to estimate the nature of the coefficient, the t-test can be used.

$$\text{Degrees of freedom: } n - k - 1 = 74 - 1 - 1 = 72 \approx 60$$

The critical t-value at 10 % level of significance (for a two sided test) is 1,671, which is less than the t-statistic for the coefficient of SP20, 1,685. Thus, at 90% level of confidence, it can be stated that the coefficient of SP20 is significantly different from zero and statistically significant in the hypothesised direction.

Especially with econometric time series, the variance of the error term is often not constant. This phenomenon is called heteroscedasticity. To test for heteroscedasticity, the White heteroskedasticity test is used, as illustrated in figure 4 below:

Figure 4: The White Heteroscedasticity test

Heteroskedasticity Test: White

F-statistic	0.738443	Prob. F(2,71)	0.4815
Obs*R-squared	1.507923	Prob. Chi-Square(2)	0.4705
Scaled explained SS	2.685448	Prob. Chi-Square(2)	0.2611

Test Equation:  
 Dependent Variable: RESID^2  
 Method: Least Squares  
 Date: 01/22/09 Time: 14:09  
 Sample: 1 74  
 Included observations: 74

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.001616	0.000714	2.262021	0.0268
SP20	0.033478	0.048863	0.685150	0.4955
SP20^2	4.078669	3.657470	1.115161	0.2685

R-squared	0.020377	Mean dependent var	0.002225
Adjusted R-squared	-0.007218	S.D. dependent var	0.004345
S.E. of regression	0.004360	Akaike info criterion	-7.992839
Sum squared resid	0.001350	Schwarz criterion	-7.899431
Log likelihood	298.7350	Hannan-Quinn criter.	-7.955577
F-statistic	0.738443	Durbin-Watson stat	2.074380
Prob(F-statistic)	0.481491		

The overall significance can be calculated with the chi-square test. The test statistic is calculated by multiplying the sample size with the coefficient of determination,  $R^2$ .

$$\text{Test statistic} = N \times R^2 = 74 \times 0,020377 = 1,507898 \approx 1,508$$

$$\text{Degrees of freedom} = \text{Number of slope coefficients} = 1$$

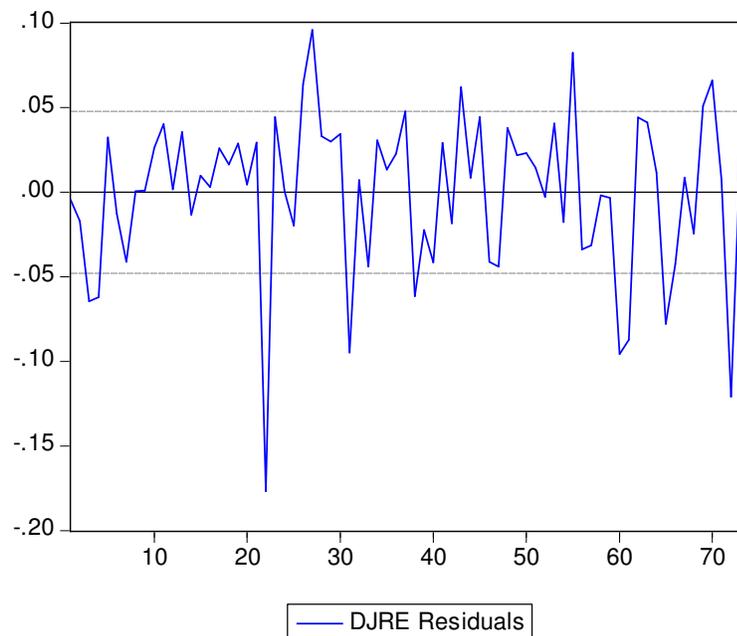
$$\text{The critical } \chi_2 \text{ value at 5 \% level of significance} = 3,84$$

Since the  $NR^2$  is not larger than the critical test value, we cannot reject the null hypothesis of homoscedasticity. In other words, at the 95 % level of confidence, there is no heteroscedasticity. (Studenmund 2006)

One needs to also test for serial correlation, which occurs frequently in time series data sets. Serial correlation is defined as follows: "If the expected value of the simple correlation coefficient between any two observations of the error term is not equal to zero, then the error term is said to be serially correlated" (Studenmund 2006, 314). Pure serial correlation may be caused by an underlying distribution of the error term, which is an issue which cannot be changed. On the other hand, impure serial correlation is most often caused by some kind of a specification error which can usually be corrected. While pure serial correlation does not cause bias in the coefficient estimates, serial correlation does cause the OLS to no longer be the minimum variance estimator and the standard error estimates will be biased. (Studenmund 2006)

Serial correlation can be detected by either visually or with the help of the Durbin-Watson test statistic. The residual graph can be examined in attempts to find serial correlation.

Figure 5: Serial correlation - residual graph



Looking at the residual graph, it would seem that there is no serial correlation – if the error term were to be systematically determined by the earlier error term, there

would be serial correlation. To accurately estimate serial correlation, the Durbin-Watson statistic is used.

Durbin-Watson statistic = 1,996

$k' = 1$

$n = 74 \approx 75$

The critical upper and lower values for the d statistic at 10 % two-sided level of significance are:

$d_L = 1,60$

$d_U = 1,65$

The d statistic falls into the non-rejection region, and therefore at 90 % level of confidence, there is no serial correlation (at first lag).

The results of this regression analysis are, thus, that there is a positive relationship between the monthly return of the Dow Jones U.S. Real Estate Index and the S&P/Case-Shiller Composite-20 Home Price Index. Looking at the estimated equation, illustrated again below, one can see that a one unit change in the monthly return of the S&P/Case-Shiller Composite-20 Home Price Index leads to a 0,886 unit change in the monthly return of the Dow Jones U.S. Real Estate Index.

$$DJRE = 0,004 + 0,886(SP20) \quad (20)$$

Table 11: Example calculation from the regression equation

<b>Monthly return - Example with imaginary figures</b>		
<b>Month</b>	<b>S&amp;P/Case-Shiller Composite 20</b>	<b>Dow Jones U.S. Real estate</b>
1	-2,0 %	-1,4 %
2	-1,0 %	-0,5 %
3	1,0 %	1,3 %
4	3,0 %	3,1 %
5	5,0 %	4,8 %
6	4,0 %	3,9 %

In order to complete this empirical analysis on the relationship between the monthly returns of these two indices, one more issue needs to be examined – the stationarity of the time series.

#### 4.2.4 Testing for the stationarity in the time series

Standard statistical models assume time series to come from a stationary process, where means and variances are constant over time (Hendry et al 1999, 1). Granger explains, however, that “many series in economics, particularly in finance and macroeconomics, do not have this property and can be called integrated or, sometimes incorrectly, non-stationary” (Granger 2004, 421). The problems with integrated time series are multiple, as outlined by Hendry et al (1999) in their discussion paper “Explaining Cointegration Analysis: Part I”. First, when the means and variances of the data are non-constant, “the observations come from different distributions over time, posing difficult problems for empirical modelling” (Hendry et al 1999, 2). Secondly, making false assumptions about means and variances being constant can cause statistical mistakes. Also, it should be noted that economic and financial time series are susceptible to non-stationarity due to “evolutions of the economy, legislative changes, technological change and political turmoil” (Hendry et al 1999, 2). With economic and financial data, the issue normally is stochastic non-stationarity, which is induced by the cumulation of past effects (called unit-root processes). Stochastic refers to a process which can be interpreted as allowing a different trend at every point in time. Stochastic trends are not the only cause of integrated or non-stationary time series, but also a variable exhibiting a shift in its mean or having a heteroscedastic variance over time. (Hendry et al 1999, 4). The unit-root assumption implies “an ever increasing variance to the time series..., violating the constant-variance assumption of a stationary process” (Hendry et al 1999, 8). This phenomenon is caused by cumulating random errors, and also results in successive observations in the time series being highly interdependent. (Hendry et al, 1999)

It has, nonetheless, been discovered that some forms of non-stationarity may be eliminated by transformations. More specifically, when data are non-stationary purely due to unit roots, a linear transformation can be used to make the data stationary. In other words, the difference between a pair of integrated time series can be stationary – this is a property called cointegration. In a stationary process, the mean and variances are constant. A stationary process may be autocorrelated, but as long as the influence of any past shock dies out. Here, the augmented Dickey-Fuller test is used to test whether the time series used for above empirical

analysis are stationary. For the following test statistics, the decision rule is as follows:

If test critical value > ADF test statistic => Do not reject null hypothesis  
Unit root exists  
Time series is non-stationary

If test critical value < ADF test statistic => Reject null hypothesis  
Unit root does not exist  
Time series is stationary

Figure 6: The augmented Dickey-Fuller test for the Dow Jones U.S. Real Estate

### Index

Null Hypothesis: DJRE has a unit root  
Exogenous: Constant  
Lag Length: 0 (Automatic based on SIC, MAXLAG=11)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-8.038395	0.0000
Test critical values:		
1% level	-3.522887	
5% level	-2.901779	
10% level	-2.588280	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
Dependent Variable: D(DJRE)  
Method: Least Squares  
Date: 01/22/09 Time: 15:37  
Sample (adjusted): 2 74  
Included observations: 73 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
DJRE(-1)	-0.953037	0.118561	-8.038395	0.0000
C	0.007158	0.005803	1.233464	0.2215
R-squared	0.476462	Mean dependent var		0.000233
Adjusted R-squared	0.469088	S.D. dependent var		0.067291
S.E. of regression	0.049031	Akaike info criterion		-3.165733
Sum squared resid	0.170683	Schwarz criterion		-3.102981
Log likelihood	117.5493	Hannan-Quinn criter.		-3.140725
F-statistic	64.61579	Durbin-Watson stat		2.000110
Prob(F-statistic)	0.000000			

The null hypothesis of the time series having a unit root can be rejected at the 99 % level of confidence because the test statistic is smaller than the critical values at all levels. Therefore, this time series is stationary, next step is to evaluate the stationarity of the SP20 time series.

Figure 7: The augmented Dickey-Fuller test for S&P/Case-Shiller Composite-20

Null Hypothesis: SP20 has a unit root  
 Exogenous: Constant  
 Lag Length: 1 (Automatic based on SIC, MAXLAG=11)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-1.514010	0.5210
Test critical values:		
1% level	-3.524233	
5% level	-2.902358	
10% level	-2.588587	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(SP20)  
 Method: Least Squares  
 Date: 01/22/09 Time: 15:38  
 Sample (adjusted): 3 74  
 Included observations: 72 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
SP20(-1)	-0.041637	0.027501	-1.514010	0.1346
D(SP20(-1))	0.365855	0.097233	3.762668	0.0003
C	-0.000119	0.000303	-0.393784	0.6950
R-squared	0.177884	Mean dependent var		-0.000311
Adjusted R-squared	0.154054	S.D. dependent var		0.002644
S.E. of regression	0.002432	Akaike info criterion		-9.159639
Sum squared resid	0.000408	Schwarz criterion		-9.064778
Log likelihood	332.7470	Hannan-Quinn criter.		-9.121875
F-statistic	7.464869	Durbin-Watson stat		1.535199
Prob(F-statistic)	0.001162			

Here, on the other hand, the null hypothesis cannot be rejected because the test statistic is larger than the critical values at all levels. Thus, the S&P/Case-Shiller Composite-20 home price index monthly return time series is non-stationary.

One needs to now test whether the S&P/Case-Shiller Composite-20 time series is non-stationary also after taking the first difference.

Figure 8: The ADF test for S&P/Case-Shiller Composite-20, 1<sup>st</sup> difference

Null Hypothesis: D(SP20) has a unit root  
 Exogenous: Constant  
 Lag Length: 0 (Automatic based on SIC, MAXLAG=11)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-6.823193	0.0000
Test critical values:		
1% level	-3.524233	
5% level	-2.902358	
10% level	-2.588587	

\*MacKinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation  
 Dependent Variable: D(SP20,2)  
 Method: Least Squares  
 Date: 01/22/09 Time: 16:09  
 Sample (adjusted): 3 74  
 Included observations: 72 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
D(SP20(-1))	-0.659516	0.096658	-6.823193	0.0000
C	-0.000269	0.000289	-0.929473	0.3558
R-squared	0.399430	Mean dependent var		-0.000188
Adjusted R-squared	0.390851	S.D. dependent var		0.003144
S.E. of regression	0.002454	Akaike info criterion		-9.154736
Sum squared resid	0.000422	Schwarz criterion		-9.091496
Log likelihood	331.5705	Hannan-Quinn criter.		-9.129560
F-statistic	46.55596	Durbin-Watson stat		1.510371
Prob(F-statistic)	0.000000			

Taking the first difference of the time series corrects the non-stationarity problem. The functional form of the regression should thus be changed so that the first difference of the SP20 time series should be used in estimating.

## 4.2.5 Empirical analysis after adjusting for non-stationarity

The Ordinary Least Squares analysis was run again on EViews. First, it was run using the first difference of both time series. The results gave a strongly negative coefficient for the independent variable, and it was clear that the relationship between the two time series was destroyed by the alteration. Taking the first difference only of the SP20 time series gave slightly more sensible results, and they are presented here. Model specification remained constant except for the SP20, which was turned into a first difference time series. The new functional form is:

$$DJRE = \beta_0 + \beta_1(\Delta SP20) + e \quad (21)$$

Where  $\Delta SP20$  represents the first difference of the SP20 time series.

The regression output is illustrated in figure 9:

Figure 9: Regression output, adjusted for non-stationarity

Dependent Variable: DJRE  
 Method: Least Squares  
 Date: 01/22/09 Time: 16:31  
 Sample: 1 74  
 Included observations: 74  
 DJRE=C(1)+C(2)\*SP20

	Coefficient	Std. Error	t-Statistic	Prob.
C(1)	0.007398	0.005649	1.309522	0.1945
C(2)	1.202725	1.768536	0.680068	0.4986
R-squared	0.006383	Mean dependent var		0.007398
Adjusted R-squared	-0.007418	S.D. dependent var		0.048415
S.E. of regression	0.048595	Akaike info criterion		-3.183951
Sum squared resid	0.170024	Schwarz criterion		-3.121679
Log likelihood	119.8062	Hannan-Quinn criter.		-3.159110
F-statistic	0.462493	Durbin-Watson stat		1.898488
Prob(F-statistic)	0.498642			

The new estimated equation is:

$$DJRE = 0,007 + 1,203(\Delta SP20) \quad (22)$$

Where:

Standard error = 1,769

t = 0,680

n = 74

$\bar{R}^2 = -0,007$

Again, the fit of the equation seems to be poor when looking at R-squared and adjusted R-squared. The F-statistic is evaluated next:

$$\begin{aligned} \text{Degrees of freedom:} \quad v_1 &= k = 1 \\ v_2 &= n - k - 1 = 74 - 1 - 1 = 72 \end{aligned}$$

Again, the critical F-statistic at 10 % level of significance is 2,78. The F-statistic of this regression is significantly lower at 0,462, meaning that the null hypothesis that the coefficients of the equation are zero cannot be rejected. It is often the case, that while taking the first difference may correct the problem of non-stationarity, it may also destroy the relationship between the dependent and independent variables. It is, nonetheless, worth also testing the significance of the coefficients with the t-test.

$$\text{Degrees of freedom: } n - k - 1 = 74 - 1 - 1 = 72 \approx 60$$

The critical t-value at 10 % level of significance (for a two sided test) is 1,671, which is more than the t-statistic for the coefficient of SP20, 0,680. Thus, at 90% level of confidence the coefficient of SP20 is not significantly different from zero. Thus, taking the first difference of the SP20 time series has eroded the statistical relationship that seemed to hold in earlier analysis. Because correcting for such an advanced problem is beyond the scope of this paper, the choice of index will rely on the earlier analysis. This defect should, however, be kept in mind by the reader and it also does present an interesting topic for further study.

#### 4.3 Results of the empirical analysis

Even though the comparisons between the correlations of the stock market indices and the S&P/Case-Shiller composite indices show that the Philadelphia Housing Sector Index correlates more with the composites in most cases than the Dow Jones U.S. Real Estate Index, in no case are the correlations significantly high. The annual return correlations are, not surprisingly, the highest. However, at the level of annual returns, the observation group is very small and thus, this analysis cannot be deemed reliable. As the number of observations increases, for example when moving to quarterly returns, the correlations fall significantly. The results of comparing the simple correlations between the different indices are therefore not solid, and decision over the index cannot be made based on this analysis.

In this empirical analysis, it was then argued that the Dow Jones U.S. Real Estate Index should be chosen for the regression analysis because it is the more viable solution for the actual product. The reasons, which are to do with the pricing and the availability of options, have been explained in more detail earlier. It should be remembered also, that surely banks such as Nordea who have issued Real Estate index-linked notes would have preferred to use an index which through its name and its contents is more relevant to the product. They, and other issuers, have nonetheless chosen the Dow Jones U.S. Real Estate Index. As will be illustrated in the next section, explaining the construction of the product, the pricing of the final product depends on the prices of the options as well as on the projected interest rate for the maturity of the product. Thus, since market behaviour and the aforementioned arguments show that the prices of the options on iShares Dow Jones U.S. Real Estate ETF are significantly cheaper and more liquid; this product should also be constructed upon it as opposed to upon the Philadelphia Housing Sector Index.

All that was left was to show that there is a significantly strong positive relationship between the Dow Jones U.S. Real Estate Index and the S&P/Case-Shiller Composite-20. The results of the regression were that there is a positive relationship, with one unit change in the S&P/Case-Shiller Composite-20 monthly return leads to a 0,89 unit change in the Dow Jones U.S. Real Estate Index monthly return. It was then discovered that the S&P/Case-Shiller time series was not

stationary, which may jeopardise these regression results. Fixing this problem was, however, out of the reach for this paper. Moreover, the study may have been complicated by the fact that the two time series in the regression analysis are different in nature. The calculation of the S&P/Case-Shiller index uses a 3-month moving algorithm, which smooths out the time series. The Dow Jones US Real Estate Index, on the other hand, calculates each index value based only on the latest values. This method in the calculation of the S&P/Case-Shiller indices may in fact be the contributing factor to the non-stationarity of the SP20 time series. This problem, as well fixing the non-stationarity of the S&P/Case-Shiller time series are issues for further study. However, for the purposes of this paper, the initial regression results were the ones used for choosing the Dow Jones U.S. Real Estate Index for the case study index-linked note.

## 5 CONSTRUCTING THE US RESIDENTIAL REAL ESTATE CAPITAL GUARANTEED INDEX-LINKED NOTE

### 5.1 Background

#### 5.1.1 Structured products in general

Structured products, or structured notes, are “fixed income debentures linked to derivatives” (Peng 1995, 2). They are packaged financial instruments or investment strategies. The invested funds are invested partly in a less-risky note, and partly in a more risky asset such as a single security, currency or commodity, a basket of securities, currencies or commodities, an option or some other kind of derivative. They can be issued by banks, financial institutions, corporations, municipals, government agencies and governments, and their maturities may range from 3 months up to 10 years or more. The investor is only exposed to issuer risk. (Peng 1995). The risks to the issuer are interest rate risk and counterparty risk.

According to Commerzbank (2008), structured products are “investments which, generally speaking, have a risk – return profile situated between traditional fixed income investments and pure (delta one) equity investments depending on the riskiness of the investors exposure.” Fixed income investments, or debt securities, have two main types of return profiles. A zero coupon bond is sold at a discount, and trades at that until maturity. As the bond reaches the maturity date, and there is no threat of default by the issuer, the value of the bond increases to nominal value. At maturity, the investor receives the full nominal, and the return is the change from the discounted issue price and the nominal price. A fixed or floating coupon bond may also be sold at a discount, or it may also be sold at a premium, depending on the coupon, among other things. During the lifetime of the bond, the holder of the bond receives a coupon, which may be fixed, or may change according to a payment schedule or may be linked to for example inflation. The coupon may be quarterly, semi-annual or annual. At maturity, the return from the bond is a combination of the price of the bond at issue and the coupons that have been received over the lifetime of the bond, at maturity the investor receives the nominal.

The return that may be received from a debt instrument is normally more modest than what can be received from equity or other investments, and this is due to the more conservative risk profile. Unless there is issuer default, the investor will receive the coupons, if there are any, and the nominal from the bond investment at maturity. For equity investments, however, there is no guarantee of return – if the company's value goes to zero, the investor will have lost her whole investment. There is, however, the benefit that the equity investor reaps the full benefits from the upside as well. If the company's value rockets, the equity investor's return rockets as well. Meanwhile, the debt investor receives the same coupon as before. Thus, with higher volatility, or risk, come higher potential returns.

Thus, structured products offer an opportunity to combine these two return profiles in a way that suits the investor's needs. Commerzbank (2008) explains that structured products "allow investors with specific return needs and/or market view to take specific risks." There may be a variety of reasons behind the investor's risk aversion; Commerzbank (2008) lists the most common reasons behind choosing a structured product as opposed to a direct investment in one asset class: regulatory reasons, hedging purposes, yield enhancement and speculation purposes. First, some, especially institutional investors such as pension funds, may have rather strict risk limits. In these cases, especially capital guaranteed structured products may provide good opportunities to aim for some extra returns while maintaining low risk profile.

Hedging purposes refer to using a structured product to hedge away some existing exposure, for example currency exposure. Yield enhancement is similar to hedging, but in this case the asset that is already held is expected to increase in value. If the investor for, for example regulatory purposes, cannot hold any more of the pure asset, it may be possible for them to hold the structured product that has the asset as an underlying.

The speculation purpose is probably the most obvious rationale behind investing in a structured product. If an investor has a specific market view, but wants to at the same time guarantee the nominal of the investment, they can use a capital guaranteed structured product to do so. One may ask, why the investor doesn't simply invest most of his money in a zero coupon bond, and use the rest on

derivative on the suitable underlying. In theory, the cost and return should be the same, but in practice, unless the option is extremely common and liquid, and unless the investor is a large institution, it is much more costly for an investor to make the derivative contract themselves than it is for the issuing bank. The issuer of the structured product gathers the funds from all the purchasers of the structured product at issue, and makes one larger derivative trade, giving the bank more pricing power than the individual investor has.

The aforementioned criticism over why the investor doesn't simply directly invest in a diversified portfolio of stocks in the same proportion as the basket or index has been studied by several academics. The criticism has been responded to by, among others, Gorton and Pennachi (1993). They show in their paper "Security baskets and index-linked securities" that "when investors have immediate needs to trade, and prices are not fully revealing, the return of these composite securities cannot be replicated by holding the individual underlying assets in the same proportions." (Gorton 1993, 2) Gorton and Pennachi go on to explain that when prices are not fully revealing, "informed agents can take advantage of lesser informed agents who have urgent needs to trade, and thus cannot wait for information to be revealed" (Gorton 1993, 24). The urgent need the less informed agent has would commonly be the need to rebalance the portfolio to match the basket or index. Informed agents, such as insiders or market makers, will have less 'camouflage' with which to disguise their trades with as liquidity traders move on to using composite securities. Moreover, taking part in a composite security would also reduce the investor's trading losses that are created when trading with less knowledge than the market counterparties such as insiders or market makers. Thus, a composite security such as a structured product is a trading vehicle with a lower rate of return variance and therefore, has less information asymmetries. Gorton and Pennachi (1993) conclude that not only does the introduction of composite securities increase the expected utility of lesser informed agents but also enhance the process for reaching equilibrium prices in the individual primitive securities.

### 5.1.2 Different types of structured products

There are two main types of structured products, as defined by the type of return that the derivative offers to the investor. The deposit or zero coupon bond part of the

product is a standard building block for a structured product, and the character of the structured product is defined by the derivative. Commerzbank (2008) groups structured products into income products and growth products. Income products provide “an alternative to debt instruments where the interest rate linked coupon is replaced by an equity-linked coupon” (Commerzbank 2008, 6). Growth products, on the other hand, aim at maximising leverage via giving the investor “a direct and proportional exposure to the underlying security” (Commerzbank 2008, 6). The case study product will be a growth product. As opposed to the product paying a coupon throughout its maturity, it seeks to retain all funds in the derivative until the maturity. This strategy is chosen to maximise the leverage on the expected boom in the real estate prices.

Structured products may also be classified according to the type of asset class to which the investor is exposed to when investing in the product. There are single asset class products, which offer exposure to one asset class such as foreign exchange, interest rate, equity or commodities. There are also multi-asset class products, which are called hybrids, offering exposure to a more diversified portfolio. For example, the hybrid may provide exposure to global equity and gold, assuming that they may have a low or even negative correlation. Ideally, if the correlation holds, the investor holds a product which yields well in bear and bull equity markets. (Commerzbank 2008, 6)

There is a vast variety of structured products. One of the most straightforward types of structured products is a basket product. A basket product consists of a zero coupon bond and a basket of stocks. The return structure depends on the price development of a basket of stocks. In addition to a simple basket product, a basket product's return can depend for example only on the return of the worst or the best stock in the basket. These products are called worst of option and best of option. Since the worst-of and best-of products have higher payoff potential than more straightforward products, they are also more expensive. The idea behind these kinds of worst-of or best-of products is that the underlying assets are chosen so that they have a low or negative historical correlation. For example, if we hold a worst-of structure with gold and oil as the underlying asset. The structure withholds a call option on gold and a put option on oil, based on some historical period of negative correlation. At maturity, the price of gold has increased and the price of oil has

decreased, continuing the historical correlation trend. Another straightforward and common structured product is a reverse convertible, the return of which depends on the underlying not falling below a certain level. A slightly more complicated structured product is a cliquet – a product whose return is calculated based on several lock-in dates throughout the maturity of the product. In a way, the product's return is “clicked” in on each lock-in date. Reverse cliquets are often used in high volatility environments.

These are only a few examples of structured products common in the market, and it is not in the scope of this thesis to explain them all in detail. According to the structured product professional, for a long time, products with a full capital guarantee and a basic call structure were the most popular in the Finnish market. However, the current investment climate is not so favourable to this basic structure since “options are extremely expensive due the current market volatility and interest rates are low” (Interview, 10.2.2009). According to the professional, the new products are more often partly guaranteed products, and products with autocallable, reverse convertible or soft protection features are becoming more popular. (Interview, 10.2.2009) Now the paper turns to examine factors affecting the construction of structured products as well as the qualities of the case study product in more detail.

### 5.1.3 Factors affecting the construction of a structured product

There are two main factors which affect the structure and pricing of a structured product: the risk-free rate and the cost of derivatives, most commonly options. Commerzbank publication (2008) outlines some examples of different market situations, which illustrate the effect of these two factors. Firstly, if interest rates are at relatively high levels, “the investor will give up more riskless return and therefore he'll be able to increase the upside participation of the option” (Commerzbank 2008, 5). This can be illustrated with a simple calculation. The example product is a capital guaranteed index-linked note, with maturity of 2 years and participation rate of 100 %. The minimum investment is 1 000 euros, which is used for these calculations. We have two scenarios, a low interest rate and high interest rate scenario. In scenario 1, the risk-free interest rate at which the deposit is made is 2 % p.a., and in scenario 2 the interest rate is 5 % p.a.

In the following table, the sum invested in the zero coupon bond is derived from the bond pricing formula. The figures for the example calculation are those of the low interest rate scenario, and there is no coupon as the deposit is a zero coupon bond:

$$\begin{aligned} \text{Price}_{\text{Bond}} &= \sum \frac{\text{Coupon}}{(1 + \text{Risk free interest rate})^{\text{Time}}} + \frac{\text{Nominal}}{(1 + \text{Risk free interest rate})^{\text{Time}}} \\ \text{Price}_{\text{Bond}} &= \sum \frac{0}{(1 + 0,02)^2} + \frac{1000}{(1 + 0,02)^2} \\ \text{Price}_{\text{Bond}} &= 961,17 \end{aligned} \quad (23)$$

The table 12 below illustrates how in a high interest rate environment less of the investment has to be left as a deposit, and thus more is available for the derivative investment.

Table 12: Low and high interest rate scenarios

	Scenario 1	Scenario 2
<b>Risk-free interest rate, p.a.</b>	2,00 %	5,00 %
<b>Maturity, years</b>	2	2
<b>Invested sum</b>	1 000,00 €	1 000,00 €
<b>Nominal to be paid back at maturity</b>	1 000,00 €	1 000,00 €
<b>Sum invested in the zero-coupon bond</b>	961,17 €	907,03 €
<b>Sum invested in the option</b>	38,83 €	92,97 €

Thus, higher interest rates increase the financing available for the derivative, as does freeing up some of the principal for the derivative investment, as outlined before. Naturally, the product is no longer fully capital guaranteed in the latter case.

The price of the option is affected by a variety of factors, but one of the main factors at work is volatility of the underlying. In short: the lower the volatility, the lower the price of the option. This is due to the fact that as volatility of the underlying increases, the probabilities of higher and lower values in the underlying increase as well. The price of the option is in effect a probability-weighted, discounted average of all possible outcomes in the value of the underlying, as defined by volatility. In "Options, futures and other derivatives", Hull (2006, 206), explains the effect of volatility on option pricing: "As volatility increases, the chance that the stock will do very well or very poorly increases... The owner of a call benefits from price

increases but has limited downside risk in the event of price decreases because the most the owner can lose is the price of the option.” The Black-Scholes option pricing formula (Hull, 2006, 295) illustrates how the price of the option is not only affected by volatility but also by the interest rate.

The Black-Scholes pricing formula for a European option on a non-dividend paying stock (or other instrument) is as follows:

$$\begin{aligned}
 C &= S_0 N(d_1) - Ke^{-rT} N(d_2) \\
 \text{and} \\
 P &= Ke^{-rT} N(-d_2) - S_0 N(-d_1)
 \end{aligned}
 \tag{24}$$

where

$$\begin{aligned}
 d_1 &= \frac{\ln(S_0 / K) + (r + \sigma^2 / 2)T}{\sigma\sqrt{T}} \\
 d_2 &= d_1 - \sigma\sqrt{T}
 \end{aligned}$$

Definitions:

C = Price of a call option

P = Price of a put option

$S_0$  = Price of the underlying at time 0

$N(x)$  = Cumulative probability distribution for a standardised normal distribution

K = Exercise price of the option

r = continuously compounded risk-free interest rate

$\sigma$  = Volatility of the underlying

T = Maturity of the option

Payoffs from European call and put options, ignoring the cost of the option, are:

$$\begin{aligned}
 \text{Call option payoff} &= \max(S_0 - Ke^{-rT}, 0) \\
 \text{Put option payoff} &= \max(Ke^{-rT} - S_0, 0)
 \end{aligned}
 \tag{25}$$

European options can only be exercised at maturity, the pricing of American options is slightly more complicated as they can be exercised also before maturity. The underlying is the instrument or asset to which the option contract is linked. The payoff conditions show when the option is used and when it expires worthless. For example, if one holds a call option, and at maturity the exercise price is higher than the stock market price, the investor will be better off buying the stock at the market. In this case, the payoff is zero and the option expires worthless. On the other hand, if the exercise price is again higher than the stock market price at the time of maturity, but this time the investor holds a put option, the investor can use the option, sell the stock at the exercise price and buy the shares back at the lower market price. The holder of this option gets a payoff that is the difference between the stock market price and exercise price of the option. Commerzbank (2008, 5) summarises in their publication on structured products that since in a low volatility environment, the option price is lower, "it is therefore cheaper to buy an extra percentage of participation on the upside."

From the pricing formula as well as the payoff conditions, we can see that the price of an option is affected by, among other things, the volatility of the underlying as well as the risk-free interest rate. The higher the interest rates, the lower the present value of future cash flows. The discounting effect can be directly seen from the option payoff conditions. An increase in the risk-free interest rate makes the present value of the exercise price smaller. This increase in the risk-free rate increases the payoff of the call option, and decreases the payoff the put option. Naturally, an increase in the payoff of the option causes an increase in its price, and vice versa.

The value of the option can be also examined from another point of view. The payoff conditions above show the payoff of the option at maturity. This difference between the stock price and exercise price is called the intrinsic value of the option, and it is always equal or more than zero. (No investor would exercise the option in the case the payoff is negative). At maturity, the payoff of the option is certain. However, the further away in time we are from the maturity date, the more chance there is that the payoff may increase, and also decrease, due to volatility of the underlying. The closer we get to maturity, the less probable large deviations from the current price of the underlying become. This phenomenon is called the time value of the option. The time value is highest when the option is written, decreases the closer we get to

maturity, and is zero at maturity. The total option value can be seen as consisting of the time value and the intrinsic value together.

As can be noted from this analysis of option pricing, there are a multitude of factors that need to be taken into account when constructing a structured product that entails options. Commerzbank (2008, 5) states that all in all, "capital protected structures work best in a high interest rate and low equity volatility environment." One must note that this statement does not bode well for the case study product: volatilities in the market have been at all time highs for the past year or year and a half, and also it seems that at least for now interest rates have peaked and Western central banks have begun to drastically cut them in response to the oncoming recession. On the other hand, the turn to the economic slowdown may come quickly, which could be followed by sharp changes in interest rate policy and market returns. Even though right at this moment the situation may seem bad, in order to reap the highest rewards, the best course of action is to be well prepared for the turn in the economy.

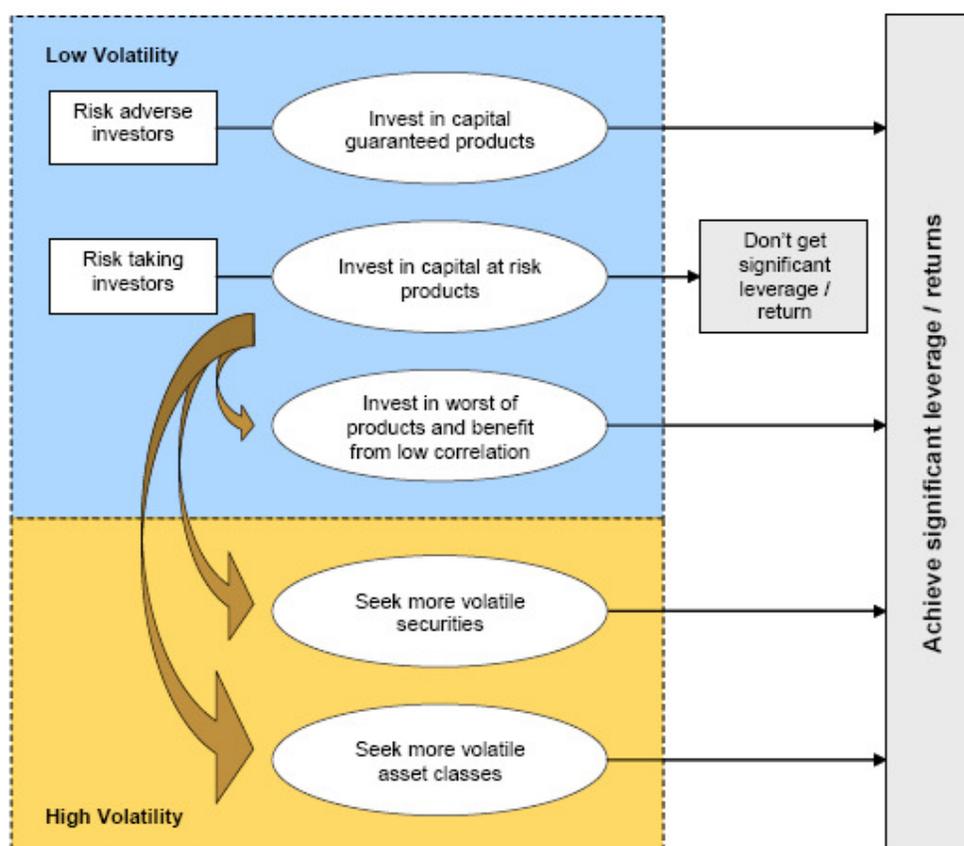
#### 5.1.4 Typical investor's characteristics

Investors choose to invest in structured products because they have specific needs or requirements. A major motivation behind choosing a structured product results from the investor being an institutional investor with relatively strict investment regulation. As already explained before, especially pension funds may have rather strict risk limits. If the pension fund's equity limit is full, it may be possible for the institution to invest in a structured product with equity as the underlying, upon which the derivative is built. Another common need among institutional investors is that they already hold a specific risk position in their portfolio. As explained before, the investor may choose to use a structured product either to hedge against or to enhance this position. For example, an international company may want to hold some amount of US dollars always in an account for liquidity, even if their major currency is euro. This company may benefit from purchasing a capital guaranteed structured product, which yields if the US dollar depreciates. (Commerzbank, 2008)

The one factor affecting all investors is the degree of risk aversion. Economic agents can be risk averse, risk neutral or risk loving. For a risk neutral agent, the fact that

the future rewards from an action are unsure does not affect their expected utility. A risk averse agent's expected utility, on the other hand, is decreased by uncertainty in the outcome. Similarly, for a risk loving or risk taking agent, uncertainty in the outcome increases the agent's expected utility. According to Commerzbank (2008), the kinds of investment decisions investors make depend on their level of risk aversion as well as the market conditions. The figures below illustrate the typical choices made by risk averse and risk taking investors in both low and high volatility environments.

Figure 10: Typical choices made by investors in a low volatility market

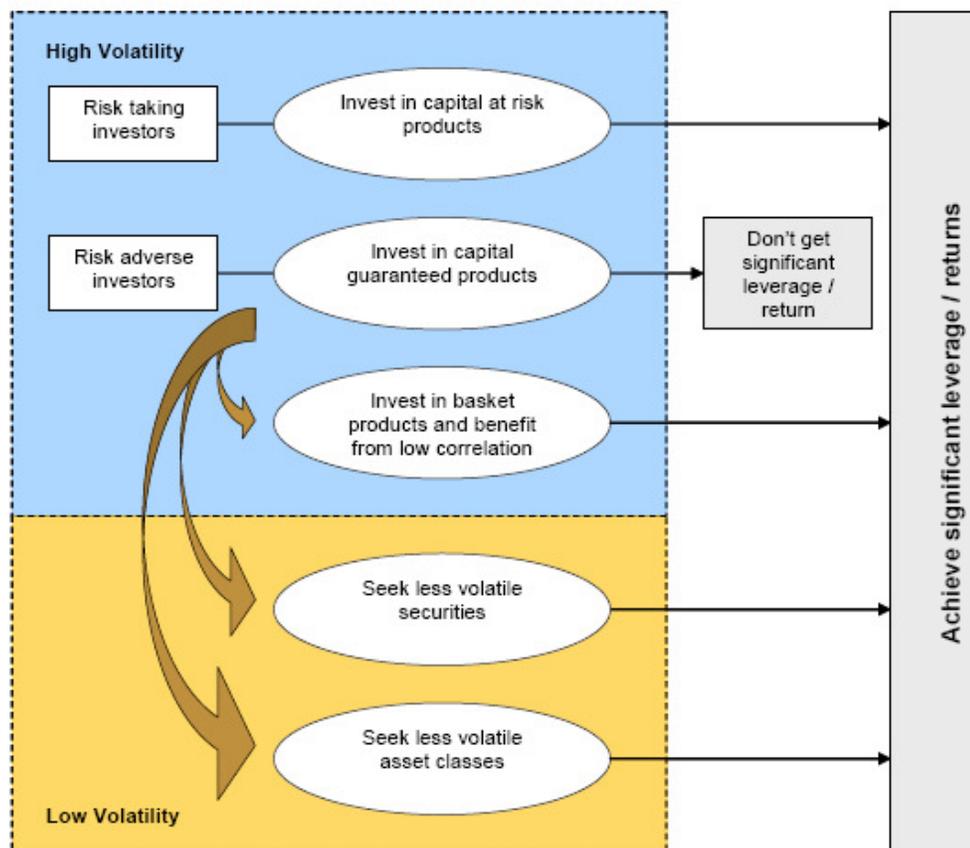


(Source: Commerzbank 2008, 7)

The figure 10 above shows some of the typical choices made by different kinds of investors at a time when market volatility is relatively low. The top region entails products which are low risk or volatility, as compared to other products which are high volatility and are grouped in the bottom half of the figure. Thus, in a low volatility environment, according to Commerzbank (2008), risk averse investors

typically choose to invest in capital guaranteed products as they allow them to get significant leverage. When volatilities are low in the market, the derivative structures are usually cheaper, and thus the investors get more for their money, so to speak. Risk taking investors would be willing to invest in capital at risk, as opposed to capital guaranteed products, but these products “may not give a significant increase in returns compared to the risk free rate” (Commerzbank 2008, 7) This is because in a low volatility environment, the upside potential of the derivatives is more limited than in a high volatility environment. The risk-return ratio is thus worse in a low volatility environment, and therefore it may be more profitable for a risk taking investor to invest in more volatile products such as products linked directly to more volatile securities or asset classes, or invest directly in for example emerging market equities or oil. The risk taking investor may also choose to invest still in low volatility products such as explained before. Even though the market volatility is low, there may be returns in these kinds of arbitrage products.

Figure 11: Typical choices made by investors in a high volatility market

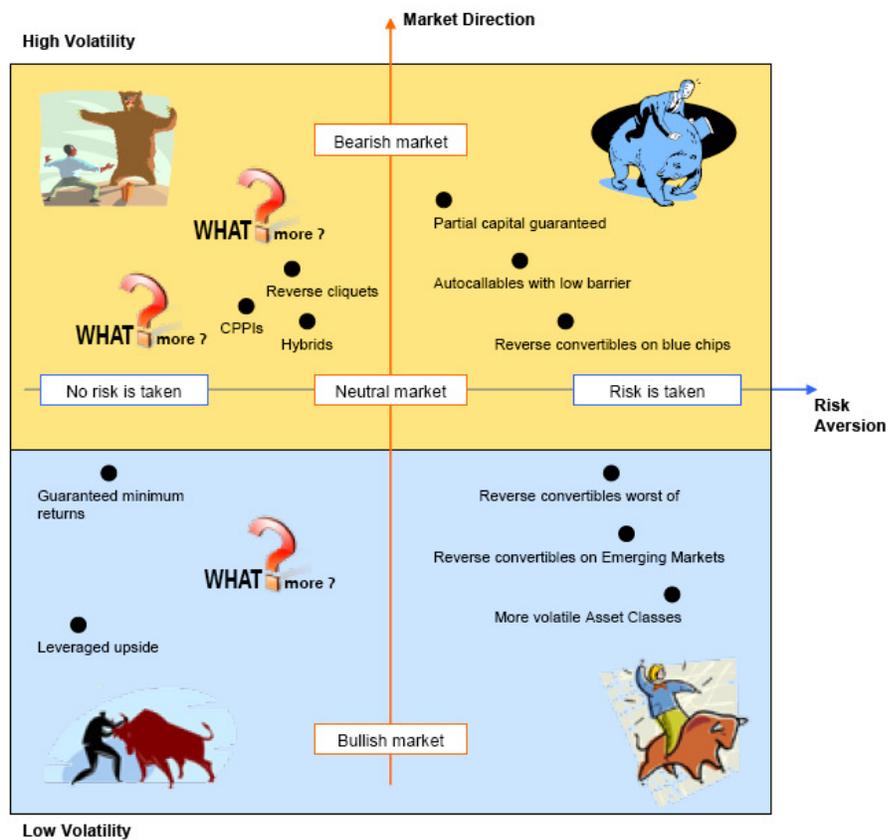


(Source: Commerzbank 2008, 8)

When market volatility is high, the risk taking investors may see enough return opportunities in capital at risk products that they choose to use them. Risk averse investors, on the other hand, face a situation where they cannot achieve sufficient return from capital guaranteed products as in a high volatility environment the derivatives become quite costly. Also, because of the high volatility environment, the probability of downside outcomes increase, which may decrease the risk averse investor's utility too much. Rather, the risk averse investors may choose to use low volatility products, as described by the lower section in the above figure. They may also choose to use basket products, which have been described earlier.

Commerzbank (2008, 9) also shows how market direction affects the choice of product for the investor. Figure 12 below summarises these different strategies.

Figure 12: Effect of market volatility and market direction on investor's choice of product



(Source: Commerzbank 2008, 9)

Considering the current situation, one could see us positioning ourselves on the left side of the figure. The case study product is capital guaranteed because it is expected that in Finland at this moment there are many risk averse investors such as large institutions who would like the capital guarantee. Also the aim is to attract wealthy private investors, who would like to carefully begin to increase the level of risk in their portfolio, but do not yet or at all wish to take on direct investments in derivatives. It is also expected that the global economy is beginning to reach the bottom of the bear market, and we are slowly beginning to move to a bull market, even though it is expected that volatilities remain high. The Commerzbank figure shows that indeed, in this state of the market, a structured product providing guaranteed minimum returns is could be what investors are looking for.

#### 5.1.5 Overview of the pros and cons of capital guaranteed index-linked notes

After discussing the product from the investor's point of view, it is important to note also that structured products, and especially notes like a capital guaranteed index-linked note are a valuable source of financing for the issuers. Majority of the funds that are invested in such a product remain with the issuer in the form of the zero coupon bond. Many of the advantages of structured products in general and capital guaranteed index-linked notes in specific for the investor have already been outlined in the sections above. The table below shows gathers together the advantages and disadvantages of these products that have already been brought up as well as adds a few more aspects.

Table 13: The pros and cons of capital guaranteed index-linked notes

Point of view	PROS	CONS
<b>Investor</b>	<ul style="list-style-type: none"> <li>+ Low risk profile and potentially very good risk-return ratio</li> <li>+ Only issuer risk</li> <li>+ Easy access to exotic derivative structures</li> <li>+ May be a good addition to existing portfolio (hedging or yield enhancement)</li> <li>+ Tool for investing in another asset class in the guise of a debt instrument</li> <li>+ Gain access to certain markets, which are otherwise hard to reach or involve too much risk</li> </ul>	<ul style="list-style-type: none"> <li>- Expensive or impossible to redeem investment before maturity</li> <li>- Mostly long maturities</li> <li>- Complicated structures make costs non-transparent</li> <li>- Some structures may be illiquid - the only available secondary market counterparty is often the issuer, which gives the issuer unproportionate pricing power</li> </ul>
<b>Issuer</b>	<ul style="list-style-type: none"> <li>+ Bring valuable financing for the medium and long-term</li> <li>+ May attract large institutional investors with conservative risk profiles</li> <li>+ Possibility of making private placement contracts (a good service for wealthy individuals or institutional investors)</li> <li>+ Ability to construct tailor made products to suit investor needs (especially relevant to large investors)</li> </ul>	<ul style="list-style-type: none"> <li>- Difficulty of finding new, good investment ideas may be a problem especially if the issues are driven more by need for financing than a good investment idea</li> <li>- Threat of misunderstandings between the issuer and investor especially in the case of complicated structures (a compliance issue)</li> <li>- Counterparty risk with the derivative counterparty</li> <li>- Volatility in interest rates affect the return from the deposited funds (the cost of this deposit is set for the maturity of the product from the issuer's point of view)</li> </ul>

Some of the points above have been included from the interview with the structured product professional. He outlined, among other things, the positive aspect of structured products providing access to more exotic markets (Interview, 10.2.2009). To add to his point, structured products may utilise combinations from different market or specific subsectors of markets as well, where ever a beneficial or predictable price development can be foreseen. A serious disadvantage for the investor is, however, the common illiquidity of structured products. (Interview, 10.2.2009) Structured products need to be seen as a long-term investment. The cost for the investor from the illiquidity can not be necessarily directly measured, but there definitely is a cost. However, in exchange for this cost, it receives a capital guaranteed investment solution in a market that may not be accessible to them otherwise, and a good risk-return ratio. It is true, nonetheless, that an investor who has a shorter investment horizon may not find structured products suitable for their needs, despite all the pros included in the investment.

## 5.2 The case study product

The name of the case study product is the US Residential Real Estate Capital Guaranteed Index-Linked Note. The structure of the case study is chosen to be a simple structure, consisting of a call option in the underlying and a zero coupon bond. The case study product also has a full capital guarantee, and the participation factor is 100%. There is no possibility of early redemption. All of these properties were chosen to keep the construction of the product simple. Making the product more complicated would be out of reach for this thesis. There is a vast variety of different kinds of structures that have been used in the market, and new structures are developed each day. Explaining the return possibilities and other characteristics of such products may become extremely complicated. This product should be viewed as a basic example, which can serve as a basis for further development.

### 5.2.1 Properties of the case study product

A capital guarantee refers to the investor being guaranteed the originally invested capital back at maturity, in addition the return may include profits from the derivative investment. The capital guarantee does not eliminate issuer risk for the investor, and neither does it usually hold if the structured note is redeemed before maturity. The capital guarantee is generated by depositing part of the invested funds for the maturity of the product at a fixed interest rate, so the future value of the invested funds equals the total invested funds. How much of the whole investment is placed in the deposit depends on the current interest rate and the maturity of the product. From this deposit arises the interest rate risk for the issuer. If the deposit has been made at for example 4 % a., for the maturity of 3 years, the issuer needs to invest the deposit so that it gains a return of 4 % a. in order break even. If, during the next 3 years, there is a decline in the market interest rates across the board, it becomes more challenging for the issuer to invest the funds from the deposit in order to be able to break even. The original deposit interest has naturally been calculated taking into account predictions of future interest rates, but of course, unexpected changes may happen.

Capital protected or guaranteed products consists of a zero coupon bond and an option structure, as do majority of structured products. The return of the zero coupon

bond is considered risk-free return. In order to get better return than the risk free return, the investor must give up some of the riskless income to enhance the return from the product. Since capital guaranteed products are built on zero coupon bonds, as opposed to coupon bonds, the investor is in essence giving up the coupon. Instead, the funds are invested in a derivative which may potentially yield more than the risk-free interest rate. (Commerzbank 2008)

Depending on the projected risk-free interest rate for the lifetime of the product and on investor preferences, the participation factor on the derivative may vary. If the interest from the deposit is not enough to buy the whole option, for example in the case that interest rates are expected to be low are become lower during the lifetime of the product, the investor may buy a smaller than 100 % share of the derivative. This share is called the participation factor. If an investor chooses to do so, she may also increase the participation factor buy putting some of the principal at risk. For example, the investor may agree that an additional 10 % of the principal is also invested in the derivative. In this case if at maturity the derivative expires worthless, the investor only receives 90 % of the principal, not 100 % as in the full capital guarantee case. (Commerzbank 2008, 3)

Index-linked notes, or equity linked notes as called by Peng and Dattatreya (1995), are from the investor's point of view products that "obtain equity linked returns with fixed income instruments." In other words, an index-linked structured note is a product, whose return is dependent on the performance of an index. They were first created in the 1980's to provide capital arbitrage possibilities and to improve asset allocation. They have been from the beginning very popular among investors due to their good risk-return ratio, especially as the only major risk that remains with the capital guaranteed products is the issuer risk. The issuer risk refers to the issuer bank not being able to pay back the deposited capital due to a liquidity crisis or bankruptcy. The position of the investor, when investing in an index-linked note, is long in fixed or floating rate note and long or short in equity index forwards and options. (Peng 1995, 290) Connecting the return of a structured product into an index gives the benefit of diversification. The product could be linked to a basket of stocks, but, with using index-linked derivatives, transaction costs are lower. There is no need to trade the actual shares in the index, as the aim is not to own the shares

but to benefit from the change in the value of the whole basket of shares, as measured by the index.

The derivatives used for index-linked notes are usually options on exchange traded funds. An option is a derivative and a contract between two counterparties, which gives its holder the right to buy or sell an asset (Hull 2006, 753). Call options give the right to buy the asset, and put options give the right to sell some specified quantity. An exchange traded fund is a fund, whose composition and therefore also performance mimics some existing index. Exchange traded funds can be created on any kind of an index, but in reality they only exist for indices that create enough interest to raise sufficient funds for it to be profitable to start an exchange traded fund. The underlying indices may be anything from broad market indices to sector specific indices. An option on an exchange traded fund, ETF, therefore, gives the holder of the call option the right to buy a specified number of shares of the ETF, and the holder of a put option the right to sell a specified quantity of ETF shares. One could also use options on the index itself, but market practice is to use options on an ETF as they are more liquid and available. The counterparty risk of the issuer arises from the use of derivatives, which are in essence contracts between two parties. For example, the option counterparty may go bankrupt just before the maturity of the product, and thus fail to at least immediately fulfil their part of the bargain, so to speak.

Thus, a capital guaranteed index-linked note is a structured product, where some percentage of the invested fund is invested in a deposit, which during the maturity of the product grows into 100 %. The remaining share of the funds is used to buy options on an ETF, which reflects the performance of a market or sector, whichever is the focus of the product. If it is believed that the underlying market or sector, and therefore the index, will be bullish during the maturity, call options will be used. On the other hand, if the forecast is that the index will be bearish put options will be used. The options will be chosen naturally so that the strike prices of the options reflect the market view – for example, in expectation of a bullish market, a call option with a lower strike price than expected future price will be chosen. The strike price of the option is defined as “the price at which the asset may be bought or sold in an option contract (also called the exercise price)” (Hull 2006, 757)

## 5.2.2 The purpose of the product

The case study product, the US residential real estate capital guaranteed index-linked note is a product aimed at large institutional investors and wealthy private investors. The target group is chosen because large, more professional investors may be more interested in such exotic products, and also because they are a profitable client group for the issuer. The common factor with the target investors for this product is that they are expecting a turn in the market conditions in the near future. The main idea behind this product is that the US residential real estate sector has become undervalued in the recent subprime crisis and the market turmoil that has ensued. As the global and US economy begins recover from the bear market and financial and economic crises, the supply and demand conditions in the US real estate market will begin to stabilise and the house prices will begin to recover from the largest slump since the early 1990's. As demand for housing begins to pick up and house prices increase, the returns of house construction companies and real estate Investment Trusts, among others, will begin to increase. This increase should be reflected in the stock prices of these companies, which will lead to an appreciation in the chosen index. Thus, this product is a bullish product, aiming to benefit from an expected increase in the value of the US residential real estate sector.

There is, nonetheless, high volatility still in all markets across the board. Most investors have experienced negative returns for the past year, and either because of risk appetites or regulations, they may not be yet willing or able to take large risks. For this reason the product is constructed as a fully capital guaranteed product. In practice, many banks may offer investors private placements, where the level of risk and capital guarantee may be tailored to individual needs. However, the basic product is established as a fully capital guaranteed note.

The index chosen for this product is the Dow Jones U.S. Real Estate Index. The reasons behind this decision have been explained in full detail in previous chapters. The choice of this index means that the pricing of the option is easier to obtain or estimate – as was established before, there are relatively long maturity options available in the market for the ETF linked to this index. Thus, a commercial real estate index is used for the case study product, which attempts to reflect the price

development of the US residential real estate sector. As contradictory as this may sound, the decision relies on major pricing issues and on the finding that there seems to be a positive and rather strong correlation between the commercial and residential real estate markets.

According to the structured product professional, the first issue to note when choosing an index for an index-linked product is whether the index is a price index or a return index. (Interview, 10.2.2009) In the case of a price index, dividends are not included in the calculation of the index, which according to the professional also makes the options linked to the index cheaper. An important aspect to consider in choosing the index for a structured product is the careful consideration of the sales story for the index. In order not to provide misleading or incorrect information the underlying index needs to suit the product description. An example given by the structured product professional is a product called US Export Basket, aiming to benefit from the positive consequences of the weak US dollar. The basket of stocks includes companies such as McDonalds, Coca Cola and Wal-Mart. It is arguable whether these companies are American export companies, or whether they in fact are international companies who are selling their products locally across the globe, and in the case of Wal-Mart, completely domestic company selling their products domestically. Thus, it can be questioned whether the sales story and product description and the actual contents of the product match. This issue is in fact central to the thesis: since the Dow Jones U.S. Real Estate Index does not include residential real estate, it needs to be clearly and reliably explained to the client that it is nonetheless a suitable index for the product, as proved by the thesis earlier.

### 5.3 Putting together the product

#### 5.3.1 The maturity

The maturity of the product should mainly depend on the projected development of the underlying. In other words, on what is the time frame during which the forecasted price development is supposed to take place. This is naturally impossible to forecast especially in high accuracy – it could take anything in between a year and 10 years for the US residential real estate sector to recover from the crisis and the current economic recessionary trend. The equilibrium level of the market is also a

mystery. It can be questioned whether the peak levels of the years before the housing market crash were sustainable and what will the peak levels of the future. Considering the current situation, where the US economy is still in a recessionary trend, the market sentiment seems to be that the residential real estate sector will take some years to recover. The longest options available on the underlying ETF are approximately 2 years, and thus the maturity of the product is chosen to be 2 years. It should be noted that this is still a rather short product in maturity, and as a result, the participation levels might be low, as pointed out by the structured product professional (Interview, 10.2.2009).

### 5.3.2 The structure and pricing

First, the zero coupon bond needs to be established. According to the structured product professional, the pricing of the zero coupon bond is based on spot interest rates (Interview, 10.2.2009) He continues to state that at the moment, when issuers such as banks and financial institutions are in great need of financing, it may be possible to get a good funding premium on top of the spot rate. This way there is more money available to spend on the option. However, spot interest rates as well as other interest rates are at long time lows at the moment. According to Bloomberg (2009) the spot interest rate (mid price) for 2 years is at the moment 2,1243 % p.a. This is a rather low interest rate, and thus, there is a lot less room to purchase options than lets say, a year ago when the interest rates were 3-4 % p.a. However, if one adds on a premium of for example 50 basis points, the situation is at least slightly improved. Thus, at the rate of 2,6243 % p.a., the price of the zero coupon bond is calculated.

$$\begin{aligned} \text{Price}_{\text{Bond}} &= \sum \frac{0}{(1 + 0,026243)^2} + \frac{100}{(1 + 0,026243)^2} \\ \text{Price}_{\text{Bond}} &= 94,95101 \approx 94,95 \end{aligned} \quad (26)$$

The price of the zero coupon bond indeed ends up being rather high, leaving only 5,05 EUR for each 100 EUR invested to the purchasing of options, when excluding any other costs such as trading or administration fees.

This product will utilise a very simple option strategy of purchasing American call options. Nonetheless, the pricing is the same as for European options, since the price index is used and therefore, dividends are ignored. The bet is that the price of the underlying will increase significantly over the next two years, making the payoff profile of the option unlimited. If the strike price of the option is lower than the price of the underlying at the maturity date, the option will be in the money. The option holder can then buy the underlying at the strike price and immediately sell it on the market at the current, higher market price, and the payoff is their positive difference. The higher the market price, the higher the payoff - therefore, there is potential for very high returns.

The longest maturity for options linked to the iShares Dow Jones U.S. Real Estate Index ETF is January 2011, providing almost fully for our 2 year maturity note. There are two strikes for this maturity for which there is volume in the market at the moment. These options are described in the table below; the information is from Bloomberg (2009).

Table 14: The iShares Dow Jones U.S. Real Estate Index ETF call options

Option	IYR JAN 2011 ZPE+AT	IYR JAN 2011 ZPE+AD
<b>Exercise type</b>	American	American
<b>Strike</b>	20,00 USD	30,00 USD
<b>Contract size</b>	100	100
<b>Implied volatility</b>	65,91 %	64,62 %
<b>Volume (millions)</b>	33	24
<b>Bid price</b>	12,90 USD	8,60 USD
<b>Ask price</b>	13,50 USD	9,20 USD
<b>Ask price (EUR)</b>	10,46 EUR	7,13 EUR
<b>Last traded price</b>	12,20 USD	8,60 USD
<b>iShares US Dow Jones Index ETF price 11.02.2009</b>	30,67 USD	
<b>EUR/USD 11.02.2009</b>	1,29015	
<b>Option payoff 11.02.2009</b>	10,67 USD	0,67 USD
<b>State of the option 11.02.2009</b>	In the money	In the money

The option with the 30,00 USD strike price is very close to the current price and would thus suit a product issued at this moment better. Moreover, due to the higher

strike price, it is a cheaper option, which works well since there is a limited amount of money available for investing in the derivative. Thus, the option strategy chosen for the product is a long position in American call option with strike 30,00 USD, maturing in January 2011.

The pricing of the product is affected by the short maturity of the product. On the other hand, the shorter the maturity, the cheaper the call options are. However, the short maturity makes the zero coupon bond more expensive, leaving less money for the options. Moreover, as outlined by the structured product professional, one can make the product cheaper for the investor by decreasing the degree of capital guarantee (Interview, 10.2.2009). To keep this case study product simple, nonetheless, the full capital guarantee is sustained.

### 5.3.3 Other characteristics and possible limitations

One of the main issues to consider is the position on the US dollar – should it be hedged against, or should the exposure be sustained? According to the structured product professional, normally the products especially on the retail side are hedged against the USD, or in other words, quanto Euro. He continues, however, that with this specific product, taking a bet on the USD might be a good idea since it is rather weak at the moment. If the US dollar was to strengthen over the maturity, the possible payoff from the product at maturity would also include an exchange rate gain. (Interview, 10.2.2009)

A serious limitation to the construction of this product is that, in the current high volatility, low interest rate environment it is almost impossible to construct the product in its current form (Interview, 10.2.2009). Possible solutions would be to sell the product with a premium or make it a partial capital guarantee product. Moreover, the structured product professional points out that it may be that we have not seen the bottom of the US housing market yet. A possible solution would be a more complicated derivative structure where the final payoff depends on an average of the initial valuation of the underlying and the valuation on a later date or dates. Also, the structured product professional suggests that it might be more suitable to take a view on some specific area in the US which has been hit the most in the market turmoil, and which has more upside potential than the overall country's market.

These are all issues that should be considered in further study or if the product was to be actually issued. (Interview, 10.2.2009)

#### 5.3.4 The final structure and projections of future performance

The final structure of the US Residential Real Estate Capital Guaranteed Index-linked Note consists of a zero coupon bond and a long position in American call options, and it matures in 2 years. For simplicity, there is no possibility of early redemption. The zero coupon bond is linked to the spot interest rate, with a premium of 50 basis points, resulting in a yield of 2,62 % and a price of 94,95 EUR, when the nominal is 100EUR, for the bond. The option used is a call option on the iShares Dow Jones U.S. Real Estate Index ETF, with a strike price of 30,00 USD and maturity in January 2011. The lock-in level for the payoff is the maturity date, and there are no earlier lock-in dates. In other words, the price of the underlying at the date of maturity for the options determines the payoff of the product. The US dollar exposure is not hedged against, and thus, any appreciation in the USD against EUR will lead to a higher return profile, and any depreciation will lead to a lower return profile. All other costs except for the cost of the options are also for the purposes of this thesis excluded. These costs may include administration costs, and subscription and redemption fees, among others. Now, some final estimations will be made on the future performance of the product

As already calculated, the price of the zero coupon bond now for the 2-year-maturity note is 94,95 EUR for a nominal of 100 EUR. There is, therefore, 5,05 EUR for each 100 EUR available to purchase options. After choosing the option with the strike of 30,00 USD, we can now make a set of future projections. The starting point is the current price of the underlying, the iShares Dow Jones U.S. Real Estate Index ETF, which is 30,67 USD on 11.02.2009. There are 5 scenarios for the product: the price of the underlying by maturity decreases by 15%, decreases by 5%, is the same at maturity, increases by 5%, and increases by 15% by maturity. These calculations as well as the final product structure are made ignoring any other costs except for the cost of the option. For example, there is usually a subscription fee, and perhaps also a redemption fee if it is possible to redeem the note early – these are nonetheless excluded here.

Table 15: Projected performance for five price scenarios

IYR JAN 2011 ZPE+AD					
<b>Strike</b>	30,00 USD				
<b>Spot rate + premium</b>	2,62 %				
<b>Price of bond</b>	94,95 EUR				
<b>Cost of one option EUR</b>	7,13 EUR				
<b>Cost of one option USD</b>	9,20 USD				
<b>Invested in options EUR</b>	5,05 EUR				
<b>Invested in options USD</b>	6,52 USD				
<b>Number of options</b>	0,7082				
<b>Quantity of underlying</b>	70,82				
<b>Price of underlying now</b>	30,67 USD				
<b>Scenarios</b>	<b>Decrease 15 %</b>	<b>Decrease 5 %</b>	<b>No change</b>	<b>Increase 5 %</b>	<b>Increase 15 %</b>
<b>Price at maturity</b>	26,07 USD	29,14 USD	30,67 USD	32,20 USD	35,27 USD
<b>State of the option</b>	Out of the money	Out of the money	In the money	In the money	In the money
<b>Total payoff from option at maturity</b>	0,00 USD	0,00 USD	40,93 USD	149,53 USD	366,73 USD
<b>Total payoff from option EUR</b>	0,00 EUR	0,00 EUR	31,73 EUR	115,90 EUR	284,25 EUR
<b>Total payoff from the product</b>	100,00 EUR	100,00 EUR	131,73 EUR	215,90 EUR	384,25 EUR
<b>Total return</b>	0 %	0 %	32 %	116 %	284 %

The total payoffs from the option at maturity show the payoff from the option at maturity if it is in the money, also the option premium has been deducted. As can be seen in table 15, the projected returns from the product seem very lucrative. If, over the next 2 years, the price of the underlying ETF increases by 15%, the return from investing into this case study product would be 284 % for the 2 year period. Even if the price of the underlying is the same as at issue on the maturity date, the return on investment will be 32%. These return figures are in part a result of the leverage an option provides. One call option contract gives the right to purchase a certain quantity of the underlying, here the contract size is 100. Even if the return is 0,67 USD per share, the contract gives the right to purchase approximately 71 shares, and thus the total return is leveraged. The projected returns have been calculated with the current EUR/USD exchange rate. However, if the USD begins to appreciate against the dollar, the returns will be even higher.

However, these lucrative returns may not be enough to attract investors. In the past years, some US real estate products have been issued and they have been relatively popular. Nonetheless, according to the structured product professional, the current market situation may be such that now is not the time to try to sell the case study product. There may be some institutions, which hold the same view as this product and thus could be interested. Nevertheless, especially private investors may

be still looking to bet on markets closer to them, but also have good return potential such as the Finnish or European equity markets. (Interview, 10.2.2009) It should also be noted that, no doubt, the attractiveness of these kinds of structures has been damaged by the recent bailouts and bankruptcies of financial institutions as well as the poor performance of hedge funds. Even though the product is capital guaranteed, in today's volatile financial markets, it is in the back of every investor's mind that anything can happen.

## 6 CONCLUSION

### 6.1 Overview of the thesis and answers to the research questions

This thesis illustrated the process behind choosing a suitable index for an index-linked structured product. A multitude of issues were dealt within this process. First, the theoretical background to indices and the most common index formulae were introduced. The construction of an index from a mathematical viewpoint is in itself a rather simple task – what is required is understanding of mathematics and, more importantly, a view of what the index number is meant to show us about reality. The construction of an index is drastically complicated by introducing the variety of viewpoints, which, in the end, determine the criteria for choosing the correct index. Some index theorists see that an index needs to correspond to an economic theory, which is assumed to represent reality. Other index theorists decide on the goodness of an index based on statistical testing – the best index formula can consistently and accurately condense all the individual data points into one index number reflecting the underlying reality. An overview of the academic discussion on the best index was given, upon which it was concluded that currently, the field of index theory does not provide a united solution to the choice of the best index. This result was acceptable especially since all the indices under examination for the case study used a moderated version of the Laspeyres index formula – a comparison of pure formulae was not even possible in the frames of this paper. Another approach was needed to discover what kind of an index is the best choice for the case study product, the US residential real estate capital guaranteed index-linked note.

The thesis therefore moved on to examine the choice of the best index from the statistical point of view described above. It was decided that the best index out of the Dow Jones U.S. Real Estate Index and the Philadelphia Housing Sector Index would be the one which statistically represented reality better. Reality was represented by the S&P/Case-Shiller Home Price Composite indices. Basic correlations and other figures were calculated on both of the two stock market indices in question, and it seemed that the Philadelphia Housing Sector Index followed the S&P/Case-Shiller Home Price Composite indices better. For the sake of the construction of product, it was nonetheless assumed that the Dow Jones index

would be from a pricing point of view much more feasible than the Philadelphia index.

Thus, the Dow Jones U.S. Real Estate Index, regardless of the fact that it seemingly does not represent the residential real estate sector, was chosen for the regression analysis in the empirical section. Due to a variety of reasons, such as the existence of an exchange traded fund and the established name of the Dow Jones indices, there is a significantly larger quantity of longer maturity options available in the market for this index than for the Philadelphia Housing Sector Index. A more liquid market is a more efficient market, which leads to there being less risk premium or mispricing in the options - neither desired by the issuer or the investor. Ideally, the Philadelphia Housing Sector Index should be the most suitable index because of its contents. However, this thesis wanted test whether the Dow Jones index, even though concentrating on the commercial real estate sector, would have a strong enough relationship with the residential real estate sector to justify its use in the case study product. The results of the empirical testing were deemed to being supportive of this proposal – there is in fact a strong positive correlation between the Dow Jones U.S. Real Estate Index and the S&P/Case-Shiller Home Price Composite-20 Index. The reader should be reminded that even though this paper draws this conclusion, the data set was not perfectly reliable for the purposes of the testing and decision making. Further study would be needed to solidify the results of this thesis.

After choosing the Dow Jones U.S. Real Estate Index for the case study product, the paper turned to look at the structured products in more detail. The basic structures were explained, as well as investor's and issuer's points of view were covered. Factors affecting the pricing and construction of structured products and especially capital guaranteed index-linked notes were discussed. This section relied partly on academic literature, partly on marketing material from financial institutions and most importantly, on an interview with a structured product professional. The final structure of the US Residential Real Estate Capital Guaranteed Index-Linked Note was simple. The product has a maturity of 2 years, and the only lock-in date is at maturity. The product consists of a basic zero coupon bond and a long position in American call options on the iShares Dow Jones U.S. Real Estate Index ETF with the strike of 30,00 USD. The USD position of the product is not hedged. The return

potential is rather good as illustrated by the projected return calculation. However, the current financial market environment with low interest rates, high volatilities and low levels of trust between investors and financial institutions does not bode well for issuing this product.

## 6.2 Issues for further consideration

As was discussed in conjunction with the empirical analysis, the study should be taken further with more advanced statistical testing methods in order to solidify the results. Moreover, similar tests could be done on the Philadelphia Housing Sector Index. Also, it could be useful to study the relationship between the residential and commercial real estate sectors to further evaluate the suitability of the Dow Jones U.S. Real Estate Index for this kind of a product. The problem of small data sets can really only be fixed by time passing by.

The product also took a rather simple form, and if it were in fact issued, it should be considered whether the product should be made partly capital guaranteed, whether the participation ratio should be increased and whether the derivative structure should be more complicated. This way the product may be made more suitable to the current market environment. Moreover, if the product were to be realised, the issuer would have a better understanding of the final cost structure, which of course affects the returns gained by the investor. Costs decrease returns, and a product which in theory may seem profitable, may not be so after adding on the variety of costs, such as administration and trading costs.

Furthermore, a study could be done, in an interview form, on the nature and effect of these kinds of costs. Another study could be made on how the liquidity really affects the pricing of options, an issue which became central in this study. In fact, this study opened up a variety of issues, which all could be extended into a study of their own. The author wishes that this study at least brought up, if didn't examine in detail, the most important issues that should be taken into account when studying, constructing, issuing or purchasing index-linked structured products.

## BIBLIOGRAPHY

Anonymous professional in the structured products field. 10.2.2009. Interview (E-Mail).

Bloomberg Online Database. 2008-2009. [Accessed 1.10.2008-12.02.2009]

Case, Karl. E, Shiller, Robert J. 2003. "Is There a Bubble in the Housing Market". Brookings Papers on Economic Activity, Vol. 2003, No. 2, pp. 229-342. The Brookings Institution.

Case, Karl. E, Shiller, Robert J. March 1989. "The Efficiency of the Market for Single Family Homes". The American Economic Review, Vol. 79, No. 1, pp. 125-137. American Economic Association.

Commerzbank. 2008. The "W's" of the Structured Products World. Commerzbank Corporates & Markets (marketing publication).

Diewert W.E. and A.O. Nakamura, editors. 1993. Essays in Index Number Theory, Volume I.. Elsevier Science Publishers.

<<http://www.econ.ubc.ca/diewert/inbook.htm>> [21.10.2008]

Dow Jones Indexes. October 2008. Guide to the Dow Jones Global Indexes. Dow Jones & Company.

<[http://www.djindexes.com/mdsidx/downloads/rulebooks/DJGI\\_Rulebook.pdf](http://www.djindexes.com/mdsidx/downloads/rulebooks/DJGI_Rulebook.pdf)>

[02.11.2008]

Dow Jones Indexes. 2008. Dow Jones U.S. Real Estate Index Factsheet. Dow Jones & Company.

<[http://www.djindexes.com/mdsidx/downloads/fact\\_info/USRealEstate\\_info.pdf](http://www.djindexes.com/mdsidx/downloads/fact_info/USRealEstate_info.pdf)>

[03.11.2008]

Eichhorn, W. March 1976. "Fisher's Tests Revisited". *Econometrica*, Vol. 44, No. 2, pp. 247-256. The Econometric Society.

Fabozzi, Frank J. 1998. Overview in *Handbook of Structured Financial Products*. Ed. Fabozzi, Frank J. Frank J. Fabozzi Associates: New Hope, Pennsylvania.

Gastineau, Gary L.. 2002. *The Exchange Traded Funds Manual*. John Wiley & Sons: New York, USA.

Gorton, Gary B., Pennachi, George G. January 1993. "Security Baskets and Index-Linked Securities". *The Journal of Business*, Vol. 66, No. 1, pp. 1-27. The University of Chicago Press.

Granger, Clive W. J. June 2004. "Time Series Analysis, Cointegration, and Applications": *The American Economic Review*, Vol. 94, No. 3, pp. 421-425. American Economic Association.

Hendry, David F., Juselius Katarina. September 1999. *Explaining Cointegration Analysis: Part I*. Discussion paper. <<http://www.econ.ku.dk/okokj/>> [05.12.2008]

Hendry, David F., Juselius Katarina. November 2000. *Explaining Cointegration Analysis: Part II*. Discussion paper. <<http://www.econ.ku.dk/okokj/>> [05.12.2008]

Himmelberg, Charles, Mayer Christopher and Sinai, Todd. Autumn 2005. "Assessing High House Prices: Bubbles, Fundamentals and Misperceptions". *The Journal of Economic Perspectives* Vol. 19, No. 4, pp. 67-92. The American Economic Association.

Hull, John C. 2006. *Options, Futures and Other Derivatives*. 6<sup>th</sup> ed. Pearson Prentice Hall: Upper Saddle River, New Jersey, USA.

Johansen, Søren. 1988. "Statistical Analysis of Cointegration Vectors". *Journal of Economic Dynamics and Control*, Vol. 12, pp. 231-254. North-Holland.

- Landis, David. October 2006. "Reinventing The Index". Kiplinger's Personal Finance, Vol. 60, Issue 10, pp. 44-47. Kiplinger Washington Editors Inc.
- Nasdaq OMX. 2008. Nasdaq Indexes – Philadelphia Housing Sector Index. <<http://www.nasdaqtrader.com/trader.aspx?id=NASDAQIndexes>> [10.12.2008]
- Nordea Markets. 2008. Indeksilaina USA Kiinteistö – Perus ja Ekstra. Marketing material. Nordea Bank.
- Peng, Scott Y. and Ravi E. Dattatreya. 1995. The Structured Note Market. Irwin Professional Publishing: Chicago, USA.
- Perold, André F. (2007) "Fundamentally Flawed Indexing". Financial Analysts Journal, Vol. 63, No. 6. CFA Institute.
- Samuelson, P.A., Swamy S. September 1974. "Invariant Economic Index Numbers and Canonical Duality: Survey and Synthesis". The American Economic Review, Vol. 64, No. 4, pp. 566-593. American Economic Association.
- Silver, M.S. December 1982. "Appropriate Index Number Formulae for Productivity Measurement at the Plant/Organisational Level". Managerial and Decision Economics, Vol. 3, No. 4, pp. 244-250. John Wiley & Sons.
- Standard & Poors. 2008. S&P/Case-Shiller Home Price Indices Factsheet. <[http://www2.standardandpoors.com/spf/pdf/index/SP\\_CS\\_Home\\_Price\\_Indices\\_Factsheet.pdf](http://www2.standardandpoors.com/spf/pdf/index/SP_CS_Home_Price_Indices_Factsheet.pdf)> [15.12.2008]
- Standard & Poors. 2008. S&P/Case-Shiller Home Price Indices Methodology. <[http://www2.standardandpoors.com/spf/pdf/index/SP\\_CS\\_Home\\_Price\\_Indices\\_Methodology\\_Web.pdf](http://www2.standardandpoors.com/spf/pdf/index/SP_CS_Home_Price_Indices_Methodology_Web.pdf)> [15.12.2008]
- Studenmund, A.H. 2006. Using Econometrics – A practical guide. Pearson Education: Boston, USA.

Säynevirta, Maarit. 1991. The Index Theory and Wage and Salary Index.  
Tilastokeskus Studies Nr. 185. Valtion Painatuskeskus: Helsinki.

The Office of Federal Housing Enterprise Oversight. .July 2008. Mortgage Markets  
and The Enterprises in 2007. <<http://www.ofheo.gov/media/research/MME2007.pdf>>  
[28.11.2008]

## APPENDIX 1

**FISHER'S SEVEN TESTS**

## 1) The unit of measurement test

If the unit of measurement for the quantity of good  $i$  in the index is altered, then naturally the price for the good will also change.  $P(p^1, p^0)$  represents here and also further on a price index for when the prices change from time period 0 to the next time period 1 from  $p^0$  to  $p^1$ . The new price will be  $\beta_i p_i$ .  $\beta_i$  is a diagonal matrix, where  $p^1$  and  $p^0$  are vectors. If the change in the unit of measurement is taken into account in both periods and for all goods affected, the index number should not be affected. The index will then be of the form:

$$P(p^1, p^0) = P(\beta p^1, \beta p^0) \quad (1)$$

## 2) The proportionality test

Here  $\beta$  is a scalar representing a change in prices. If prices change over time so that  $p^1 = \beta p^0$ , the index number will also change in the same proportion. The test here is thus:

$$P(p^1, p^0) = P(\beta p^1, \beta p^0) = \beta \quad (2)$$

## 3) The identity test

If the price of for example the base period is compared to itself, there is no change in prices.

$$P(p^0, p^0) = 1 \quad (3)$$

## 4) The circular test

The use of an intermediate period should not affect the index number. Thus, even if the comparison between the base period  $t^0$  and comparison period, here  $t^2$ , is made

using an intermediate time period  $t^1$ , the resulting index number should be the same as when calculated without the intermediate period. Here is the same represented in mathematical terms:

$$P(p^2, p^1)P(p^1, p^0) = P(p^2, p^0) \quad (4)$$

This test, however, quite often does not hold because adding more time periods leads to an increase in the probability of the weights  $w_i(t)$  changing. When the weights change, the resulting index number changes as well.

#### 5) The time reversal test

If one changes the time periods around, the index becomes its inverse. This can be illustrated more clearly by a mathematical formula:

$$P(p^1, p^0) = 1/P(p^0, p^1) \quad (5)$$

This is one of the two tests which Fisher himself thought to be the most important in determining a good index formula.

#### 6) The factor reversal test

Resulting from the index basics illustrated in appendix 2, the following condition has to hold: Price and quantity indices calculated with the same formula have to fulfil the condition that the product of the price and quantity indices is the total value ratio.

$$P(p^1, p^0) \times Q(q^1, q^0) = V^1/V^0 \quad (6)$$

This is the second test that Fisher underlined as most important tests in determining a good index formula.

7) The determinateness test

Even if one of the prices or quantities in the index approaches either zero or infinity, the index will not approach zero or infinity. In other words, one or few abnormal goods should not affect the index number in a significant magnitude. (Säynevirta 1991)

## APPENDIX 2

**INDEX BASICS**

According to Säynevirta (1991), the starting point for index calculations is a group of goods  $a_1, a_2, a_3, \dots, a_n$ . For each of these goods, the prices,  $p_i^t$  and quantities,  $q_i^t$ , are observed at time periods  $t^0$  and  $t^1$  (Säynevirta 1991, 7). Below are listed some of the basic building blocks of indices.

Group of goods:	$A = \{a_1, a_2, a_3, \dots, a_n\}$
Goods:	$a_1, a_2, a_3, \dots, a_n$
Prices ( $t^0$ ):	$p_1^0, p_2^0, p_3^0, \dots, p_n^0$
( $t^1$ ):	$p_1^1, p_2^1, p_3^1, \dots, p_n^1$
Quantities ( $t^0$ ):	$q_1^0, q_2^0, q_3^0, \dots, q_n^0$
( $t^1$ ):	$q_1^1, q_2^1, q_3^1, \dots, q_n^1$
Price vector ( $t^0$ ):	$p^0 = (p_1^0, p_2^0, p_3^0, \dots, p_n^0)$
( $t^1$ ):	$p^1 = (p_1^1, p_2^1, p_3^1, \dots, p_n^1)$
Quantity vector ( $t^0$ ):	$q^0 = (q_1^0, q_2^0, q_3^0, \dots, q_n^0)$
( $t^1$ ):	$q^1 = (q_1^1, q_2^1, q_3^1, \dots, q_n^1)$

Below are some ratio figures. These are necessary as prices and quantities cannot be added together as they are. Using price- and quantity ratios eliminates the effect of units.

Price ratio:	$\Pi_0^1(i) = p_i^1/p_i^0$
Quantity ratio:	$K_0^1(i) = q_i^1/q_i^0$
Value:	$v = p \times q$
Value weight:	$w_i = v_i/V$
Value ratio:	$v_i^1/v_i^0$
Total value ratio:	$V^1/V^0 = \sum v_i^1/\sum v_i^0$

## APPENDIX 3

**MAJOR NON-FIXED-BASE INDICES**

## 1) Divisia-Törnqvist index

The Divisia-Törnqvist index is the first chain index, an index calculated with the chained method, presented in this list of indices. All prices, quantities and weights are constantly changing functions of time. This quality is well represented by the index formula:

$$\log\left(\frac{P^b}{P^a}\right) = \sum_a^b \int w_i^t d(\log p_i^t) \quad (1)$$

In this formula, the value weight of each good changes over time. The formula may however be simplified if the index number is calculated for shorter periods. (Säynevirta 1991, 27-28) In other words, the whole period, for example a year, may be divided into days. The index number is calculated for day one with average weights for each good for that day. For day two, the average weights for the day are calculated again and they are used for calculating the index number for day two. This simplified version of the index formula may be represented as follows:

$$\log(P^1/P^0) = \sum w_i \log(p_i^1/p_i^0) \quad (2)$$

The Divisia-Törnqvist index is more accurate than the previous fixed base indices exactly because it is not a fixed base index. In Säynevirta's words, the index does not "grow old" as it is constantly adjusting for the new prices as well as quantities (Säynevirta 1991, 28). There are, however, issues that need to be taken into account with this index. First of all, it is not consistent when aggregated. Secondly, it requires a lot more up to date information than the previously presented indices. This may not be such an issue for stock market indices as the information on prices and quantities is real time, but when calculating indices that are based on statistical data, one may have to wait for weeks or months for the necessary information. Also, as the index is calculated with the chained method, an error in one period's index

number affects all the future index numbers, and the error is also cumulated.  
(Säynevirta 1991)

## 2) Vartia I

The extensive research of Yrjö Vartia on index theory in the 1970's lead to him establishing an index called Vartia I. As a starting point, this index has the logarithmic mean, which is similar to arithmetic and geometric means. It is expressed as follows:

$$L(x, y) = \frac{y - x}{\log(y/x)} = \frac{y - x}{\log y - \log x} \quad (3)$$

The aforementioned definition of value is also used to derive the index. The value condition stated that:

$$V = \sum v_i = \sum p_i q_i \quad (4)$$

To derive the Vartia I index, we need to make the following derivations:

$$\log(V^1 / V^0) = \sum \left( \left[ \frac{L(v_i^1, v_i^0)}{L(V^1 / V^0)} \right] \log(v_i^1 / v_i^0) \right) = \sum [w_i \log(v_i^1 / v_i^0)] \quad (5)$$

Since  $\log(v_i^1 / v_i^0) = \log(p_i^1 / p_i^0) + \log(q_i^1 / q_i^0)$ , then :

$$\log(V^1 / V^0) = \log P_0^1 + \log Q_0^1 \quad (\text{where } Q_0^1 \text{ is a quantity index})$$

Finally, the Vartia I price index may be represented as follows:

$$\log P_0^1 = \sum \left[ \frac{L(v_i^1, v_i^0)}{L(V^1, V^0)} \right] \log(p_i^1 / p_i^0) = \sum w_i \log(p_i^1 / p_i^0) \quad (6)$$

Usually the index weights,  $w_i$ , sum up to 1 for all the goods,  $a_i$ . This is natural since in the fixed basket indices the weights used are constant. For the Vartia I index, however, the sum of the weights is 1 if and only if:

$$w_i^0 = w_i^1 \text{ for all goods } a_i \quad (7)$$

If and when even one of the goods weight in the basket changes when the time period changes, which is highly realistic, the sum of the weights for all the goods is either more or less than one. (Säynevirta 1991)

### 3) Sato-Vartia

Vartia I index's largest problem is the fact that the weights used for calculation do not add up to one. This problem can be fixed, however, with using a different formulation for calculating the weights, as suggested by economist Henri Theil (Säynevirta 1991, 30) In this calculation, the weights are calculated via the logarithmic averages for each good, which are divided by the sum of all logarithmic averages:

$$w_i = \frac{L(w_i^1, w_i^0)}{\sum L(w_i^1, w_i^0)} \quad \text{and} \quad \sum w_i = 1 \quad (8)$$

The Sato-Vartia index formula otherwise takes the same form as the Vartia I formula:

$$\log P_0^1 = \sum \left[ \frac{L(w_i^1, w_i^0)}{\sum L(w_i^1, w_i^0)} \right] \log(p_i^1/p_i^0) = \sum w_i \log(p_i^1/p_i^0) \quad (9)$$

With this change, one might assume that this makes the Sato-Vartia index better than the Vartia I index. Säynevirta lists, however, that these both indices satisfy the same Fisher tests, the time and factor reversal tests. There are also differences between these two indices: Vartia I is consistent when aggregated, Sato-Vartia is not, and on the other hand, Vartia I does not satisfy the proportionality test while Sato-Vartia does. Therefore, it is difficult to distinguish between the two, especially when they give approximately the same values when used with data. Similarly to the Divisia-Törnqvist index, these two indices require up to date data. (Säynevirta 1991, 30-31)